Four Essays in Applied Microeconomics

Inaugural-Dissertation

zur Erlangung des Grades eines Doktors
der Wirtschafts- und Gesellschaftswissenschaften

durch

die Rechts- und Staatswissenschaftliche Fakultät der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Radost Holler

aus Tübingen

Bonn

2023

Dekan:	Prof. Dr. Jürgen von Hagen
Erstreferent:	Prof. Dr. Thomas Dohmen
Zweitreferent:	Prof. Dr. Sebastian Kube
Tag der mündlichen Prüfung:	05.05.2023

Acknowledgements

I am very grateful for my main advisor Thomas Dohmen. We now know each other for more than seven years and he always supported my academic development. He supported me despite our first or second interaction being me complaining via E-Mail about not having enough paper in his exam (and some other let's call it atypical behavior that I do not need to elaborate) which shows his kind and forgiving character – a true rarety in this profession. He further greatly influenced how I think about certain problems – more than he probably knows. I always hear his voice in my head when I think and discuss the relationship how economic preferences are shaped by the world and shape the world.

I am also very grateful for the support my second advisor Sebastian Kube and Hans-Martin von Gaudecker granted me. I always felt unconditionally supported by both of you, which is more important than you may realize in a world that is full of ambition and competition and where I often felt anxious and discouraged.

This thesis would have definitely not been possible without my co-authors. I thank them, in particular, for enduring my, let's call it bluntness, and my general conviction that I am right. For that I have to particularly thank Lenard Simon who never more than eye-rolled while enduring my condescention.

I am grateful to Lena Janys, who, first of all, continuously provided excellent econometric advice, and second of all is just very fun to hang out with. Susann Adloff deserves a special place in this acknowledgements – we shared it all Bachelor, Master, PhD. Although we were in places for the PhD, I always felt we are still doing this together. Thanks for your incredibly helpful and thoughtful comments, your positive spirit and creative mind, and your unmatched sense of humor. In addtion, thank you goes out to Pamela Qendrai and Lucie Stoppok. We laughed, we cried and we drank a lot of coffee. Thank you for your support!

I am also grateful to Simone Jost, Holger Gerhardt, Janos Gabler, and Tobias Raabe who always were ready to provide help with anything I needed, but particularly with advise on how to solve technical problems in Latex (Holger), in Python (Janos, Tobias) and Simone with basically everything else.

I also would like to thank the Excellence-Cluster Econtribute and everyone who contributed to that for the very conducive atmosphere it provides to us young researchers. Last but not least, I want to thank my colleagues at BGSE, at the Cluster, and in the Macro History Group (in particular Cathrin Mohr and Chi Hyun Kim) for making this time fun and for giving feedback and advise whenever I needed it.

Contents

List of Figures

List of Tables

Introduction

Culture – as defined by the collection of beliefs, norms and preferences shared within social groups – interacts in complex ways with social and economic outcomes. It, both, shapes the economic and social environment of societies as well as adapts as a consequence to changes in it, such as economic and environmental shocks as well as social policy.

In this thesis, I explore the complex relationship between culture, the social environment, policies, and economic shocks. I investigate what determines specific aspects of culture, how government policy can enhance potentially desirable cultural traits, how culture adapts to economic shocks and how this, in turn, translates into economic and social outcomes.

The thesis consists of four independent chapters. Chapter 1 (joint work with Paul Schäfer) can be seen as foundational work on the relationship of culture as measured by community-level social norms and the economic and social environment. We present a newly composed data set on local social norms containing restrictions for women after giving birth as well as obligations to mutual assistance among neighbors covering around 16,000 German-speaking localities in Central Europe. We, first, show that communities can have vastly different norms, even if they are located in a small distance to each other. We, then, set out to explain this intra-regional heterogeneity by investigating the relationship between social norms and locally varying factors such as political institutional history, religious composition, population size and remoteness. We find that localities do not generally become more similar in social norms if they share institutional history, but that they become more similar in norms of cooperation which are conceptually more tied to political institutions than our other social norm measures. Community-characteristics, such as, population size, local population density, and religious heterogeneity are all joint predictors of social norms. We argue that our results align with theories that explain how social norms vary based on a community's ability to enforce those norms and the value they provide to community members.

Chapter 2 (joint work with Thomas Dohmen and Uwe Sunde) zooms in on another aspect of culture, namely economic preferences and, more particularly, patience. Patience measures the willingness to trade-off immediate benefits for future returns and, thus, governs almost every economic-decision making. In their foundational study, Falk, Becker, Dohmen, Enke, Huffman, et al. (2018) show that patience displays large cross-country and cross-regional variation. Further, Sunde, Dohmen, Enke, Falk, Huffman, et al. (2022) show that variation in patience at the aggregate level has a large impact on the accumulation of human capital, physical capital, and the stock of knowledge and, thereby, positively influences economic development. In our study, we adress the question whether governments can influence the patience of their constituents by increasing their level of education through compulsory schooling reforms. We, thereby, also adress the long-standing question of whether there is a reverse causality between education and patience, which has been theorized but not coherently tested (see e.g. Becker and Mulligan, 1997; Cutler and Lleras-Muney, 2006; Oreopoulos and Salvanes, 2011). The analysis combines data on patience from the Global Preferences Survey (GPS) with newly constructed data on changes in compulsory schooling laws between 1947 and 2003 in 48 countries around the globe. Using within-country variation in compulsory years of schooling, we find that respondents who were affected by a schooling reform exhibit a significantly higher level of patience than respondents who were not subject to the same reform. The effect is sizable but exhibits substantial heterogeneity and is mainly driven by reforms that target secondary education.

Chapter 3 (joint work with Christian Zimpelmann, Hans-Martin von Gaudecker, Lena Janys and Bettina Siflinger) and Chapter 4 (joint work with Hans-Martin von Gaudecker, Lenard Simon and Christian Zimpelmann) focus on, how, the largest economic and social shock in recent times – the CoVid-19 pandemic – affected economic and family life. Chapter 3, among other things, sets the scene for chapter 4, by showing how the CoVid-19 pandemic affected work culture by shifting the norm to (more) remote work instead of work at the workplace in jobs in which this is possible. Chapter 4 exploits this shift in remote work to show how increases in job flexibility affect the gendered pattern in parental division of labor. We, thereby, suggest avenues for how to make the division of labor between parents more equal.

Using a customized panel data set, Chapter 3, additionally, investigates whether the CoVid-19 pandemic exacerbated pre-existing inequalities in the Netherlands and how this interacts with government policies. We show that socioeconomic status is strongly related to changes in working hours, especially when strict economic restrictions are in place. In contrast, household income is equally unaffected for all socioeconomic groups. Examining the drivers of these observations, we find that pandemic-specific job characteristics (the ability to work from home and essential worker status) help explain the socioeconomic gradient in total working hours. Household income is largely decoupled from shocks to working hours for employees. We provide suggestive evidence that large-scale labor hoarding schemes have helped insure employees against shocks to their employers. We argue that this may explain the large differences in the socio-economic gradient in the effects of the CoVid-19 pandemic across countries.

In chapter 4, we analyze how the parents of young children react to the shift to remote using representative panel data from the Netherlands spanning four waves from 2019 to 2021 combined with administrative records on working hours from 2013 to 2021. We find that parents with large potential to work remotely strongly increase their remote work hours and reduced their commuting hours. We further show that the remote work potential among fathers and mothers is asymmetrically distributed. Fathers have much more potential to work remotely. Thus, they asymmetrically gain more job flexibility through the shift to remote work than mothers. We then turn to the question of how this affects the division of labor within households. We show that parents who can work from home increase their childcare provision in response to the shift to remote work. Given the asymmetric distribution of the potential to work from home, this leads to a decrease in the gender gap in childcare provision – increasing the childcare provision of fathers while decreasing the provsion of mothers. Using large scale administrative data coupled with an event-study difference-in-differences design, we evaluate whether the shift in remote work also affected the labor supply of parents. We find that, indeed, parents whose partners profit more from the shift to remote work, increase their labor supply in response to this shift. This study shows that residual gender differences in the division of labor are not just caused by gender-related preferences or social norms, but also by the difficulties of combining full-time jobs with childcare needs. This often leads to the male breadwinner model after the birth of the first child, where the father works full-time and the mother opts for no or part-time work.

References

Becker, Gary S, and Casey B Mulligan. 1997. “The Endogenous Determination of Time Preference.” *Quarterly Journal of Economics* 112 (3): [2]

Cutler, David M, and Adriana Lleras-Muney. 2006. “Education and Health: Evaluating Theories and Evidence.” *NBER Working Paper 12352*, [2]

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde. 2018. “Global Evidence on Economic Preferences.” *Quarterly Journal of Economics* 133 (4): 1645–92. [2]

Oreopoulos, Philip, and Kjell G. Salvanes. 2011. “Priceless: The nonpecuniary benefits of schooling.” *Journal of Economic Perspectives* 25 (1): 159–84. [2]

Sunde, Uwe, Thomas Dohmen, Benjamin Enke, Armin Falk, David Huffman, and Gerrit Meyerheim. 2022. “Patience and Comparative Development.” *Review of Economic Studies* 89 (5): 2806–40. [2]

Chapter 1

Local Variation in Social Norms: Evidence From a Large-Scale Historical Survey of German-Speaking Localities

Joint with Paul Schäfer

1.1 Introduction

Economic and social science research has shown that social norms[1] are an important factor in explaining cross-cultural differences in economic and political outcomes.[2] However, limited data availability has partly prevented researchers from exploring the determinants of variation in norms within cultural groups, along, for example, urbanization and community size, market access, religious denomination or historic political borders, as existing sources have little within-culture coverage. To overcome this issue, we present a newly digitized data set. This data set contains information on particular norms concerning gender and cooperation for up to 16,500 Central

* We thank Georg Kehren for providing us with the data from his Doctoral Dissertation. We are grateful for the valuable research assistance of Lena Michaelis, and Luis Wardenbach. Part of this project is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2126/1–390838866 and the Economic History Association through the Exploratory Data and Travel Grant.

1. There are multiple ways to define social norms (see Legros and Cislaghi, 2020, for a cross-disciplinary review). We view social norms as informal standards of behavior within a community to which individuals in a community conform even in the presence of deviating incentives on the individual level. This definition is similar to Burke and Young (2011) and Bicchieri, Muldoon, and Sontuoso (2018). As opposed to suggestions of Bicchieri (2006) and Bicchieri, Muldoon, and Sontuoso (2018), we will not distinguish between expected and actual conformity because we cannot disentangle them empirically. However, we will implicitly assume that a norm's existence also implies that it is adhered to at least to some degree.

2. See, for instance, McCloskey (1991), Gelfand, Raver, Nishii, Leslie, Lun, et al. (2011), Alesina, Giuliano, and Nathan (2013), Gelfand, Harrington, and Jackson (2017), Buggle (2020), Jackson, Gelfand, and Ember (2020), and Buggle and Durante (2021).

European, German speaking localities around 1930s and is based on the German Ethnographic Atlas (GEA)[3].

The GEA contains two different social norms. First, it contains cooperation norms in the form of the obligation to mutual assistance among neighbors also including in which domain of private and economic life these are practiced. Second, childbed norms which contain rules for women in the weeks subsequent to giving birth. The latter cover a large range of behaviors. We, thus, focus on the two most dominant categories: rules that confine the mother to her property, house or apartment (norms of seclusion), rules that are related to the belief that contact to the respective woman harms others because she is impure (impurity norms).

We start by showing that these norms were highly prevalent within the sampling region. More than two thirds of localities have at least one childbed norm and more than half display any type of neighborhood obligation. Exploring the spatial distribution of these norms, we find that these norms exist everywhere within our sampling region. There, however, exist localized clusters in which they are almost universal.

We then set out to explain this spatial pattern in social norms by exploring their relationship to political institutions, population size, remoteness of the locality, and religious denomination. We find that the spatial pattern is not well aligned with political and administrative boundaries around the 1930s. Using the data of Abramson (2017) and Abramson, Carter, and Ying (2022), we then examine their relationship to historical political institutions between 1200 and 1790 exploiting the constantly changing borders within the sampling region during this time span. Using a cultural dyadic gravity equation, we find that indeed localities are more similar in their social norms when they have spent a longer time together in the same state in the past. This effect is, however, almost exclusively driven by the similarity in neighborhood help obligations. Childbed norms seem to be unrelated to historical political institutions.

Population size of the localities varies strongly even within regions. It, thus, may also explain the vast intra-regional variations in norms. Indeed, we find that localities with larger populations are less likely to exhibit childbed norms or reciprocal assistance among neighbors even when controlling for various levels of regional fixed effects (province, district, county). The relationship between remoteness as measured by distance to the nearest town and prevalence of social norms is more complex. Localities that are closer to towns or that are towns themselves are less likely to display neighborhood obligations. The effect of remoteness on childbed norms is less clear. In tendency, localities that are closer to towns are less likely to display childbed norms but the effect is not completely robust across specifications indicating a weaker relationship than the relationship to neighborhood obligations.

3. German: Atlas der deutschen Volkskunde

In our sampling region, most localities have either almost only Catholic inhabitants, almost only Protestant inhabitants or some mixture of both while other religious groups are rarely present. In terms of religious denomination, we find that, in particular, denominational heterogeneity plays a large role in explaining the prevalence of social norms. Comparing localities with predominantly Protestant inhabitants with predominantly Catholic localities, we only find a strong gradient for norms of seclusion with predominantly Catholic localities being much more likely to display these rules.

Our results are related to the literature on the effect of historical political institutions on cultural traits. This literature argues that political institutions have a profound impact on culture through the provision of formal rules and their implementation, as well as the distribution of political power, which shape incentives and experiences of individuals and groups in society (see e.g. Tabellini, 2008; Grosjean, 2011; Alesina and Giuliano, 2015; Becker, Boeckh, Hainz, and Woessmann, 2016; Buggle, 2016). The empirical literature on this topic primarily focuses on the effect of institutions on cultural traits related to cooperation, such as social and political trust (Grosjean, 2011; Becker, Boeckh, et al., 2016; Buggle, 2016), civic attitudes (see e.g. Guiso, Sapienza, and Zingales, 2016), or preferences for redistribution (see e.g. Alesina and Fuchs-Schündeln, 2007). However, there is limited research on the effect of institutions on social norms in general. Our findings suggest that not all social norms are equally influenced by political institutions. Instead, norms related to cooperation tend to be more sensitive to historical institutions. This may be because reciprocal exchange is more easily influenced by public governance, such as the provision of public goods, or mechanisms for social insurance, and therefore depend more on the institutional structure than other social norms such as in our case, for instance, rules for women after giving birth which concern more the private life of individuals and may be more closely tied to religious beliefs and intra-generationally passed down superstitions.

Further, our results are related to the work of Henrich, Boyd, Bowles, Camerer, Fehr, et al. (2001) and Henrich, Ensminger, McElreath, Barr, Barrett, et al. (2010) who study the evolution of cooperation in societies. They argue that societies relying on markets and building larger communities (towns, cities, larger settlements) coevolved with the ability to sustain cooperation in anonymous interactions. Studying variation across small-scale societies, they find that market integration or market interaction is positively associated with high levels of cooperation and fairness in anonymous games, while the willingness to engage in third party punishment increases with community size. Using being close to a town or being a town as a proxy for market access and the population size of localities, we complement this research by suggesting the opposite effect for norms of cooperation among neighbors. We, thereby, provide first evidence of an inverse relationship of community size and market access for cooperative behavior within existing social groups. We will argue that this is partly driven by the substitutability between markets and reciprocal exchange and partly driven by inducing a change in the social structure of localities.

Additionally, a large empirical literature investigates the effect of the German Reformation and Protestantism (vs. Catholicism) on economic and social outcomes (see Becker, Rubin, and Woessmann, 2021, for a comprehensive literature review). This literature causally examines the effect of Protestantism and the Reformation on literacy (Becker and Woessmann, 2009), economic growth (Cantoni, 2015), sectoral reallocation (Cantoni, Dittmar, and Yuchtman, 2016) or even more social outcomes such as suicides (Becker and Woessmann, 2018). As opposed to Becker and Woessmann (2018), we do not find a consistent negative effect of Protestantism on social cohesion. Instead, we find that religious homogeneity (Catholic or Protestant) is most conducive for the existence of social norms in a locality. We, thus, contribute to this literature by highlighting that the reformation has not only affected society within Germany or Prussia through replacing one denomination by the other, but also by making communities more heterogeneous in their religious beliefs and providing a social cleavage which may itself affect economic and social outcomes in affected regions. In the discussion, we also show how this finding relates to the commonly used distance to Wittenberg-instrument (see e.g. Becker and Woessmann, 2008; Becker and Woessmann, 2009, 2018). In particular, we argue that our results indicate that this may not be a suitable instrument for outcomes that are also plausibly related to denominational heterogeneity.

We also contribute to the literature on cultural tightness, which studies the prevalence of norms across domains (Gelfand, Raver, et al., 2011; Gelfand, Harrington, and Jackson, 2017; Jackson, Gelfand, and Ember, 2020). Tight cultures have a higher prevalence of norms and higher levels of conformity. We complement this literature by investigating intra-cultural variation in the prevalence of social norms. Our results suggest that cultural tightness and its determinants may vary very locally, i.e. communities directly next to each other are different in the prevalence of social norms.

Finally, our data addresses a lack of data lamented in the literature on collective action and ethnographic data in historical economics (Poteete and Ostrom, 2008; Lowes, 2020). Lowes (2020) notes that while ethnographic datasets can be very useful for economic historians, currently available datasets have several shortcomings. Existing datasets are compiled from many ethnographies that might use differing definitions. These data sets' patchwork nature makes the resulting data less systematic and hides variation within pre-defined cultural boundaries. Such data sets also include very few European data-points. The data, we digitized, contributes towards filling those gaps. In particular, we complement Satyanath, Voigtländer, and Voth (2017) who digitized membership numbers in associations for certain German cities in the 1930s by providing a new measure of historical social capital in the form of neighborhood help obligations for more rural localities.

This paper is structured as follows: Section 2 gives an overview over the GEA and its social norm questions in particular, Section 3 explores the spatial distribution of the social norms, Section 4 examines their relationship to (historical) political institutions, Section 5 their relationship with population size and remoteness, Section 6 their relationship to religious denomination. Section 7 discusses potential mechanisms and the implications of our findings with respect to the previously named literatures.

1.2 Data

We present a newly digitized data set containing the results of the German Ethnographic Atlas (GEA) collected by anthropologists between 1930 and 1933 as well as 1935. It contains a different set of questions each year of collection. The aim of the GEA was to capture rural culture before its transformation caused by industrialization (Schmoll, 2009, p. 236-238). The question, thus, do not ask about individual behavior but about what is customary or commonly done in the specific locality.

The topics of the survey range from agricultural production, food, festivities, folklore, religious and profane rituals, to marriage customs, and social norms. In total, the GEA contains 243 questions of which two questions are related to social norms and are presented in this paper: neighborhood obligations, and restrictions for women after birth. Both of these questions where asked in questionnaire four in 1933[4] and religious composition of the locality was asked in every questionnaire between 1930 and 1933. In addition, in every questionnaire respondents were asked to indicate the religious composition of the locality, which we additionally digitized.[5] For each sample village, researchers recruited volunteers to fill out the questionnaires for one or multiple localities (Kehren, 1994).[6]

The target population of the GEA consists of German speaking localities that have at least one school. Localities vary strongly in size and type. We were able to merge a large part of localities to the registry of communities of the German Reich of 1910. According to the registry, 91% are rural communities (Landgemeinden, Gutsbezirke or similar), 2% are so-called 'Marktgemeinden', and 7% are towns or cities or part of towns and cities.[7] The data was mainly collected within the German Reich in its inter-war borders (1919-1939) including the Saar Region. On top, localities were also sampled from Gdansk, the Czech part of Czechoslovakia, Luxembourg,

4. For an indepth treatment of the relationship of the GEA and the Nazi Regime, consult Schmoll (2009).

5. The process and sources of digitization are detailed in Appendix 1.C.

6. We display characteristics of respondents in the Rhine Province digitized by Kehren (1994) in Appendix 1.C.4.

7. Note that those not matched to the registry are most likely even smaller localities that are part of rural communities or new settlements and thus are not listed in the registry.

Liechtenstein, the First Austrian Republic and Poland as well as more fragmented attempts of data collection in the German-speaking enclaves of Slovakia, Transylvania, Bessarabia, Banat, Lorraine, Klaipėda Region, Switzerland, Belgium, Hungary and Denmark. The full sample contains more than 20,000 localities, however, not every locality answered every questionnaire. We only include localities if they answered either social norms question.[8] As we attempt to analyze spatial variation, we exclude data points that are geographically comparably isolated because of the fragmented sampling within the region. In consequence, we only include answers from the German Reich, the first Austrian Republic, the Czech part of Czechoslovakia, Gdansk and the border regions of Poland in this study and that were included in the 1933 sample. This leaves us with upto 16,800 observations.

8. Some questionnaires were only asked in some regions. Switzerland and Lorraine were only included in the questionnaire 1, Luxembourg only in the questionnaire 1 and 2. Border region with Denmark is only available in questionnaire four.

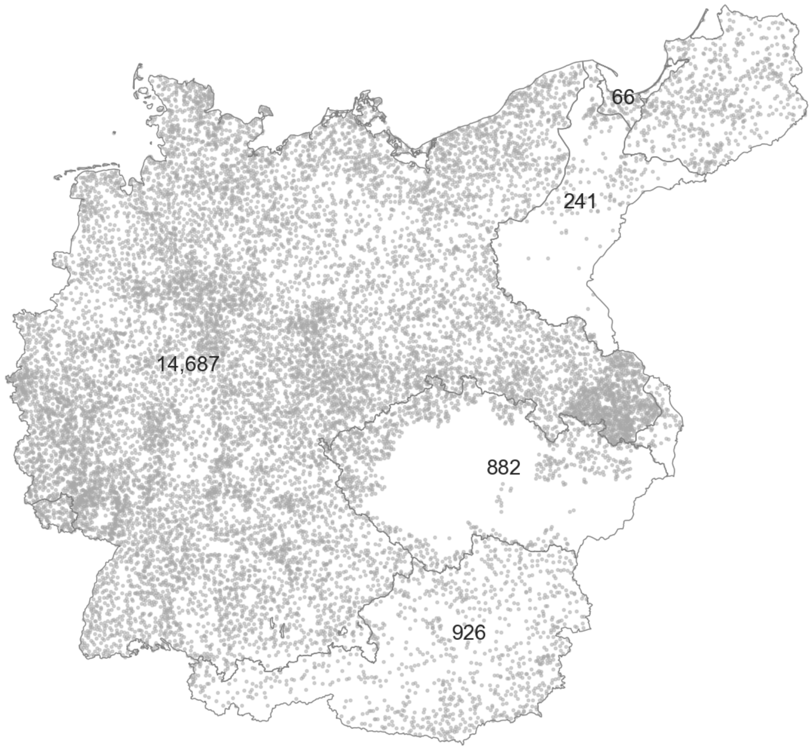


Figure 1.2.1. Localities in the sample.

Notes: Sample restricted to localities that answered any of the social norms question and lie in either the German Reich, Saar Region, Austria, the Czech part of Czechoslovakia, Gdansk and the border region of Poland. Borders represent administrative boundaries in Europe, 1930 obtained from MPIDR and CGG (2013). Numbers represent numbers of observations per country.

Each locality is (uniquely) represented by a coordinate that defines the spatial position of the locality on a custom designed grid. We digitally reconstructed this grid to geocode the localities (see Appendix 1.C for more details).[9] Figure 1.2.1 displays the resulting geographic distribution. The 14,687 localities that are located within the German Reich (including the Saar Region) are densely distributed over its full area. The second largest number of observations falls with 926 on Austria which is less densely but still fully covered. 882 localities are located in Czechia. These localities are mostly concentrated in the border region to the German Reich,

9. We additionally improve the precision of the spatial location by comparing it to the historical locality database of the Association of Computer Genealogy available under http://gov.genealogy.net which contains the location of a variety of historical localities, communities and towns.

which corresponds to the area with a large or often even a majority share of German-speaking population (see Figure 1.A.1 in Appendix). The localities in the center of Czechia are distributed around Prague. Less densely covered is the border region of Poland which at least partly reflects the timing of the data collection and its close temporal proximity to World War I. According to Schmoll (2009), the data collection was only possible covertly in Poland. With 66 localities, Gdansk is well represented relative to its area and population size.

1.2.1 Measuring Norms

The GEA contain two questions regarding social norms: neighborhood obligations and restrictions for women after giving birth.

Cooperation norms: Neighborhood help obligations

The cooperation norm contained in the GEA concerns the obligation to mutual assistance among neighbors.[10] The question in the GEA is phrased in the following manner:

a) In your village, are neighbors traditionally obligated to mutual assistance?

b) At which occasions of family life, like childbirth (e.g. care for women in childbed), weddings (e.g. help in the kitchen), illness (e.g. night watch), death (carrying the coffin, digging the grave)?

c) For which economic tasks, like harvests, building a house (transporting wood) etc.?[11]

The level of detail of answers to b) and c) varies greatly between respondents. Some respondents just affirm or deny, some only name the event at which the neighborhood help activity is performed (e.g. help at wedding), some detail the exact activity that is performed or underline the activities in the example on the questionnaire if they apply (e.g. help in the kitchen), and a few write small essay like answers (see Barruzi-Leicher and Frauenknecht, 1966, for detailed examples). All major coding decisions are explained in Appendix 1.C.1. Given the varying level of detail, we code answers to b) and c) into five meta categories, that are structured around the major events named in the question, – help upon death of a neighbor, help at a neighbor's wedding, help after birth, help during sickness, help with building a house, help with agricultural activities, help in emergencies such as fire and

10. Note that neighborhood help as a common cooperative activity of historic village communities is also documented in Weber (1922), Kramer (1954), and Wurzbacher (1961).

11. In the German original the survey question is given by: a) Sind in ihrem Ort die Nachbarn noch von alters her zu gegenseitiger Hilfeleistung verpflichtet? b) Bei welchen Anlässen des Familienlebens , wie Geburt (z.B. Pflege der Wöchnerin), Hochzeit (z.B. Hilfe in der Küche), Krankheit (z.B. Nachtwache), Tod (Tragen des Sargs, Graben des Grabs)? c) Bei welchen wirtschaftlichen Arbeiten, wie Ernte, Hausbau (Anfahren des Bauholz) usw.? (Zender and Wiegelmann, 1959, p. 30)

flood – and one residual category.[12] All of these categories are coded as indicator variables if help at this event or a related activity is mentioned. That is, for instance, transporting material for building a house is coded as help with house building. Most respondents, that name specific activities, name a subset of activities of the examples in the question. If additional events or activities are named, they typically fall either in the agricultural domain (help with the livestock, in particular birthing of calves, help flailing etc) or concern help at other family festivities, such as baptism or confirmation festivities. The latter is coded in our residual category 'other help' For a detailed account of recorded subcategories and coding choices consult Appendix 1.C.1.

To reduce the number of outcome variables, we use our main outcomes to the answer to question a) – whether there is a general obligation to help the neighbors – as well as the sum of events at which neighborhood help is performed to measure the intensity of neighborhood obligations, which measures the intensity of neighborhood help, as out main outcomes. Note that these two variables do not perfectly align. That is, there are respondents that affirm the existence of a general obligation, and do not name any specific activities, and there are cases in which respondents name activities but are very explicit that these are performed (commonly) but voluntarily. These differences may reflect the different perception of what respondents consider 'obligatory'. For instance, the respondent of Liegetrocken (East Prussia) mentions that it is obligatory out of commandment of Christian charity[13] but does not name any specific activity, hinting more towards an abstract moral obligation. The respondent of Reesen (Prussia-Saxony) denies an obligation by stating: "One helps in all cases without being obligated by law"[14]. Thus, the respondent implies that it would only be obligatory if it was a written law. The respondent of Freiberg (Saxony), however, affirms the obligation but notes that neighborhood help is morally obligatory and not obligatory by law. Given these differences in perception, we do not code the activities differently by whether they are considered obligatory or voluntarily by the respondent.

Table 1.2.1 displays the summary statistics of the aggregate measures as well as the subcategories. 41% of respondents indicate that neighbors are traditionally obligated to mutual assistance, while 64% perform any activity some of which respondents do not consider obligatory. Looking at the sum of neighborhood obligations reveals that neighborhood obligations typically span from one to two domains. Among the specific neighborhood activities, activities surrounding death are with 51% most common. 31% of respondents named help at wedding, 24% help after

12. With this categorization, we largely follow Barruzi-Leicher and Frauenknecht (1966) but additionally code agricultural activities, and fully code help during sickness, and after birth, which is only partially published on their maps.

13. German: "[...] Gebot christlicher Nächstenliebe"

14. German: "Man hilft in allen Fällen ohne dazu gesetzlich verpflichtet zu sein."

birth, and 28% help upon a neighbors sickness as a neighborhood help activity. Among the economic activities, help with building a house is with 39% most frequent. In 18% of localities, neighbors help each other with agricultural tasks and in 7% in case of emergencies. 2% of localities name some other activities that do not fall in any of the previous categories.

Table 1.2.1. Summary statistics: norms

	N	Share	Past	Mean	Std	p25	p50	p75	Max
Neighborhood obligations									
Obligated	15,999	0.41	0.01						
Any nb act	15,999	0.62	0.04						
Sum nb help	15,999			2.00	2.00	0	1	4	8
Type of Help									
Nb help at death	15,999	0.51							
Nb help at wedding	15,999	0.31							
Nb help at birth	15,999	0.24							
Nb help when sick	15,999	0.28							
Nb help with house building	15,999	0.39							
Nb help with agriculture	15,999	0.18							
Nb help in emergencies	15,999	0.07							
Nb other help	15,999	0.02							
Childbed norms									
Any rule (mother)	15,785	0.78	0.21						
Seclusion	15,785	0.61							
Impurity	15,785	0.30							
Only other rule	15,785	0.05							

Note: For categorical and dichotomous variables, table displays shares. For continuous variables mean, standard deviation, 25th, 50th, 75th percentile and maximum are reported. Past: share of localities where norm does not apply but it applied in the past. Childbed norms have a lower number of observation because part of the data was destroyed or deemed unreadable in the war (see also Grober-Glück (1966b)). In addition, a larger part of answers could not be categorized because their meaning was ambiguous, for more details on the coding and digitization see Appendix 1.C.2. Sum nb help = Number of neighborhood help obligations; Any rule (mother) is an indicator variable that is 1 if there is any type of rule that applies for the woman after birth. Childbed norms are divided into two subcategories 'seclusion' and 'impurity'. Seclusion: do not leave the house, do not go over the street, do not pass the roof border of your house, stay in your living area, do not go anywhere. Impurity rules are rules that are associated with a notion of impurity of women after birth. The column *Past* contains the share of localities which had the norm in the past but do not have it today anymore. Samples do not fully overlap.

Gender Norms: Restrictions for women after birth

The GEA asks about norms that apply to women in the weeks following birth – also called the woman in childbed. The birth of a child used to be surrounded with behavioral rules for the new mothers in Europe (see e.g. Ploss, 1876; Nowottnick, 1935; Labouvie, 1992). The GEA enables the first quantitative assessment of the prevalence of these norms. The question reads as follows:

> *a) Where is the woman in childbed not allowed to go before her first churchgoing? (e.g. basement, attic, barn, well, neighbor)*
>
> *b) Which boundary is she not allowed to pass? (e.g. gutter, street, crossroad, village border)*
>
> *c) Which other traditional precautions does the women in childbed follow?*[15]

The question additionally asks to give a date if rules do not apply anymore but have applied in the past. As before answers vary largely in the level of details provided. In general the answer structure is, however, a bit more complex than in the previous question. In particular, the variation in the type of restrictions named is large, yielding us to code the answers in 79 categories and one residual category in which we group all restrictions that are named to rarely (less than 5 times) to form an additional category. All categories are described in Appendix 1.C.2. [16]

To have a bit more homogeneity on the types of norms covered by these variables, we focus on behavioral restrictions after birth, and thus do not consider rules that only apply on the way of the first churchgoing, means of protection against bad spirits, bad luck or similar, such as sprinkling holy water, putting a wholy book under the pillow etc., and rules that do not focus on the mother herself but on someone or something else. An example of such a rule that we exclude is 'not hanging the children cloth outside'. We, thus, solely focus on rules that restrict the mother's behavior in the period after birth.

The, by far, most prevalent rules confine the woman to her home or the area around her home. 41% of respondents state that the women are restricted to their

15. Original German: a) Wohin darf die Wöchnerin vor dem ersten Kirchgang (Aussegnung) nicht gehen? (z.B. Keller, Boden, Stall, Brunnen, Nachbar) b) Welche Grenze darf sie nicht überschreiten? (z.B. Dachtraufe, Gosse, Straße, Kreuzweg, Dorfgrenze) c) Welche besonderen altherkömmlichen Vorsichtsmaßnahmen beachtet die Wöchnerin sonst vor dem ersten Kirchgang?

16. In addition, some respondents indicate one or more of the following dimensions (1) how long after birth the rule applies, (2) whether the rule only applies at specific times (of day), such as after sundown, (3) whether the rule is adhered to only 'if possible' or similar, (4) whether the rule is only followed by some, (5) the reasoning behind a rule, typically in form of a consequence that happens if the woman violates the rule, (6) whether there are means of protection that enable women to violate the rule without consequences (e.g. covering their head when leaving the house, putting salt in the well etc.). We resolve some of this complexity by defining a restriction to only apply if it still applies to all women in the locality, it applies at all times (of day), and there is no means to protect against consequences upon violation. Note that this decision is not particularly consequential as the number of cases in which this additional information is given is typically below 1%.

property, house, appartment, 19.5% indicate that she shall not cross any streets, roads, gutters inherently restricting her mobility to her own home or its surroundings, 9% indicate that she shall not pass her roof's eave confining her to her house, 1.5% indicate that she shall not go 'anywhere', and another 1% confine her to stay in specific rooms of her house or apartment (mostly her parlor). The second most prevelant rules restrict a woman's social interactions. Most prominently, in 15% of localities, she is not allowed to visit her neighbor, and 5% restrict any kind of visit or/and receiving visits. 0.5% of localities mix seclusion and social rules by restricting women's public appearance ('the woman should not be seen publicly') – so she can go out as long as no one sees her. Another very common rule, prohibits the woman to get water at the well or from the water pump (9%).

The length of application of the restriction after birth mostly depends on the time of the first churchgoing of the mother, which typically is two to six weeks after giving birth (Grober-Glück, 1985).[17]

It is difficult to judge whether these rules are bad, good or neutral for the mother. Being secluded after birth may protect mothers from having to work too early after birth and thus may protect her health (see also Grober-Glück, 1966b). Thus, they may constitute some form of early maternity leave in times where there was no legal protection. On the other hand, if it was only about protecting a mother's health, then localities may have implemented restrictions that target work and health more explicitly. However, only approximately 2% explicitly reference restrictions to protect a mother's health, and 2% explicitly prohibit working or performing specific kinds of work, such as heavy lifting (see Table 1.C.2 for details).

Some respondents indicate the believed consequence a violation of a restriction yields. If the woman goes to the neighbor, the neighbor, for instance, 'gets rats' (in Bellin, Mecklenburg), the neighbor house 'burns down' (in Wilkau, Niederschlesien) or gets struck by lightening (in Nagel, Bavaria). In Lindenkreuz, Thuringia, the woman brings bad luck everywhere she goes, and in Rosenthal I, Bohemia, the neighbor throws a broom after her. If she went to the well in Wellheim, Bavaria, lice grow in it. If she went in Langenwetzendorf, Thuringia, the well rots. If she leaves her child alone in Krompusch, Niederschlesien, or in Marienberg, Bohemia, it gets exchanged for a changeling. If she goes to a dance floor in Giesmannsdorf, Niederschlesien, the devil gets her and in Stüde, Hannover, she does not show herself publicly so that she will not be bewitched. The most extreme consequences named are physical harm of the woman as punishment (rarely named). For instance in Etterwinden, Thuringia, the woman should not cross the street in front of a moving cart, otherwise the coach-

17. Grober-Glück (1985) shows that the length depends on the region as well as on the denomination. Catholics tend to be more likely to hold the first churchgoing after two weeks, Protestants after six. The length tends to be longer for both denominations in the Northern half as opposed to the Southern half of the sampling region.

man gives her whiplashes. Similarly in Neukirchen, Hessen-Nassau, every coachman was allowed to hit her when seeing her.

These answers indicate that at least some rules are more directly related to a general belief of impurity of women after birth that is also manifested in the old testament and not an early form of labor protection.[18] Because these rules may be different from other rules, we create an additional category that we call 'Impurity'. A rule is classified as 'Impurity' norm if misfortune for others is named in at least 1% of cases where this rule applies (for reference: only 4% percent of observations name a consequence and mostly also only for a subset of rules they name). We do not count consequences for herself and for the child (mostly that the child gets a bad character or is exchanged), because these are less clearly related to the impurity notion.[19] The most prevalent impurity categories are, as is already evident from the examples above, not going to the neighbor or visiting anyone and not getting water from the well or water pump. Most of the other rules that can be classifieed as strongly entangled with a belief about impurity also restrict social interactions, such as prohibiting participation in specific festivities including weddings (causes a fight), not interacting with specific women (e.g. other women that just gave birth), not showing herself publicly as well as not going to the pub. Other impurity rules include: staying away from/not touching certain foods (it will perish), not preserving food, not participating in slaughters, not going into stables (animals will die), not borrowing anything, not crossing water and not showing herself in the window.[20]

In contrast to the neighborhood help obligations, the sum of rules is not a meaningful summary measure as some rules imply others. For instance, if you do not leave the house, you also cannot visit the neighbor. In order to avoid the use of too many outcomes, we focus on the two dominant rules, namely rules related to seclusion which confine women to their home, and rules more strongly related to the belief that women after birth are impure and bring harm to others.

78% of all localities name at least one childbed norm (see Table 1.2.1). 61% of localities name a rule related to seclusion, 30% a rule more strongly related to a belief about impurity, and only 5% name exclusively other rules which fall in neither category. Note that 21% of localities indicate that while there is no norm any more,

18. "A woman who [..] gives birth to a son will be ceremonially unclean for seven days, just as she is unclean during her monthly period. [...] 4 Then the woman must wait thirty-three days to be purified from her bleeding. She must not touch anything sacred or go to the sanctuary until the days of her purification are over. If she gives birth to a daughter, for two weeks the woman will be unclean, as during her period. Then she must wait sixty-six days to be purified from her bleeding." (Leviticus 12)

19. We do not classify a rule as impurity category in case the most dominant consequence is harm for the child, as these consequences do not have to do anything with impurity.

20. The idea to classify the rules in this manner goes back to Grober-Glück (1966b) who also separately considers rules that are related to the idea that the woman brings harm. Her classification is very similar but not fully equivalent. Most notably, she does not classify going to the well in that manner.

there was a restriction in place in the past, making them almost universally existent (99%) at some point in time. This is not the case for the neighborhood obligations, where only 4% of localities indicate that it applied in the past, but not today.

1.2.2 Interrelationship of the norm measures

Table 1.2.2 displays Pearson's correlation coefficients among our norm variables. Note that our norm variables are binary with the exception of the sum of neighborhood help activities. For binary variables, the Pearson's Correlation Coefficient is numerically equivalent to typical measures of statistical association among nominal variables, such as the ϕ-coefficient or Cramer's V and thus suitable to quantify the association among our norm variables.[21] p-values of tests of independence are obtained using the modified t-test as suggested by Clifford and Richardson (1985) which adjusts for spatial autocorrelation in the respective variables.

The lower panel of Table 1.2.2 contains the correlation among different neighborhood help categories. All neighborhood categories are highly and significantly correlated implying that conditional on having one, a locality is more likely to have any other. This correlation structure suggests that these are not distinct norms, but different expressions of the same norm to help your neighbor. Most strongly interrelated with a correlation of above 60% are help at wedding, birth and sickness. The correlation of the other neighborhood help activities with help at agriculture is weakest potentially suggesting that there are some distinct factors underlying this particular neighborhood help activity. Looking at the relationship of our two summary measures of neighborhood obligations in the upper panel of Table 1.2.2, namely the indicator variable indicating that the respondent perceives the neighborhood help as obligatory and the number of neighborhood help activities named, we see a strong relationship among these two answers. This is unsurprising, because respondents only rarely indicate an obligation without specifying at least one task where this obligations applies. The incomplete overlap between these variables mostly stems from respondents insisting that the neighborhood help is not obligatory but commonly done. As opposed to neighborhood help norms, our two childbed norms are statistically orthogonal to each other. The correlation between norms of seclusion after birth and norms related to an impurity notion is -0.01 and statistically insignificant suggesting that these are indeed distinct norms that may have different origins or different mechanisms of persistence.

By looking at the inter-relationship of our two norm domains, we find that restrictions of women after birth are more likely to occur in localities that also display neighborhood help obligations and vice versa. However, the cross norm relationships

21. Strictly speaking, the norms are not binary for all locations, because as mentioned above, when we have multiple answers for a coordinate, we take the mean across answers, potentially yielding values between 0 and 1. However, this is rare, typically less than 1% of cases.

are with less than 10% lower than the intra-neighborhood help norm relationships with a typical correlation of above 40%. Further, this relationship is almost entirely driven by norms of seclusion after birth. Norms related to impurity have with 3% an only small raw correlation to neighborhood help obligations which is also only borderline statistically significant. The lack of relationship is rather surprising in light of typical mechanism underlying the evolution of social norms. In particular, one may suspect that localities that are able to maintain cooperative norms are also more able to maintain any other norm given that the maintenance norms in general is thought to require some level of community level social interaction or community tightness. The result, thus, hints to one of two possible explanations: Either impurity norms have a distinct maintenance mechanism, or they have a distinct origin that is negatively associated with factors influencing the origin or value of neighborhood obligations. We begin exploring this question by exploring spatial variation of the different norms in the next section.

Table 1.2.2. Raw correlation among norms

	Obligated	Sum nb help	Any rule (mother)	Seclusion	Impurity	
Obligated	1	0.586***	0.069***	0.073***	0.035*	
Sum nb help		1	0.093***	0.104***	0.031	
Any rule (mother)			1	0.67***	0.349***	
Seclusion				1	−0.01	
Impurity					1	
	..at death	..at wedding	..at birth	..when sick	..with house building	..with agriculture
..at death	1	0.528***	0.431***	0.476***	0.427***	0.244***
..at wedding		1	0.629***	0.623***	0.421***	0.303***
..at birth			1	0.614***	0.338***	0.311***
..when sick				1	0.37***	0.323***
..with house building					1	0.308***
..with agriculture						1

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Entries represent Pearson's correlation coefficients. p-values are adjusted for spatial autocorrelation and obtained using the modified t-test of Clifford and Richardson (1985) as implemented by Osorio, Vallejos, and Cuevas (2016). For readability, the table does not contain the correlation between the two smaller categories help in emergencies and other help.

1.3 Spatial Dependence and Spatial Co-Location

The raw geographic distribution of the aggregate measures of neighborhood obligations and restrictions of mothers after giving birth are displayed in Figure 1.A.2 and Figure 1.A.3 in the Appendix, respectively. They show that none of the norms are restricted to a particular region, instead all norms are widespread over the whole sampling region. Other than that, the geographic patterns are difficult to discern given the lack of restriction of a norm to a particular region and the non-uniform of the underlining sample. To quantify the degree of spatial dependence, we first examine the overall spatial autocorrelation in each variable (see Table 1.3.1) and then examine the existence and location of local clusters of particular high frequency in each norm (see Figures 1.3.1 and 1.3.2).

Table 1.3.1 quantifies the degree of global spatial autocorrelation in the social norms. That is, it quantifies the degree to which localities that are closer together have more similar values. The table displays Moran's I for (standardized) continuous values and the Φ coefficient[22] of the contingency table of Spatial Joint Counts for binary variables. The contingency table is constructed using each locations own value in the row vs. its neighbors values as columns yielding a 2x2 contingency table. The p-values are based on the expected frequency under spatial independence of this contingency table as proposed by Iyer (1949). The p-value of the test of spatial independence for the continuous variable are based on a permutation test. Both the Φ-coefficient as well as Moran's I can take values between -1 and 1 equivalent to normal correlation coefficients, where negative values correspond to negative spatial dependence, (close to) zero[23] correspond to no spatial dependence, and positive values correspond to positive spatial dependence.

As can be seen in Table 1.3.1, all social norms measures display positive and significant global spatial autocorrelation. All of our aggregate measures display with around 0.13-0.15 remarkably similar levels of spatial autocorrelation with the exception of the indicator variable 'Any rule (mother)' which only displays a spatial autocorrelation as measured by Φ of 0.08. Despite the high statistical significance, the size of the spatial autocorrelation can be interpreted as moderate. To set a reference of what large spatial autocorrelation would look like: it is approximately one sixth of the spatial autocorrelation in majority religious denomination in our sample which displays a Φ-coefficient of 0.83. As a robustness check, we estimate the spatial autocorrelation using distance-based spatial weighting matrix (uniform, within 20 km radius) in Table 1.A.1 of the Appendix. Results are very similar, though the esti-

22. For the reader unfamiliar with the Φ-coefficient: $\Phi := \frac{n_{11}n_{00}-n_{10}n_{01}}{\sqrt{n_{1.}n_{.1}n_{0.}n_{.0}}}$ of a typical 2x2 contingency table, where n_{ij} refers to number of observations in row i, column j and $.i$ refers to the margin of row i etc. It is thus -1 if all observation lie on the cross-diagonal of the contingency table and 1 if all observations lie on the diagonal, 0 if proportions are equal on diagonal and cross-diagonal (independence) and thus has a similar interpretation to a Pearson's correlation coefficient.

23. Note that $E(I) = \frac{-1}{N-1} \approx 0$ under spatial independence.

mated spatial autocorrelation is somewhat smaller than when using the four nearest neighbors. Given that the four nearest neighbors are typically closer than 20 km, this hints to a strong decay of autocorrelation with distance.

Table 1.3.1. Spatial autocorrelation in social norms

	Moran's I	Phi	p-value
Obligated		0.13	$p < 0.01$
Sum nb help	0.15		$p < 0.01$
Nb help at death		0.15	$p < 0.01$
Nb help at wedding		0.14	$p < 0.01$
Nb help at birth		0.09	$p < 0.01$
Nb help when sick		0.10	$p < 0.01$
Nb help with house building		0.15	$p < 0.01$
Nb help with agriculture		0.05	$p < 0.01$
Nb help in emergencies		0.06	$p < 0.01$
Any rule (mother)		0.08	$p < 0.01$
Seclusion		0.13	$p < 0.01$
Impurity		0.15	$p < 0.01$

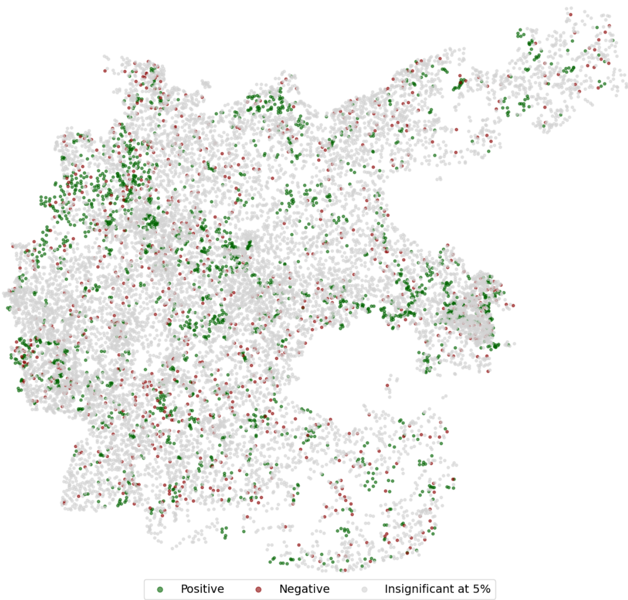
Notes: Table displays Moran's I for (standardized) continuous variable and Φ-coefficients of the contingency tables of the Spatial Joint Counts. The table reports p-values of tests testing the H_0 of spatial independence. For continuous variables this is tested via permutation tests, for binary variables it is tested via the Jount Count test developed by Iyer (1949). All statistics as implemented by Rey and Anselin (2010) using four nearest neighbor spatial weighting. Observations that do not have a neighbor within 10 km are dropped. As a robustness test, we alternatively report distance-based weighting in Table 1.A.1 in the Appendix.

1.3.1 Local clusters

Despite moderate levels of global autocorrelation, there still may exist local "hot" regions, i.e. regions in which the respective norm is particularly universal. To investigate this, we estimate Local Joint Counts for binary variables (see Anselin and Li, 2019) as well as Local Moran's Is for the sum of neighborhood activities. Figure 1.3.1a is colored according to Local Joint Count based on four nearest neighbors spatial weighting (excluding points with no neighbors within 10 km) for the indicator variable of whether neighbors are obligated to mutual assistance. Local Joint Count is at its maximum of four if the locality itself displays the norm as well as all its neighbors. It is three, two and one if three, two, or one of its neighbors displays the norm, respectively. It is zero if the locality itself does not display the norm or none of its neighbors displays the norm or neither. Accordingly, the bluer the region, the more universal the norm is in this region. As we can see in Figure 1.3.1a, neighborhood obligations are almost universal in the North-West of the German Reich (Prussian-Province Hannover, Munster Region of Province Westphalia, Rhine Province) skipping the middle (industrialized areas) of the Rhine Province as well as largerly the Grand Duchy of Oldenburg. Other clusters are very local and are widely distributed over the sampling area. Looking at Local Moran's *I* of the number of neighborhood help activities yields similar results, however, the results are less pronounced (see Figure 1.3.1b).

Figure 1.3.2 displays the "hot" regions of the restriction of women after birth. The figures show that seclusion norms are almost universal in the South, Center-West, and North East of the German Reich, including respective border region of Czech Republic, and Austria while skipping the Center-North and Center East regions of the German Reich (see in Figure 1.3.2a). Impurity norms are almost spatially orthogonal and are more universal only in the Center-East (Upper Bavaria, Saxony, Thuringia, some part of Lower Silisia, see Figure in Figure 1.3.2b).

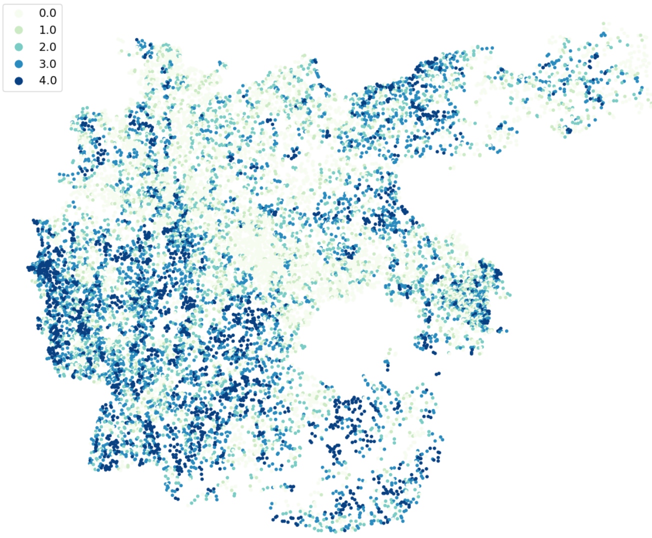
(a) Obligated



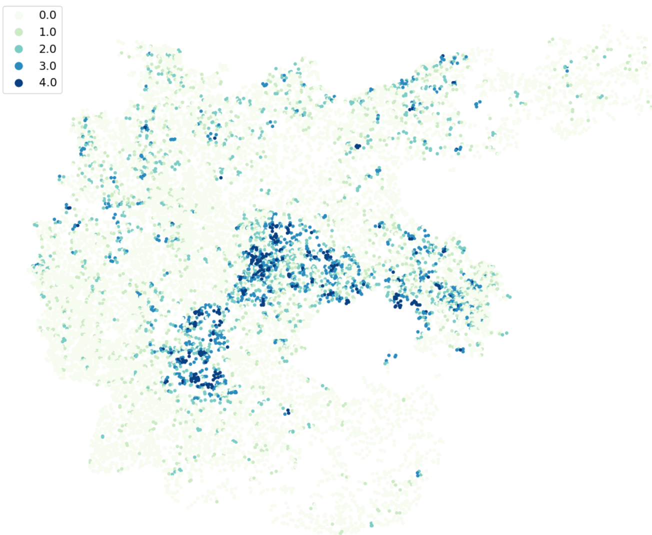
(b) Sum nb help

Figure 1.3.1. Local spatial cluster: Neighborhood help obligations

Notes: Figure 1.3.1a: Colored according to Local Joint Count based on four nearest neighbors spatial weighting (excluding points with no neighbors within 10 km). Local Joint Count is at the maximum of four if the locality itself displays the norm as well as all its neighbors, it is three, two and one if three of four, two of four, and one of four neighbors displays the norm, respectively. It is 0 if none of its neighbors displays the norm or itself does not display the norm or both (see Anselin and Li, 2019, for more information). Figure 1.3.1b: Displays local spatial autocorrelation according to local Moran's *I*. Points that do not have a neighbor within 10 km are excluded.



(a) Seclusion norms



(b) Impurity norms

Figure 1.3.2. Local spatial cluster: Restrictions for mothers after giving birth

Notes: Colored according to Local Joint Count based on four nearest neighbors spatial weighting (excluding points with no neighbors within 10 km). Local Joint Count is at the maximum of four if the locality itself displays the norm as well as all its neighbors, it is three, two and one if three of four, two of four, and one of four neighbors displays the norm, respectively(see Anselin and Li, 2019, for more information). It is 0 if none of its neighbors displays both norms or itself does not display both norms or both. Points that do not have a neighbor within 10 km are excluded.

(a) Co-location regions: Obligated to neighborhood help & rules of seclusion after birth

(b) Co-location regions: Obligated to neighborhood help & rules of related to impurity

Figure 1.3.3. Joint Spatial Distribution of Social Norms: Co-Location Areas of Neighborhood help obligations and Childbed norms

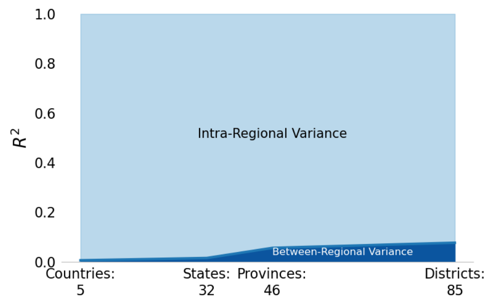
Notes: Colored according to Join Count of co-locations based on four nearest neighbors spatial weighting (excluding points with no neighbors within 10 km). Joint Count of Co-Location is at the maximum of four if the locality itself displays both norms as well as all its neighbors, it is three, two and one if three of four, two of four, and one of four neighbors displays both norms, respectively. It is 0 if none of its neighbors displays both norms or itself does not display both norms or both.

To inverstigate whether the inter-relationship between norm domains is at least partly driven by regional overlap or lack thereof, we investigate the co-location patterns of norms. Figure 1.3.3 displays local clusters of co-location between the existence of neighborhood help obligations and existence of restrictions for women after birth. The maps are colored according to the Joint Count of Co-Locations based on Anselin and Li (2019) relying on the four nearest neighbor weighting matrix. Joint Count of Co-Location is at the maximum of four if the locality itself displays both norms as well as all its neighbors, it is three, two and one if three of four, two of four, and one of four neighbors displays both norms, respectively. It is zero if none of its neighbors displays both norms or itself does not display both norms or both. The Join Count of Co-Location, thus, measures the degree of regional overlap between norms. As we can see Figures 1.3.3a and 1.3.3b dominant regions of general restrictions for seclusion norms overlap in the most dominant region of neighborhood obligations (North-West of German Reich) described above. Accordingly, there is almost no regional overlap between neighborhood obligations and impurity norms, as the latter are most universal in the Center-East, partially explaining the lack of correlation between neighborood help obligations and impurity norms. Figure 1.A.4 in the Appendix replicates this results for the indicator variable whether any neighborhood help activity is performed in the locality.

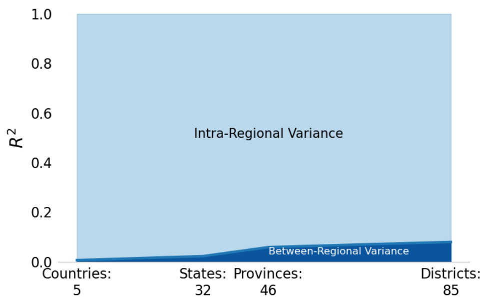
Thus, neighborhood help obligations and childbed norms moderately, but statistically significantly, spatially autocorrelated. Despite this moderate spatial autocorrelation, all norms display some regions in which they are almost universal. These regions partially overlap between neighborhood help obligations and norms of seclusion after birth, however, they are almost orthogonal for impurity norms and any other type of norms.

1.3.2 Alignment of spatial patterns and poltical/administrative boundaries

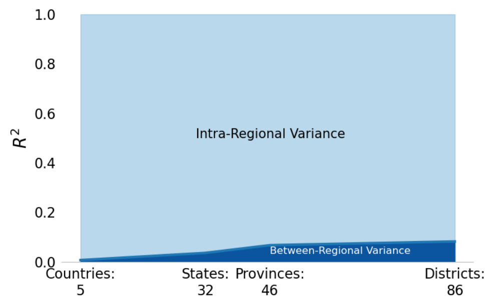
To explore how well the previously laid out spatial patterns line up with institutional boundaries around 1930, we can regress each of our norm measures on the institution the locality belongs to at varying level: country, state or state-like region, province and district. Our localities are distributed over five countries, 32 states, 46 provinces and 85 districts. The R^2 of these regressions gives us the between regional variance, while $1-R^2$ gives us the unexplained variance, namely the intra-regional variance. Figure 1.3.4 displays the explanatory power of these institutional divisions at different levels. The explanatory power of institutional units across all levels is lowest for the existence of obligatory neighborhood help, and highest for impurity norms across all levels of institutional boundaries. Country indicators explain between 0.5% (obligated to neighborhood help) and 1.1% (impurity norms) of the overall variance, state indicators between 1.4% and 4.1%, province indicators between 5.5% and 6%, and districts indicators between 7.6% and 9.6%.



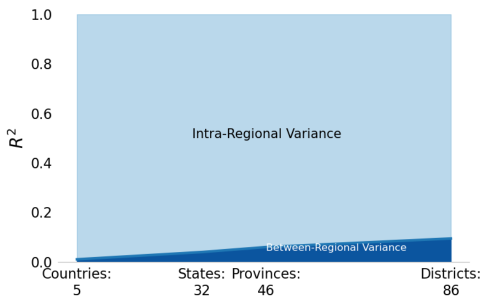
(a) Obligated to neighborhood help.



(b) Intensity of neighborhood obligations



(c) Seclusion

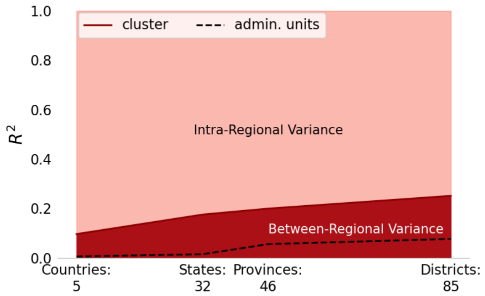


(d) Impurity

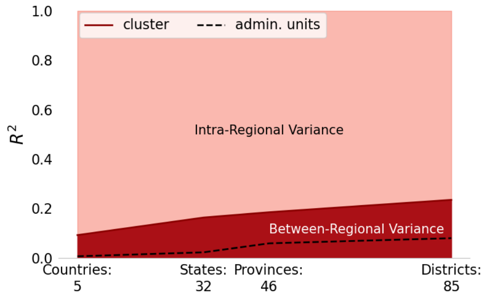
Figure 1.3.4. Explanatory power of institutional boundaries.

Notes: Figures display R^2 of a regression with dependent variable norm, independent variable institutional boundary dummies. Countries=German Reich, Austria, Czechoslovakia, Gdansk, Poland. States correspond to states for the German Reich and Austria, the Czech lands Bohemia and Moravia, and Silesia, Poznan, former West Prussia. Provinces are the same as states, except for Prussia, where they are replaced with Provinces. Districts align with *Regierungsbezirke* in the German Reich, *Voivodeships* in Poland, for Czechia and Austria, these are equivalent to provinces (see Figure 1.A.5). Districts are obtained from MPIDR and CGG (2013). Intensity of neighborhood obligations: sum across all neighborhood activities performed.

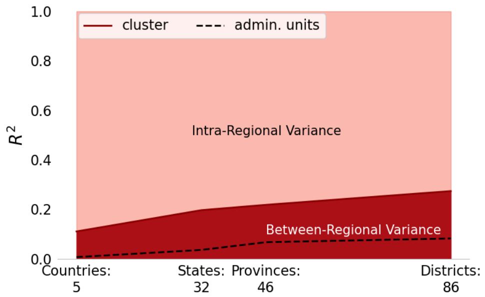
Given that the spatial correlation in the underlying data is far from perfect, as well as potential measurement error problems, it is difficult to judge whether this is large or small. To get a better grasp of the size of the potential explanatory power of institutional boundaries, we can flip the question and asked how regions should be constructed in order to minimize intra-regional variance, and thus maximize the explanatory power of regions. This gives us the maximum explanatory power any collections of contiguous localities can achieve given the spatial distribution of the underlying data. We can obtain these artificial regions via regionalization, i.e. clustering under spatial connectivity constraints for the same number of clusters as there are administrative regions.



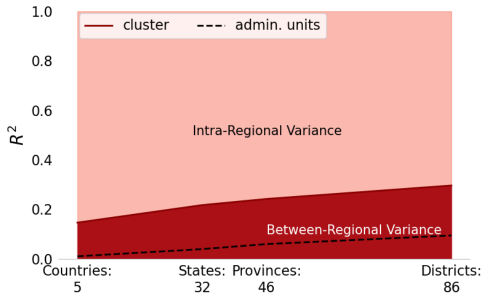
(a) Obligated to neighborhood help



(b) Intensity of neighborhood obligations



(c) Seclusion



(d) Impurity

Figure 1.3.5. Explanatory power of artificial regions vs. administrative units

Figures display R^2 of a regression with dependent variable norm, independent variable artificial regions obtained by geographically constraint clusters compared to administrative boundaries. Intensity of neighborhood obligations: sum across all neighborhood activities performed.

We generate clusters by using a agglomerative hierarchical clustering algorithm under connectivity constraints (four nearest neighbors).[24] Given the number of clusters, the algorithm chooses clusters such that the weighted sum of intra-cluster variance given the number of cluster and the connectivity constraint is minimized.[25] The connectivity constraint is given by the geographic distribution of observations. We define each locality to be connected to its four nearest neighbors in our data set. The connectivity constraint ensures that a locality can only be added to a cluster if it is directly connected to another locality in that cluster. These yields collections of localities which are contiguous to each other.

Figure 1.3.5 compares the explanatory power of contemporary administrative boundaries to the explanatory power of these artificial regions. Artificial regions explain from approximately 9%-14% (5 artificial regions) to 22%-29% of the overall

24. Note that we exclude points that have no neighbors within a 10 km radius, as well as disconnected components, i.e. points that are only connected among themselves and not with the main sampling region.

25. To be more precise: the algorithm chooses clusters such that the sum of squared differences in the input variables within all clusters is minimized.

variance. Thus, even artificial regions that are constructed to minimize the intra-regional variation capture less than one third of the overall variance in social norms. If we look at the ratio of explanatory power between artificial regions and institutional units, we find that countries only explain up to one twelfth of what could be potentially explained, states upto one fifth, provinces and districts upto one third. In general, across norms this picture is rather similar, with the exception that countries and states explain an exceptionally low share of the variance in neighborhood obligations.

The above analysis reveals that the spatial patterns in the norm variables are likely a result of a rather complex data generating process. All norms can be found almost in every region in our sampling area inducing large intra-regional heterogeneity even when regions are endogenously defined to be variance minimizing. However, for each norm there are some regions in which they are more universal than in others. These only slightly overlap across norm domains. Contemporary political institutional boundaries (1930) such as countries and states do not explain a large share of this inter-regional variation. In the remaining paper we study different reasons for this complex spatial pattern.

1.4 Historical Political Institutions

Social norms may be related not only to current but also to past political institutions (see e.g. Grosjean, 2011; Becker, Boeckh, et al., 2016; Buggle, 2016; Guiso, Sapienza, and Zingales, 2016). Given the frequent institutional border shifts within the sampling region due to the various petty states in the German lands,[26] this is not well summarized by current institutions and may even yield vast intra-regional variation even if political institutions matter for the evolution and prevalence of social norms.

To investigate this, we use a cultural dyadic gravity equation as put forth by Grosjean (2011). For this, we first calculate the (weighted) length each pair of localities i, j belonged to the same independent state. In comparison to Grosjean (2011), we do not only focus on the large empires but additionally use variation that comes from the territorial shifts through the numerous petty states, in particular, within the area of the German Reich. To construct the variable, we rely on a new GIS-data set constructed by Abramson (2017) and published in the replication files of Abramson, Carter, and Ying (2022). We thus also follow Abramson (2017) state definition. The data covers European borders in 5-year intervals from 1200 to 1790. We merge this data to our geocoded localities. We drop observations that are ever within 5 kms of a border to avoid false assignment due to some potential imprecision in location or border or both. In total, we retain about 18 million dyads. As a robustness check, we also report results for a 10 km buffer where we only have around five million dyads left.

Because the influence of an institutions should at least partially diminish over time (see e.g. Abramson, Carter, and Ying, 2022), we assign higher weights to years spent together in the nearer past than in the farther past. We construct our main dependent variable as follows:

$$\text{Yrs. same state (wgt.)}_{ij} = \sum_{t \in T} I(i,j \text{ same state in } t) \times \exp\left(\frac{-(1790 - t)}{1790 - 1200}\right) \tag{1.4.1}$$

where $T = \{1200, 1205, ..1785, 1790\}$ and $I(.)$ is an indicator function. The weighting function is inspired by Abramson, Carter, and Ying (2022) and is increasing in the year assigning a maximum weight of 1 for the year 1790. We additionally divide the result by 10. Given that the data is in 5-year intervals, the scaling is equivalent to scaling it by 50 years. The resulting variable varies between 0 and 7, where 7 is equivalent to being in the same state for the full 590 years.

26. These shifts are mostly due to strategic marriages across different dynasties, deaths of sovereigns and inheritance rules within dynasties, and sometimes also wars and resulting treaties.

To delineate effects that come from a shared institutional history from effects that come through a persistence in institutional borders, we additionally control for being in the same (federal) state or state like region in 1930. Identification, thus, relies on the fact that past territorial divisions do not fully align with 1930's borders. In addition, shared institutional history is naturally correlated with physical distance and localities that are physically close together also may share other shocks that influence social norms. To account for this, we additionally flexibly control for the physical distance of the two localities. Thus, our resulting dyadic gravity equation looks as follows:

$$d(n_i, n_j) \;=\; \alpha + \beta \text{Yrs. same state (weighted)}_{ij} + \gamma \text{Same state}_{ij} + f(d(s_i, s_j)) + \epsilon_{ij} \tag{1.4.2}$$

where $d(.)$ is the euclidean distance function. n_i, n_j are the norm vectors at locations i, j with $i \neq j$. Yrs. same state (weighted)$_{ij}$ is as defined above. Same state is an indicator variable that is one if both localities are in the same (federal) state or state-like unit around 1930. $d(s_i, s_j)$ represents the physical distance between i and j and $f(.)$ represents a quadratic function in the main specification in addition to an interaction between the distance and an indicator variable whether the localities are closer together than 100 km as well as an indicator variable that is one if the localities farther than 800km apart to match the pattern in the relationship between distance in social norms and physical distance perfectly (see Figure 1.A.6). In a robustness specification, we also vary these cutoffs but it does not change the results (see Table 1.A.3).

We report the results for three different outcome variables. First the cultural norm distance as measured by all social norms: being obligated to neighborhood help general, being obligated to help at wedding, death, upon sickness, upon birth, with building a house, upon an emergency, or with some agricultural work, as well as rules of seclusion, impurity norms, and a residual indicator indicating whether any other rule applies for the mother after birth. We then report the results for each norm domain separately, i.e. the cultural norm distance of neighborhood help obligations and childbed norms. To be able to compare coefficients across outcome variables, we standardize the outcome variables. Results are displayed in Table 1.4.1. Column (1) shows that the overall social norm distance is significantly negatively related to the weighted years spent in the same state controlling for physical distance and being part of the same (federal) state in 1930. To give an idea of the effect size: one unit increase in the weighted years same state measure is approximately equivalent to spending 50 years in the same state in the end of 18th century, 60 years in the end of the 17th century, 75 years in the end of the 16th century, ..., more than the full 13th century. Thus, spending 50 years in the same state in the 18th century, etc., decreases the cultural distance in norms by approximately .02 standard deviations, which is the same effect as being currently in the same state ceteris paribus.

Column (4) and (7) show the result separately for neighborhood help obligations and childbed norms. It shows that the effect is largely driven by decreasing the distance in neighborhood obligations. In fact, the decrease in the difference in childbed norms is with .002 standard deviations one sixth of the decrease in the difference in neighborhood obligations which amounts to .012 and statistically insignificant.[27]

One could now argue that this result is not due to being exposed to a similar institutional environment but more due to borders hindering diffusion of social norms by decreasing the interaction between localities. If this were true, we would expect that the effect should be mainly driven by localities being rather close together rather than those being farther away. To investigate these potential channels, we additionally interact the years spent together with a distance threshold of 50 km. The results are displayed in column (2), (4), (8). It shows that the effect is not driven by particularly close-together localities. This result is robust to shifting the distance threshold to 100 km (see Table 1.A.2 in the Appendix).

Columns (3), (7), (9) investigate whether it is actually the length that plays a role, or just ever being in the same state between 1200 and 1790 by including an indicator variable that is 1 if the two localities ever were in the same state. The results suggest that both extensive margin and intensive margin play a role. The results are also robust to using a buffer of 10 km around borders instead of 5 km (see Table 1.A.4).

27. Note that this result does not depend on how we define the distance in childbed norms, i.e. whether we instead use having any childbed norm, or whether we use the full set of different childbed norms instead of only focusing on norms of seclusion and norms of impurity.

Table 1.4.1. Length of being part of the same state & similarities in social norms

	Distance All			Distance Nbh. Obligations			Distance Childbed Norms		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Yrs same state (wgt.)	-0.019^{***}	-0.019^{***}	-0.013^{***}	-0.012^{***}	-0.012^{***}	-0.007^{***}	-0.002	-0.002	-0.003^{*}
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
Same state, 1930	-0.02^{**}	-0.02^{**}	-0.016	-0.013	-0.013	-0.01	0.024^{***}	0.024^{***}	0.023^{***}
	(0.01)	(0.01)	(0.01)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)
Distance < 50km		0.003			0.009			-0.02	
		(0.017)			(0.016)			(0.013)	
Yrs same state (wgt.) × Distance < 50km		-0.004			-0.005			0.003	
		(0.003)			(0.003)			(0.002)	
Ever same state			-0.059^{***}			-0.05^{***}			0.011^{*}
			(0.007)			(0.007)			(0.006)
Distance (km)	-0.01^{***}	-0.01^{***}	-0.01^{***}	0.002	0.003	0.003	0.021^{***}	0.021^{***}	0.021^{***}
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Distance (km) sq	0.001^{***}	0.001^{***}	0.001^{***}	0.000^{**}	0.000^{**}	0.000^{*}	-0.002^{***}	-0.002^{***}	-0.002^{***}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Intercept	0.147^{*}	0.146^{*}	0.162^{**}	-0.046	-0.046	-0.032	-0.247^{***}	-0.247^{***}	-0.25^{***}
	(0.076)	(0.076)	(0.074)	(0.1)	(0.1)	(0.099)	(0.054)	(0.054)	(0.054)
Observations	18,699,670	18,699,670	18,699,670	18,699,670	18,699,670	18,699,670	18,699,670	18,699,670	18,699,670
R^2	0.005	0.005	0.006	0.003	0.003	0.003	0.002	0.002	0.002

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors obtained using two-way clustering on each dyads localities. All=euclidean distance of all norm meausures, Distance Nbh. Obligations=euclidean distance across all neighborhood norms (being obligated, performing neighborhood help upon death, at weddings, upon birth, upon sickness, with house building, in emergencies (fire/flood), with agricultural activities). Distance in childbed norm = distance across the three indicator variables: seclusion, impure, any other childbed norm. Yrs same state is a contiunuous measure that varies between 0 and 7 with 7 indicating that the two localities were always together between 1200 and 1790. All specifications additionally include a linear interaction between distance and an indicator variable of being closer than 100 km, as well as, an indicator variable being farther away than 800 km to match the pattern in the relationship to physical distance displayed in Figure 1.A.6. In addition all specifications additionally control for the distance in an indicator variable of being part of an imperial city or independent city.

1.5 Local population and distance to towns

Being closer to a town may allow the inhabitants of the locality to migrate to said town, access the market of that town, and expose them to new ideas from the town. The opportunity to migrate to a city or access the market in the town may undermine social sanctions upon norm violation (Aoki (2001, p. 51) and Kranton (1996)), and may increase substitutability of neighborhood help through access to specialized services (Wurzbacher, 1961, p. 114). Similarly, population size is related to market access or size of markets available to members of a community. Population size may, however, also affect the existence of social norms by increasing the availability of alternative social relationships and networks potentially decreasing the reliance on specific community members and, thus, the efficacy of social sanctions. Both of these factors strongly vary within regions and, hence, may explain part of the spatial pattern in social norms, we observe.

To investigate these potential determinants, we matched names of localities and administrative units to the Gemeindeverzeichnis (the official list of communities in the German Reich), 1900, and Uli Schubert's collection of matched population sizes from 1910.[28] Given the nature of the data, namely that the match is based on names of localities and counties, it is not a perfect match. However, we were able to match 75% of localities ($= 12,537$) that lie in the German Reich to the database. We were also able to match 193 of localities that lie in Poland and Gdansk but where part of the German Reich prior to World War I and 11 localities located in Czech-Silesia that were part of Prussia until 1920. In case, the locality is a district of a city/town, and this district does not constitute its own community according to the list of villages, we assign the population of the city/town. For 788 unmatched localities (614 on the territory of the German Reich, 141 in Czechia, 33 in Austria), we were able to obtain population data through other sources.[29] To measure the level of remoteness of a locality, we calculate a localities distance to the nearest town/city.[30]

28. Content and sources can be found on Uli Schubert's Website https://gemeindeverzeichnis.de/ (last accessed on 24/11/2022).

29. For the State Salzburg of Austria, we used Klein (2016), for unmatched in Czechia and the German Reich, we relied on the collection "Historisches Ortsverzeichnis" of the Association of Computer-genealogy available at http://gov.genealogy.net/ (last accessed on 24/11/2022) from which we also collected alternative spellings of all available localities to ensure better name-based matching. For these matches, we use any population data available between 1900 and 1914 but generally the closest available to 1910.

30. For this purpose geocode the location of towns and cities listed in the Gemeindeverzeichnis with the contemporary nominatim geocoder https://nominatim.openstreetmap.org/. For implausible results and for east Prussia (because of the change in language) we manually checked the results of this automated geocoding by hand using Wikipedia and Google Maps. For Austria, we use towns as classified by Census, 2001 of Austria Statistics, for the Czech Republic we use municipalities with a population of more than 10,000 inhabitants as of January 2021 to proxy past towns. We set the distance to zero if the locality itself is a town/city (it is otherwise rarely completely zero).

Figure 1.5.1a and Figure 1.5.1b display the distribution of population and distances, respectively. The population distribution is highly right skewed and long-tailed. By far the majority of localities have less than 1,000 inhabitants (approx. 75%), but 5% of localities have between 4,334 and 2,071,257 inhabitants (see Table 1.A.5 in the Appendix). Similarly, around 75% of villages are located within 12 km of a town, but 5% are located between 22 km and 48 km far away from the next town. Given this spread out support after 95th percentile in each variable, we focus on the first 95% of each variable in our main analysis. Note that roughly 6.7% of the sample are towns themselves, and, thus, display a distance to the next town of zero.

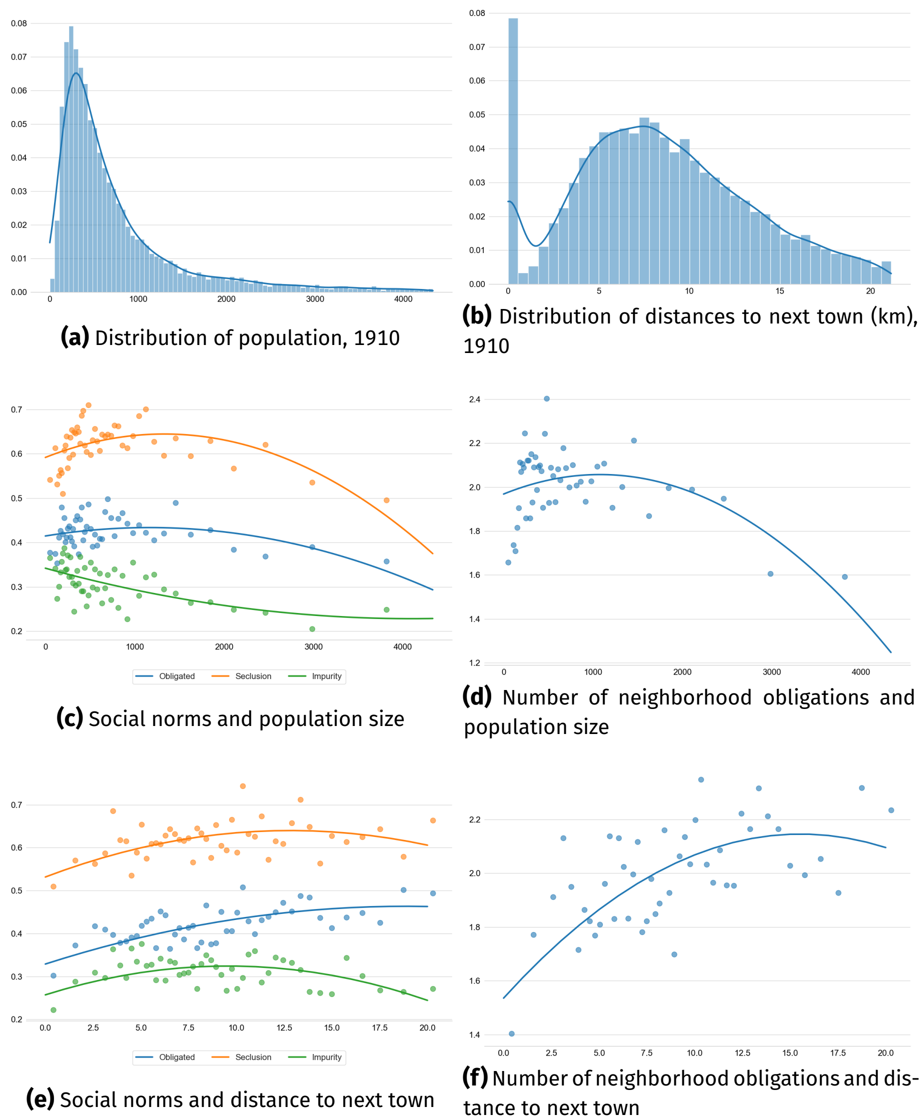
Figure 1.5.1c and Figure 1.5.1d display the relationship between social norms on the locality level and population size. They show that the prevalence of all norms are decreasing with population and that this is mostly driven by a sharper decline of norms above approx 2,000 inhabitants. The prevalence of norms (with the exception of impurity norms) is rather flat or even (weakly) increasing for population sizes between 0 and 1,000 inhabitants, which also tend to display a higher variance.

Distance to the nearest town is a less pronounced mirror image to population. The farther away from a town/city the more prevelant our norm measures. However, this relationship is only pronounced for neighborhood help obligations. In particular, impurity norms seem almost unrelated to the distance to the next town.

Table 1.5.1 and 1.5.2 further explore these relationships additionally controlling for the locality being itself a town, and investigating within district and within county relationships. It shows that population size is negatively related to all norms, in particular within regions and within districts. However, the concavity of the relationship is mostly a result of cross regional comparison (see Tables 1.A.8 and 1.A.9 for the polynomial specification).

The relationship between remoteness and social norms is somewhat more complex. The likelihood of existence of neighborhood help obligations is increasing in the distance to the nearest town and this is not driven by localities themselves being towns, while the gradient in the intensity of neighborhood help activities seems to be only consistently lower in towns than in rural communities but does not consistently depend on the distance to the nearest town. Chilbed norms tend to also be less prevalent in towns but the effect is only marginally significant (at the 10% level) and the significance is not robust to alternative standard error definition (see Table 1.A.5)

The results are robust to using standard errors that are adjusted for spatial autocorrelation instead of clustering (see Tables 1.A.10 and 1.A.11). We can also use an alternative definition of remoteness instead of distance to the next town, namely, localized population density (1939) calculated based on data from the German Population Database (Roesel, 2022). This has the disadvantage that the data is only available for the Federal Republic of Germany in its borders after 1990, thus, lacks data on territories in present day Poland, Czechia, and Austria. The database contains



(a) Distribution of population, 1910

(b) Distribution of distances to next town (km), 1910

(c) Social norms and population size

(d) Number of neighborhood obligations and population size

(e) Social norms and distance to next town

(f) Number of neighborhood obligations and distance to next town

Figure 1.5.1. Relationship between remoteness, population size and social norms on the locality level.

Notes: Figures (a) and (b) display histograms and kernel density estimates of the population of each locality, and the the distance to the next town, respectively. Figures (c)-(f) display bin scatter plot with observations binned in 50 quantiles (2% each) and a regression line using a quadratic fit. Because, both, population size as well as distance to next town are right tailed, Figures ignore the 95th percentile and above of each variable, respectively. For an overview of percentiles and distribution of these variables, consult Table 1.A.5 in the Appendix.

population in 1939 on the level of present-day German municipalities (referenced to 2011). It is, thus, on a more aggregated level than the community level data – the median population size in the former is 4,404 while the median of th latter is 524 (see Table 1.A.5 in the Appendix). The results suggest that local population density is highly predictive for the existence of all norms even conditional on locality's population size (see Tables 1.A.12 and 1.A.13). As opposed to distance to the next town, this also holds within region.

Table 1.5.1. Social norms and population size: Neighborhood obligations

	Dependent variable:									
	Obligated					Sum nb help				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Pop 1910	−0.001		0.0004	−0.003***	−0.002***	−0.008**		−0.001	−0.014***	−0.012***
	(0.001)		(0.001)	(0.001)	(0.001)	(0.003)		(0.003)	(0.003)	(0.003)
Dist town (km)		0.007***	0.005***	0.002**	0.002**		0.032***	0.021***	0.008*	0.006
		(0.001)	(0.001)	(0.001)	(0.001)		(0.005)	(0.006)	(0.005)	(0.005)
Town			−0.064**	−0.042*	−0.041			−0.300***	−0.247***	−0.308***
			(0.026)	(0.024)	(0.025)			(0.102)	(0.094)	(0.100)
Constant	0.432***	0.352***	0.374***			2.063***	1.688***	1.825***		
	(0.010)	(0.011)	(0.014)			(0.044)	(0.048)	(0.065)		
Mean dep. var.	0.418	0.418	0.418	0.418	0.418	1.978	1.978	1.978	1.978	1.978
District				✓					✓	
County FE					✓					✓
Observations	12,993	12,993	12,993	12,990	12,993	12,993	12,993	12,993	12,990	12,993
R^2	0.003	0.006	0.008	0.083	0.167	0.005	0.008	0.011	0.093	0.190

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors clustered on the county level. Results using standard errors adjusted for spatial autocorrelation according to Conley (1999) is displayed in table 1.A.10. Column (1), (3)-(6), and (8)-(10) additionally contain an indicator variable that is one if the locality is in the 95th percentile of the population size distribution interacted with population size in 1910. Column (2)-(4), (7)-(10) additionally contain an indicator variable that is one if the locality is in the 95th percentile of the distance to next town distribution interacted with distance to next town distribution. The full set of coefficients of each regression are displayed in Table 1.A.6. Local population per 100 inhabitants. Town is an indicator variable that is one if the locality is defined as 'Stadt' (City/Town) in the official list of communities in the German Reich, 1910 (Gemeindeverzeichnis). Localities that could not be matched to this list, e.g. because they lie outside the German Reich, are classified as Towns if they lie within 2 km of a town / city and have more than 1,000 inhabitants.

Table 1.5.2. Social norms and population size: Childbed norms

	Dependent variable:									
	Seclusion					Impurity				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Pop 1910	−0.001*		−0.001	−0.003***	−0.002***	−0.003***		−0.003***	−0.003***	−0.002***
	(0.001)		(0.001)	(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)
Dist town (km)		0.005***	0.002*	0.003***	0.002		0.0005	−0.003**	0.001	0.002
		(0.001)	(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)	(0.001)
Town			−0.032	−0.010	−0.045*			−0.035	−0.025	−0.038*
			(0.026)	(0.024)	(0.025)			(0.022)	(0.021)	(0.021)
Constant	0.629***	0.572***	0.606***			0.334***	0.302***	0.365***		
	(0.010)	(0.013)	(0.017)			(0.010)	(0.013)	(0.018)		
Mean dep. var.	0.611	0.611	0.611	0.611	0.611	0.303	0.303	0.303	0.303	0.303
District				✓					✓	
County FE					✓					✓
Observations	12,784	12,784	12,784	12,781	12,784	12,784	12,784	12,784	12,781	12,784
R^2	0.005	0.002	0.006	0.085	0.174	0.006	0.001	0.007	0.100	0.189

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors clustered on the county level. Results using standard errors adjusted for spatial autocorrelation according to Conley (1999) is displayed in table 1.A.11. Column (1), (3)-(6), and (8)-(10) additionally contain an indicator variable that is one if the locality is in the 95th percentile of the population size distribution interacted with population size in 1910. Column (2)-(4), (7)-(10) additionally contain an indicator variable that is one if the locality is in the 95th percentile of the distance to next town distribution interacted with distance to next town distribution. The full set of coefficients of each regression are displayed in Table 1.A.7. Local population per 100 inhabitants. Town is an indicator variable that is one if the locality is defined as 'Stadt' (City/Town) in the official list of communities in the German Reich, 1910 (Gemeindeverzeichnis). Localities that could not be matched to this list, e.g. because they lie outside the German Reich, are classified as Towns if they lie within 2 km of a town / city and have more than 1,000 inhabitants.

1.6 Religion

Religion may affect norms by providing the source of the norm through religious dogma. Further, religious differences within localities may constitute a social cleavage that may hamper with social connections across groups potentially affecting a communities ability to sanction norm violations and thus affecting the prevalence of social norms on a community level (see Fearon and Laitin, 1996; Miguel and Gugerty, 2005; Alexander and Christia, 2011, for this mechanism in relation to cooperation).

In every questionnaire, every respondent was asked about the religious composition of the locality. The level of details of answers to this question greatly varies, from exact numbers and percentage of every religious group present in the locality, to 'almost all Catholic'. Further, oftentimes a respondent referred to the answer given in a previous questionnaire, or answers vary across questionnaires, requiring to go through all of the more than 60,000 answers to get an accurate picture of the answer to the religious composition question. That is why we rely on the maps published by Grober-Glück (1966a),[31] which we digitized, to retain the religious composition of localities (for more information see Appendix 1.C).

The major religious groups in our sampling region are Protestants[32] and Catholics. Minor religious groups are comprised by non-denominational, indicated to be present in 7.6% of localities, Jewish, present in 10.5% of localities, and other Christian groups, present in 6.8% of localities. Given that for these groups, respondents often did not name the precise number but instead indicated something like 'some' or similar, we only use indicator varibales that are one if any of the group are present in the respective locality. Grober-Glück (1966a) classified localities according to their share of Protestants of the total number of Catholics and Protestants in a locality in 7 groups: less than 5% Protestants, 5-30% Protestants, 30-50% Protestants, equal split, 50-70% Protestants, 70-95% Protestants, more than 95% Protestants. Note that the inverse relationship holds for Catholics given the definition.

We reduce these seven categories to four coarser categories: 5% Protestants (homogeneous Catholic), 5-50% Protestants (heterogeneous Catholic), 50-95% Protestants (heterogeneous Protestant), more than 95% Protestants (homogeneous Protestant). The reason for this further discretezation is that the groups 30-50% Protestants, equal split, 50-70% Protestants are small relative to our sample size. Together,

31. As well as her unpublished material available in the archive of the GEA, in particular to improve the matching between localities and maps.

32. Grober-Glück (1966a) classified all religious groups as Protestants that were named under Protestant churches in Germany ["evangelische Kirchen Deutschlands"] in the official list of names of religious groups of 1970 by the German Statistical Agency [Verzeichnis der Religionsbenennungen, Ausgabe 1970] as well as under Protestant free churches. The former contain the Lutheran Protestants as well as the Calvinists which are the most dominant groups, the ladder rarely matter for the classification (Grober-Glück, 1966a).

they only make only 7% of our sample. We are thus not able to statistically distinguish between the effects of different degrees of denominational heterogeneity.

Figure 1.6.1a displays the spatial distribution of this categorization (see Figure 1.A.7 in the Appendix for results using the finer categorization). The data generating process behind the inter-regional pattern in religious denomination is well-understood. It goes back to sovereigns' choices in the 16th and 17th century as well as the distribution of borders and regions around 1618-1624 (see e.g. Herzig, 2000; Becker and Woessmann, 2009; Cantoni, 2015; Becker, Pfaff, and Rubin, 2016, and references therin). After the Reformation – with the Peace of Augsburg (1555) – the principle *'Cuius regio, eius religio'* was established that allowed the sovereign of the territory to choose the denomination of his constituents yielding rather homogeneous distributions within former territories. The Westpahlian Peace of 1648, which marked the end of the Thirty Year War, partially put an end to changes in denominations induced by changes in sovereigns' denomination by constituting that territories of the German Holy Empire should carry the denomination that they had on the first of January 1624 with the exception of the hereditary lands of the Habsburgs, in which the Habsburgs where free to choose the denomination of their constituents (see e.g. Herzig, 2000).[33] Accordingly, the North-East and Center-North consists mostly of homogeneously Protestant localities, with the exception of smaller Catholic Enclaves, such as the former Prince-Bishopric Ermland in East Prussia, as well as Eichsfeld and the Hochstift Hildesheim in the Center. As the Reformation started in Wittenberg, it has been argued that Protestantism is particularly clustered in the North/Center East because sovereigns of territories closer to Wittenberg were more likely to convert to Protestantism because the Reformation started diffusing from Wittenberg (Becker and Woessmann, 2009). Localities in the South-East of the sampling region predominantly consist of localities with more than 95% Catholics carrying the legacy of the Electorate of Bavaria, and the Catholic Habsburgs. The very West of the German Reich is more mixed, with, in particular, the former Duchy of Württemberg, Nassau, and Palatinate being predominantly Protestant, and most of the remaining larger territories being Catholic.

Most localities are either almost exclusively Catholic or Protestant. 54.2% are homogeneous Protestant and 26.4% are homogeneous Catholic. The excessively larger share of Protestant villages in our sample is at least partly due to restricting the sample to localities for which we have population size. These are rather scarce in the Catholic Czechia and Austria (see the previous section). However, 19.1% are indeed heterogeneous with respect to either denomination (12.6% majority Protestant, 6.8% majority Catholic, see left axis of Figure 1.6.1b). Interestingly, these seem to be partly clustered as well but less than the majority denomination with partic-

33. It was stipulated that in Silesia and Lower Austria, Protestants needed to be tolerated, however, the later Habsburg did not fully honor this agreement (Herzig, 2000). Another exception was made for the Electorate Palatinate which should go back to the denomination it had in 1618.

ular large clusters Palatinate territories, and Silesia – regions that were subject to not fully successful re-catholization efforts by the Habsburgs and the Catholic Wittelsbacher after the Westphalian Peace which was interrupted by the intervention of foreign Protestant powers (Herzig, 2000).

Figure 1.6.1b and Figure 1.6.1c display the relationship between denominational composition and neighborhood obligations and childbed norms controlling for the presence of other religious groups in the locality (Jewish, Christian Sects, Non-Denominational). The number of neighborhood help activities in a locality as well as the existence of a general obligation follow a U-shaped pattern. Obligations are significantly more prevelant and more intense in religiously homogeneous localities than in religiously heterogeneous localities. In terms of majority religion, neighborhood obligations are slightly more prevalent in Catholic localities, however, this difference statistically not significant (p-value=0.144 and 0.135). Norms of seclusion after giving birth seem to be more strongly related to Catholicism. They are most prevalent in homogeneous Catholic localities and least prevalent in majority Protestant localities (homogeneous or heterogeneous). Impurity norms slightly follow a U-shaped form as well but with a slightly but statistically significantly higher prevalence in homogeneous Protestant localities (compared to homogeneous Catholic). Respective regressions are displayed in columns (1) and (6) of Tables 1.6.1 and 1.6.2.

To better illustrate the U-Shaped form, we drop the intercept of each regression and instead use the full set of indicators of religious denomination. To test differences between Catholicism and Protestantism, we provide the p-value of the Waldtest testing the equality of coefficients between homogeneous Catholic and homogeneous Protestant localities. To test whether religious heterogeneity plays a role, we report the p-value of a Waldtest testing the Hypothesis(H_0) that prevalence of norm in homogeneous Catholic localities is equal to the prevalence of norm in heterogeneous Catholic localities and prevalence of norm in homogeneous Protestant localities is equal to the prevalence of the norm in heterogeneous Protestant localities. Note that the interpretation of the latter as actually rejecting the H_0 that heterogeneity does not have an effect on the relative size of the coefficients as this H_0 would also be rejected if there is a continuous decrease e.g. in the share of Catholics. Only if the coefficients indeed suggest a U-shaped structure, this actually indicates that norms are less prevalent in localities which are heterogeneous.

In columns (2) and (7) of Tables 1.6.1 and 1.6.2, we additionally include the population size of a locality, the distance to the next town and an indicator variable that is one if the locality is itself a town. As has been also noted in the literature, towns and cities tend to be more heterogeneous in terms of their religious composition (see e.g. Becker and Cinnirella, 2020). This does not change the U-Shaped structure of the effect on neighborhood obligations, nor the relationship between seclusion and Catholicism but it reduces the relationship between impurity norms and Protestantism, rendering it marginally insignificant (p-value=0.132).

Even though the inter-regional pattern of denomination is largely determined by past institutional boundaries, there may be a lot of other differences between the more heterogeneous Western strip, the Protestant North-East and the Catholic South-East that may be correlated with the diffusion of social norms. To account for that, we additionally include Province Fixed Effects in columns (3) and (8) of Tables 1.6.1 and 1.6.2. In terms of the effect of denomination, it only changes the results for the intensity of neighborhood help activities, which within provinces is stronger in homogeneous Protestant than in homogeneous Catholic localities. However, within provinces the effect on norms of seclusion and impurity norms become more U-shaped with heterogeneous villages being less likely to display these norms.

Given the granularity of our data, we can restrict regions even further by first including district fixed effects, and then county fixed effects. However, a note of caution is in need when interpreting the results. First, the smaller the region the less variation in the majority denomination, so it becomes difficult to detect differences between Catholic localities and Protestant localities. Further, it is not clear in this context that within region variation is less endogenous than cross-regional variation in religious composition. As argued above, the cross regional variation is almost exclusively driven by historic borders (that also did not fully persist), while it is unclear where the intra-regional variation (conditional on remoteness and population size) comes from. Nevertheless, we report the results for narrower fixed effects in column (4)-(5) and (9)-(10) to see whether effects hold within narrow regions. The results are mostly equivalent to previous specification.

The presence of other religious groups is, if anything, negatively related to the prevalence of social norms. However, the effect is not consistently statistically significant across specifications.

Tables 1.A.16 and 1.A.17 in the Appendix report results with standard errors adjusted for spatial autocorrelation instead of clustering. For computational efficiency, we do not include the full set of indicators, because we fit the reprojected model freed of the relevant fixed effects. All coefficients are relative to being homogeneous Catholic. The interpretation of results stays the same. Note that also significance levels are largely unaffected with one major exception - the U-shape in impurity norms turns insignificant at conventional levels ($p=0.149$).

To sum up: Whether social norms are related to denomination depends on the norm. Catholic villages tend to be more likely to display norms prescribing seclusion after birth, Protestants tend to display more intense neighborhood obligations. The latter is, however, less pronounced than the former. More consistently though: Heterogeneous localities are less likely to display any kind of norm. This is most pronounced for the cooperative norms, and least pronounced for the impurity norms.

Table 1.6.1. Social norms and religious composition: Neighborhood obligations

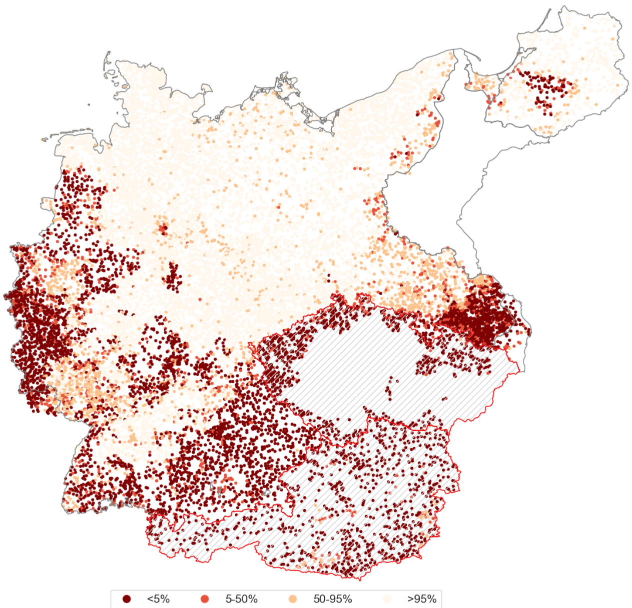
	Dependent variable:									
	Obligated					Sum nb help				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<5%	0.462***	0.407***	0.328***	0.527***	0.061**	2.169***	1.968***	1.383***	2.152***	0.329***
	(0.016)	(0.021)	(0.042)	(0.073)	(0.029)	(0.063)	(0.085)	(0.176)	(0.202)	(0.118)
5-50%	0.392***	0.356***	0.297***	0.494***	0.022	1.811***	1.708***	1.235***	1.999***	0.168
	(0.021)	(0.024)	(0.043)	(0.075)	(0.037)	(0.082)	(0.097)	(0.180)	(0.213)	(0.146)
50-95%	0.371***	0.331***	0.286***	0.495***	0.031	1.769***	1.639***	1.298***	2.105***	0.289*
	(0.015)	(0.020)	(0.039)	(0.075)	(0.038)	(0.067)	(0.089)	(0.164)	(0.218)	(0.150)
>95%	0.434***	0.386***	0.334***	0.548***	0.081**	2.055***	1.880***	1.552***	2.409***	0.581***
	(0.011)	(0.016)	(0.035)	(0.075)	(0.036)	(0.048)	(0.070)	(0.149)	(0.216)	(0.147)
Any sects	−0.025	−0.020	−0.010	−0.007	−0.002	0.012	0.046	0.064	0.082	0.095
	(0.019)	(0.019)	(0.019)	(0.019)	(0.020)	(0.079)	(0.079)	(0.079)	(0.078)	(0.078)
Any non-denominational	−0.083***	−0.074***	−0.020	−0.023	−0.006	−0.422***	−0.368***	−0.089	−0.084	−0.020
	(0.018)	(0.018)	(0.018)	(0.018)	(0.019)	(0.068)	(0.068)	(0.064)	(0.064)	(0.068)
Any Jewish	−0.029*	−0.006	−0.028*	−0.026	−0.019	−0.127**	0.009	−0.121*	−0.124**	−0.076
	(0.016)	(0.016)	(0.017)	(0.016)	(0.018)	(0.062)	(0.066)	(0.066)	(0.062)	(0.069)
p-value Prot vs. Cath	0.139	0.279	0.803	0.291	0.422	0.141	0.259	0.063	0.002	0.011
p-value Heterogeneity	0	0	0.008	0.002	0.008	0	0	0	0	0
Mean dep. var.	0.42	0.42	0.42	0.42	0.42	1.989	1.989	1.989	1.989	1.989
Controls		✓	✓	✓	✓		✓	✓	✓	✓
Province FE			✓					✓		
District FE				✓					✓	
County FE					✓					✓
Observations	12,410	12,410	12,410	12,407	12,410	12,410	12,410	12,410	12,407	12,410
R^2	0.007	0.012	0.062	0.085	0.169	0.009	0.016	0.070	0.097	0.192

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors clustered on the county level. Results using standard errors adjusted for spatial autocorrelation according to Conley (1999) is displayed in Table 1.A.16. Columns (2)-(5) and (7)-(10) additionally control population size of the locality as measured in 1910, distance to next town as well as an indicator whether the locality itself is a town. The full set of coefficients (excluding FE) of each regression are displayed in Table 1.A.14. Test Prot vs. Cath gives the p-value of the Wald Test with H_0: Prevalence of norm in localities with < 5% Protestants (vs. Catholics) = Prevalence of norm in localities with > 95% Protestants (vs. Catholics). Test Heterogeneity gives the p-value of the Wald Test with H_0: Prevalence of norm in localities with < 5% Protestants (vs. Catholics) = prevalence of norm in localities with 5 − 50% Protestants (vs. Catholics) and prevalence of norm in localities with 95% Protestants (vs. Catholics) = prevalence of norm in localities with 50 − 95% Protestants (vs. Catholics). Province FE implies controlling for Prussian Provinces, states of Austria and German Reich (except Prussia), voivodships of Poland, and the Regions Bohemia and Moravia for Czechia.

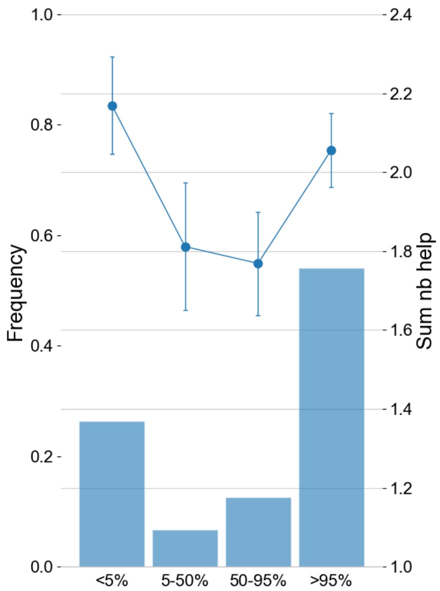
Table 1.6.2. Social norms and religious composition: Childbed norms

	Dependent variable:									
	Seclusion					Impurity				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<5%	0.750***	0.765***	0.650***	0.900***	0.731***	0.295***	0.347***	0.470***	0.324***	0.116***
	(0.010)	(0.018)	(0.063)	(0.028)	(0.031)	(0.014)	(0.021)	(0.053)	(0.038)	(0.031)
5-50%	0.669***	0.704***	0.604***	0.833***	0.681***	0.280***	0.340***	0.450***	0.283***	0.073**
	(0.020)	(0.024)	(0.064)	(0.033)	(0.032)	(0.018)	(0.024)	(0.053)	(0.040)	(0.033)
50-95%	0.584***	0.607***	0.538***	0.751***	0.624***	0.298***	0.348***	0.442***	0.275***	0.096***
	(0.017)	(0.022)	(0.063)	(0.035)	(0.025)	(0.016)	(0.022)	(0.050)	(0.042)	(0.024)
>95%	0.565***	0.576***	0.550***	0.757***	0.637***	0.327***	0.370***	0.478***	0.292***	0.114***
	(0.011)	(0.017)	(0.060)	(0.033)	(0.015)	(0.011)	(0.019)	(0.047)	(0.042)	(0.015)
Any sects	−0.020	−0.003	−0.011	−0.014	−0.027	−0.034**	−0.021	−0.018	−0.003	0.009
	(0.018)	(0.019)	(0.018)	(0.018)	(0.019)	(0.017)	(0.017)	(0.017)	(0.017)	(0.018)
Any non-denominational	−0.134***	−0.119***	−0.076***	−0.069***	−0.046**	0.050**	0.058***	0.008	0.006	0.009
	(0.017)	(0.018)	(0.018)	(0.018)	(0.018)	(0.020)	(0.020)	(0.017)	(0.016)	(0.017)
Any Jewish	0.013	0.056***	0.008	0.008	−0.001	−0.087***	−0.057***	−0.011	−0.026*	−0.031*
	(0.016)	(0.018)	(0.017)	(0.016)	(0.017)	(0.015)	(0.016)	(0.016)	(0.015)	(0.017)
p-value Prot vs. Cath	0	0	0	0	0	0.062	0.184	0.725	0.106	0.93
p-value Heterogeneity	0	0.001	0.061	0.003	0.04	0.199	0.409	0.056	0.057	0.079
Mean dep. var.	0.612	0.612	0.612	0.612	0.612	0.305	0.305	0.305	0.305	0.305
Controls		✓	✓	✓	✓		✓	✓	✓	✓
Province FE			✓					✓		
District FE				✓					✓	
County FE					✓					✓
Observations	12,216	12,216	12,216	12,213	12,216	12,216	12,216	12,216	12,213	12,216
R^2	0.035	0.040	0.075	0.094	0.177	0.006	0.010	0.061	0.102	0.192

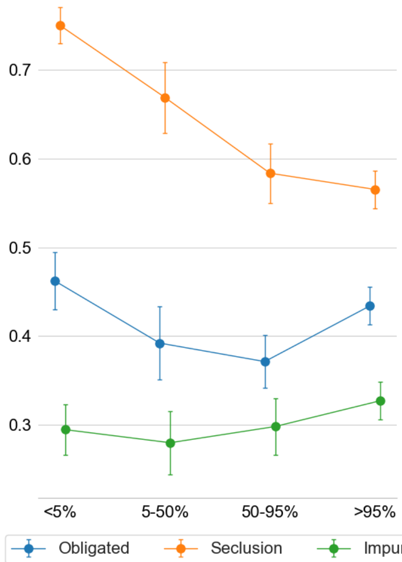
Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors clustered on the county level. Results using standard errors adjusted for spatial autocorrelation according to Conley (1999) is displayed in Table 1.A.17. Columns (2)-(5) and (7)-(10) additionally control population size of the locality as measured in 1910, distance to next town as well as an indicator whether the locality itself is a town. The full set of coefficients (excluding FE) of each regression are displayed in Table 1.A.15. Test Prot vs. Cath gives the p-value of the Wald Test with H_0: Prevalence of norm in localities with $<5\%$ Protestants (vs. Catholics) = Prevalence of norm in localities with $>95\%$ Protestants (vs. Catholics). Test Heterogeneity gives the p-value of the Wald Test with H_0: Prevalence of norm in localities with $<5\%$ Protestants (vs. Catholics) = prevalence of norm in localities with $5-50\%$ Protestants (vs. Catholics) and prevalence of norm in localities with 95% Protestants (vs. Catholics) = prevalence of norm in localities with $50-95\%$ Protestants (vs. Catholics). Province FE implies controlling for Prussian Provinces, states of Austria and German Reich (except Prussia), voivodships of Poland, and the Regions Bohemia and Moravia for Czechia.



(a) Spatial distribution of religious denomination: Share of protestants as opposed to Catholics.



(b) Distribution of religiuos denomination in the sampling region (left axis), effect on number of neighborhood activities (right axis)



(c) Effect of religious denomination on neighborhood obligations, seclusion after birth, and norms related to the impurity of mothers after giving birth.

Figure 1.6.1. Religious denomination and social norms.

Notes Figure 1.6.1(a) shows the share of protestants relative to the number of Catholics and Protestants in a given locality. Shaded regions represent regions that have low number of observations for which, we have population size and are thus excluded from the remaining analysis. Coefficients ploted in Figure 1.6.1(b) and 1.6.1(c) controlling for the presence for other religious groups (Jewish, Christian sects, non-denominational). Lines represent 95% Confidence Intervals based on county-level clustered standard errors. Religious composition based on Grober-Glück (1966a) categorization. Sum nb help = Number of neighborhood activities performed in a locality.

1.7 Discussion

Shared Political Institutional history. In the literature on political institutions and culture, political institutions are held to have an effect on culture by providing formal rules and enforcement mechanisms and by affecting the distribution of political power. This shapes the incentives and experiences that individuals and groups in a society face (see e.g. Alesina and Fuchs-Schündeln, 2007; Tabellini, 2008; Grosjean, 2011; Alesina and Giuliano, 2015; Becker, Boeckh, et al., 2016). The empirical strands of this literature mostly focus on the effect of historical political institutions on cultural traits related to cooperation between individuals, such as social and political trust (see e.g. Grosjean, 2011; Becker, Boeckh, et al., 2016; Buggle, 2016), prevalence of non-profit organizations and other measures of civic capital (Guiso, Sapienza, and Zingales, 2016), or preferences for redistribution (Alesina and Fuchs-Schündeln, 2007). However, the effect of institutions on social norms more generally is rarely explored. Our results suggest that not all social norms are generally sensitive to political institutions. However, in line with the literature we find that social norms related to cooperation are affected by past historical institutions. One reason for the discrepancy between norms related to cooperation and those related to beliefs about impurity may be that reciprocal exchange is much more substitutable through public governance and public good provision. As a result, social norms related to reciprocal exchange may depend more on the institutional structure than rules of seclusion after birth or rules related to beliefs of impurity. These latter rules may be more closely tied to the role of motherhood in a community, which in turn may be more related to religious beliefs and superstitions.

A potential threat to this interpretation is the high inter-relationship between political boundaries and religious denomination. One may, thus, be worried that the effect of time spent together in the same state does not operate through any features of institutions but through religion instead. To investigate that point, we split up our measure of years spent in the same state into years spent in the same state prior to the reformation and years spent together in post reformation time. As argued in the previous section, only post reformation borders should affect the religious composition. This is confirmed in column (1)of Table 1.7.1 which displays the result of regressing length spent in the same state pre and post reformation on the difference in the share of Protestants as opposed to Catholics between the localities as measured by the mid-points of the original intervals.[34] It shows that localities are more similar in their religious composition if they spent more time together in the same state after the reformation, but not before the reformation. Column (2) does the same exercise using the distance in neighborhood obligations between localities as an outcome. It shows that the effect of historical political institutions on neigh-

34. These bin the share of Protestants in the following intervals: 0-5, 5-30, 30-50, 50, 50-70, 70-95, and 95-100.

borhood obligations is driven by pre-reformation years and, thus, are unlikely to be driven by the effect of political institutions on religious composition. Column (3) shows the results for childbed norms. It shows that the small effect of political institutions on childbed norms is entirely driven by post-reformation periods. The results thus suggest that any relation between political institutions and childbed norms is likely driven by the former affecting religious denomination and not by something inherent in the institutions, while, as argued above, there may be some more direct link between historical political institutions and cooperative norms.

Table 1.7.1. Is the effect of institutional history driven by religious denomination?

	Dist Share Prot (1)	Dist Nbh. Obl (2)	Dist Childbed (3)
Yrs post Reformation	-3.794^{***}	-0.005	-0.008^{***}
	(0.205)	(0.004)	(0.003)
Yrs prior Reformation	2.625^{***}	-0.027^{***}	0.01^{***}
	(0.2)	(0.003)	(0.002)
Same state (1930)	-12.393^{***}	-0.017^{*}	0.022^{***}
	(0.661)	(0.01)	(0.008)
Distance (km)	0.102^{***}	-0.013^{***}	0.023^{***}
	(0.004)	(0.003)	(0.003)
Distance (km) sq	-0.142^{***}	0.001^{***}	-0.002^{***}
	(0.008)	(0.000)	(0.000)
Intercept	17.642^{***}	0.158^{**}	-0.255^{***}
	(3.815)	(0.076)	(0.054)
Observations	18,699,670	18,699,670	18,699,670
R^2	0.117	0.006	0.002

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors obtained using two-way clustering on each dyads localities. Distance Share Prot= Difference in the share of protestants between localities using the mid-point of each bin, Distance Nbh. Obligations=euclidean distance across all neighborhood norms (being obligated, performing neighborhood help upon death, at weddings, upon birth, upon sickness, with house building, in emergencies (fire/flood), with agricultural activities). Distance in childbed norm = distance across the three indicator variables: seclusion, impure, any other childbed norm. Years prior (post) Reformation is the sum of years between 1200 and 1515 (1515 and 1790) spent together in the same state divided by 50.

Local and joint predictors of social norms. In terms of strongly locally varying factors, we find that in particular population size, local population density as well as denominational homogeneity are highly predictive for the existence of norms in a community. While remoteness as measured by the distance to the nearest town is only strongly related to neighborhood obligations.

Population size and population density can, in principle, affect social norms through different channels. First, the respective norms may be more or less valuable in densely populated areas. In the case of neighborhood help obligations, this may, for instance, be due to increased access to markets in the form of specialized services.[35] Instead of relying on their neighbors for services, such as help in the kitchen during wedding festivities, or digging the grave, families may be more able to hire professionals to take over these services. That this type of market specialization crowded out specific forms of neighborhood help in certain localities is evidenced by spontaneous comments of respondents. For instance, in four localities respondents indicate that digging the grave is now done by a professional gravedigger.[36]

In the case of childbed norms, the relationship to markets is not as obvious. However, also for them some respondents noted that, for instance, going to the well was forbidden until the well was replaced by water pipes, evidencing the susceptibility of these norms to environmental factors potentially correlated with urbanization.

Second, population size and population density can affect social norms through changing the structure of social interactions by providing more options (see Wirth, 1938, for a seminal analysis of how urbanity shapes social structure). Members of more rural communities are more limited in their choices with whom to socialize and, thus, may interact more frequently. They may rely more on the other community members because of limited outside options increasing the ability of members in the community to exert social control over each other which enables the maintaining of social norms.[37] In cities and densely populated areas, however, individuals have a larger choice set of individuals to interact with.[38] Thus, they can more easily escape any type of social control by undermining effective social sanctioning of norm violations and thereby the ability of a community to maintain any social norm on the community-level.

We cannot strictly separate the two mechanisms and both of them are most likely at play at the same time. For neighborhood help obligations, we can, however, inves-

35. Note that this is the reverse effect of what Henrich, Boyd, et al. (2001) and Henrich, Ensminger, et al. (2010) argue for fairness, punishment of uncooperative behavior and cooperation in anonymous situations. They show that market interaction is positively associated to levels of cooperation and fairness in anonymous interaction and community size is positively associated with the willingness to engage in third party punishment. They argue that this squares well with a theory that argues that market interaction and cooperation in anonymous exchanges are complements. Our results complement theirs by suggesting the opposite effect for norms of cooperation in non-anonymous situations.

36. In Schönau im Gebirge (Lower Austria), Herbsleben (Thuringia), Biberachzell (Bavaria), and Kirchhaslach, (Bavaria)

37. The idea that social network closure and strong social ties enable more social control, e.g. through better monitoring and social sanctioning (gossip, shaming, exclusion), by others in a community is inherent in the social disorganization theory of crime (see e.g. Kubrin and Weitzer, 2003, and references therin) as well as in social capital theory (see e.g. Coleman, 1988)

38. In fact, Hawley (2012) investigates the effect of population density within urban areas in the United States on social interaction using an instrumental variable approach. He finds that local population density increases social interactions among friends, but decreases interactions among neighbors.

tigate if in more populated areas individuals have the type of reciprocal exchange with friends instead of neighbors because part e) of the neighborhood obligation question asks whether any other group is obligated to help. The results indicate that, indeed, friends serve partly as substitutes for neighbors in more populous localities (see Table 1.7.2). This suggests that reciprocal exchange among neighbors is not merely becoming less valuable through e.g. the access to markets but at least partly shifts to a self-chosen social group.[39]

Remoteness as measured by the distance to the nearest town or being a town does not relate to the likelihood that friends take over (part) of the reciprocal exchange indicating that the relationship between remoteness and norms does not operate through an increased reliance on the local community but through other factors such as market access or the spread of new ideas through towns.

39. Note that, however, the sample size drops substantially because we condition on having answered anything to question e) to ensure that people just did not answer the question because they did not read it because they thought it does not apply because the locality has no neighborhood obligations.

Table 1.7.2. Reciprocal help among friends

	Dependent variable:	
	Friends	Friends (no neighbors)
	(1)	(2)
Local pop 1910	0.005***	0.002*
	(0.002)	(0.001)
Dist town (km)	0.00002	−0.001
	(0.002)	(0.001)
Town	−0.003	−0.012
	(0.056)	(0.036)
Mean dep. var.	0.191	0.056
District FE	✓	✓
Observations	2,320	2,316
R^2	0.057	0.061

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors clustered on the county level. Results using standard errors adjusted for spatial autocorrelation according to Conley (1999) is displayed in Table 1.A.18. All specifications additionally contain an indicator variable that is one if the locality is in the 95th percentile of the population size distribution interacted with population size in 1910 and an indicator variable that is one if the locality is in the 95th percentile of the distance to next town distribution interacted with distance to next town distribution, the share of protestants binned as in the previous section, an indicator variables whether there are any Jewish, any Christian sects, any non-denominational present in the locality as well as an indicator of whether the locality is a town or city. Friends=Respondent indicated that friends help with the kind of activities that are typically comprised in neighborhood obligations. Friends (no neighbors)=Respondent indicated that friends and not neighbors help with the kind of activities that are typically comprised in neighborhood obligations.

In addition to population size and remoteness, our results show a strong relationship between social norms and denominational heterogeneity. Heterogeneity in denomination represented a major social cleavage historically. For instance, mixed-marriages between Protestants and Catholics were discouraged by the churches and were only permissible under strict constraints (Bendikowski, 2016). The social cleavage was also reflected in superstitions. People believed that the remains of people that were part of a mixed marriage were cursed (Hoffmann-Krayer and Bächtold-Stäubli, 1974, p. 179). Further, some people believed that if Catholics and Protestants met after church service a person in the village was going to die the next day (Hoffmann-Krayer and Bächtold-Stäubli, 1974, p.181). Protestants and Catholics disliked each other and used slurs for each other even centuries after the reformation (see Hoffmann-Krayer and Bächtold-Stäubli (1974, p.177-178) as well as Lorentz (2002-08-12)). A large part of outside family social life was happening in church or clubs aligned with the corresponding religious denominations (Bendikowski, 2016, p.208) – potentially also as a consequence of these animosities. This lack of social connections across-groups may lead to a reduction in the power to enforce norms across groups and, thus, prevalence of norms on a community-level (see Fearon and Laitin, 1996; Miguel and Gugerty, 2005; Alexander and Christia, 2011, for this mechanism in relation to cooperation).

An alternative channel through which denominational heterogeneity may affect the existence of certain social norms on a community level is through competing norms inherent in Catholicism and Protestantism impeding the coordination on one particular social norm on the community level (Jackson, Gelfand, and Ember, 2020). Given that we do not see strong differences in neighborhood help obligations between homogeneous Catholic and homogeneous Protestant localities, it seems unlikely that Protestants and Catholics inherently display differing norms with regard to neighborhood obligations. Thus, it is unlikely that competing norms explain the relationship between heterogeneity and neighborhood obligations. Competing norms may, however, explain the gradient in childbed norms, as norms of seclusion are consistently more prevalent in Catholic localities than in Protestant localities. Some, but few, respondents even explicitly reference this divide by naming different rules for Catholic women in childbed than for Protestant.[40]

Of course, our results may also be driven by unobserved heterogeneity. In terms of unobserved but relevant factors, migration seems to be the most likely. Religious heterogeneity may be caused by migration to the respective locality, for instance, at the border to Poland or other religious borders. In localities with a large share of new residents, it may be more difficult to sustain historically inherited norms. Indeed, spontaneous comments of respondents indicate that there is some divide between long-term residents and newer residents. In three localities neighborhood obligations only apply among long-term residents.[41] However, only one out of these three is heterogeneous with respect to religious denomination.

Implications for estimating the effect of Protestantism. A large body of literature examines the effect of Protestantism on different economic and social outcomes using the distance to Wittenberg as an instrument for the share of Protestants in a Prussian county (see e.g. Becker and Woessmann, 2008, 2009, 2018). Using our more granular data, we can re-investigate the validity of this instrument. Figure 1.7.1 displays the share of homogeneous Protestant localities in blue and the share of heterogeneous localities in orange. It shows that while indeed, within Prussia, the likelihood of a locality being homogeneous Protestant is strongly related to the distance to Wittenberg, it also shifts the likelihood of a locality being heterogeneous in terms of religious denomination.

40. For instance, in Kietz and Schönlanke (Grenzmark Posen-Westpreußen) as well as Südlohne (Oldenburg) only Catholic women in childbed are not allowed to leave the house. In Rieste (Hannover), Catholic women are not allowed to pass a crossroad, while no rules apply for Protestant women.

41. in Alxnupönen (East Prussia), Ösede (Hannover), Dürrenberg (Prussia-Saxony)

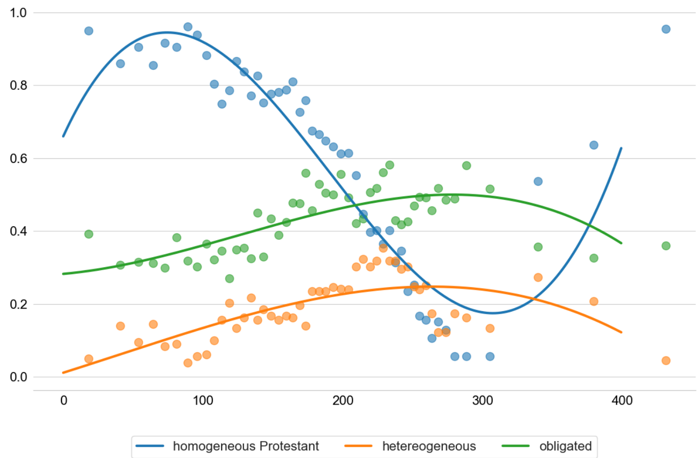


Figure 1.7.1. Religious composition and distance to Wittenberg

Notes: *x*-axis displays the distance to Wittenberg. Sample restricted to Prussia. Observations binned in 2% quantiles. Lines represent cubic fit of the data. homogeneous Protestant = more than 95% Protestant, heterogeneous= between 5 and 95% Protestants.

Our results thus stress the importance of being cautious about using widely used strategies (see e.g. Becker and Woessmann, 2008, 2009, 2018) to identify the effect of Protestantism if these identification strategies do not account for the effect on the overall religious composition. For instance, Becker and Woessmann (2018) use the distance to Wittenberg to identify the effects of Protestantism on suicides finding that Protestanism indeed increased suicide rates in Prussia. They argue that the effect works through Protestanism negatively affecting social cohesion suggesting that Protestantism may display "[...] a more individualistic and less community-oriented nature than Catholicism" (see Becker and Woessmann, 2018, p. 389). We do not find any strong evidence for this. The green line in Figure 1.7.1 represents the share of localities in which neighborhood help is obligatory as a function of the distance to Wittenberg. It perfectly co-moves with the share of heterogeneous localities but not with the share of homogeneously Protestant localities. However, if we were to estimate the share of Protestants on neighborhood help obligations using distance to Wittenberg as a linear instrument, we would have come to the same conclusion: the likelihood of neighborhood obligations decreases with the share protestants.

Origins of Childbed Norms, Religion and Gender Roles. Due to limitations in our data, we are unable to investigate the relationship between childbed norms and other gender-related values. Another question that remains unresolved is why norms of seclusion correlate with Catholicism, while what we define as impurity norms are not strongly related to religion at all. Yet, the GEA contains the first comprehensive documentation of these rules, which according to the ethnographic literature appear to have existed across cultures and religions for several centuries (Heller and Carrière, 2015, p. 6ff). Ploss (1876, p. 48) argues that already the old

Greek as well as Romans considered the woman in childbed to be impure. The old Greek forbade her to go into the temple or participating in any holy rituals without taking a purification bath; the Romans thought that the house of the woman in childbed is impure and individuals leaving the house needed to wash themselves. In her first ethnographic analysis of the GEA, Grober-Glück (1966b) also takes the stance of Ploss (1876) that the origin of these rules lies in the believe of the impurity of women after giving birth, which she says is, both, rooted in Christianity as well as folk superstitions but similarly to us distinguishes between rules that have protective motive and rules that have the purpose to circumvent harm (what we call for simplicity impurity norms).[42] Thus, one can speculate that the commonness of the impurity belief across religions and denominations is the reason behind the lower gradient in impurity norms than norms related to seclusion.

In general, however, these norms seem to be a complex mixture of profane rituals reinforced by Christian believes (see also Grober-Glück, 1966b). Given that the end of the restrictions for the women is typically marked by her first churchgoing, a connection to religion is clear also for rules whose content does not have a clear connection to Christian faith. We are not aware of any differences in religious beliefs between Protestantism and Catholicism itself that would explain differences in the prevalence of restrictions for women after birth. A potential explanation for the robust finding that Catholic localities are more likely to have restrictions after birth, in particular rules of seclusion, in 1930 than Protestant localities may instead stem from differences in religiousness and degree they still practice their faith between denominations. Protestants, at least, nowadays go to church less frequently than Catholics (Becker and Woessmann, 2018). We cannot judge to what degree this already holds for the 1930s, however, if indeed Protestants became faster detached from their church and church related customs than Catholics, which may explain a faster disappearance of these rules in Protestant localities. Another potential channel that could explain differences between Catholic communities and Protestant communities may be through its relationship to the share of females in formal employment. Wyrwich (2019) argues that Protestantism is related to industrialization which is in turn related to female labor force participation which makes it harder for women to seclude themselves after birth or give them more power and independence in the community to battle any type cultural discrimination. We lack granular data to really investigate this channel. However, we think it is unlikely that this strongly drives the results. First, the results hold when controlling for county level fixed effects, thus controlling for a wide-range of other regional differences. Second, when we compare the estimate of the effect of Catholicism on the prevalence of any restriction or seclusion norms, more specifically, with and without controlling for

42. Given the lack of context given by respondents on why the women should stay secluded, it is not completely clear that they always serve a protective motive.

industrialization measured on the county level, we do not find strong differences in estimated effect sizes across specifications (see Table 1.A.19 in the Appendix).

1.8 Conclusion

In this study, we presented a newly digitized dataset of social norms in around 16,500 German-speaking localities in Central Europe (the German Reich, Austria, Czechia, Poland, and Gdansk) around 1930. The data contains two sets of social norms, namely cooperation norms in the form of reciprocal neighborhood help obligations, and childbed norms that restrict women's behavior after giving birth. We show that these norms are not restricted to specific regions, but are instead widely distributed throughout the sampling region. However, there are some regions in which some norms are more universal than in others. We then set out to explain this spatial pattern in social norms. We found that local clusters are not well aligned with political institutional boundaries, and are instead likely the result of a more complex data-generating process that also depends on the specific norm type. While neighborhood obligations depend on institutional history, religious composition and remoteness of a locality, they are not as strongly influenced by the majority religious denomination. Childbed norms, on the other hand, are not strongly related to neither institutional history nor remoteness, but are more related to religious denomination. Population size, local population density, and religious heterogeneity are all joint and particularly localized predictors of social norms. Our results are in line with theories explaining the variation of social norms with variations in the ability to social sanction norm violations together with variation in the value a norm provides to a community, for instance, through variation in the access to markets, religious and political institutions.

References

Abramson, Scott F. 2017. "The Economic Origins of the Territorial State." *International Organization* 71 (1): 97–130. [6, 31]

Abramson, Scott F, David B. Carter, and Luwei Ying. 2022. "Historical Border Changes, State Building, and Contemporary Trust in Europe." *American Political Science Review* 116 (3): 875–95. [6, 31]

Alesina, Alberto, and Nicola Fuchs-Schündeln. 2007. "Good-Bye Lenin (or Not?): The Effect of Communism on People's Preferences." *American Economic Reivew* 97 (4): 1507–28. [7, 48]

Alesina, Alberto, and Paola Giuliano. 2015. "Culture and Institutions." *Journal of Economic Literature* 53 (4): 898–944. [7, 48]

Alesina, Alberto, Paola Giuliano, and Nunn Nathan. 2013. "On the Origins of Gender Roles: Women and the Plough." *Quarterly Journal of Economics* 128 (2): 469–530. [5]

Alexander, Marcus, and Fotini Christia. 2011. "Context Modularity of Human Altruism." *Science (New York, N.Y.)* 334 (6061): 1392–94. [41, 53]

Anselin, Luc, and Xun Li. 2019. “Operational Local Join Count Statistics for Cluster Detection.” *Journal of Geographical Systems* 21 (2): 189–210. [22, 24, 25, 27]

Aoki, Masahiko. 2001. *Toward a Comparative Institutional Analysis*. Cambridge, Massachusetts: MIT Press, 560. [35]

Barruzi-Leicher, Renate, and Gertrud Frauenknecht. 1966. “Nachbarschaft.” In *Atlas Der Deutschen Volkskunde/Erläuterungen Bd. 2. Zu Den Karten NF 37 - 48 Und 54 - 56c, 65 - 69d, 70 - 72c, 73 - 76a,b*. Edited by Matthias Zender. Düsseldorf: Droste Verlag. Chapter XVII, 279–337. [12, 13]

Becker, Sascha O, and Francesco Cinnirella. 2020. “Prussia Disaggregated: The Demography of Its Universe of Localities in 1871.” *Journal of Demographic Economics* 86 (3): 259–90. [43]

Becker, Sascha O, Jared Rubin, and Ludger Woessmann. 2021. “Religion in Economic History: A Survey.” In *The Handbook of Historical Economics*. Edited by Alberto Bisin and Giovanni Federico. Academic Press, 585–639. [8]

Becker, Sascha O., Katrin Boeckh, Christa Hainz, and Ludger Woessmann. 2016. “The Empire Is Dead, Long Live the Empire! Long-run Persistence of Trust and Corruption in the Bureaucracy.” *Economic Journal* 126 (590): 40–74. [7, 31, 48]

Becker, Sascha O., Steven Pfaff, and Jared Rubin. 2016. “Causes and Consequences of the Protestant Reformation.” *Explorations in Economic History* 62: 1–25. [42]

Becker, Sascha O., and Ludger Woessmann. 2008. “Luther and the Girls: Religious Denomination and the Female Education Gap in Nineteenth-century Prussia.” *Scandinavian Journal of Economics* 110 (4): 777–805. [8, 53, 54]

Becker, Sascha O., and Ludger Woessmann. 2009. “Was Weber Wrong? A Human Capital Theory of Protestant Economic History.” *Quarterly Journal of Economics* 124 (2): 531–96. [8, 42, 53, 54]

Becker, Sascha O., and Ludger Woessmann. 2018. “Social Cohesion, Religious Beliefs, and the Effect of Protestantism on Suicide.” *Review of Economics and Statistics* 100 (3): 377–91. [8, 53–55]

Bendikowski, Tillmann. 2016. *Der Deutsche Glaubenskrieg: Martin Luther, Der Papst Und Die Folgen*. First. München: c. Bertelsmann Verlag. [52]

Bicchieri, Cristina. 2006. *The Grammar of Society*. Cambridge University Press. [5]

Bicchieri, Cristina, Ryan Muldoon, and Alessandro Sontuoso. 2018. “Social Norms.” In *Stanford Encyclopedia of Philosophy*. Edited by Edward N Zalta. Winter 201. Metaphysics Research Lab, Stanford University. [5]

Buggle, Johannes C. 2020. “Growing Collectivism: Irrigation, Group Conformity and Technological Divergence.” *Journal of Economic Growth* 25 (2): 147–93. [5]

Buggle, Johannes C, and Ruben Durante. 2021. “Climate Risk, Cooperation and the Co-Evolution of Culture and Institutions.” *Economic Journal* 131 (637): 1947–87. [5]

Buggle, Johannes C. 2016. “Law and Social Capital: Evidence from the Code Napoleon in Germany.” *European Economic Review* 87 (8): 148–75. [7, 31, 48]

Burke, Mary A., and H. Peyton Young. 2011. “Chapter 8 - Social Norms.” In *Handbook of Social Economics*. Edited by Jess Benhabib, Alberto Bisin, and Matthew O. Jackson. Vol. 1, North-Holland, 311–38. [5]

Cantoni, Davide. 2015. “The Economic Effects of the Protestant Reformation: Testing the Weber Hypothesis in the German Lands.” *Journal of the European Economic Association* 13 (4): 561–98. [8, 42]

Cantoni, Davide, Jeremiah Dittmar, and Noam Yuchtman. 2016. "Reformation and Reallocation: Religious and Secular Economic Activity in Early Modern Germany." *CESifo Working Paper, No. 6218*, [8]

Clifford, P., and S. Richardson. 1985. "Testing the Association between Two Spatial Processes." *Statistics and Decisions* 2 (supp. issue): 155–60. [18, 20]

Coleman, James S. 1988. "Social Capital in the Creation of Human Capital." *American Journal of Sociology* 94 (Supplement): S95–S120. [50]

Conley, T.G. 1999. "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics* 92 (1): 1–45. [39, 40, 45, 46, 52]

Fearon, James D., and David D. Laitin. 1996. "Explaining Interethnic Cooperation." *American Political Science Review* 90 (4): 715–35. [41, 53]

Gelfand, Michele J., Jesse R. Harrington, and Joshua Conrad Jackson. 2017. "The Strength of Social Norms across Human Groups." *Perspectives on Psychological Science* 12 (5): 800–9. [5, 8]

Gelfand, Michele J., Jana L. Raver, Lisa Nishii, Lisa M. Leslie, Janetta Lun, Beng Chong Lim, Lili Duan, Assaf Almaliach, Soon Ang, Jakobina Arnadottir, Zeynep Aycan, Klaus Boehnke, Pawel Boski, Rosa Cabecinhas, Darius Chan, Jagdeep Chhokar, Alessia D'Amato, Montse Ferrer, Iris C. Fischlmayr, Ronald Fischer, Marta Fülöp, James Georgas, Emiko S. Kashima, Yoshishima Kashima, Kibum Kim, Alain Lempereur, Patricia Marquez, Rozhan Othman, Bert Overlaet, Penny Panagiotopoulou, Karl Peltzer, Lorena R. Perez-Florizno, Larisa Ponomarenko, Anu Realo, Vidar Schei, Manfred Schmitt, Peter B. Smith, Nazar Soomro, Erna Szabo, Nalinee Taveesin, Midori Toyama, Evert Van De Vliert, Naharika Vohra, Colleen Ward, and Susumu Yamaguchi. 2011. "Differences between Tight and Loose Cultures: A 33-Nation Study." *Science* 332 (6033): 1100–4. [5, 8]

Grober-Glück, Gerade. 1985. "Der Erste Kirchgang der Wöchnerin." In *Atlas Der Deutschen Volkskunde. Erläuterungen Bd. 3, Zu Den Karten NF 50 - 53, 56d - 58, 64, 72d, 76c - d, 77 - 84a - d*. Edited by Matthias Zender. Marburg: Elwert. Chapter XXVII, 105–44. [16]

Grober-Glück, Gerda. 1966a. "Verbreitung religiöser Gruppen." In *Atlas Der Deutschen Volkskunde. Erläuterungen Bd. 2, Zu Den Karten NF 37 - 48 Und 54 - 56c, 65 - 69d, 70 - 72c, 73 - 76a,b*. Chapter XVIII, 318–38. [41, 47]

Grober-Glück, Gerda. 1966b. "Volksglaubenvorstellung über die Wöchnerin." In *Atlas Der Deutschen Volkskunde. Erläuterungen Bd. 2, Zu Den Karten NF 37 - 48 Und 54 - 56c, 65 - 69d, 70 - 72c, 73 - 76a,b*. Edited by Matthias Zender. Chapter XXII, 457–523. [14, 16, 17, 55]

Grosjean, Pauline. 2011. "The Weight of History on European Cultural Integration: A Gravity Approach." *American Economic Review* 101 (3): 504–8. [7, 31, 48]

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2016. "Long-Term Persistence." *Journal of the European Economic Association* 14 (6): 1401–36. [7, 31, 48]

Hawley, Zackary B. 2012. "Does Urban Density Promote Social Interaction? Evidence from Instrumental Variable Estimation." *Review of Regional Studies* 42 (3): [50]

Heller, Angela, and Beate Carrière. 2015. *Nach Der Geburt: Wochenbett Und Rückbildung*. 2. Auflage. Thieme. [54]

Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, and Richard McElreath. 2001. "In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies." *American Economic Review* 91 (2): 73–78. [7, 50]

Henrich, Joseph, Jean Ensminger, Richard McElreath, Abigail Barr, Clark Barrett, Alexander Bolyanatz, Juan Camilo Cardenas, Michael Gurven, Edwins Gwako, Natalie Henrich, Carolyn Lesorogol, Frank Marlowe, David Tracer, and John Ziker. 2010. "Markets, Religion, Community Size, and the Evolution of Fairness and Punishment.” *Science (New York, N.Y.)* 327 (5972): 1480–84. arXiv: 1011.1669v3. [7, 50]

Herzig, Arno. 2000. “Die Rekatholisierung in Deutschen Territorien Im 16. Und 17. Jahrhundert.” *Geschichte Und Gesellschaft* 26 (1): 76–104. [42, 43]

Hoffmann-Krayer, Eduard, and Hanns Bächtold-Stäubli, editors. 1974. *Handwörterbuch des deutschen Aberglaubens. Band 5: Knoblauch - Matthias*. Berlin, New York: De Gruyter. [52]

Iyer, P. V. Krishna. 1949. “The First and Second Moments of Some Probability Distributions Arising from Points on a Lattice and Their Application.” *Biometrika* 36 (1/2): 135. [21, 22]

Jackson, Joshua Conrad, Michele Gelfand, and Carol R. Ember. 2020. “A Global Analysis of Cultural Tightness in Non-Industrial Societies.” *Proceedings of the Royal Society B: Biological Sciences* 287 (1930): [5, 8, 53]

Kehren, Georg. 1994. “Möglichkeiten und Grenzen der computativen Auswertung von Daten des Atlas der Deutschen Volkskunde (ADV).” In *Bonner Kleine Reihe Zur Alltagskultur*. Edited by H.L. Cox, Hildegard Mannheims, Peter Oberem, and Adelheid Schrutka-Rechtenstamm. First. Bonn: Rheinische Vereinigung für Volkskunde, 277. [9]

Klein, Kurt. 2016. “Salzburg.” In *Historisches Ortslexikon:* Statistische Dokumentation Zur Bevölkerungs- Und Siedlungsgeschichte. [35]

Kramer, Karl-Sigismund. 1954. *Die Nachbarschaft als bäuerliche Gemeinschaft: Ein Beitrag zur rechtlichen Volkskunde mit besonderer Berücksichtigung Bayerns*. Bayerische Heimatforschung. München-Pasing: Verl. Bayer. Heimatforschung. [12]

Kranton, Rachel E. 1996. “Reciprocal Exchange: A Self-Sustaining System.” *American Economic Review* 86 (4): 830–51. [35]

Kubrin, Charis E., and Ronald Weitzer. 2003. “New Directions in Social Disorganization Theory.” *Journal of Research in Crime and Delinquency* 40 (4): 374–402. [50]

Labouvie, Eva. 1992. “Selbstverwaltete Geburt: Landhebammen Zwischen Macht Und Regrlementierung (17. - 19. Jahrhundert).” *Geschichte und Gesellschaft 18*, 477–506. [15]

Legros, Sophie, and Beniamino Cislaghi. 2020. “Mapping the Social-Norms Literature: An Overview of Reviews.” *Perspectives on Psychological Science* 15 (1): 62–80. [5]

Lorentz, Frank. 2002-08-12. “Die geteilte Kleinstadt.” *Welt am Sonntag*, [52]

Lowes, Sara. 2020. “Ethnographic and Field Data in Historical Economics.” Cambridge, MA. arXiv: 1011.1669v3. [8]

McCloskey, Donald N. 1991. “The Prudent Peasant: New Findings on Open Fields.” *Journal of Economic History* 51 (2): 343–55. [5]

Miguel, Edward, and Mary Kay Gugerty. 2005. “Ethnic Diversity, Social Sanctions, and Public Goods in Kenya.” *Journal of Public Economics* 89: 2325–68. [41, 53]

MPIDR, and CGG. 2013. “MPIDR Population History GIS Collection – Europe (Partly Based on EuroGeographics for the Administrative Boundaries).” [11, 28]

Nowottnick, Georg. 1935. *Geburt, Hochzeit, Tod in Sitte, Brauch und Volksdichtung*. Berlin: Weidmann. [15]

Osorio, Felipe, Ronny Vallejos, and Francisco Cuevas. 2016. “SpatialPack: Computing the Association Between Two Spatial Processes.” arXiv: 1611.05289 [stat]. [20]

Ploss, Hermann Heinrich. 1876. *Das Kind in Brauch und Sitte der Völker*. Vol. 1, Stuttgart: August Auerbach. [15, 54, 55]

Poteete, Amy R., and Elinor Ostrom. 2008. “Fifteen Years of Empirical Research on Collective Action in Natural Resource Management: Struggling to Build Large-n Databases Based on Qualitative Research.” *World Development* 36 (1): 176–95. [8]

Rey, Sergio J., and Luc Anselin. 2010. “PySAL: A Python Library of Spatial Analytical Methods.” *Review of Regional Studies* 37 (1): 5–27. [22]

Roesel, Felix. 2022. “The German Local Population Database (GPOP), 1871 to 2019.” *Jahrbücher für Nationalökonomie und Statistik*, (8): [36]

Satyanath, Shanker, Nico Voigtländer, and Hans-Joachim Voth. 2017. “Bowling for Fascism: Social Capital and the Rise of the Nazi Party.” *Journal of Political Economy* 125 (2): 478–526. [8]

Schmoll, Friedemann. 2009. *Die Vermessung Der Kultur. Der “Atlas Der Deutschen Volkskunde” Und Die Deutsche Forschungsgemeinschaft 1928-1980*. Edited by Rüdiger vom Bruch, Ulrich Herbert, and Patrick Wagner. First. Vol. 5, Stuttgart: Franz Steiner Verlag, 331. [9, 12]

Tabellini, Guido. 2008. “The Scope of Cooperation: Values and Incentives.” *Quarterly Journal of Economics* 123 (3): 905–50. [7, 48]

Weber, Max. 1922. “Nachbarschaftsgemeinschaft, Wirtschaftsgemeinschaft Und Gemeinde.” In *Wirtschaft Und Gesellschaft - Grundriss Der Verstehenden Soziologie*. Fifth. Tübingen: Mohr Siebeck. Chapter 2.2. [12]

Wirth, Louis. 1938. “Urbanism as a Way of Life.” *American Journal of Sociology* 44 (1-24): [50]

Wurzbacher, Gerhard. 1961. *Das Dorf Im Spannungsfeld Industrieller Entwicklung. Untersuchung an Den 45 Dörfern Und Weilern Einer Westdeutschen Ländlichen Gemeinde*. Second. Stuttgart: Enke, 307. [12, 35]

Wyrwich, Michael. 2019. “Historical and Current Spatial Differences in Female Labour Force Participation: Evidence from Germany.” *Papers in Regional Science* 98 (1): 211–39. [55]

Zender, Matthias, and Günter Wiegelmann. 1959. “Technische Einweisung in die neue Folge des ADV.” In *Atlas Der Deutschen Volkskunde/Erläuterungen Bd. 1. Zu Den Karten NF 1 - 36*. Edited by Matthias Zender. Düsseldorf: Droste Verlag. Chapter II, 17–21. [12]

Appendix 1.A Additional Material

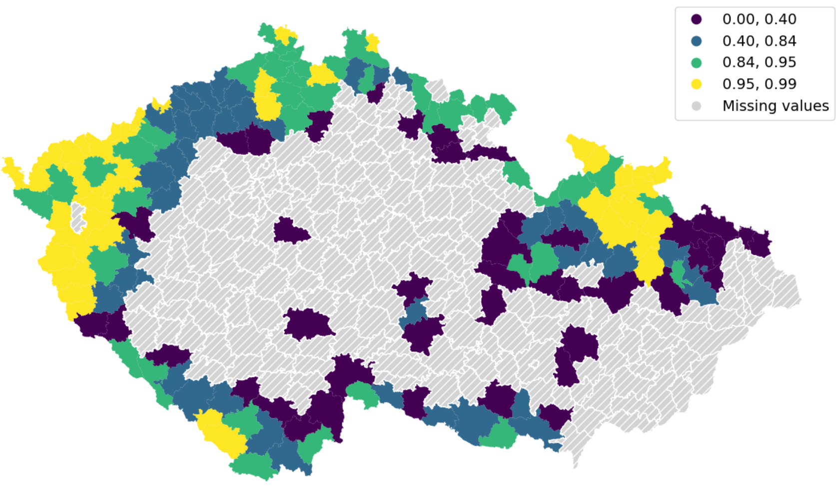
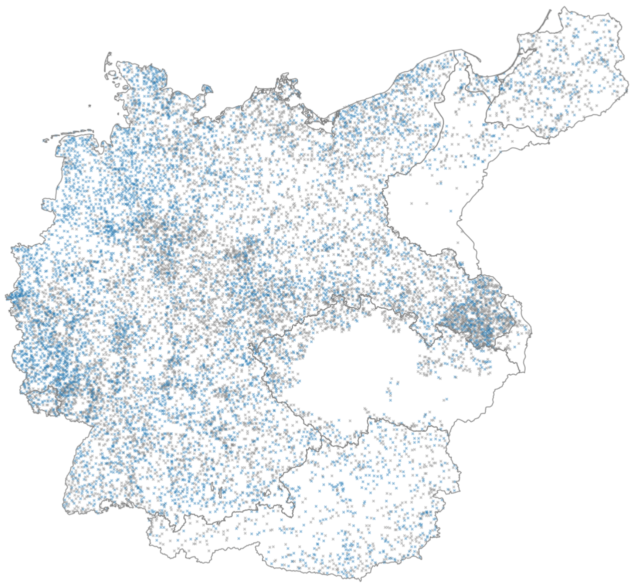
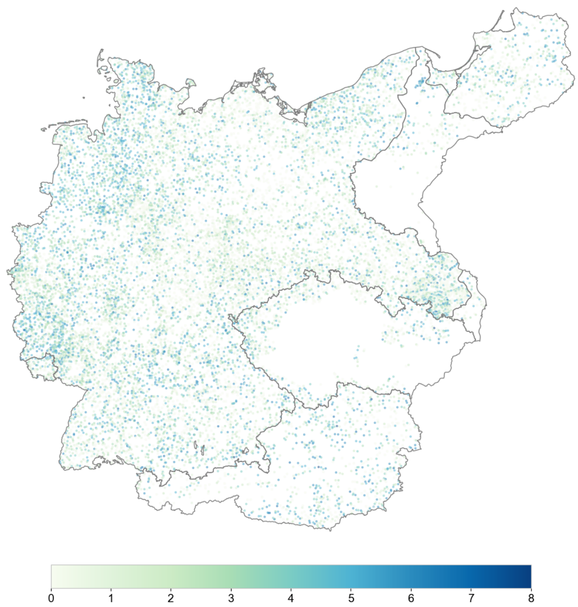


Figure 1.A.1. Share of ethnic Germans in Czechia county level, 1930.

Notes: Share of ethnic Germans in Czechia county level, 1930. Sample restricted to sampling area of the GEA. Data obtained from Jíchová, Soukup, Nemeskal, Pospísilová, and Svoboda (2014).



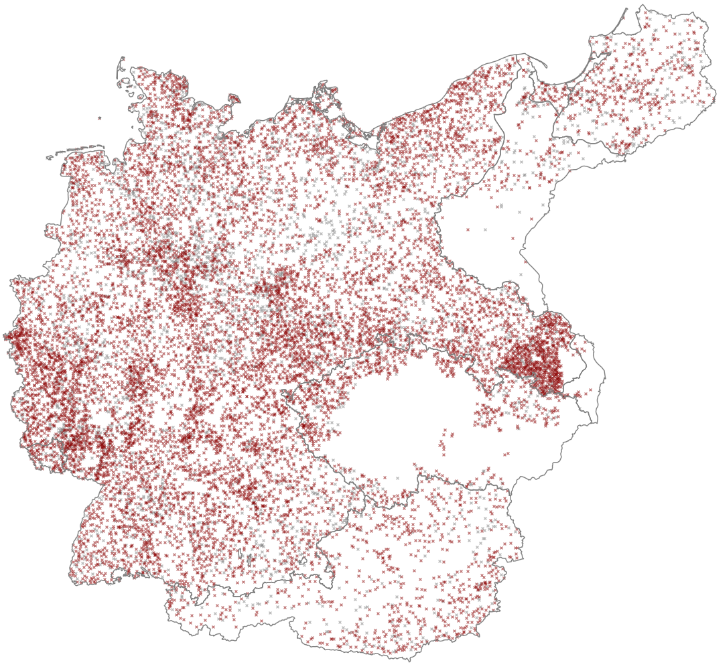
(a) Obligated to help the neighbors



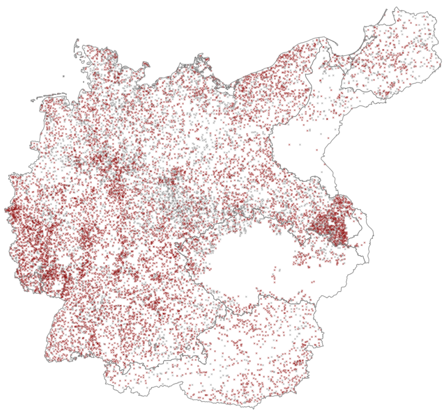
(b) Sum of neighborhood help activities

Figure 1.A.2. Spatial Distribution of Social Norms: Neighborhood help obligations

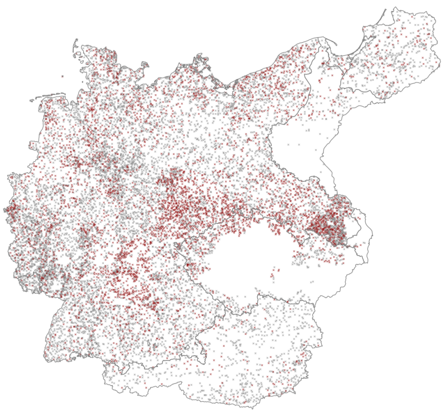
Notes: In Figure a: Blue crosses indicate the existence of an obligation in the locality; grey crosses indicate that the norm does not exist in the locality.



(a) Any restriction after birth



(b) Seclusion norms



(c) Impurity norms

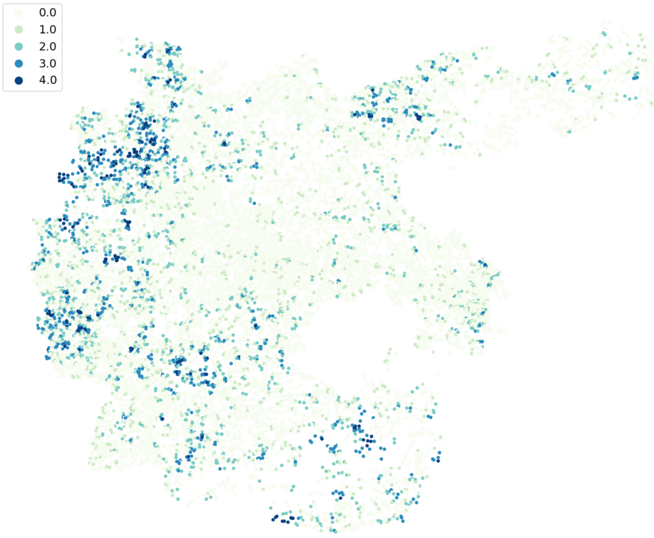
Figure 1.A.3. Spatial Distribution of Social Norms: Restrictions for mothers after birth

Notes: Red crosses indicate the existence of the childbed norm in the locality; grey crosses indicate that the norm does not exist in the locality.

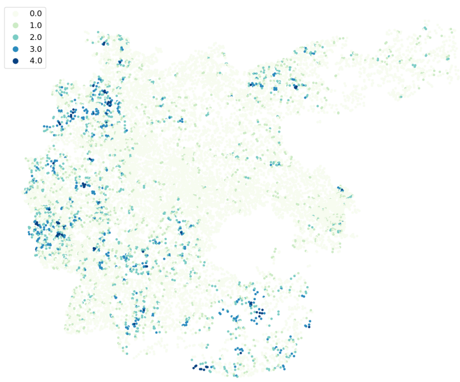
Table 1.A.1. Spatial autocorrelation in social norms (Distance-based weighting)

	Moran's I	Phi	p-value
Obligated		0.11	$p < 0.01$
Sum nb help	0.12		$p < 0.01$
Nb help at death		0.12	$p < 0.01$
Nb help at wedding		0.12	$p < 0.01$
Nb help at birth		0.06	$p < 0.01$
Nb help when sick		0.07	$p < 0.01$
Nb help with house building		0.12	$p < 0.01$
Nb help with agriculture		0.03	$p < 0.01$
Nb help in emergencies		0.04	$p < 0.01$
Any rule (mother)		0.06	$p < 0.01$
Seclusion		0.11	$p < 0.01$
Impurity		0.13	$p < 0.01$

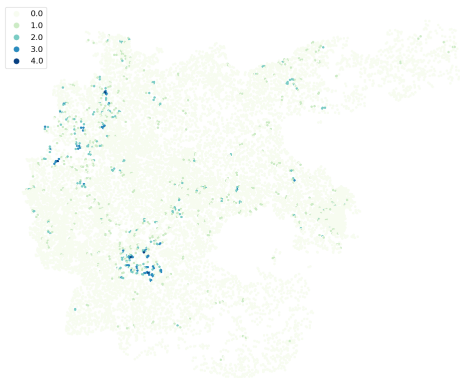
Notes: Table displays Moran's I for (standardized) continuous variable and Φ-coefficients of the contingency tables of the Spatial Joint Counts. The table reports p-values of tests testing the H_0 of spatial independence. For continuous variables this is tested via permutation tests, for binary variables it is tested via the Jount Count test developed by Iyer (1949). All statistics as implemented by Rey and Anselin (2010) using distance-based weighting (uniform, within 20km radius).



(a) Co-location regions: Any neighborhoood help activities & any restriction after birth



(b) Co-location regions: Any neighborhood help activities & rules of seclusion after birth



(c) Co-location regions: Any neighborhood help activities & rules of related to impurity

Figure 1.A.4. Joint Spatial Distribution of Social Norms: Co-Location Areas of Any Neighborhood help activity and childbed norms

Notes: Colored according to Join Count of co-locations based on four nearest neighbors spatial weighting (excluding points with no neighbors within 10 km). Joint Count of Co-Location is at the maximum of four if the locality itself displays both norms as well as all its neighbors, it is three, two and one if three of four, two of four, and one of four neighbors displays both norms, respectively. It is 0 if none of its neighbors displays both norms or itself does not display both norms or both.

Figure 1.A.5. Districts, 1930

Notes: Districts are obtained from MPIDR and CGG (2013).

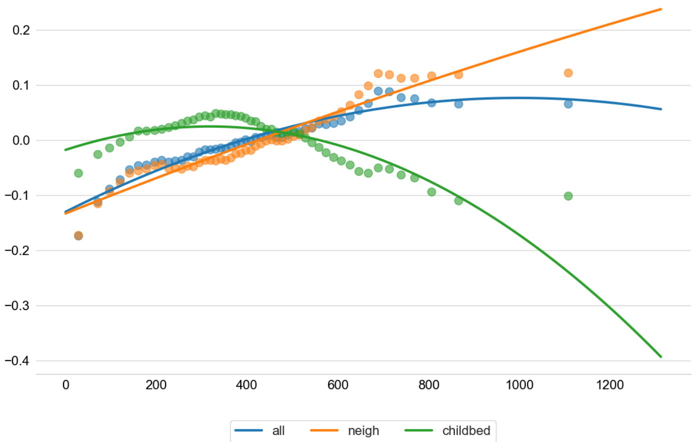


Figure 1.A.6. Binscatter of cultural distance as a function of physical distance

Notes: Values are binned in 2% quantiles of the physical distance. Lines represent a quadratic fit. all=euclidean distance of all norm meausures, neigh=euclidean distance across all neighborhood norms (being obligated, performing neighborhood help upon death, at weddings, upon birth, upon sickness, with house building, in emergencies (fire/flood), with agricultural activities). childbed = distance across the three indicator variables: seclusion, impure, any other childbed norm.

Table 1.A.2. Length of being in the same country & similarities in Norms – interaction with less than 100 km away

	Distance All			Distance Nbh. Obligations			Distance Childbed Norms		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Yrs same state (wgt.)	-0.019^{***}	-0.018^{***}	-0.013^{***}	-0.012^{***}	-0.011^{***}	-0.007^{***}	-0.002	-0.003	-0.003^{*}
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
Same state, 1930	-0.02^{**}	-0.02^{**}	-0.016	-0.013	-0.013	-0.01	0.024^{***}	0.024^{***}	0.023^{***}
	(0.01)	(0.01)	(0.01)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)
Distance $<$ 100km	-0.054^{**}	-0.02	-0.081^{***}	-0.087^{***}	-0.052^{***}	-0.109^{***}	-0.06^{***}	-0.072^{***}	-0.055^{***}
	(0.023)	(0.021)	(0.023)	(0.021)	(0.019)	(0.021)	(0.014)	(0.016)	(0.014)
Yrs same state (wgt.)$\times$dist $<$ 100km		-0.005			-0.005^{*}			0.002	
		(0.003)			(0.003)			(0.002)	
Ever same state			-0.059^{***}			-0.05^{***}			0.011^{*}
			(0.007)			(0.007)			(0.006)
Distance (km)	-0.01^{***}	-0.01^{***}	-0.01^{***}	0.002	0.003	0.003	0.021^{***}	0.021^{***}	0.021^{***}
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Distance (km) sq	0.001^{***}	0.001^{***}	0.001^{***}	0.000^{**}	0.000^{**}	0.000^{*}	-0.002^{***}	-0.002^{***}	-0.002^{***}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Intercept	0.147^{*}	0.144^{*}	0.162^{**}	-0.046	-0.049	-0.032	-0.247^{***}	-0.246^{***}	-0.25^{***}
	(0.076)	(0.076)	(0.074)	(0.1)	(0.1)	(0.099)	(0.054)	(0.054)	(0.054)
Observations	18,699,670	18,699,670	18,699,670	18,699,670	18,699,670	18,699,670	18,699,670	18,699,670	18,699,670
R^2	0.005	0.005	0.006	0.003	0.003	0.003	0.002	0.002	0.002

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors obtained using two-way clustering on each dyad's localities. All=euclidean distance of all norm meausures, Distance Nbh. Obligations=euclidean distance across all neighborhood norms (being obligated, performing neighborhood help upon death, at weddings, upon birth, upon sickness, with house building, in emergencies (fire/flood), with agricultural activities). Distance in childbed norm = distance across the three indicator variables: seclusion, impure, any other childbed norm. Yrs same state is a contiunuous measure that varies between 0 and 7 with 7 indicating that the two localities were always together between 1200 and 1790. All specifications additionally include a linear interaction between distance and an indicator variable of being closer than 100 km, as well as, an indicator variable being farther away than 800 km to match the pattern in the relationship to physical distance displayed in Figure 1.A.6. In addition all specifications additionally control for the distance in an indicator variable of being part of an imperial city or independent city.

Table 1.A.3. Length of being in the same country & similarities in Norms – different physical distance function

	Distance All			Distance Nbh. Obligations			Distance Childbed Norms		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Yrs same state (wgt.)	−0.02***	−0.019***	−0.014***	−0.013***	−0.012***	−0.008***	−0.003	−0.003	−0.004**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
Same state, 1930	−0.02**	−0.02**	−0.017*	−0.014	−0.013	−0.011	0.024***	0.024***	0.023***
	(0.01)	(0.01)	(0.01)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)
Ever same state			−0.058***			−0.048***			0.012*
			(0.007)			(0.007)			(0.006)
Distance < 50km		−0.011			−0.023			−0.052***	
		(0.016)			(0.015)			(0.013)	
Yrs same state (wgt.)×dist < 50km		−0.004			−0.004			0.005**	
		(0.003)			(0.003)			(0.002)	
Distance (km)	−0.008***	−0.009***	−0.006**	0.006**	0.005*	0.008***	0.024***	0.023***	0.024***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Distance (km) sq	0.001***	0.001***	0.001***	0.000	0.000*	0.000	−0.002***	−0.002***	−0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Intercept	0.137*	0.141*	0.146**	−0.062	−0.056	−0.055	−0.26***	−0.255***	−0.262***
	(0.075)	(0.075)	(0.074)	(0.1)	(0.1)	(0.099)	(0.054)	(0.054)	(0.054)
Observations	18,699,670	18,699,670	18,699,670	18,699,670	18,699,670	18,699,670	18,699,670	18,699,670	18,699,670
R^2	0.005	0.005	0.006	0.003	0.003	0.003	0.002	0.002	0.002

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors obtained using two-way clustering on each dyad's localities. All=euclidean distance of all norm meausures, Distance Nbh. Obligations=euclidean distance across all neighborhood norms (being obligated, performing neighborhood help upon death, at weddings, upon birth, upon sickness, with house building, in emergencies (fire/flood), with agricultural activities). Distance in childbed norm = distance across the three indicator variables: seclusion, impure, any other childbed norm. Yrs same state is a contiunuous measure that varies between 0 and 7 with 7 indicating that the two localities were always together between 1200 and 1790. All specifications additionally include a linear interaction between distance and an indicator variable of being closer than 20 km, as well as, an indicator variable being farther away than 800 km to match the pattern in the relationship to physical distance displayed in Figure 1.A.6. In addition all specifications additionally control for the distance in an indicator variable of being part of an imperial city or independent city.

Table 1.A.4. Length of being in the same country & similarities in Norms – buffer of 10 km around borders

	Distance All			Distance Nbh. Obligations			Distance Childbed Norms		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Yrs same state (wgt.)	−0.017***	−0.016***	−0.012***	−0.009**	−0.008**	−0.005	−0.002	−0.002	−0.002
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
Same state, 1930	−0.032**	−0.032**	−0.029**	−0.022*	−0.022*	−0.019	0.031***	0.031***	0.032***
	(0.014)	(0.014)	(0.014)	(0.013)	(0.013)	(0.013)	(0.011)	(0.011)	(0.011)
Distance < 50km		0.13***			0.124***			0.004	
		(0.037)			(0.035)			(0.032)	
Yrs same state (wgt.)×dist < 50km		−0.026***			−0.024***			0.001	
		(0.005)			(0.005)			(0.005)	
Ever same state			−0.053***			−0.046***			−0.002
			(0.012)			(0.011)			(0.009)
Distance (km)	−0.023***	−0.023***	−0.022***	−0.008	−0.008	−0.007	0.024***	0.024***	0.024***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Distance (km) sq	0.002***	0.002***	0.002***	0.001***	0.001***	0.001***	−0.002***	−0.002***	−0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Intercept	0.041	0.038	0.073**	−0.004	−0.006	0.023	−0.043*	−0.043*	−0.042
	(0.029)	(0.029)	(0.031)	(0.027)	(0.027)	(0.028)	(0.024)	(0.024)	(0.027)
Observations	5,250,420	5,250,420	5,250,420	5,250,420	5,250,420	5,250,420	5,250,420	5,250,420	5,250,420
R^2	0.008	0.009	0.009	0.004	0.004	0.005	0.002	0.002	0.002

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors obtained using two-way clustering on each dyads localities. All=euclidean distance of all norm meausures, Distance Nbh. Obligations=euclidean distance across all neighborhood norms (being obligated, performing neighborhood help upon death, at weddings, upon birth, upon sickness, with house building, in emergencies (fire/flood), with agricultural activities). Distance in childbed norm = distance across the three indicator variables: seclusion, impure, any other childbed norm. Yrs same state is a contiunuous measure that varies between 0 and 7 with 7 indicating that the two localities were always together between 1200 and 1790. All specifications additionally include a linear interaction between distance and an indicator variable of being closer than 100 km, as well as, an indicator variable being farther away than 800 km to match the pattern in the relationship to physical distance displayed in Figure 1.A.6.

Table 1.A.5. Summary statistics: Local population variables

	Pop 1910	Dist. to next town (km)	Muni pop 1939	Local pop density 1939
5%	135	0.0	398.5	26.28
10%	178	2.47	643	32.74
25%	290	5.21	1474.5	48.56
50%	524	8.37	4405	79.48
75%	1043	12.44	10341	147.07
80%	1278	13.61	12681	177.33
90%	2374	17.38	23441	320.29
95%	4334	21.14	43424	573.4
max	2071257	48.47	4338756	4868.88
mean	2231.19	9.4	22408.38	167.59
std	30541.46	6.4	150737.34	324.19
N	13,531	13,531	10,171	10,171

Notes: Distribution of population and remoteness variabely. Local pop 1910 = Localities population 1910, Muni pop 1939=population 1939 in the area of 2011 municipality's boundary, local pop density 1939 = population per square kilometer 1939.

Table 1.A.6. Social norms and population size: Neighborhood obligations

	Dependent variable:									
	Obligated					Sum nb help				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Pop 1910	−0.001		0.0004	−0.003***	−0.002***	−0.008**		−0.001	−0.014***	−0.012***
	(0.001)		(0.001)	(0.001)	(0.001)	(0.003)		(0.003)	(0.003)	(0.003)
Pop 1910 (95th)	−0.114***		−0.030	−0.131***	−0.126***	−0.570***		−0.204*	−0.567***	−0.498***
	(0.023)		(0.029)	(0.027)	(0.029)	(0.099)		(0.123)	(0.110)	(0.120)
Pop 1910 × Pop 1910 (95th)	0.001		−0.0004	0.003***	0.002***	0.007**		0.001	0.014***	0.012***
	(0.001)		(0.001)	(0.001)	(0.001)	(0.003)		(0.003)	(0.003)	(0.003)
Dist town (km)		0.007***	0.005***	0.002**	0.002**		0.032***	0.021***	0.008*	0.006
		(0.001)	(0.001)	(0.001)	(0.001)		(0.005)	(0.006)	(0.005)	(0.005)
Dist town (km) (95th)		0.129	0.107	0.034	−0.024		1.029**	0.912**	0.518	0.294
		(0.079)	(0.079)	(0.074)	(0.085)		(0.404)	(0.407)	(0.403)	(0.400)
Town			−0.064**	−0.042*	−0.041			−0.300***	−0.247***	−0.308***
			(0.026)	(0.024)	(0.025)			(0.102)	(0.094)	(0.100)
Dist town (km) × Dist town (km) (95th)		−0.007**	−0.005	−0.001	0.001		−0.046***	−0.035**	−0.017	−0.011
		(0.003)	(0.003)	(0.003)	(0.003)		(0.016)	(0.016)	(0.016)	(0.016)
Constant	0.432***	0.352***	0.374***			2.063***	1.688***	1.825***		
	(0.010)	(0.011)	(0.014)			(0.044)	(0.048)	(0.065)		
Mean dep. var.	0.418	0.418	0.418	0.418	0.418	1.978	1.978	1.978	1.978	1.978
District				✓					✓	
County FE					✓					✓
Observations	12,993	12,993	12,993	12,990	12,993	12,993	12,993	12,993	12,990	12,993
R^2	0.003	0.006	0.008	0.083	0.167	0.005	0.008	0.011	0.093	0.190

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors clustered on the county level. Local pop 1910 (95th) is an indicator variable that is one if the locality is in the 95th percentile of the population size distribution. Dist to next town (km) (95th) is an indicator variable that is one if the locality is in the 95th percentile of the distance to next town distribution. Local population per 100 inhabitants. Town is an indicator variable that is one if the locality is defined as ‘Stadt’ (City/Town) in the official list of communities in the German Reich, 1910 (Gemeindeverzeichnis). Localities that could not be matched to this list, e.g. because they lie outside the German Reich, are classified as Towns if they lie within 2 km of a town / city and have more than 1,000 inhabitants.

Table 1.A.7. Social norms and population size: Childbed norms

	Dependent variable:									
	Seclusion					Impurity				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Pop 1910	−0.001*		−0.001	−0.003***	−0.002***	−0.003***		−0.003***	−0.003***	−0.002***
	(0.001)		(0.001)	(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)
Pop 1910 (95th)	−0.143***		−0.105***	−0.134***	−0.101***	−0.129***		−0.128***	−0.116***	−0.085***
	(0.024)		(0.031)	(0.029)	(0.028)	(0.019)		(0.026)	(0.023)	(0.026)
Pop 1910 × Pop 1910 (95th)	0.001*		0.001	0.002***	0.002***	0.003***		0.003***	0.003***	0.002***
	(0.001)		(0.001)	(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)
Dist town (km)		0.005***	0.002*	0.003***	0.002		0.0005	−0.003**	0.001	0.002
		(0.001)	(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)	(0.001)
Dist town (km) (95th)		0.115	0.093	0.128	0.128		−0.092	−0.123	−0.007	−0.029
		(0.094)	(0.095)	(0.083)	(0.093)		(0.085)	(0.085)	(0.078)	(0.091)
Town			−0.032	−0.010	−0.045*			−0.035	−0.025	−0.038*
			(0.026)	(0.024)	(0.025)			(0.022)	(0.021)	(0.021)
Dist town (km) × Dist town (km) (95th)		−0.008**	−0.005	−0.007**	−0.006*		0.001	0.004	−0.001	−0.0001
		(0.004)	(0.004)	(0.003)	(0.004)		(0.003)	(0.003)	(0.003)	(0.004)
Constant	0.629***	0.572***	0.606***			0.334***	0.302***	0.365***		
	(0.010)	(0.013)	(0.017)			(0.010)	(0.013)	(0.018)		
Mean dep. var.	0.611	0.611	0.611	0.611	0.611	0.303	0.303	0.303	0.303	0.303
District				✓					✓	
County FE					✓					✓
Observations	12,784	12,784	12,784	12,781	12,784	12,784	12,784	12,784	12,781	12,784
R^2	0.005	0.002	0.006	0.085	0.174	0.006	0.001	0.007	0.100	0.189

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors clustered on the county level. Local pop 1910 (95th) is an indicator variable that is one if the locality is in the 95th percentile of the population size distribution. Dist to next town (km) (95th) is an indicator variable that is one if the locality is in the 95th percentile of the distance to next town distribution. Local population per 100 inhabitants. Town is an indicator variable that is one if the locality is defined as 'Stadt' (City/Town) in the official list of communities in the German Reich, 1910 (Gemeindeverzeichnis). Localities that could not be matched to this list, e.g. because they lie outside the German Reich, are classified as Towns if they lie within 2 km of a town / city and have more than 1,000 inhabitants.

Table 1.A.8. Social norms and population size: Neighborhood obligation - population squared

	Dependent variable:					
	Obligated			Sum nb help		
	(1)	(2)	(3)	(4)	(5)	(6)
Pop 1910	0.004**	−0.002	−0.001	0.021**	−0.006	−0.003
	(0.002)	(0.002)	(0.002)	(0.009)	(0.008)	(0.008)
Pop 1910(sq)	−0.0001**	−0.00003	−0.00005	−0.001***	−0.0002	−0.0003
	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0002)
Dist town (km)	0.005***	0.002**	0.002**	0.021***	0.008*	0.005
	(0.001)	(0.001)	(0.001)	(0.005)	(0.005)	(0.004)
Town	−0.055**	−0.038	−0.037	−0.259**	−0.231**	−0.293***
	(0.026)	(0.024)	(0.025)	(0.102)	(0.094)	(0.100)
Constant	0.357***			1.733***		
	(0.016)			(0.073)		
Mean dep. var.	0.418	0.418	0.418	1.978	1.978	1.978
District		✓			✓	
County FE			✓			✓
Observations	12,993	12,990	12,993	12,993	12,990	12,993
R^2	0.009	0.083	0.167	0.012	0.093	0.190

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors clustered on the county level. All columns additionally contain an indicator variable that is one if the locality is in the 95th percentile of the population size distribution interacted with population size in 1910 as well as an indicator variable that is one if the locality is in the 95th percentile of the distance to next town distribution interacted with distance to next town distribution. Local population per 100 inhabitants. Town is an indicator variable that is one if the locality is defined as 'Stadt' (City/Town) in the official list of communities in the German Reich, 1910 (Gemeindeverzeichnis). Localities that could not be matched to this list, e.g. because they lie outside the German Reich, are classified as Towns if they lie within 2 km of a town / city and have more than 1,000 inhabitants.

Table 1.A.9. Social norms and population size: Childbed norms - population squared

	Dependent variable:					
	Seclusion			Impurity		
	(1)	(2)	(3)	(4)	(5)	(6)
Pop 1910	0.008***	0.001	0.002	−0.006***	−0.006***	−0.004**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Pop 1910(sq)	−0.0003***	−0.0001**	−0.0001*	0.0001	0.0001**	0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.00005)	(0.00005)
Dist town (km)	0.002*	0.003***	0.002	−0.003**	0.001	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Town	−0.017	−0.005	−0.042*	−0.038*	−0.028	−0.039*
	(0.027)	(0.025)	(0.025)	(0.022)	(0.021)	(0.021)
Constant	0.568***			0.375***		
	(0.019)			(0.019)		
Mean dep. var.	0.611	0.611	0.611	0.303	0.303	0.303
District		✓			✓	
County FE			✓			✓
Observations	12,784	12,781	12,784	12,784	12,781	12,784
R^2	0.008	0.086	0.175	0.008	0.101	0.189

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors clustered on the county level. All columns additionally contain an indicator variable that is one if the locality is in the 95th percentile of the population size distribution interacted with population size in 1910 as well as an indicator variable that is one if the locality is in the 95th percentile of the distance to next town distribution interacted with distance to next town distribution. Local population per 100 inhabitants. Town is an indicator variable that is one if the locality is defined as 'Stadt' (City/Town) in the official list of communities in the German Reich, 1910 (Gemeindeverzeichnis). Localities that could not be matched to this list, e.g. because they lie outside the German Reich, are classified as Towns if they lie within 2 km of a town / city and have more than 1,000 inhabitants.

Table 1.A.10. Social norms and population size: Neighborhood norms - Conley Standard Errors

	Dependent variable:					
	Obligated			Sum nb help		
	(1)	(2)	(3)	(4)	(5)	(6)
Pop 1910	0.0004	-0.003^{**}	-0.002^{**}	-0.001	-0.014^{***}	-0.012^{***}
	(0.002)	(0.001)	(0.001)	(0.006)	(0.004)	(0.004)
Dist town (km)	0.005^{***}	0.002^{*}	0.002^{**}	0.021^{***}	0.008	0.006
	(0.002)	(0.001)	(0.001)	(0.008)	(0.005)	(0.004)
Town	-0.064^{*}	-0.042	-0.041	-0.300^{**}	-0.247^{**}	-0.308^{***}
	(0.037)	(0.030)	(0.026)	(0.129)	(0.100)	(0.108)
Constant	0.374^{***}			1.825^{***}		
	(0.023)			(0.113)		
Mean dep. var.	0.418	0.418	0.418	1.978	1.978	1.978
District FE		✓			✓	
County FE			✓			✓

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors adjusted for spatial autocorrelation according to Conley (1999) using a distance cutoff of 50 km and a uniform kernel. All columns additionally contain an indicator variable that is one if the locality is in the 95th percentile of the population size distribution interacted with population size in 1910 as well as an indicator variable that is one if the locality is in the 95th percentile of the distance to next town distribution interacted with distance to next town distribution. The full set of coefficients of each regression are displayed in Table 1.A.7. Local population per 100 inhabitants. Town is an indicator variable that is one if the locality is defined as 'Stadt' (City/Town) in the official list of communities in the German Reich, 1910 (Gemeindeverzeichnis). Localities that could not be matched to this list, e.g. because they lie outside the German Reich, are classified as Towns if they lie within 2 km of a town / city and have more than 1,000 inhabitants.

Table 1.A.11. Social norms and population size: Childbed norms - Conley Standard Errors

	Dependent variable:					
	Seclusion			Impurity		
	(1)	(2)	(3)	(4)	(5)	(6)
Pop 1910	−0.001	−0.003***	−0.002**	−0.003***	−0.003***	−0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Dist town (km)	0.002	0.003***	0.002	−0.003	0.001	0.002*
	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Town	−0.032	−0.010	−0.045	−0.035	−0.025	−0.038
	(0.038)	(0.025)	(0.028)	(0.032)	(0.024)	(0.023)
Constant	0.606***			0.365***		
	(0.035)			(0.041)		
Mean dep. var.	0.611	0.611	0.611	0.303	0.303	0.303
District FE		✓			✓	
County FE			✓			✓

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors adjusted for spatial autocorrelation according to Conley (1999) using a distance cutoff of 50 km and a uniform kernel. All columns additionally contain an indicator variable that is one if the locality is in the 95th percentile of the population size distribution interacted with population size in 1910 as well as an indicator variable that is one if the locality is in the 95th percentile of the distance to next town distribution interacted with distance to next town distribution. The full set of coefficients of each regression are displayed in Table 1.A.7. Local population per 100 inhabitants. Town is an indicator variable that is one if the locality is defined as 'Stadt' (City/Town) in the official list of communities in the German Reich, 1910 (Gemeindeverzeichnis). Localities that could not be matched to this list, e.g. because they lie outside the German Reich, are classified as Towns if they lie within 2 km of a town / city and have more than 1,000 inhabitants.

Table 1.A.12. Social norms and population size: Neighborhood norms - Including local population density, 39

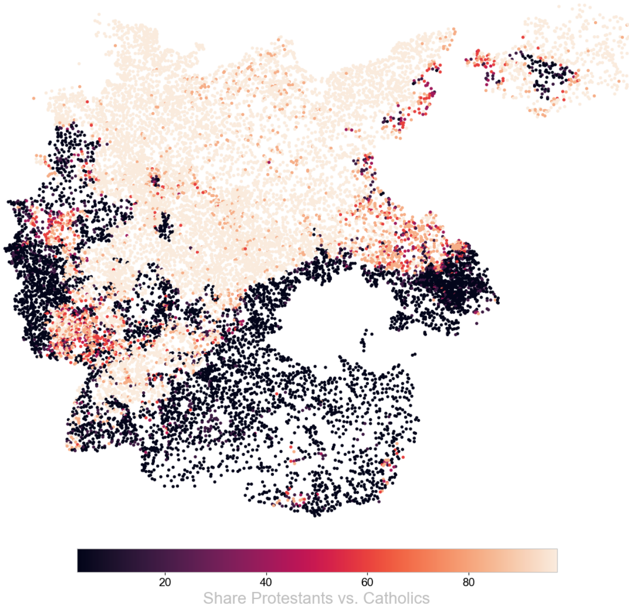
	Dependent variable:							
	Obligated				Sum nb help			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pop 1910	−0.0001	0.001	-0.002^{**}	-0.002^{***}	−0.002	0.006^{*}	-0.010^{***}	-0.011^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.004)	(0.004)	(0.003)	(0.003)
Dist town (km)	0.008^{***}	0.005^{***}	0.001	0.001	0.030^{***}	0.018^{***}	0.003	0.003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.007)	(0.006)	(0.006)	(0.006)
Town	-0.058^{**}	-0.096^{***}	-0.075^{***}	-0.051^{*}	-0.268^{**}	-0.475^{***}	-0.375^{***}	-0.325^{***}
	(0.028)	(0.028)	(0.027)	(0.027)	(0.114)	(0.110)	(0.103)	(0.110)
Local pop density 1939		-0.031^{***}	-0.030^{***}	-0.027^{***}		-0.180^{***}	-0.169^{***}	-0.124^{***}
		(0.008)	(0.007)	(0.007)		(0.031)	(0.025)	(0.025)
Constant	0.387^{***}	0.431^{***}			1.849^{***}	2.100^{***}		
	(0.016)	(0.019)			(0.074)	(0.084)		
Mean dep. var.	0.445	0.445	0.445	0.445	2.064	2.064	2.064	2.064
District			✓				✓	
County FE				✓				✓
Observations	9,728	9,728	9,727	9,728	9,728	9,728	9,727	9,728
R^2	0.013	0.018	0.095	0.184	0.016	0.024	0.118	0.211

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors clustered on the county level. All columns additionally contain an indicator variable that is one if the locality is in the 95th percentile of the population size distribution interacted with population size in 1910 as well as an indicator variable that is one if the locality is in the 95th percentile of the distance to next town distribution interacted with distance to next town distribution as well as an indicator variable that is one if the locality is in the 95th percentile of the local population density distribution interacted with population density. Local population (1910, 1939) per 100 inhabitants. Town is an indicator variable that is one if the locality is defined as 'Stadt' (City/Town) in the official list of communities in the German Reich, 1910 (Gemeindeverzeichnis). Localities that could not be matched to this list, e.g. because they lie outside the German Reich, are classified as Towns if they lie within 2 km of a town / city and have more than 1,000 inhabitants. Local population density, 1939, is defined as a 100 inhabitants per square kilometer within the boundaries of municipality of Germany from 2011 obtained from Roesel (2022).

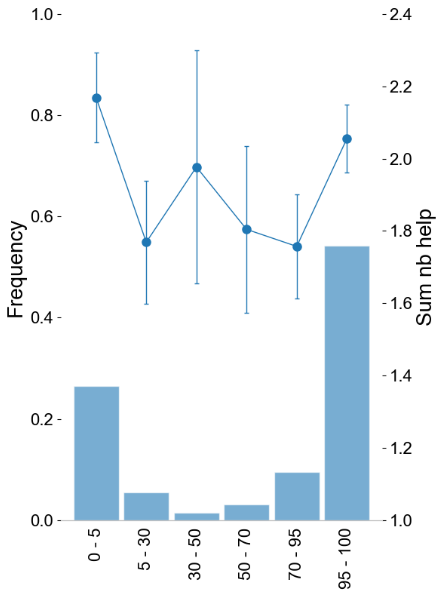
Table 1.A.13. Social norms and population size: Childbed norms - Including local population density, 39

	Dependent variable:							
	Seclusion				Impurity			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pop 1910	−0.0004	0.001	−0.002**	−0.002**	−0.004***	−0.004***	−0.003***	−0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Dist town (km)	0.002	0.001	0.002*	0.001	−0.005***	−0.005***	−0.0002	0.0002
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Town	−0.037	−0.071**	−0.031	−0.055**	−0.033	−0.040*	−0.040*	−0.048**
	(0.030)	(0.030)	(0.027)	(0.027)	(0.024)	(0.023)	(0.023)	(0.023)
Local pop density 1939		−0.018**	−0.025***	−0.015**		−0.002	−0.016***	−0.014**
		(0.008)	(0.006)	(0.006)		(0.008)	(0.006)	(0.006)
Constant	0.604***	0.632***			0.375***	0.379***		
	(0.020)	(0.020)			(0.021)	(0.021)		
Mean dep. var.	0.608	0.608	0.608	0.608	0.292	0.292	0.292	0.292
District				✓				
County FE					✓			
Observations	9,716	9,716	9,715	9,716	9,716	9,716	9,715	9,716
R^2	0.006	0.010	0.108	0.195	0.009	0.010	0.117	0.210

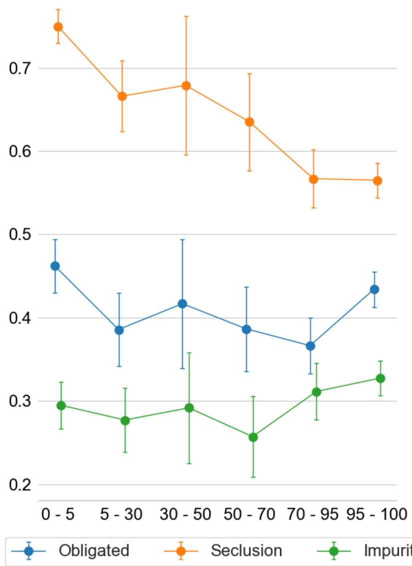
Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors clustered on the county level. All columns additionally contain an indicator variable that is one if the locality is in the 95th percentile of the population size distribution interacted with population size in 1910 as well as an indicator variable that is one if the locality is in the 95th percentile of the distance to next town distribution interacted with distance to next town distribution as well as an indicator variable that is one if the locality is in the 95th percentile of the local population density distribution interacted with population density. Local population (1910, 1939) per 100 inhabitants. Town is an indicator variable that is one if the locality is defined as 'Stadt' (City/Town) in the official list of communities in the German Reich, 1910 (Gemeindeverzeichnis). Localities that could not be matched to this list, e.g. because they lie outside the German Reich, are classified as Towns if they lie within 2 km of a town / city and have more than 1,000 inhabitants. Local population density, 1939, is defined as a 100 inhabitants per square kilometer within the boundaries of municipality of Germany from 2011 obtained from Roesel (2022).



(a) Spatial distribution of religious denomination: Share of protestants as opposed to Catholics.



(b) Distribution of religiuos denomination in the sampling region (left axis), effect on number of neighborhood activities (right axis)



(c) Effect of religious denomination on neighborhood obligations, seclusion after birth, and norms related to the impurity of mothers after giving birth.

Figure 1.A.7. Religious denomination and social norms, six groups

Figure (a) shows the share of protestants relative to the number of Catholics and Protestants in a given locality. Coefficients ploted inf Figure (b) and (bc) controlling for the presence for other religious groups (Jewish, Christian sects, non-denominational). Lines represent 95% Confidence Intervals.

Table 1.A.14. Social norms and religious composition: Neighborhood obligations

	Dependent variable:									
	Obligated					Sum nb help				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<5%	0.462***	0.407***	0.328***	0.527***	0.061**	2.169***	1.968***	1.383***	2.152***	0.329***
	(0.016)	(0.021)	(0.042)	(0.073)	(0.029)	(0.063)	(0.085)	(0.176)	(0.202)	(0.118)
5-50%	0.392***	0.356***	0.297***	0.494***	0.022	1.811***	1.708***	1.235***	1.999***	0.168
	(0.021)	(0.024)	(0.043)	(0.075)	(0.037)	(0.082)	(0.097)	(0.180)	(0.213)	(0.146)
50-95%	0.371***	0.331***	0.286***	0.495***	0.031	1.769***	1.639***	1.298***	2.105***	0.289*
	(0.015)	(0.020)	(0.039)	(0.075)	(0.038)	(0.067)	(0.089)	(0.164)	(0.218)	(0.150)
>95%	0.434***	0.386***	0.334***	0.548***	0.081**	2.055***	1.880***	1.552***	2.409***	0.581***
	(0.011)	(0.016)	(0.035)	(0.075)	(0.036)	(0.048)	(0.070)	(0.149)	(0.216)	(0.147)
Any sects	−0.025	−0.020	−0.010	−0.007	−0.002	0.012	0.046	0.064	0.082	0.095
	(0.019)	(0.019)	(0.019)	(0.019)	(0.020)	(0.079)	(0.079)	(0.079)	(0.078)	(0.078)
Any non-denominational	−0.083***	−0.074***	−0.020	−0.023	−0.006	−0.422***	−0.368***	−0.089	−0.084	−0.020
	(0.018)	(0.018)	(0.018)	(0.018)	(0.019)	(0.068)	(0.068)	(0.064)	(0.064)	(0.068)
Any Jewish	−0.029*	−0.006	−0.028*	−0.026	−0.019	−0.127**	0.009	−0.121*	−0.124**	−0.076
	(0.016)	(0.016)	(0.017)	(0.016)	(0.018)	(0.062)	(0.066)	(0.066)	(0.062)	(0.069)
Pop 1910		0.001	−0.002**	−0.002**	−0.002*		0.002	−0.012***	−0.011***	−0.010***
		(0.001)	(0.001)	(0.001)	(0.001)		(0.004)	(0.003)	(0.003)	(0.003)
Pop 1910 95th		0.024	−0.073**	−0.071**	−0.082***		0.022	−0.364***	−0.325***	−0.334**
		(0.031)	(0.029)	(0.029)	(0.031)		(0.132)	(0.116)	(0.118)	(0.135)
Pop 1910 × Pop 1910 95th		−0.001	0.002**	0.002**	0.002*		−0.002	0.011***	0.010***	0.009***
		(0.001)	(0.001)	(0.001)	(0.001)		(0.004)	(0.003)	(0.003)	(0.003)
Dist town (km)		0.005***	0.004***	0.002**	0.002**		0.018***	0.013***	0.007	0.005
		(0.001)	(0.001)	(0.001)	(0.001)		(0.006)	(0.005)	(0.005)	(0.005)
Dist town (km) 95th		0.054	−0.038	−0.010	−0.057		0.700*	0.328	0.352	0.006
		(0.080)	(0.072)	(0.075)	(0.089)		(0.397)	(0.390)	(0.391)	(0.433)
Dist town (km) × Dist town (km) 95th	−0.003	0.001	0.001	0.002		−0.028*	−0.014	−0.012	−0.0003	
	(0.003)	(0.003)	(0.003)	(0.004)		(0.016)	(0.016)	(0.016)	(0.018)	
Town		−0.065**	−0.024	−0.040	−0.035		−0.343***	−0.170*	−0.252**	−0.285***
		(0.027)	(0.025)	(0.025)	(0.027)		(0.106)	(0.099)	(0.098)	(0.109)
p-value Prot vs. Cath	0.139	0.279	0.803	0.291	0.422	0.141	0.259	0.063	0.002	0.011
p-value Heterogeneity	0	0	0.008	0.002	0.008	0	0	0	0	0
Mean dep. var.	0.42	0.42	0.42	0.42	0.42	1.989	1.989	1.989	1.989	1.989
Controls		✓	✓	✓	✓		✓	✓	✓	✓
Province FE			✓					✓		
District FE				✓					✓	
County FE					✓					✓
Observations	12,410	12,410	12,410	12,407	12,410	12,410	12,410	12,410	12,407	12,410
R^2	0.007	0.012	0.062	0.085	0.169	0.009	0.016	0.070	0.097	0.192

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors clustered on the county level. Results using standard errors adjusted for spatial autocorrelation according to Conley (1999) is displayed in Table 1.A.16. Columns (2)-(5) and (7)-(10) additionally control population size of the locality as measured in 1910, distance to next town as well as an indicator whether the locality itself is a town. Test Prot vs. Cath gives the p-value of the Wald Test with H_0: Prevalence of norm in localities with $< 5\%$ Protestants (vs. Catholics) = Prevalence of norm in localities with $> 95\%$ Protestants (vs. Catholics). Test Heterogeneity gives the p-value of the Wald Test with H_0: Prevalence of norm in localities with $< 5\%$ Protestants (vs. Catholics) = prevalence of norm in localities with $5-50\%$ Protestants (vs. Catholics) and prevalence of norm in localities with 95% Protestants (vs. Catholics) = prevalence of norm in localities with $50-95\%$ Protestants (vs. Catholics). Province FE implies controlling for Prussian Provinces, states of Austria and German Reich (except Prussia), voivodships of Poland, and the Regions Bohemia and Moravia for Czechia.

Table 1.A.15. Social norms and religious composition: Childbed Norms

	Dependent variable:									
	Seclusion					Impurity				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<5%	0.750***	0.765***	0.650***	0.900***	0.731***	0.295***	0.347***	0.470***	0.324***	0.116***
	(0.010)	(0.018)	(0.063)	(0.028)	(0.031)	(0.014)	(0.021)	(0.053)	(0.038)	(0.031)
5-50%	0.669***	0.704***	0.604***	0.833***	0.681***	0.280***	0.340***	0.450***	0.283***	0.073**
	(0.020)	(0.024)	(0.064)	(0.033)	(0.032)	(0.018)	(0.024)	(0.053)	(0.040)	(0.033)
50-95%	0.584***	0.607***	0.538***	0.751***	0.624***	0.298***	0.348***	0.442***	0.275***	0.096***
	(0.017)	(0.022)	(0.063)	(0.035)	(0.025)	(0.016)	(0.022)	(0.050)	(0.042)	(0.024)
>95%	0.565***	0.576***	0.550***	0.757***	0.637***	0.327***	0.370***	0.478***	0.292***	0.114***
	(0.011)	(0.017)	(0.060)	(0.033)	(0.015)	(0.011)	(0.019)	(0.047)	(0.042)	(0.015)
Any sects	−0.020	−0.003	−0.011	−0.014	−0.027	−0.034**	−0.021	−0.018	−0.003	0.009
	(0.018)	(0.019)	(0.018)	(0.018)	(0.019)	(0.017)	(0.017)	(0.017)	(0.017)	(0.018)
Any non-denominational	−0.134***	−0.119***	−0.076***	−0.069***	−0.046**	0.050**	0.058***	0.008	0.006	0.009
	(0.017)	(0.018)	(0.018)	(0.018)	(0.018)	(0.020)	(0.020)	(0.017)	(0.016)	(0.017)
Any Jewish	0.013	0.056***	0.008	0.008	−0.001	−0.087***	−0.057***	−0.011	−0.026*	−0.031*
	(0.016)	(0.018)	(0.017)	(0.016)	(0.017)	(0.015)	(0.016)	(0.016)	(0.015)	(0.017)
Pop 1910		−0.002***	−0.002**	−0.002**	−0.001		−0.003***	−0.003***	−0.003***	−0.002***
		(0.001)	(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)	(0.001)
Pop 1910 (95th)		−0.128***	−0.087***	−0.094***	−0.073**		−0.093***	−0.112***	−0.089***	−0.065**
		(0.031)	(0.032)	(0.030)	(0.031)		(0.028)	(0.026)	(0.026)	(0.029)
Dist town (km)		0.0004	0.002	0.002**	0.001		−0.002*	−0.0001	0.001	0.002
		(0.001)	(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)	(0.001)
Dist town (km) (95th)		0.016	0.035	0.086	0.073		−0.098	−0.082	0.009	−0.006
		(0.100)	(0.088)	(0.085)	(0.101)		(0.092)	(0.089)	(0.085)	(0.108)
Town		−0.029	−0.021	−0.012	−0.039		−0.017	−0.010	−0.019	−0.028
		(0.027)	(0.026)	(0.025)	(0.027)		(0.022)	(0.022)	(0.022)	(0.023)
Pop 1910 × Local pop 1910 (95th)		0.002**	0.002**	0.002**	0.001		0.003***	0.003***	0.003***	0.002***
		(0.001)	(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)	(0.001)
Dist town (km) × Dist town (km) (95th)		−0.002	−0.003	−0.004	−0.004		0.003	0.001	−0.002	−0.001
		(0.004)	(0.003)	(0.003)	(0.004)		(0.004)	(0.003)	(0.003)	(0.004)
p-value Prot vs. Cath	0	0	0	0	0	0.062	0.184	0.725	0.106	0.93
p-value Heterogeneity	0	0.001	0.061	0.003	0.04	0.199	0.409	0.056	0.057	0.079
Mean dep. var.	0.612	0.612	0.612	0.612	0.612	0.305	0.305	0.305	0.305	0.305
Controls		✓	✓	✓	✓		✓	✓	✓	✓
Province FE			✓					✓		
District FE				✓					✓	
County FE					✓					✓
Observations	12,216	12,216	12,216	12,213	12,216	12,216	12,216	12,216	12,213	12,216
R^2	0.035	0.040	0.075	0.094	0.177	0.006	0.010	0.061	0.102	0.192

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors clustered on the county level. Results using standard errors adjusted for spatial autocorrelation according to Conley (1999) is displayed in Table 1.A.16. Columns (2)-(5) and (7)-(10) additionally control population size of the locality as measured in 1910, distance to next town as well as an indicator whether the locality itself is a town. Test Prot vs. Cath gives the p-value of the Wald Test with H_0: Prevalence of norm in localities with $< 5\%$ Protestants (vs. Catholics) = Prevalence of norm in localities with $> 95\%$ Protestants (vs. Catholics). Test Heterogeneity gives the p-value of the Wald Test with H_0: Prevalence of norm in localities with $< 5\%$ Protestants (vs. Catholics) = prevalence of norm in localities with $5-50\%$ Protestants (vs. Catholics) and prevalence of norm in localities with 95% Protestants (vs. Catholics) = prevalence of norm in localities with $50-95\%$ Protestants (vs. Catholics). Province FE implies controlling for Prussian Provinces, states of Austria and German Reich (except Prussia), voivodships of Poland, and the Regions Bohemia and Moravia for Czechia.

Table 1.A.16. Social norms and religious composition: Neighborhood obligations (Conley Errors)

	Dependent variable:									
	Obligated					Sum nb help				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
5-50%	−0.070**	−0.050*	−0.031	−0.033*	−0.039*	−0.358***	−0.260**	−0.148	−0.154*	−0.161*
	(0.027)	(0.027)	(0.020)	(0.020)	(0.023)	(0.120)	(0.115)	(0.094)	(0.088)	(0.095)
50-95%	−0.091**	−0.075**	−0.043	−0.032	−0.030	−0.401***	−0.329**	−0.085	−0.047	−0.040
	(0.035)	(0.035)	(0.028)	(0.023)	(0.023)	(0.154)	(0.155)	(0.128)	(0.094)	(0.098)
>95%	−0.028	−0.021	0.006	0.021	0.020	−0.114	−0.088	0.169	0.257***	0.251**
	(0.039)	(0.039)	(0.032)	(0.021)	(0.023)	(0.160)	(0.160)	(0.143)	(0.098)	(0.106)
Any sects	−0.025	−0.020	−0.010	−0.007	−0.002	0.012	0.046	0.064	0.082	0.095
	(0.021)	(0.021)	(0.019)	(0.019)	(0.019)	(0.095)	(0.094)	(0.081)	(0.078)	(0.072)
Any non-denominational	−0.083***	−0.074***	−0.020	−0.023	−0.006	−0.422***	−0.368***	−0.089	−0.084	−0.020
	(0.022)	(0.022)	(0.017)	(0.016)	(0.018)	(0.093)	(0.090)	(0.065)	(0.061)	(0.063)
Any Jewish	−0.029*	−0.006	−0.028*	−0.026*	−0.019	−0.127	0.009	−0.121	−0.124*	−0.076
	(0.017)	(0.015)	(0.016)	(0.014)	(0.016)	(0.080)	(0.084)	(0.080)	(0.073)	(0.075)
Constant	0.462***	0.407***				2.169***	1.968***			
	(0.035)	(0.041)				(0.133)	(0.163)			
p-value Prot vs. Cath	0.463	0.59	0.862	0.318	0.39	0.476	0.582	0.239	0.009	0.018
p-value Heterogeneity	0.004	0.037	0.046	0.009	0.012	0.002	0.027	0.015	0	0
Mean dep. var.	0.42	0.42	0.42	0.42	0.42	1.989	1.989	1.989	1.989	1.989
Controls		✓	✓	✓	✓		✓	✓	✓	✓
Province FE			✓					✓		
District FE				✓					✓	
County FE					✓					✓

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Reference Category: Locality with less than 5% Protestants (as opposed to Catholics). Standard errors adjusted for spatial autocorrelation according to Conley (1999) using uniform kernel and a distance cutoff of 10 km. Columns (2)-(5) and (7)-(10) additionally control population size of the locality as measured in 1910, distance to next town as well as an indicator whether the locality itself is a town. Test Prot vs. Cath gives the p-value of the Wald Test with H_0: Prevalence of norm in localities with < 5% Protestants (vs. Catholics) = Prevalence of norm in localities with > 95% Protestants (vs. Catholics). Test Heterogeneity gives the p-value of the Wald Test with H_0: Prevalence of norm in localities with < 5% Protestants (vs. Catholics) = prevalence of norm in localities with 5 − 50% Protestants (vs. Catholics) and prevalence of norm in localities with 95% Protestants (vs. Catholics) = prevalence of norm in localities with 50 − 95% Protestants (vs. Catholics). Province FE implies controlling for Prussian Provinces, states of Austria and German Reich (except Prussia), voivodships of Poland, and the Regions Bohemia and Moravia for Czechia.

Table 1.A.17. Social norms and religious composition: Childbed norms (conley errors)

	Dependent variable:									
	Seclusion					Impurity				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
5-50%	-0.081^{***}	-0.061^{**}	-0.046^{*}	-0.067^{***}	-0.051^{**}	-0.015	-0.007	-0.020	-0.041^{*}	-0.044^{*}
	(0.025)	(0.025)	(0.024)	(0.023)	(0.022)	(0.024)	(0.025)	(0.026)	(0.025)	(0.024)
50-95%	-0.166^{***}	-0.158^{***}	-0.113^{***}	-0.149^{***}	-0.107^{***}	0.004	0.0004	-0.028	-0.049^{**}	-0.020
	(0.029)	(0.029)	(0.024)	(0.023)	(0.022)	(0.037)	(0.037)	(0.038)	(0.021)	(0.024)
>95%	-0.185^{***}	-0.189^{***}	-0.101^{***}	-0.143^{***}	-0.094^{***}	0.033	0.023	0.009	-0.032	-0.002
	(0.027)	(0.026)	(0.026)	(0.024)	(0.024)	(0.035)	(0.034)	(0.043)	(0.027)	(0.031)
Any sects	-0.020	-0.003	-0.011	-0.014	-0.027	-0.034	-0.021	-0.018	-0.003	0.009
	(0.020)	(0.021)	(0.020)	(0.020)	(0.020)	(0.021)	(0.021)	(0.018)	(0.018)	(0.018)
Any non-denominational	-0.134^{***}	-0.119^{***}	-0.076^{***}	-0.069^{***}	-0.046^{**}	0.050	0.058	0.008	0.006	0.009
	(0.023)	(0.022)	(0.020)	(0.019)	(0.019)	(0.038)	(0.035)	(0.018)	(0.017)	(0.018)
Any Jewish	0.013	0.056^{***}	0.008	0.008	-0.001	-0.087^{***}	-0.057^{**}	-0.011	-0.026	-0.031^{*}
	(0.021)	(0.021)	(0.016)	(0.016)	(0.014)	(0.025)	(0.027)	(0.023)	(0.017)	(0.017)
Constant	0.750^{***}	0.765^{***}				0.295^{***}	0.347^{***}			
	(0.019)	(0.027)				(0.028)	(0.042)			
p-value Prot vs. Cath	0	0	0	0	0	0.349	0.502	0.84	0.239	0.941
p-value Heterogeneity	0.003	0.01	0.123	0.011	0.037	0.631	0.786	0.219	0.172	0.111
Mean dep. var.	0.612	0.612	0.612	0.612	0.612	0.305	0.305	0.305	0.305	0.305
Controls		✓	✓	✓	✓		✓	✓	✓	✓
Province FE			✓					✓		
District FE				✓					✓	
County FE					✓					✓

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Reference Category: Locality with less than 5% Protestants (as opposed to Catholics). Standard errors adjusted for spatial autocorrelation according to Conley (1999) using uniform kernel and a distance cutoff of 50 km. Columns (2)-(5) and (7)-(10) additionally control population size of the locality as measured in 1910, distance to next town as well as an indicator whether the locality itself is a town. Test Prot vs. Cath gives the p-value of the Wald Test with H_0: Prevalence of norm in localities with $< 5\%$ Protestants (vs. Catholics) = Prevalence of norm in localities with $> 95\%$ Protestants (vs. Catholics). Test Heterogeneity gives the p-value of the Wald Test with H_0: Prevalence of norm in localities with $< 5\%$ Protestants (vs. Catholics) = prevalence of norm in localities with $5 - 50\%$ Protestants (vs. Catholics) and prevalence of norm in localities with 95% Protestants (vs. Catholics) = prevalence of norm in localities with $50 - 95\%$ Protestants (vs. Catholics). Province FE implies controlling for Prussian Provinces, states of Austria and German Reich (except Prussia), voivodships of Poland, and the Regions Bohemia and Moravia for Czechia.

Table 1.A.18. Reciprocal help among friends

	Dependent variable:	
	Friends	Friends (no neighbors)
	(1)	(2)
Local pop 1910	0.005***	0.002**
	(0.002)	(0.001)
Dist town (km)	0.00002	−0.001
	(0.002)	(0.001)
Town	−0.003	−0.012
	(0.054)	(0.032)
Mean dep. var.	0.191	0.056
District FE	✓	✓

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors clustered adjusted for spatial autocorrelation according to Conley (1999) using a 50 km distance cutoff and a uniform kernel. All specifications additionally contain an indicator variable that is one if the locality is in the 95th percentile of the population size distribution interacted with population size in 1910 and an indicator variable that is one if the locality is in the 95th percentile of the distance to next town distribution interacted with distance to next town distribution. Friends=Respondent indicated that friends help with the kind of activities that are typically comprised in neighborhood obligations. Friends (no neighbors)=Respondent indicated that friends and not neighbors help with the kind of activities that are typically comprised in neighborhood obligations.

Table 1.A.19. Childbed Norms, Catholicism, and Industrialization

	Dependent variable:			
	Any rule (mother)		Seclusion	
	(1)	(2)	(3)	(4)
$> 95\%$	-0.116^{***}	-0.111^{***}	-0.144^{***}	-0.139^{***}
	(0.022)	(0.022)	(0.023)	(0.023)
$5-50\%$	-0.031	-0.026	-0.061^{**}	-0.057^{**}
	(0.019)	(0.018)	(0.027)	(0.026)
$50-95\%$	-0.120^{***}	-0.113^{***}	-0.157^{***}	-0.152^{***}
	(0.021)	(0.020)	(0.025)	(0.025)
Ind. occ. 1925		-0.132^{**}		-0.107
		(0.055)		(0.069)
District FE	✓	✓	✓	✓
Observations	11,970	11,970	11,970	11,970
R^2	0.062	0.063	0.094	0.095

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors clustered are clustered on the district level. Reference category is displaying less than 5% Protestants. Ind. occ. 1925 is the share of the population whose occupation is in industry and craft in 1925 on the county level in the German Reich as obtained from Falter and Hänisch (1990). All specifications additionally contain an indicator variable that is one if the locality is in the 95th percentile of the population size distribution interacted with population size in 1910 and an indicator variable that is one if the locality is in the 95th percentile of the distance to next town distribution interacted with distance to next town distribution. They also include indicator variables to whether any other religious group is present in the locality (Christian sects, Jewish, Non-denominational)

Appendix 1.B External Data

An essential challenge of working with the GEA is merging relevant external data sets to the GEA. The GEA itself contains the GEA-code giving the approximate location of the data point, the localities name, the county name in the 1930's, the province or state, and the country. Giving the variety of administrative boundary changes, alternative naming of administrative units and localities and the imprecision of the location, a lot of work needs to be invested when merging data sets to the GEA. We provide code and ids for merging relevant publicly available historical data sets (see Table 1.B.1).

Table 1.B.1. External data sets

Country	Level	Content	Year	Source
German Reich	County	Population, religious affiliation, occupation	1925, 1933	Falter and Hänisch (1990)
Czechoslovakia	County	Population, religious affiliation, occupation, ethnicity	1930	Jíchová, Soukup, Nemeskal, Pospísilová, and Svoboda (2014)
Germany	Municipality 2011	Population	1871 - 2019	Roesel (2022)
German Reich	County	Crime Rates, demography, industrialization	1882-1900	Thome (2002)
German Reich	Communities	Population	1910	Uli Schubert's collection
German Reich, Czechoslovakia, Austria	Locality	Precise location, alternative naming, municipality, type of locality		Verein für Computer Genealogy (last accessed: 2022-06-15)

Merge keys to Falter and Hänisch (1990) and Thome (2002) are created using the work of Rahlf (2020) in combination with county shapefiles the Census Mosaic Project (Hubatsch and Klein, 1978; MPIDR and CGG, 2011).

Appendix 1.C Structure of Data and Digitization

The original data available in the archive of the German Ethnographic Atlas is located in the Cultural Anthropology Department of the University of Bonn. It contains the official list of locations that participated in the first four waves of the survey including the locality's name, the county, province or/and country as well as its unique coordinate on the custom designed geographic grid. It also includes the waves the locality participated in. It is important to note that this list is not fully accurate. First, there are some localities that are not on the list, even though the archive contains an answer for this locality. Second, there are some localities that are on the list, but the answer for this locality cannot be found in the archive.[43] Nonetheless, we digitized the list to get a condensed background of all localities.

43. It is important to note here that the latter may be also because the answer was lost at some point in time.

To geocode the list of villages, we recreated the customized geographic grid digitally by following the verbal description of the grid that is contained in the official list of localities. This customized grid divided the map of Cental Europe into rectangles and each locality was assigned a four to five part coordinate. The first coordinate divides the map into 287 large rectangles. The rectangles are displayed in Figure 1.C.1 The second coordinate divides the large rectangles into 36 smaller rectangles, the third coordinate divides these smaller rectangles again into 25 small rectangles, and the fourth coordinate divides each of these 25 rectangles into another four rectangles indicated by letters a-d. The fifth coordinate is only assigned when two villages are directly next to each other, i.e. to ensure uniqueness. It is indicated by the letters l,r,o,u, where l is left, r is right, o is up, u is down. So an example coordinate is, thus, given by 105 2 25 al.

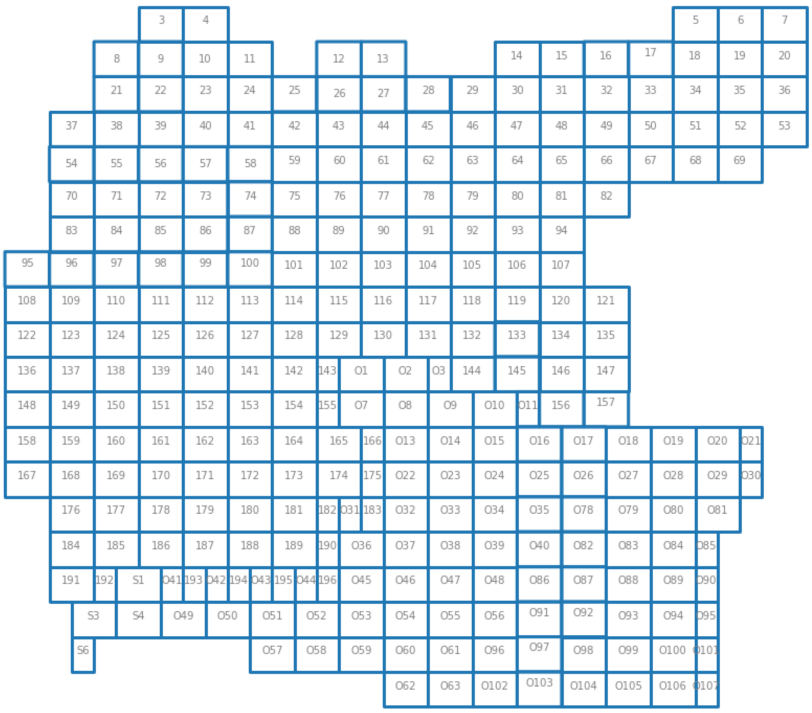


Figure 1.C.1. Large rectangles

The answers of to the questions are written on so called answer cards and separated by question. The background of each village and some characteristica of the respondent can be found on the so-called 'Personenfragebögen' separated by wave.

These contain some background for each village and the respondent, such as name, age, profession. The exact information on the answer card varies by wave. Each wave, however, contains the name, coordinate, county, and information on the religious composition of the locality.

Answers to the norms question were already previously processed, sorted, and partially pre-coded by Barruzi-Leicher and Frauenknecht (1966) and Grober-Glück (1966b). Their complete material is also available in the archive of the GEA. For digitization, we mostly rely on their pre-processed lists (on which they based their published maps but that contains the complete answers, also those not drawn). We ensured their accuracy by partially checking the data against the original answers. The advantage of this procedure is threefold. First, the some answer cards have become increasingly difficult to read due to the fact that they are handwritten, and for instance partially using pencil, which gets more mushed overtime. Second, Barruzi-Leicher and Frauenknecht (1966) and Grober-Glück (1966b) are ethnographic experts and their pre-coded lists already contain translation of answers that are written in dialect. Third, answers are not only contained in the boxes of answer cards but also in so-called 'Anlagen' – additional documents respondents sent – which are already incorporated in their lists, but not available in a sorted manner in the archive nowadays. Comparing the answer cards to the lists reveals no losses of relevant details, allowing for a different coding than that published in Barruzi-Leicher and Frauenknecht (1966) and Grober-Glück (1966b).

The digitization of the religious composition is particularly challengings. First, not all localities answered this question in each wave, some only answered it in early waves, some only answered them in later waves. Given the unbalancedness of the waves – not every locality is available in all waves – it is a difficult task to collect religious composition of relevant localities. Second, the answers are sometimes conflicting across waves. Third, the level of details and the format strongly varies across answers. Some answer the precise share of Lutheran, Calvinist, Catholics, Jewish, and other religious minorities, some only say something like "almost all Catholic", some only share the absolute numbers belonging to each denomination/religion. Hence, we rely on the published material of Grober-Glück (1966a) and Zender (1966). They went through all of the more the 60,000 records, and categorized each locality in one of five categories: less than 5% Protestants / more than 95% Catholics, between than 5% and 25% Protestants / between 75% and 95% Catholics, between than 25% and 50% Protestants / between 50% and 75% Catholics, 50-50 Protestants / Catholics, between than 5% and 25% Catholics / between 75% and 95% Protestants, less than 5% Catholics / more than 95% Protestants. We checked approximately 1,000 records, and found that there categorization is the most feasible given the varying level of detail of the answers. We digitized the map by first georeferencing it, then vectorizing the data points, and then using a linear assignment algorithm to match the vectorized data to the geocoded list of villages. We then checked the matches, in particular, we went through all points that were matched to a point more than 2 km away. We removed respective points, if it was clear or suspected that they were not drawn on the map and reiterated the procedure.

1.C.1 Coding Neighborhood Obligations

To quantify the degree of neighborhood help obligations, we use answers to the following survey question.[44]

> *a) Are neighbors traditionally obligated to mutual assistance in your village?*
>
> *b) At which occasions of family life, like childbirth (e.g. care for women in childbed), weddings (e.g. help in the kitchen), illness (e.g. night watch), death (carrying the coffin, digging the grave)*
>
> *c) At which economic tasks, like harvests, building a house (transporting wood) etc.?*

Table contains the list of variables including a variable description resulting from the coding of the original answers.

Obligation

Question a) asks about traditionally inherited obligations towards neighbors. Even though, it is a yes or no question, a lot of answers are not merely yes or no. In particular, the answers suggest that there is at least partly confusion/disagreement what obligation means (e.g. answers such as 'Yes but not legally' and 'No not legally but morally', 'Yes it is tradition/customary' vs. 'No, but it is tradition'). To homogeneize the answers, we decided to code answers 'No not legally but morally' and 'No, but it is tradition' as obligation. However, we also record whether they mention moral or customary reasons in their answer in the variable 'obl_type'. We code conditional obligations such as 'Yes if they have a good relationship' as no obligation. If it is obligatory but only at specific tasks, we code it as meaning there is obligatory neighborhood help. In case the answer to a) is just a line, we interpret this as not applicable, i.e. as 'no', if the answer to b) and c) is missing, explicitly marked as voluntary or crossed, and as 'yes' if they are named neighborhood activities in answers to b) and c). However, we also record those cases in the variable 'obligated_raw'.

44. In the German original the survey question is given by:
a) Sind in ihrem Ort die Nachbarn noch von alters her zu gegenseitiger Hilfeleistung verpflichtet? b) Bei welchen Anlässen des Familienlebens , wie Geburt (z.B. Pflege der Wöchnerin), Hochzeit (z.B. Hilfe in der Küche), Krankheit (z.B. Nachtwache), Tod (Tragen des Sargs, Graben des Grabs)? c) bei welchen wirtschaftlichen Arbeiten, wie Ernte, Hausbau (Anfahren des Bauholz) usw.? (Zender and Wiegelmann, 1959, p. 30)

Major and minor categorizations

We divide answers to question b) into four main categories: help at death, help with the wedding, help after birth, help in case of illness. Additionally we group every other activity that surrounds the family life into a residual category ‘other’ (typically help with the baptism, communion or confirmation of children). The level of detail of the answers varies greatly with a large part of answers being unspecific about the exact tasks performed at each occasions. In case a greater level of detail is provided, answers typically contain a subset of the tasks named in the example. On top of the main categories, we thus provide the subcategories: help at death by carrying the casket, help at death by digging the grave, help in the kitchen at the wedding, caring for the woman in childbed, doing nightwatch upon illness of a neighbor. All other activities as well as unspecific answers are grouped in a residual other category of the main categories (e.g. ‘nb_death_other). We ignore activities that are only of a ceremonial nature: praying in the case of death, visiting the woman in childbed, carrying a candle at the funeral, giving a gift to the child at the baptism.

Answers to question c) can be divided into agricultural activities, help with house building, as well as help in case of fire or floods. On top of these three major categories, we provide several common subcategories. The recorded subcategories in the main category help with agricultural activities are: help with the harvest (as named in the question), help with potato farming, help with flailing, help with slaughtering, help surrounding livestock, help with processing flax, machine sharing and moving, help surrounding poultry, general hand- and horse/cart services, and help surrounding ships. Agricultural activities that do not belong in one of these categories are counted and recorded in the variable ‘cl_other’. For the meta-category help with house building, we additionally record the subcategory: transporting material as named in the question.

Obligation vs. voluntary help

Sometimes respondents note down when a subtask or a whole category is commonly done but not considered obligatory. Hence, we provide for each category that indicates whether a specific tasks is obligatory or voluntarily done (‘_vol’, ‘_obl’). Tasks are categorized as voluntary if either the answer to a) is no, or the answer states that this specific task is voluntary.

Qualifiers

Sometimes respondents note certain conditions (if necessary, if no relatives are around, if they have good relationship, on request etc.) under which neighborhood activities are performed, specific subgroups within which neighborhood help activities are done (e.g. farmers, on the countryside), or that they are performed infrequently, sometimes, in some cases. We record these qualifiers in the respective 'quali' variable together with the respective tasks if applicable. We chose to set the respective tasks to 0 in these cases.

Coding in case of unspecific and ambiguous answers

Affirmative but unspecific answers. 5% of answers to question b) or question c) or both are affirmative (e.g. 'Yes', 'Always', 'Yes but voluntarily' or similar) but do not name any specific events at which the help occurs. In this case, we classify the answer as affirmative unspecific and note it down in column 'aff_unspecific'. If an answer to question b) is affirmative unspecific, we set all categories that are named in question b) (help at wedding, death, birth, sickness) to 1. If the answer to question c) is affirmative unspecific, we set help with house building to 1, as well as help with harvest to 1 as these two categories are explicitly named in the question.

Fortunes and Misfortunes. 0.1% of answers to question b) contain the phrase "[...] in fortunes and misfortunes" (German: 'in Freud und Leid') or similar without a reference to a specific activity. Given that the examples named in question b) cover fortunes and misfortunes, we chose to code this and similar phrases to mean that neighbors help at death, wedding, birth and sickness as named as examples in question b) as well as other emergencies (fire, flood).

Hardship and Misfortunes. 1% of answers to question b) or c) name help in cases of misfortune or emergencies (German: 'Unglücksfälle' and 'Notfälle') as a neighborhood help activity. We code this as help at emergencies together with help in case of fire or flood.

Family occasions or festivities. 0.5% of answers name any help with family festivities/help with family occasions/help at family days without being specific about the event or nature of the help. We code family festivities as meaning at wedding and at other family events. We code family occasions and family days as help at wedding, birth, death, and other family events.

Partly ambiguous answers. 0.3% answers are partly ambiguous (in a different way then the aforementioned) we note down our coding decision in the variable ambiguous_setting. The original answer can be found in the variable 'comment'.

Partly unreadable answers/fully ambiguous answers. Sometimes answers are not readable or their meaning is not comprehensible. In this case, the respective category or subquestion is set to missing.

Given that we provide all relevant columns and descriptions, our coding choices can be easily changed by other users of the data, if they disagree on the coding choices.

Neighborhood help in the past

Sometimes respondents note that there were (obligatory) neighborhood help activities in the past. In some cases, they provide an approximate year at which they disappeared, and/or which activities they covered. We record this in the '_past' and '_past_year' variables. In cases, where they do not state the year but indicate how many years ago, we code this using 1930 as the base year, i.e. '50 years ago', we code as 1880. If answers state something like 'not since the World War', we code them as 1914.

Payment

Sometimes respondents note down that certain activities are paid or done for a wage. In these cases we do not count them as neighborhood help activities. We record mentions of payment in the respect variable. It is important to note that this variable is not fully reliable and should be treated with caution as the digitizing researchers and research assistants were not explicitly instructed to record this cases.

Table 1.C.1. Variable Overview Neighborhood

variable	description	category description	categories & shares
obligated	indicates to whether neighbors are obligated to reciprocal help (cleaned answer a)	1: yes, 0: no	0: 0.582, 1: 0.418
obligated_raw	answer to question a)	1: Yes;0: No;"-":no answer/crossed sheet	-: 0.101, 0: 0.495, 1: 0.404
obligated_past	indicates whether neighborhood help activities were obligatory in the past, if not in the present		0: 0.992, 1: 0.008
obligated_past_year	year in which neighborhood help obligation disappeared		
cl_flail	named activities surrounding flailing as neighborhood help activity		0: 0.974, 1: 0.026
cl_flail_obl	indicates whether activity is obligatory		0: 0.98, 1: 0.02

Continued on next page

Table 1.C.1. Variable Overview Neighborhood

variable	description	category description	categories & shares
cl_flail_past	indicates whether neighborhood help activities surrounding flailing were existent in the past, if not in the present		0: 1.0, 1: 0.0
cl_flail_past_year	year in which neighborhood help activities surrounding flailing disappeared		1890: 0.25, 1910: 0.5, 1914: 0.25
cl_flail_vol	indicates whether activity is voluntary		0: 0.994, 1: 0.006
cl_flax	named activities surrounding processing flax as neighborhood help activity		0: 1.0, 1: 0.0
cl_flax_obl	indicates whether activity is obligatory		0: 1.0, 1: 0.0
cl_flax_past	indicates whether neighborhood help activities surrounding flax were existent in the past, if not in the present		0: 1.0, 1: 0.0
cl_flax_past_year	year in which neighborhood help activities surrounding flax disappeared		
cl_flax_vol	indicates whether activity is voluntary		0: 1.0, 1: 0.0
cl_harvest	named activities surrounding harvest as neighborhood help activity		0: 0.88, 1: 0.12
cl_harvest_obl	indicates whether activity is obligatory		0: 0.915, 1: 0.085
cl_harvest_past	indicates whether neighborhood help activities surrounding harvest were existent in the past, if not in the present		0: 0.994, 1: 0.006
cl_harvest_past_year	year in which neighborhood help activities surrounding harvest disappeared		1848: 0.029, 1870: 0.059, 1880: 0.147, 1890: 0.088, 1900: 0.412, 1910: 0.029, 1911: 0.029, 1914: 0.206
cl_harvest_vol	indicates whether activity is voluntary		0: 0.965, 1: 0.035
cl_hcs	hand, horse and cart services was named as a neighborhood help activity (Hand- und Spanndienste)		0: 1.0, 1: 0.0

Continued on next page

Table 1.C.1. Variable Overview Neighborhood

variable	description	category description	categories & shares
cl_hcs_obl	mandatory hand, horse and cart services was named as a neighborhood help activity (Hand- und Spanndienste)		0: 1.0, 1: 0.0
cl_hcs_vol	voluntary hand, horse and cart services was named as a neighborhood help activity (Hand- und Spanndienste)		0: 1.0, 1: 0.0
cl_livestock	named activities surrounding livestock as neighborhood help activity		0: 0.961, 1: 0.039
cl_livestock_obl	indicates whether activity is obligatory		0: 0.97, 1: 0.03
cl_livestock_past	indicates whether neighborhood help activities surrounding livestock were existent in the past, if not in the present		0: 1.0, 1: 0.0
cl_livestock_past_year	year in which neighborhood help activities surrounding livestock disappeared		1890: 0.5, 1892: 0.5
cl_livestock_vol	indicates whether activity is voluntary		0: 0.991, 1: 0.009
cl_machine	named activities surrounding machines as neighborhood help activity		0: 0.999, 1: 0.001
cl_machine_obl	indicates whether activity is obligatory		0: 1.0, 1: 0.0
cl_machine_past	indicates whether neighborhood help activities surrounding agricultural machines were existent in the past, if not in the present		0: 1.0
cl_machine_past_year	year in which neighborhood help activities surrounding agricultural machines disappeared		
cl_machine_vol	indicates whether activity is voluntary		0: 1.0, 1: 0.0
cl_other	named activities surrounding other agricultural activities as neighborhood help activity		0: 0.979, 1: 0.018, 2: 0.002, 3: 0.0, 4: 0.0, 5: 0.0
cl_other_obl	indicates how many other activities are obligatory		0: 0.985, 1: 0.014, 2: 0.001, 3: 0.0, 4: 0.0

Continued on next page

Table 1.C.1. Variable Overview Neighborhood

variable	description	category description	categories & shares
cl_other_past	indicates whether neighborhood help activities surrounding other agricultural activities were existent in the past, if not in the present		0: 1.0, 1: 0.0
cl_other_past_year	year in which neighborhood help activities surrounding other agricultural activities disappeared		
cl_other_vol	indicates how many activities are voluntary		0: 0.995, 1: 0.005, 2: 0.0, 3: 0.0, 5: 0.0
cl_potato	named activities surrounding potato farming as neighborhood help activity		0: 0.993, 1: 0.007
cl_potato_obl	indicates whether activity is obligatory		0: 0.995, 1: 0.005
cl_potato_past	indicates whether neighborhood help activities surrounding potato agriculture were existent in the past, if not in the present		0: 1.0, 1: 0.0
cl_potato_past_year	year in which neighborhood help activities surrounding potato agriculture disappeared		1914: 1.0
cl_potato_vol	indicates whether activity is voluntary		0: 0.998, 1: 0.002
cl_poult	named activities surrounding poultry (sorting and cleaning feathers) as neighborhood help activity		0: 0.998, 1: 0.002
cl_poult_obl	indicates whether activity is obligatory		0: 0.998, 1: 0.002
cl_poult_past	indicates whether neighborhood help activities surrounding poultry (cleaning feathers) were existent in the past, if not in the present		0: 1.0, 1: 0.0
cl_poult_past_year	year in which neighborhood help activities surrounding poultry (cleaning feathers) disappeared		1828: 0.2, 1870: 0.2, 1900: 0.2, 1910: 0.4
cl_poult_vol	indicates whether activity is voluntary		0: 1.0, 1: 0.0
cl_ship	help with rennovating/pushing in ships was named as a neighborhood help activity		0: 1.0, 1: 0.0

Continued on next page

Table 1.C.1. Variable Overview Neighborhood

variable	description	category description	categories & shares
cl_ship_obl	mandatory help with rennovating/pushing in ships was named as a neighborhood help activity		0: 1.0, 1: 0.0
cl_ship_vol	voluntary help with rennovating/pushing in ships was named as a neighborhood help activity		0: 1.0, 1: 0.0
cl_slaughter	named activities surrounding slaughter as neighborhood help activity		0: 0.997, 1: 0.003
cl_slaughter_obl	indicates whether activity is obligatory		0: 0.998, 1: 0.002
cl_slaughter_past	indicates whether neighborhood help activities surrounding slaughter were existent in the past, if not in the present		0: 1.0
cl_slaughter_past_year	year in which neighborhood help activities surrounding slaughter disappeared		
cl_slaughter_vol	indicates whether activity is voluntary		0: 0.999, 1: 0.001
nb_agriculture	help with any agricultural activity		0: 0.821, 1: 0.179
nb_agriculture_obl	mandatory help with any agricultural activity		0: 0.87, 1: 0.13
nb_agriculture_past	indicates whether there was any help with agricultural activities in the past, but not today		0: 0.993, 1: 0.007
nb_agriculture_past_year	year at which help with agricultural activities disappeard		1828: 0.022, 1848: 0.022, 1870: 0.067, 1880: 0.111, 1890: 0.111, 1892: 0.022, 1900: 0.333, 1910: 0.089, 1911: 0.022, 1914: 0.2
nb_agriculture_vol	voluntary help with any agricultural activity		0: 0.95, 1: 0.05
nb_birth	named any neighborhood help activity surrounding the birth of a child to a neighbor		0: 0.755, 1: 0.245
nb_birth_cwc	named help caring for the woman in childbed as neighborhood activity		0: 0.898, 1: 0.102

Continued on next page

Table 1.C.1. Variable Overview Neighborhood

variable	description	category description	categories & shares
nb_birth_cwc_vol	indicates whether activity is voluntary		0: 0.975, 1: 0.025
nb_birth_obl	indicates whether activity is obligatory		0: 0.822, 1: 0.178
nb_birth_other	number of named other activities surroundingthe birth of a child to a neighbor/named neighborhood activites surrounding the birth of a child to a neighbor without being specific which activity	0: did not name other activities; 1: did not specify other activity/named one other activity;>1 named several other neighborhood activities surrounding the birth of a child	0: 0.856, 1: 0.144, 2: 0.0
nb_birth_other_vol	indicates whether activity is voluntary		0: 0.958, 1: 0.042
nb_birth_past	indicates whether neighborhood help activities surrounding birth were existent in the past, if not in the present		0: 0.998, 1: 0.002
nb_birth_past_year	year in which neighborhood help activities surrounding birth disappeared		1866: 0.056, 1870: 0.056, 1880: 0.222, 1890: 0.111, 1900: 0.222, 1910: 0.056, 1912: 0.056, 1914: 0.167, 1935: 0.056
nb_birth_vol	indicates whether activity is voluntary		0: 0.933, 1: 0.067
nb_death	named any neighborhood help activity surrounding the death of a neighbor		0: 0.493, 1: 0.507
nb_death_cc	named carrying the casket as neighborhood help activity		0: 0.646, 1: 0.354
nb_death_cc_vol	indicates whether activity is voluntary		0: 0.878, 1: 0.122
nb_death_dg	named digging the grave as neighborhood help activity		0: 0.93, 1: 0.07
nb_death_dg_vol	indicates whether activity is voluntary		0: 0.979, 1: 0.021
nb_death_obl	indicates whether activity is obligatory		0: 0.659, 1: 0.341

Continued on next page

Table 1.C.1. Variable Overview Neighborhood

variable	description	category description	categories & shares
nb_death_other	number of named other activities at death/named neighborhood activites around the death of the neighbor without being specific which specific activity	0: did not name other activities; 1: did not specify other activity/named one other activity;>1 named several other neighborhood activities upon a death of a neighbor	0: 0.789, 1: 0.202, 2: 0.009, 3: 0.0, 4: 0.0
nb_death_other_vol	indicates whether activity is voluntary		0: 0.938, 1: 0.062
nb_death_past	indicates whether neighborhood help activities surrounding death were existent in the past, if not in the present		0: 0.99, 1: 0.01
nb_death_past_year	year in which neighborhood help activities surrounding death disappeared		
nb_death_vol	indicates whether activity is voluntary		0: 0.834, 1: 0.166
nb_emerg	named any activity at emergency such as after fire/flood		0: 0.928, 1: 0.072
nb_emerg_obl	indicates whether activity is obligatory		0: 0.95, 1: 0.05
nb_emerg_past	indicates whether neighborhood help activities surrounding emergencies (fire, flood) other misfortunes were existent in the past, if not in the present		0: 0.999, 1: 0.001
nb_emerg_past_year	year in which neighborhood help activities surrounding emergencies (fire, flood) other misfortunes disappeared		1828: 0.2, 1866: 0.2, 1870: 0.2, 1900: 0.2, 1910: 0.2
nb_emerg_vol	indicates whether activity is voluntary		0: 0.978, 1: 0.022
nb_house	named any neighborhood help activity surrounding the building of a house of a neighbor		0: 0.606, 1: 0.394
nb_house_dbm	named help driving construction material to the construction of a new house as neighborhood help activity		0: 0.776, 1: 0.224
nb_house_dbm_vol	indicates whether activity is voluntary		0: 0.925, 1: 0.075

Continued on next page

Table 1.C.1. Variable Overview Neighborhood

variable	description	category description	categories & shares
nb_house_obl	indicates whether activity is obligatory		0: 0.738, 1: 0.262
nb_house_other	number of named other activities surrounding the building of a house of a neighbor /named neighborhood activites surrounding the building of a house of a neighbor without being specific which activity	0: did not name other activities; 1: did not specify other activity/named one other activity;>1 named several other neighborhood activities surrounding the birth of a child	0: 0.798, 1: 0.194, 2: 0.008
nb_house_other_vol	indicates whether activity is voluntary		0: 0.932, 1: 0.068
nb_house_past	indicates whether neighborhood help activities surrounding building a house were existent in the past, if not in the present		0: 0.983, 1: 0.017
nb_house_past_year	year in which neighborhood help activities surrounding building a house disappeared		
nb_house_vol	indicates whether activity is voluntary		0: 0.867, 1: 0.133
nb_misfortune	named unspecifc neighborhood help in case of hardship or misfortune		0: 0.986, 1: 0.014
nb_other	number of other neighborhood help activities (e.g. help surrounding baptism, confirmation, communion, and animal births)		0: 0.983, 1: 0.016, 2: 0.0
nb_other_ind	help with any other family event		0: 0.983, 1: 0.017
nb_other_ind_obl	mandatory help with any other family event		0: 0.987, 1: 0.013
nb_other_ind_past	indicates whether there was any other neighborhood help in the past		0: 1.0, 1: 0.0
nb_other_ind_past_year	year in which other neighborhood help activity disappeared		
nb_other_ind_vol	voluntary help with any other family event		0: 0.996, 1: 0.004
nb_other_obl	indicates how many activities are obligatory		0: 0.987, 1: 0.013, 2: 0.0

Continued on next page

Table 1.C.1. Variable Overview Neighborhood

variable	description	category description	categories & shares
nb_other_past	indicates whether neighborhood help activities surrounding other were existent in the past, if not in the present		0: 1.0, 1: 0.0
nb_other_past_year	year in which neighborhood help activities surrounding other disappeared		
nb_other_vol	indicates how many activities are voluntary		0: 0.996, 1: 0.004, 2: 0.0
nb_past	indicates whether any neighborhood help activities were existent in the past, if not in the present		0: 0.961, 1: 0.039
nb_past_year	year in which any neighborhood help activities disappeared		
nb_sickness	named any neighborhood help activity surrounding the illness of a neighbor		0: 0.723, 1: 0.277
nb_sickness_nw	named help nightwatch when a neighbor is ill as neighborhood help activity		0: 0.878, 1: 0.122
nb_sickness_nw_vol	indicates whether activity is voluntary		0: 0.968, 1: 0.032
nb_sickness_obl	indicates whether activity is obligatory		0: 0.799, 1: 0.201
nb_sickness_other	number of named other activities surrounding the illness of a neighbor /named neighborhood activites surrounding the illness of a neighbor without being specific which activity	0: did not name other activities; 1: did not specify other activity/named one other activity;>1 named several other neighborhood activities surrounding the birth of a child	0: 0.841, 1: 0.157, 2: 0.001, 3: 0.0, 4: 0.0
nb_sickness_other_vol	indicates whether activity is voluntary		0: 0.954, 1: 0.046
nb_sickness_past	indicates whether neighborhood help activities surrounding sickness were existent in the past, if not in the present		0: 0.997, 1: 0.003
nb_sickness_past_year	year in which neighborhood help activities surrounding sickness disappeared		1866: 0.059, 1870: 0.059, 1880: 0.176, 1890: 0.118, 1900: 0.412, 1912: 0.059, 1914: 0.118

Continued on next page

Table 1.C.1. Variable Overview Neighborhood

variable	description	category description	categories & shares
nb_sickness_vol	indicates whether activity is voluntary		0: 0.924, 1: 0.076
nb_wedding	named any neighborhood help activity surrounding the wedding of a neighbor		0: 0.686, 1: 0.314
nb_wedding_hk	named help in the kitchen (e.g. cooking, serving) during the wedding of a neighbor as neighborhood help activity		0: 0.857, 1: 0.143
nb_wedding_hk_vol	indicates whether activity is voluntary		0: 0.961, 1: 0.039
nb_wedding_obl	indicates whether activity is obligatory		0: 0.773, 1: 0.227
nb_wedding_other	number of named other activities surrounding the wedding of a neighbor/named neighborhood activites surrounding the wedding of a neighbor without being specific which activity	0: did not name other activities; 1: did not specify other activity/named one other activity;>1 named several other neighborhood activities surrounding the wedding of a neighbor	0: 0.826, 1: 0.173, 2: 0.002
nb_wedding_other_vol	indicates whether activity is voluntary		0: 0.951, 1: 0.049
nb_wedding_past	indicates whether neighborhood help activities surrounding wedding were existent in the past, if not in the present		0: 0.998, 1: 0.002
nb_wedding_past_year	year in which neighborhood help activities surrounding wedding disappeared		1866: 0.048, 1870: 0.048, 1880: 0.19, 1890: 0.143, 1900: 0.238, 1910: 0.048, 1912: 0.048, 1914: 0.19, 1935: 0.048
nb_wedding_vol	indicates whether activity is voluntary		0: 0.912, 1: 0.088
comment	raw additional information (in German)		

Continued on next page

Table 1.C.1. Variable Overview Neighborhood

variable	description	category description	categories & shares
aff_unspecific	affirmative but unspecific answers (e.g. "Yes", “always“) without additional information	’b‘: affirmative unspecific answer to question b; ‘c‘: affirmative unspecific answer to question c; ‘both‘: affirmative unspecific answer to question b and c;suffix f indicates that it is not considered obligatory	b: 0.465, bf: 0.031, both: 0.258, bothf: 0.006, c: 0.211, cf: 0.028
family_events	sometimes respondents do not note specific neighborhood help tasks but only something like ‚at all family events‘ or ‚at all family festivities‘. In these cases this variable indicates which kind of family events		family days: 0.013, family events: 0.165, family festivities: 0.608, family occasions: 0.215
hardship_misfortune	help in cases of "hardship" (general or poverty) or "misfortunes" without any additional information		a: 0.004, b: 0.215, both: 0.629, c: 0.152
fortune_misfortune	help in cases of "joy and sorrow" or similar without any additional information		0: 0.999, 1: 0.001
ambiguous_setting	how information is coded in case of ambiguous answers, the respective answer can be found in the variable comment	0: each neighborhood help category is set to 0;1: each neighborhood help category that is mentioned in the question is set to 1;NA each neighborhood category is set to NA; c: or b: indicates that this only applies to answer categories c or b, respectively	

Continued on next page

Table 1.C.1. Variable Overview Neighborhood

variable	description	category description	categories & shares
quali_a	qualifier used in the answer to question a (e.g. partly applies, applies to farmers)	good relatioship: if neighbors have a good relation-ship/friendship;partly: partly applies;on request: help only on request then it may or may not be mandatory;often: often applies;infrequently: is only infrequently done;need: in cases of need/helping poor people/if necessary;hardship: only in case of hardship;subgroup: applies only to the subgroup (formatted as „subgroup: group“;	
quali_b	qualifier used in the answer to question b (e.g. partly applies, applies to farmers)	good relatioship: if neighbors have a good relationship;partly: partly applies;on request: help only on request then it may or may not be mandatory;often: often applies;infrequently: is only infrequently done;need: in cases of need/helping poor people/if necessary;reciprocal: based on reciprocity;subgroup: applies only to the subgroup; if it is only for a specific subtask, then subtask: qualifier	

Continued on next page

Table 1.C.1. Variable Overview Neighborhood

variable	description	category description	categories & shares
quali_c	qualifier used in the answer to question c (e.g. partly applies, applies to farmers)	good relatioship: if neighbors have a good relationship;partly: partly applies;on request: help only on request then it may or may not be mandatory;often: often applies;infrequently: is only infrequently done;need: in cases of need/helping poor people/if necessary;reciprocal: based on reciprocity;subgroup: applies only to the subgroup; if it is only for a specific subtask, then subtask: qualifier	
quali_gen			
obl_type	sometimes respondents additionally note what kind obligation it is (e.g. moral or by custom) or if it is not obligated whether it may still be common (german: ‚üblich'), this variable contains the type of comment	commonly: it is not obligatory but it is common (‚üblich') or neighbors ‚do it';customary: it is a custom or convention;morally: it is a moral obligation;good will: neighbors do it out of 'good will' or similar;self-evident: it is 'self-evident' or similar to help the neighbors;other group: neighbors are not obligated but a different group is obligated to help (either other_helper) or comment	Christian duty: 0.005, commonly: 0.276, customary: 0.383, morally: 0.3, other group: 0.022, self-evident: 0.015
vill_type	sometimes respondents note the type of village		Arbeiterdorf: 0.4, Gutsdorf: 0.4, Hofdorf: 0.2
multiple	number in case of multiple answers for the same coordinate, if no multiple 0		0: 0.971, 1: 0.014, 2: 0.014, 3: 0.001, 4: 0.0

Notes: Variable description containing all digitized data not only those used in the analysis.

1.C.2 Coding: Childbed Norms

a) Where is the woman in childbed not allowed to go before her first churchgoing? (e.g. basement, attic, barn, well, neighbor) *b) Which boundary is she not allowed to pass? (e.g. gutter, street, crossroad, village border)*

c) Which other traditional precautions does the women in childbed follow?[45]

Minor Categorization

We code answers irrespective of whether there where named in a), b) or c) according to their inherent meaning. For instance, we code everything that indicates a restriction of leaving the house/property as wo_house independent of whether it was named in a) the the mother shall not go outside, or in b) shall not pass the door threshold of her house, which also means she shall not go outside. A category is created only if there are at least 5 answers can be grouped in it. One exception is the variable 'wo_visit_specific' which was only named in three cases. However, given that it is a residual category of a broader categorical spectrum 'the woman shall not receive visits or visit', we decided to keep it anyway. All other rules that are named less than five times and do not fall under residual categories of broader categories, that is that are very specific, are grouped in the residual category 'wo_other_rule'. Approximately 2% of answers fall into this category. The most common category is the restriction to leave the house which is named in 47% of answers. In total, there are 80 categories.

Time Span

If respondents indicated this, variables '_time_span' contains a string indicating how long the rule applied (e.g for the rule 'do not leave the house', N=365/5% of those naming the rule). Most of these responses just reaffirmed that this restriction holds until the first churchgoing without indicating how many days/weeks after birth the first churchgoing takes place (e.g. for not leaving the house). Some however, indicate that the rule applies only until the child's baptism, which sometimes but not always coincides with the first churchgoing (e.g. for not leaving the house 53%) Among those that indicate a time length, six weeks is the most common answer (e.g. for 'do not leave the house' 30%) and the minimum, which is however only named once, is 7 days. The absolute maximum length of application is one year after birth, which is however, only named once.

Consequences

Variables ending with "_cons" indicate the type of consequence incurred upon violation of the restriction if given. We coded this in four main categories:

45. Original German: a) Wohin darf die Wöchnerin vor dem ersten Kirchgang (Aussegnung) nicht gehen? (z.B. Keller, Boden, Stall, Brunnen, Nachbar) b) Welche Grenze darf sie nicht überschreiten? (z.B. Dachtraufe, Gosse, Straße, Kreuzweg, Dorfgrenze) c) Welche besonderen altherkömmlichen Vorsichtsmaßnahmen beachtet die Wöchnerin sonst vor dem ersten Kirchgang?

Harm/misfortune for herself, harm/misfortune for the child, harm/misfortune for third parties, and unspecified/unclear misfortune. On top of these four categories, there are three additional categories that are rarely named but cannot be grouped in the former four categories. In same cases, it says that the woman follows the restriction so that nobody can harm her or her child/the child gets bad character. Even more rarely, respondents indicated that others can punish the woman when she violates the restriction. For instance, the respondent for Etterwinden, Thuringia, indicated the woman shall not cross the street in front of a moving cart, otherwise the coachman gives her whiplashes.

Specific times

Variables ending with “_spec_times" contain specific times (during the day) the restriction applies in, e.g. after sundown. Sometimes it does not refer to specific times of day but to specific days of the week at which the rule applies. This a comparably frequent. For instance, in 2% the restriction not to leave the house only applies during specific times of day - mostly after sundown.

Means of Protection

Variables ending with “_protect" contain the means of protection if there are means of protection that enable women to violate the respective rule without consequences (e.g. covering their head, putting salt in the well etc.).

Qualifiers

Some respondents indicate a restriction of the applicability of the rule either by stating that this rule only applies to certain groups of women (e.g. Catholics, Protestants, long-term residents, some) or by indicating the rule only applies if possible or should be only infrequently violated. The former restriction on the restriction can be found in variables ending with ‘quali_part’ the latter can be found in variables ending with ‘quali_norm’.

Past

Variables ending with "_past" contain the year until which the rule applied if it ever applied but does not apply anymore.

Ambiguous answers

Answers that we were partly not able to understand are categorized in ambiguous_a, ambiguous_b, ambiguous_c depending on whether they were answers to a, b or c. In our analysis, we drop all observations that contain an ambiguous part. This is the case in less than 1% of answers.

Major categorization & Impurity

To give a more concise overview of the different rules that can apply, we can categorize the rules into two major and two interesting subcategories. The major rules are particularly frequent restrictions, namely restrictions implying seclusion (i.e. not leaving the house) and social restrictions (not visiting others, not receiving visits, not to be seen by others etc.). The two interesting categories as they potentially suggest a more protective notion are restrictions to explicitly protect the woman's health, or restrict her from working. In table 1.C.2, we can see that in a large part of localities rules implying seclusion after birth or restricting social interactions of the woman after birth are with 62% and 24%, respectively, much more common, than the potentially protective rules regarding work, and health that each only apply in just more than 2% of villages.

Table 1.C.2. Major and interesting categorization

	Any Rule	Seclusion	Social	Work	Health	Impurity
Applied in the past	0.208	0.056	0.025	0.001	0.001	0.033
Only applies partly	0.008	0.007	0.003	0.0	0.0	0.003
Can be violated if ...	0.003	0.004	0.0	0.0	0.0	0.002
Adhered to if possible	0.0	0.001	0.0	0.0	0.0	0.0
Applies at specific times (of day)	0.004	0.01	0.0	0.0	0.001	0.0
Applies	0.777	0.611	0.238	0.027	0.02	0.298
0						
N	15785	15785	15785	15785	15785	15785

Notes: Seclusion: do not leave the house, do not go over the street, do not pass the roof border of your house, stay in your living area, do not go anywhere. Social: Do not visit others, or specific others such as your neighbor, do not participate in festivities, do not go into the village, or pub, do not interact with certain types of people, do not speak, or any other social restriction. Work: do not work, do not spin, do not lift heavy things, do not make laundry, do not do other type of specific work, such as working on the field. Health: rest in bed, do not eat or drink certain types of food (e.g. alcohol), avoid draughts, don't go on walks or other health related prescriptions.

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_house	indicates that the woman should not leave her (farm)house/property or that she shall not pass the door threshold to leave the house or similar	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.528, 1: 0.044, 2: 0.006, 3: 0.002, 4: 0.001, 5: 0.011, 6: 0.408

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_str	indicates that the woman should not pass streets, gutters or similar or go ‚on the street‘	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.787, 1: 0.016, 2: 0.002, 3: 0.001, 4: 0.0, 5: 0.0, 6: 0.195
wo_visit_neighbor	indicates that the woman should not visit the neighbor/should not enter the neighbor‘s property	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.83, 1: 0.019, 2: 0.002, 3: 0.0, 5: 0.0, 6: 0.149
wo_get_water	indicates whether ‘the woman should not get water (at the well, from the pump)‘ was named as a rule	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.887, 1: 0.015, 2: 0.001, 3: 0.003, 6: 0.094
wo_roof	indicates that the woman should not pass her roof‘s eave	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.903, 1: 0.006, 2: 0.0, 3: 0.0, 5: 0.0, 6: 0.09
wo_base	indicates that the woman should not go into the basement	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.907, 1: 0.01, 2: 0.0, 3: 0.0, 5: 0.0, 6: 0.082
wo_visit_general	indicates that the woman should not visit anyone or/and should not receive visits from anyone	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.946, 1: 0.004, 2: 0.001, 3: 0.0, 4: 0.0, 5: 0.0, 6: 0.05

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_attic	indicates that the woman should not go to the attic	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.945, 1: 0.006, 2: 0.0, 3: 0.0, 5: 0.0, 6: 0.048
wo_stable	indicates that the woman should not go into (her) stable	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.952, 1: 0.006, 2: 0.0, 3: 0.0, 5: 0.0, 6: 0.041
wo_outside_vill	indicates that the woman should not leave the village	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.956, 1: 0.002, 2: 0.0, 3: 0.0, 5: 0.0, 6: 0.041
wo_crossroad	indicates that the woman should not pass a crossroad	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.975, 1: 0.001, 2: 0.0, 6: 0.023
wo_festivities_general	indicates that the woman should stay away from festivities, enjoyments, company or similar in general	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.978, 1: 0.002, 2: 0.0, 4: 0.0, 6: 0.02
wo_other_rule	indicates that other rules were name that do not fit in any of the other categories and are named too rarely to give them an extra category	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.979, 1: 0.002, 3: 0.0, 5: 0.0, 6: 0.019

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_nowhere	indicates that the woman should go 'nowhere'	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if . . ., 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.982, 1: 0.002, 2: 0.0, 3: 0.001, 5: 0.0, 6: 0.015
wo_rule_churchgoing	indicates whether a rule was named that applied only on the way to the first churchgoing or during the first churchgoing	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if . . ., 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.985, 1: 0.001, 2: 0.0, 3: 0.0, 6: 0.014
wo_room	indicates whether staying in specific rooms of the house or appartment was named as a rule	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if . . ., 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.983, 1: 0.002, 2: 0.0, 3: 0.001, 4: 0.0, 5: 0.001, 6: 0.013
wo_field	indicates that the woman should not go to /on the field	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if . . ., 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.986, 1: 0.002, 3: 0.0, 5: 0.0, 6: 0.012
wo_means_protect _other	indicates whether other means of protection against harm and misfortunes were named	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if . . ., 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.988, 1: 0.001, 2: 0.0, 5: 0.0, 6: 0.01
wo_wash	indicates whether the rule 'do not wash (cloth)' or similar was named	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if . . ., 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.99, 1: 0.0, 2: 0.0, 3: 0.0, 4: 0.0, 6: 0.01

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_work_general	indcates that the woman should not work/should only do light work	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.991, 1: 0.0, 4: 0.0, 5: 0.0, 6: 0.009
wo_festivities_specific	indicates that the woman should not participate in certain festivities (other than weddings)	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.991, 1: 0.0, 5: 0.0, 6: 0.008
wo_str_specific	indicates that the woman should not cross specific streets	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.991, 1: 0.001, 6: 0.008
wo_vill	indicates that the woman should not go into the village/town	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.991, 1: 0.001, 2: 0.0, 5: 0.0, 6: 0.008
wo_health_other	indicates other rules to preserve the woman‘s health	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.992, 1: 0.0, 6: 0.007
wo_holy_water	indicates that the woman should spray herself /the newborn with holy water or take a rosary or similar	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.992, 1: 0.0, 2: 0.0, 6: 0.007

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_food	indicates that the woman should follow a specific diet / not consume certain types of foods and drinks (alcohol)	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.992, 1: 0.001, 2: 0.0, 3: 0.0, 6: 0.007
wo_death_other	indicates that the woman should stay away from other things surrounding death, funerals (e.g. do not go to the cementary)	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.993, 1: 0.0, 3: 0.0, 6: 0.007
wo_pass_water	indicates that the woman should not pass rivers, lakes, streams, bridges or similar	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.992, 1: 0.001, 2: 0.0, 3: 0.0, 6: 0.007
wo_child_cloth	indicates whether the rule ‘do not hang (the child‘s) coth outside‘ or other rules regarding dipers and children cloth was named	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.991, 1: 0.001, 2: 0.0, 3: 0.0, 5: 0.001, 6: 0.007
wo_cover_head	indicates that the woman should cover her head (for protection)	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.993, 1: 0.0, 2: 0.0, 6: 0.006
wo_work_specific	indicates that the woman should stay away from specific types of work	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.993, 1: 0.0, 2: 0.0, 4: 0.0, 6: 0.006

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_lend	indicates that the woman should not lend things to others	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.994, 1: 0.0, 2: 0.0, 6: 0.006
wo_lift	indicates that the woman should not lift (heavy) things	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.995, 6: 0.005
wo_public	indicates whether ‘the woman should not be seen publicly‘ or similar was named as a rule	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.994, 1: 0.001, 2: 0.0, 4: 0.0, 6: 0.005
wo_alone	indicates whether ‘the woman should not be left alone‘ or similar was named	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.994, 1: 0.0, 3: 0.0, 4: 0.0, 5: 0.001, 6: 0.005
wo_child_alone	indicates that the woman should never leave the side of the newborn	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.994, 1: 0.0, 3: 0.001, 4: 0.0, 5: 0.0, 6: 0.005
wo_stairs	indicates that the woman should not climb stairs or ladders	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.995, 1: 0.0, 6: 0.004

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_borrow	indicates that the woman should not borrow (generally, or specific things)	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if ..., 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.995, 1: 0.0, 2: 0.0, 6: 0.004
wo_child_bs	indicates whether there should be put a bible or book of prayer ('Gesangsbuch') has to be put under the child's pillow or similar (for protection)	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if ..., 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.995, 1: 0.0, 2: 0.0, 5: 0.0, 6: 0.004
wo_church	indicates that the woman should not go to church	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if ..., 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.995, 1: 0.0, 2: 0.0, 3: 0.0, 6: 0.004
wo_funeral	indicates whether the woman should stay away from funerals (processions)	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if ..., 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.996, 1: 0.0, 6: 0.004
wo_bed	indicates whether the woman should stay in bed	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if ..., 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.995, 1: 0.0, 2: 0.0, 3: 0.0, 5: 0.001, 6: 0.003
wo_shopping	indicates that the woman should not go shopping (groceries etc.)	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if ..., 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.996, 1: 0.0, 2: 0.0, 3: 0.0, 6: 0.003

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_darkness	indicates whether rules regarding staying away from darkness or keeping the light on at night were named	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.997, 1: 0.0, 6: 0.003
wo_pub	indicates whether the woman should stay away from the pub or similar	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.997, 6: 0.003
wo_away_food	indicates whether a rule staying away from certain foods is named (not the consumption but the physical proximity)	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.997, 1: 0.0, 2: 0.0, 6: 0.003
wo_health_draught	indicates that the woman should avoid wind, draught	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.997, 1: 0.0, 6: 0.003
wo_wear_men_cloth	indicates whether the woman should wear (a piece of) her husband‘s clothing (for protection)	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.997, 1: 0.0, 2: 0.0, 6: 0.003
wo_child_outside	indicates whether the woman should not bring the child outside (the house/the property)	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.997, 1: 0.0, 5: 0.0, 6: 0.003

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_women_specific	indicates that the woman should not meet with particular groups of women (e.g. other women in childbed, old women, fertile women)	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.997, 1: 0.0, 6: 0.003
wo_social_other	indicates that other social restrictions were named (do not greet, do not argue with the husband etc.)	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.997, 1: 0.0, 4: 0.0, 6: 0.003
wo_cold_water	indicates that the woman should stay away from cold water	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.998, 6: 0.002
wo_window	indicates that the woman should not look outside her windows/ should not be seen in the windows	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.997, 1: 0.0, 5: 0.0, 6: 0.002
wo_line	indicates whether rules regarding not going under a (cloth) line or similar were named	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.998, 1: 0.0, 6: 0.002
wo_bake	indicates that the woman should not bake/bake bread/touch dough/go to the bakehouse or similar	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.998, 1: 0.0, 2: 0.0, 4: 0.0, 6: 0.002

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_wedding	indicates whether the woman should stay away from wedddings (processions)	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.998, 6: 0.002
wo_green	indicates that the woman should not touch/go to green gras and fields	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.998, 1: 0.0, 6: 0.002
wo_scare	indicates that the woman should avoid 'scares' (German: Schrecken) or similar	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.998, 6: 0.002
wo_economic_unit	indicates that the woman should not leave her economic unit / should not pass the border of her economic unit	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.998, 1: 0.0, 6: 0.002
wo_bedsheet	indicates that rules regarding the bedsheets ('do not make the bed') or similar were named	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.998, 1: 0.0, 2: 0.0, 6: 0.002
wo_speak	indicates that the woman faces restrictions with respects speaking (e.g. do not gossip, do not speak to strangers, do not shout, speak quietly)	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.998, 1: 0.0, 6: 0.002

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_barn	indicates that the woman should not enter barns	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.998, 1: 0.0, 2: 0.0, 6: 0.002
wo_mirror	indicates that the woman should not look into a mirror	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.998, 1: 0.0, 2: 0.0, 3: 0.0, 6: 0.002
wo_spin	indicates that the woman should not spin, knit or sew	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.999, 1: 0.0, 6: 0.001
wo_oven	indicates whether the rule ‘do not look in the oven‘ or similar was named	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.999, 2: 0.0, 6: 0.001
wo_other_stable	indicates that the woman should not enter other people‘s stables/get close to other‘s animals	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.999, 1: 0.0, 6: 0.001
wo_fire	indicates that rules regarding open fire are named (e.g. stay away from, do not give fire)	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.999, 1: 0.0, 6: 0.001

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_visit_farmer	indicates that the woman should not visit farmers / is not welcome at a farmer's home	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.999, 6: 0.001
wo_preserve	indicates that the woman should not process food (mostly for the purpose of preservation)	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.999, 6: 0.001
wo_look	indicates that the woman should not look into certain things (e.g. suitcases, wardrobes)	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.999, 1: 0.0, 6: 0.001
wo_woman_bs	indicates whether there should be put a bible or book of prayer ('Gesangsbuch') has to be put under the her pillow or similar (for protection)	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.999, 2: 0.0, 6: 0.001
wo_salt	indicates that the woman should not touch salt	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.999, 1: 0.0, 6: 0.001
wo_grap_up	indicates that the woman should not grap something above her head/a certain distance abover her head/body	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.999, 6: 0.001

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_show_child	indicates that the woman should not show the child to others (strangers)	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.999, 6: 0.001
wo_slaughter	indicates that the woman should not participate/help with slaughter	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.999, 1: 0.0, 6: 0.001
wo_bathwater	indicates whether rules regarding the pouring out of bathwater or similar are named	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 0.999, 1: 0.0, 6: 0.0
wo_walks	indicates that the woman should not go on walks	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 1.0, 1: 0.0, 6: 0.0
wo_visit_family	indicates whether the woman should not visit her relatives	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 1.0, 1: 0.0, 6: 0.0
wo_visit_friends	indicates that the woman should not visit or receive visits from friends	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 1.0, 1: 0.0, 6: 0.0

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_visit_specific	indicates that the woman should not visit or receive visits from specific other people that do not fall in the family, friends, neighbor categories (e.g. people that live far away)	0: not named, 1: named but in the past, 2: named but only applies partly, 3: named but can be violated if …, 4: adhered to if possible,5: applies at specific times (of day);6: applies	0: 1.0, 6: 0.0
a_past	contains year or indicator (‚früher') in case rules named in a do not apply anymore but applied in the past		
ambiguous_a	answer a contains an ambiguous answer (can be part or fully)		0: 0.994, 1: 0.006
ambiguous_b	answer b contains an ambiguous answer (can be part or fully)		0: 0.997, 1: 0.003
ambiguous_c	answer c contains an ambiguous answer (can be part or fully)		0: 0.991, 1: 0.009
b_past	contains year or indicator (‚früher') in case rules named in b do not apply anymore but applied in the past		
c_past	contains year or indicator (‚früher') in case rules named in c do not apply anymore but applied in the past		
multiple			0: 0.973, 1: 0.027
wo_alone_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		... so that nobody can harm her: 0.091, Harm/misfortune for child/child gets bad character: 0.273, Harm/misfortune for herself: 0.455, Harm/misfortune for others: 0.091, Other people can punish her: 0.091

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_attic_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.429, Harm/misfortune for herself: 0.286, Harm/misfortune for others: 0.143, Other people can punish her: 0.143
wo_away_food_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.182, Harm/misfortune for herself: 0.182, Harm/misfortune for others: 0.545, Unspecified/unclear misfortune: 0.091
wo_bake_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 1.0
wo_barn_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.5, Harm/misfortune for herself: 0.5
wo_base_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.194, Harm/misfortune for herself: 0.484, Harm/misfortune for others: 0.323

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_bathwater_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 1.0
wo_bed_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Unspecified/unclear misfortune: 1.0
wo_bedsheet_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.5, Harm/misfortune for herself: 0.5
wo_borrow_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.3, Harm/misfortune for herself: 0.1, Harm/misfortune for others: 0.4, Other people can punish her: 0.1, Unspecified/unclear misfortune: 0.1
wo_child_alone_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		... so that nobody can harm her: 0.05, Harm/misfortune for child/child gets bad character: 0.7, Harm/misfortune for herself: 0.1, Harm/misfortune for others: 0.05, Other people can punish her: 0.05, Unspecified/unclear misfortune: 0.05

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_child_bs_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		... so that nobody can harm the child: 0.2, Harm/misfortune for child/child gets bad character: 0.4, Harm/misfortune for herself: 0.4
wo_child_cloth_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.615, Harm/misfortune for herself: 0.077, Harm/misfortune for others: 0.308
wo_child_outside_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.75, Harm/misfortune for herself: 0.25
wo_cover_head_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.167, Harm/misfortune for herself: 0.833
wo_crossroad_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 1.0
wo_darkness_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.875, Harm/misfortune for herself: 0.125

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_death_other_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.75, Harm/misfortune for herself: 0.25
wo_festivities_general_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 1.0
wo_festivities_specific_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for herself: 0.125, Harm/misfortune for others: 0.875
wo_field_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.1, Harm/misfortune for herself: 0.1, Harm/misfortune for others: 0.8
wo_fire_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.667, Harm/misfortune for herself: 0.333
wo_food_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.667, Harm/misfortune for herself: 0.167, Harm/misfortune for others: 0.167

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_funeral_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.333, Harm/misfortune for herself: 0.5, Harm/misfortune for others: 0.167
wo_get_water_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.005, Harm/misfortune for herself: 0.011, Harm/misfortune for others: 0.973, Harm/misfortune for other-sHarm/misfortune for others: 0.005, Unspeci-fied/unclear misfortune: 0.005
wo_green_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.5, Harm/misfortune for others: 0.5
wo_health_draught_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for herself: 1.0
wo_health_other_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.6, Harm/misfortune for herself: 0.4
wo_holy_water_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.4, Harm/misfortune for herself: 0.4, Harm/misfortune for others: 0.2

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_house_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		... so that nobody can harm the child: 0.026, Harm/misfortune for child/child gets bad character: 0.256, Harm/misfortune for herself: 0.282, Harm/misfortune for others: 0.359, Harm/misfortune for others,Harm/misfortune for herself: 0.013, Other people can punish her: 0.026, Unspecified/unclear misfortune: 0.038
wo_lend_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		... so that nobody can harm the child: 0.067, Harm/misfortune for child/child gets bad character: 0.533, Harm/misfortune for herself: 0.067, Harm/misfortune for others: 0.267, Unspecified/unclear misfortune: 0.067
wo_line_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.667, Harm/misfortune for herself: 0.333

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_means_protect_other_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.55, Harm/misfortune for herself: 0.2, Harm/misfortune for others: 0.25
wo_mirror_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for herself: 1.0
wo_nowhere_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.25, Harm/misfortune for herself: 0.5, Harm/misfortune for others: 0.25
wo_other_rule_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.452, Harm/misfortune for herself: 0.194, Harm/misfortune for herself and the child: 0.032, Harm/misfortune for others: 0.258, Other people can punish her: 0.032, Unspecified/unclear misfortune: 0.032
wo_other_stable_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for others: 1.0

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_outside_vill_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.2, Harm/misfortune for herself: 0.2, Harm/misfortune for others: 0.4, Unspecified/unclear misfortune: 0.2
wo_oven_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 1.0
wo_pass_water_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.25, Harm/misfortune for others: 0.5, Other people can punish her: 0.25
wo_preserve_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for others: 1.0
wo_pub_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.25, Harm/misfortune for herself: 0.25, Harm/misfortune for others: 0.5
wo_public_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for herself: 0.333, Harm/misfortune for others: 0.333, Other people can punish her: 0.333

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_roof_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		... so that nobody can harm her: 0.042, Harm/misfortune for child/child gets bad character: 0.083, Harm/misfortune for herself: 0.25, Harm/misfortune for others: 0.583, Harm/misfortune for others,Harm/misfortune for others: 0.042
wo_room_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.125, Harm/misfortune for herself: 0.375, Harm/misfortune for others: 0.375, Unspecified/unclear misfortune: 0.125
wo_rule_churchgoing_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		... so that nobody can harm the child: 0.062, Harm/misfortune for child/child gets bad character: 0.531, Harm/misfortune for herself: 0.25, Harm/misfortune for others: 0.125, Unspecified/unclear misfortune: 0.031
wo_salt_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 1.0
wo_shopping_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for others: 1.0

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_show_child_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 1.0
wo_slaughter_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for others: 1.0
wo_social_other_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.5, Harm/misfortune for others: 0.5
wo_speak_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.5, Harm/misfortune for others: 0.5
wo_spin_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.375, Harm/misfortune for herself: 0.125, Harm/misfortune for others: 0.25, Unspecified/unclear misfortune: 0.25
wo_stable_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.158, Harm/misfortune for herself: 0.263, Harm/misfortune for others: 0.526, Unspecified/unclear misfortune: 0.053

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_stairs_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 1.0
wo_str_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		... so that nobody can harm her: 0.067, Harm/misfortune for child/child gets bad character: 0.267, Harm/misfortune for herself: 0.133, Harm/misfortune for others: 0.467, Other people can punish her: 0.067
wo_str_specific _cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.5, Harm/misfortune for others: 0.5
wo_visit_general _cons	the type of consequence of rule violation if given (harm to herself, others, the child)		... so that nobody can harm her: 0.024, Harm/misfortune for child/child gets bad character: 0.024, Harm/misfortune for herself: 0.024, Harm/misfortune for others: 0.878, Other people can punish her: 0.024, Unspecified/unclear misfortune: 0.024

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_visit_neighbor _cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.068, Harm/misfortune for herself: 0.034, Harm/misfortune for others: 0.881, Other people can punish her: 0.017
wo_wash_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.333, Harm/misfortune for herself: 0.5, Harm/misfortune for others: 0.167
wo_wear_men_cloth _cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.167, Harm/misfortune for herself: 0.5, Harm/misfortune for others: 0.333
wo_wedding_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for others: 0.857, Unspeci-fied/unclear misfortune: 0.143
wo_window_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.111, Harm/misfortune for herself: 0.111, Harm/misfortune for others: 0.333, Other people can punish her: 0.333, Unspeci-fied/unclear misfortune: 0.111

Continued on next page

Table 1.C.3. Variable Overview Childbed Norms

variable	description	category description	categories & shares
wo_woman_bs_cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for others: 1.0
wo_women_specific _cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.4, Harm/misfortune for others: 0.6
wo_work_general _cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for herself: 1.0
wo_work_specific _cons	the type of consequence of rule violation if given (harm to herself, others, the child)		Harm/misfortune for child/child gets bad character: 0.75, Harm/misfortune for herself: 0.25

Notes: Variable description containing all digitized data not only those used in the analysis. Shares for consequences are conditionally on being named.

1.C.3 Coding: Miscellaneous Comments

Sometimes there are multiple answers for the same coordinate. Most of time it seems that these represent answers of subareas of the locality. It can however also happen that multiple people answered for the same village. Given that the final answer card does only have the coordinate on it without name, and we are not able to match them two the 'Personalfragebögen' that contain the name of the locality. We resolve this issue by just assigning the mean to the coordinate for the analysis. For 231 coordinates, there exist more than one answer to the childbed norm question. For 244 coordinates, there exist more than one answer to the neighborhood question.

1.C.4 Respondents Characteristics

To get a more detailed view of the volunteer's background characteristics, we use data covering the Rhine-Province digitized by Kehren (1994). Around 90% of respondents in this region were teachers. The occupations of the remaining respondents are heterogeneous. The most common additional groups are farmers, students, craftsmen, and individuals occupied in some administrative positions (mostly local government). Each of those groups covers between 1 and 3% of respondents. Below 1% of respondents in the Rhine Province were women, and most of the respondents were between 30 and 50 years old when they answered the survey (see Figure 1.C.2a).

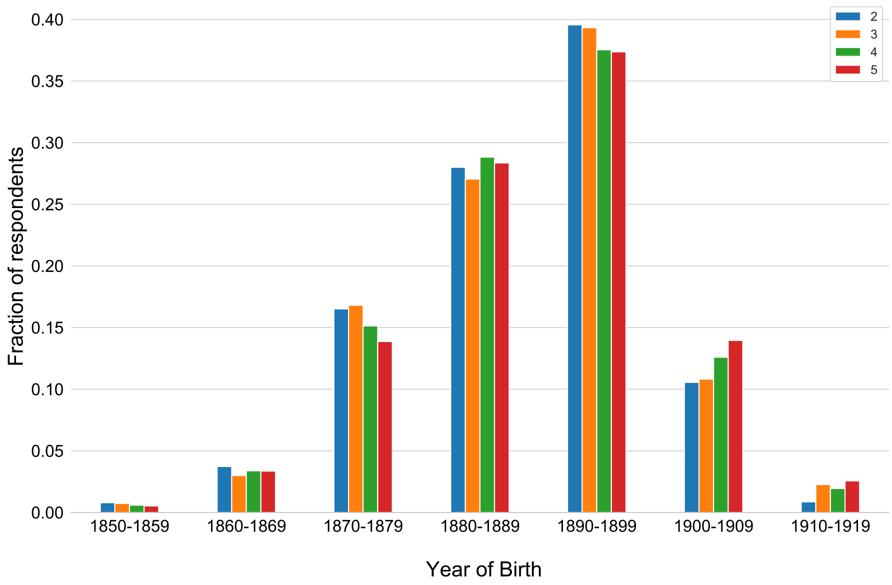
We can also use the data of Kehren (1994) to learn about the volunteer's familiarity with the village they were covering. The share of respondents born in the village they answer for varies between 11.8% and 17.4% and increases over time. In the samples of questionnaires two and four, it is 11.8% and 14.4%, respectively. The majority of the volunteers who were not born in the village moved there before or in 1920, so they spent at least 10 years in the village they answered for.[46] However, a large part also moved there only in the 1920s or even in the 1930s.

Table 1.C.4. Occupation of volunteers by wave

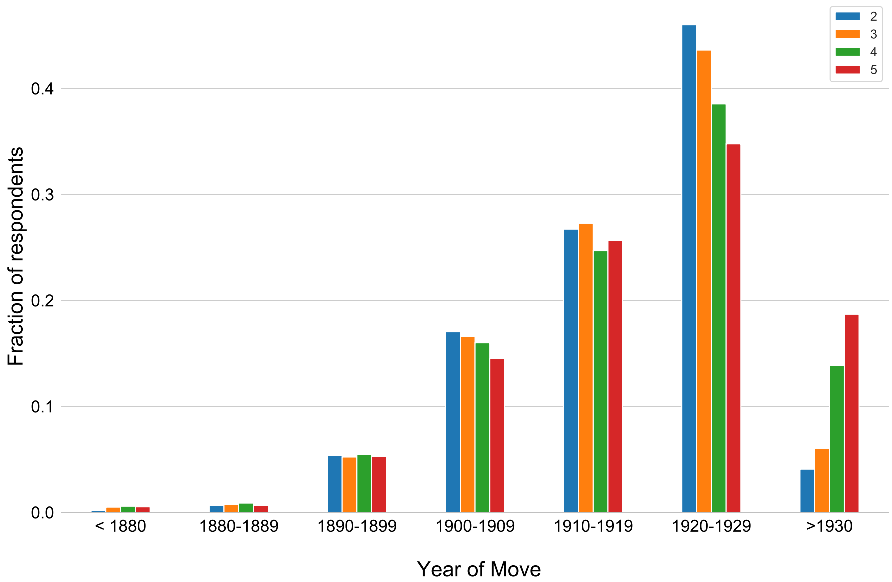
Quest. Occupation	2	3	4	5
Teacher/Principal	92.2	90.8	89.0	87.1
Other	2.1	2.5	3.5	4.3
Farmer/Winemaker	1.7	1.5	2.3	3.4
Student	1.5	2.3	1.2	1.5
Administration	0.9	1.5	1.7	1.7
Craftsman	0.9	0.8	1.3	1.0
Pastor/Chaplain	0.4	0.5	0.3	0.4
No occupation	0.2	0.1	0.5	0.3
Innkeeper	0.1	0.1	0.3	0.3

Notes: Sample restricted to the Rhine Province. Data obtained from Kehren (1994). Own calculations.

46. We know neither the locality of origin nor the year since when the respondent moved to the village for 1% of the sample in questionnaires two and three, 6.7% in questionnaire four, and 1.7% in questionnaire five. The numbers refer to the remaining sample.



(a) Volunteers year of birth.



(b) Year when volunteer moved to village.

Figure 1.C.2. Characteristics of volunteers in the Rhine Province.

Notes: Sample restricted to the Rhine Province. Data obtained from Kehren (1994). Own calculations. Part 1.C.2b is restricted to volunteers who are not born in the village they answer for.

References

Barruzi-Leicher, Renate, and Gertrud Frauenknecht. 1966. “Nachbarschaft.” In *Atlas Der Deutschen Volkskunde/Erläuterungen Bd. 2. Zu Den Karten NF 37 - 48 Und 54 - 56c, 65 - 69d, 70 - 72c, 73 - 76a,b*. Edited by Matthias Zender. Düsseldorf: Droste Verlag. Chapter XVII, 279–337. [88]

Conley, T.G. 1999. “GMM Estimation with Cross Sectional Dependence.” *Journal of Econometrics* 92 (1): 1–45. [75, 76, 80–84]

Falter, Jürgen W., and Dirk Hänisch. 1990. “Wahl- und Sozialdaten der Kreise und Gemeinden des Deutschen Reiches von 1920 Bis 1933.” *GESIS Datenarchiv, Köln. ZA8013 Datenfile Version 1.0.0*, [85, 86]

Grober-Glück, Gerda. 1966a. “Verbreitung religiöser Gruppen.” In *Atlas Der Deutschen Volkskunde. Erläuterungen Bd. 2, Zu Den Karten NF 37 - 48 Und 54 - 56c, 65 - 69d, 70 - 72c, 73 - 76a,b*. Chapter XVIII, 318–38. [88]

Grober-Glück, Gerda. 1966b. “Volksglaubenvorstellung über die Wöchnerin.” In *Atlas Der Deutschen Volkskunde. Erläuterungen Bd. 2, Zu Den Karten NF 37 - 48 Und 54 - 56c, 65 - 69d, 70 - 72c, 73 - 76a,b*. Edited by Matthias Zender. Chapter XXII, 457–523. [88]

Hubatsch, W., and T. Klein. 1978. *Grundriß der Deutschen Verwaltungsgeschichte*. Marburg. [86]

Iyer, P. V. Krishna. 1949. “The First and Second Moments of Some Probability Distributions Arising from Points on a Lattice and Their Application.” *Biometrika* 36 (1/2): 135. [64]

Jíchová, J., M. Soukup, J. Nemeskal, L. Pospísilová, and P. Svoboda. 2014. *Geodatabáze Historických Statistických a Prostorových Dat Česka Ze Sčítání Lidu, Domů a Bytů 1921-2011*. Prague: Urbánní a regionální laboratoř, Přírodovědecká fakulta Univerzity Karlovy v Praze. [61, 86]

Kehren, Georg. 1994. “Möglichkeiten und Grenzen der computativen Auswertung von Daten des Atlas der Deutschen Volkskunde (ADV).” In *Bonner Kleine Reihe Zur Alltagskultur*. Edited by H.L. Cox, Hildegard Mannheims, Peter Oberem, and Adelheid Schrutka-Rechtenstamm. First. Bonn: Rheinische Vereinigung für Volkskunde, 277. [134–136]

MPIDR, and CGG. 2011. *MPIDR Population History GIS Collection (Partly Based on Hubatsch and Klein 1975 Ff.)* Rostock. [86]

MPIDR, and CGG. 2013. “MPIDR Population History GIS Collection – Europe (Partly Based on EuroGeographics for the Administrative Boundaries).” [66]

Rahlf, Thomas. 2020. “Dokumentation zu Choroplethenkarten für Deutschland, 1882-2017Documentation of Choroplet Maps for Germany, 1882-2017.” *Historical Social Research / Historische Sozialforschung*, Transition (Online Supplement). [86]

Rey, Sergio J., and Luc Anselin. 2010. “PySAL: A Python Library of Spatial Analytical Methods.” *Review of Regional Studies* 37 (1): 5–27. [64]

Roesel, Felix. 2022. “The German Local Population Database (GPOP), 1871 to 2019.” *Jahrbücher für Nationalökonomie und Statistik*, (8): [77, 78, 86]

Thome, Helmut. 2002. “Kriminalität Im Deutschen Kaiserreich, 1883-1902.” *Geschichte und Gesellschaft* 28: 519–53. [86]

Verein für Computer Genealogy. last accessed: 2022-06-15. “Das Geschichtliche Ortsverzeichnis.” http://gov.genealogy.net/search/index. [86]

Zender, Matthias. 1966. *Atlas Der Deutschen Volkskunde/Erläuterungen Bd. 2. Zu Den Karten NF 37 - 48 Und 54 - 56c, 65 - 69d, 70 - 72c, 73 - 76a,b*. Edited by Matthias Zender. Düsseldorf: Droste Verlag. [88]

Zender, Matthias, and Günter Wiegelmann. 1959. “Technische Einweisung in die neue Folge des ADV.” In *Atlas Der Deutschen Volkskunde/Erläuterungen Bd. 1. Zu Den Karten NF 1 - 36*. Edited by Matthias Zender. Düsseldorf: Droste Verlag. Chapter II, 17–21. [89]

Chapter 2

The Effect of Education on Patience – Global Evidence from Compulsory Schooling Reforms

Joint with Thomas Dohmen and Uwe Sunde

2.1 Introduction

Almost all economic decisions involve intertemporal trade-offs. Patience is a crucial determinant of intertemporal decisions. A large body of evidence documents that greater patience is associated with a wide range of future-oriented behavioral outcomes at the individual level, such as a greater savings propensity, higher education levels, and better health. Recent research has documented that variation in patience at the aggregate level has an even larger impact on the accumulation of human capital, physical capital, and the stock of knowledge than at the individual level.

Despite its importance for individual-level and aggregate outcomes, surprisingly little is known about the determinants of variation in patience. Existing work has suggested that cultural factors, religion, geography or mortality might play a role in this context, but these determinants relate to evolutionary processes that affect societies over long periods of time and only have limited relevance for policy interventions. Recent work on the formation of non-cognitive skills has suggested that preferences and attitudes are malleable during formative years, childhood and adolescence, and therefore may be influenced by education and schooling. Conceptually, schooling

* We thank Mira Fischer, Larissa Zierow, and seminar participants at Erasmus University Rotterdam, NYU Abu Dhabi, and the University of Zurich for helpful comments. Patrizia Massner and Valerie Stotthut provided excellent research assistance. Thomas Dohmen and Radost Holler acknowledge funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC/TRR 224 (Project A01) and Germany's Excellence Strategy - EXC 2126/1- 390838866. Uwe Sunde acknowledges funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC/TRR 190 (project number 280092119).

may foster patient behavior by increasing individuals' future orientation and their ability to imagine future consequences of decisions, but evidence for such a causal link is still largely missing. Moreover, the existing evidence is not entirely conclusive and limited to particular programs or natural experiments in specific contexts.

This paper contributes to the literature by examining how variation in schooling affects individual patience across the world. In particular, we use a newly constructed data set on changes in compulsory schooling laws in 48 countries around the globe during the 20th century. Combining this with information on individual patience from the *Global Preference Survey* (Falk, Becker, Dohmen, Enke, Huffman, et al., 2018) allows us to identify the causal effect of variation in schooling on patience, at different margins, cohorts and in different countries. The identification strategy exploits the quasi-experimental variation in schooling across cohorts in the different countries and compares patience levels among cohorts that have been subject to a compulsory schooling reform to those of cohorts that have just not been subject to the reform. The estimation is based on a flexible empirical model and resembles an approach that has been widely used in the education literature. This offers a unique setting to study whether individual patience is affected by schooling that combines internal consistency and external validity. In addition, exploiting variation from a variety of schooling reforms that affect individuals at different ages and levels of eduction, and in countries with different levels of development, also allows us to investigate effect heterogeneity. The results therefore provide novel insights regarding the stability and validity of results of existing studies that studied specific interventions or reforms in particular regions or countries, and at different levels of education. Moreover, the results regarding effect heterogeneity is informative about which dimensions of education policies are most relevant for affecting patience.

The estimation results yield important insights. Having been subject to an increase in compulsory years of schooling is, on average, associated with greater patience. When accounting for potential misclassification of cohorts that have been affected by an education reform, the increase in patience that can be attributed to an increase in the compulsory years of education corresponds to approximately 0.1 standard deviations. This effect is not sensitive to the specification of the empirical model or to specific countries or reforms. Placebo estimates using randomly determined earlier reform dates reveal no effect. The results are confirmed by a mediation analysis that uses compulsory schooling reforms as an instrument for years of education. The results show that, on average, an increase of one year of education is associated with an increase in patience of approximately 0.1 standard deviations. An analysis of effect heterogeneity documents that the results are mainly driven by variation in secondary education, but we find no evidence for systematic heterogeneity in the effect with respect to the level of economic development or by the quality of the education system.

Our paper complements the literature in several dimensions. Existing research on the formation of non-cognitive skills suggests that these skills are malleable in general (Heckman, Stixrud, and Urzua, 2006; Cunha and Heckman, 2007; Cunha, Heckman, and Schennach, 2010; Kosse, Deckers, Pinger, Schildberg-Hoerisch, and Falk, 2019). While it has been suggested repeatedly that education may foster patient behavior by increasing individuals' focus on the future and their capability for imagining future outcomes (Becker and Mulligan, 1997, p. 735), or by influencing individuals' general critical thinking (Cutler and Lleras-Muney, 2006; Oreopoulos and Salvanes, 2011), evidence for a causal effect of education on patience is scarce. To our knowledge, only one study by Alan and Ertac (2018) has investigated in an experimental setting whether a schooling intervention increases patience. They show that a program in elementary school that teaches forward-looking behavior by fostering the ability to imagine and evaluate the future consequences of alternative present actions leads to an increase in patience revealed by incentivized intertemporal decision tasks. This increase persists for up to three years after the program ended. Using no incentivized measures, Bülbül and Izgar (2018) find that participation in a "patience training" among 30 students of a Turkish university increased patience. Evidence from studies using quasi-experimental variation in the quantity of schooling to identify the effect of education on patience is scarce, limited to specific contexts, and not entirely conclusive. Bauer and Chytilová (2010) exploit variation in access to schooling in ten Ugandan villages and find that education significantly reduces discount rates, but only among male students. Perez-Arce (2017) identifies the causal effect of an additional year of college education on patience from randomized delayed admission to a public college in Mexico, and finds that education significantly increases patience according to one of two measures. In contrast, the present paper provides evidence from compulsory schooling reforms in 48 countries between 1947 and 2003 and covers individuals born between 1923 and 1991. Hence, the analysis complements and extends this strand of the literature in terms of coverage, comprehensiveness and external validity, and by providing evidence for effect heterogeneity across various dimensions.

Existing evidence has shown that patience is associated with a wide range of life outcomes such as savings (Ashraf, Karlan, and Yin, 2006; Finke and Huston, 2013), income (Tanaka, Camerer, and Nguyen, 2010; Golsteyn, Grönqvist, and Lindahl, 2014; Falk, Becker, Dohmen, Enke, et al., 2018), employment (Golsteyn, Grönqvist, and Lindahl, 2014), criminal behavior (Åkerlund, Golsteyn, Grönqvist, and Lindahl, 2016), educational attainment (Golsteyn, Grönqvist, and Lindahl, 2014; Castillo, Jordan, and Petrie, 2019), skills (Hanushek, Kinne, Lergetporer, and Woessmann, 2022), and various health outcomes (Smith, Bogin, and Bishai, 2005; Khwaja, Silverman, and Sloan, 2007; Bradford, 2010; Jusot and Khlat, 2013; Kim and Radoias, 2016). Recent work suggests that the association of patience with income, savings and education prevails at the individual as well as at the aggregate level, but is substantially larger at the aggregate level (Sunde, Dohmen, Enke, Falk, Huffman, et al., 2022). By documenting a feedback effect of education on patience, our evidence provides novel insights into the feedback mechanisms that might contribute to these aggregation effects. Thereby, our results highlight the importance of viewing the returns to education from a broader perspective than is often done in the literature. Due to its crucial role for any intertemporal choices, patience might be an important mediator through which education affects other outcomes, such as income and health behaviors.

Our work also contributes to studies utilizing variation in compulsory schooling laws to identify effects of education, which are mostly limited to the 'Western World' (see, e.g., Angrist and Krueger, 1991; Oreopoulos, 2007; Black, Devereux, and Salvanes, 2008; Pischke and Wachter, 2008; Brunello, Fort, and Weber, 2009; Fort, Schneeweis, and Winter-Ebmer, 2011; Gathmann, Jürges, and Reinhold, 2015; Brunello, Fort, Schneeweis, and Winter-Ebmer, 2016; Hampf, 2019). Our data recover a wide range of compulsory schooling laws that have not been used in the existing economic literature. This data should enable other researchers to extend their scope beyond this geographical region.

The remainder of the paper is organized as follows. In Section 2.2, we describe the data and our identification strategy, Section 4.4 contains the empirical results, and Section 4.5 concludes.

2.2 Data & Methods

Our analysis combines two different data sets: the data on individual patience across the globe from the *Global Preference Survey* (GPS) and a novel collection of compulsory schooling reforms across the globe.

2.2.1 Global Preference Survey

The *Global Preference Survey* (GPS) (Falk, Becker, Dohmen, Enke, et al., 2018) contains individual-level measures of time preferences, risk preferences, and social preferences for representative samples of individuals for 76 countries around the globe. The GPS was conducted as part of the Gallup World Poll 2012. The patience measure contained in the GPS data is particularly suited for the present analysis, as it has been elicited in a comparable way, using a standardized protocol, for representative samples across 76 countries, comprising approximately 80,000 respondents.

The GPS data contain two survey items that measure patience: a qualitative self-assessment of patient behavior, and a quantitative item that involves a hypothetical choice between a fixed, immediate monetary reward, and a higher but delayed monetary reward. Based on the quantitative inter-temporal trade-off, one can elicit the Internal Rate of Return *IRR* that an individual respondent demands to defer the reward by one year, using a method of unfolding brackets. This 'staircase' method asks a series of questions concerning the following trade-off:

> *Suppose you were given the choice between receiving a payment today or a payment in 12 months. We will now present to you five situations. The payment today is the same*

in each of these situations. The payment in 12 months is different in every situation. For each of these situations we would like to know which one you would choose. Please assume there is no inflation, i.e., future prices are the same as today's prices. Please consider the following: Would you rather receive amount x today or y in 12 months?

While x is held constant, the amount of y is varied to bound the respondent's indifference point between x and y. The method allows for 32 different intervals of *IRR*s for which the respondent is indifferent between receiving the smaller amount today and the larger amount in one year. The monetary amounts for x and y are adjusted for each country so that the stake sizes are comparable across countries. In particular the monetary amounts are scaled based on a country's median income, so that the early amount corresponds to roughly the same percentage of net median household income in all countries. In Germany, x corresponded to 100 Euros, and the lowest possible delayed payment equals 103 Euros, while the highest delayed payment amounts to 215 Euros. For more details see Falk, Becker, Dohmen, Enke, et al. (2018).

For the qualitative self-assessment, respondents had to answer the question: *"how willing are you to give up something that is beneficial for you today in order to benefit more from that in the future?"*. Respondents could rate their patience on an 11-point Likert scale.

The baseline measure of patience used in the empirical analysis below is the weighted average of these two measures as constructed by Falk, Becker, Dohmen, Enke, et al. (2018).[1] To facilitate the quantitative interpretation of the results, we standardize the patience measure using the global distribution of patience.[2] This measure of patience has been shown to be strongly related to various outcomes and behaviors at the individual and aggregate level (Falk, Becker, Dohmen, Enke, et al., 2018; Sunde et al., 2022).

While the GPS covers cohorts from 1914 to 1997, we only consider individuals who are older than 20 years of age because it is not clear whether younger individuals had already finished secondary education at the time of the interview in 2012. This implies that we discard respondents in the GPS born after 1991. In addition, we discard individuals who migrated to the country in which they were interviewed within the past five years prior to the interview to avoid potentially false assignment to a reform, since these respondents received their education in a different country.[3]

1. The weights were determined using experimental validation of the survey items (Falk, Becker, Dohmen, Huffman, and Sunde, 2023).

2. Concretely, we standardize the measure according to the conventional z-score formula $z_i = \frac{x_i - \mu_X}{\sigma_X}$, where z_i is the standardized value (z-score) for respondent i, x_i denotes the baseline measure of patience, μ_X denotes the global sample mean of the baseline measure of patience, and σ_X is the global sample standard deviation of the patience measure.

3. This sample restriction is based on the only information about migration status available in the Gallup World Poll.

2.2.2 Compulsory Schooling Reforms Across the Globe

In order to scrutinize the effect of schooling reforms on patience, we constructed a novel data set that records compulsory schooling reforms in the 20th and early 21st century for the 76 countries covered by the GPS and merged it to the GPS data. To obtain information about compulsory schooling reforms, we searched through research papers, books, UNESCO reports, and other official government information, such as laws and reports. We restricted attention to reforms that where passed between 1947 and 2003 as only these reforms could possibly have affected the cohorts included in our GPS sample. This leaves us with 50 distinct reforms in 48 countries. Of the remaining 28 countries covered by the GPS, in 12 countries compulsory schooling did not change during the relevant time period.[4] For the remaining 16 countries, we decided not to use the information on compulsory schooling that we could detect. For five of these countries the literature is quite clear on a lack of enforcement of the compulsory schooling reforms.[5] For another five countries, we were not able to recover the compulsory schooling laws in a sufficiently precise manner, typically due to the lack of information about the time of the passing and of the implementation of the respective compulsory schooling reforms.[6] For the remaining set of countries, it was clear *ex ante* that the sample size would be too small as schooling laws were implemented at the local level, so that we did not pursue

4. Haiti legally introduced compulsory primary education (six years) in 1818 and has yet to fully implement it (Prou, 2009). In Suriname, the Dutch introduced compulsory schooling in 1876. There was no change in the level during the 20th century (Van Stipriaan, 1998; Godbardhan-Rambocus, 2015). Guatemala's 1879 constitution already prescribed mandatory schooling ('Constitucion de Guatemala de 1879'). We could not discover any reform after that. Nigeria launched a Universal Basic Education (UBE) program in 1999 which aimed at providing free and compulsory education until the end of 9th grade. This was passed into law in 2004 (Jaiyeoba, 2007; Ejieh, 2009). The pivotal cohort is estimated to be the 1998 cohort. Sri Lanka introduced compulsory enrollment in 1945. In 1997, it added legislation that made attendance mandatory (Bajaj and Kidwai, 2016). Saudi Arabia and Botswana lacked a universal compulsory schooling law at least until 2006 (Benavot and Resnik, 2006) and 2007 (Isaacs, 2007), respectively. Nicaragua introduced compulsory education in 2006 ('Ley núm. 582, Ley General de Educación', also see Worldbank (last accessed: 2020-06-01)). Rwanda and Uganda introduced compulsory schooling legislation in 2008. India did not introduce universal compulsory schooling laws on the country-level until 2010 ('The Right of Children to Free and Compulsory Education Act'). Some of its states introduced compulsory schooling earlier. These were, however, scarce and rarely enforced (Weiner, 1991). Pakistan introduced free and compulsory education in its 18th Amendment to the Constitution in 2010 (Ullah, 2013).

5. Yugoslavia legally introduced compulsory schooling in 1950, this was not enforced at that time (Tomich, 1963). This affects the following countries in our data set: Croatia, Bosnia Herzegovina, and Serbia. Brazil first introduced compulsory schooling in 1824 (Plank, 1990), and expanded it in 1971 (Vahl, 2018). Vahl (2018) shows that primary school enrollment decreased after this reform. Afghanistan was subject to multiple increases and decreases in compulsory education between 1931 and 1978. In 1978, four years of primary school were made compulsory (the previous reform, in 1976, increased this to eight years), this was increased to 5 years in 1986, to 6 years in 1990 (Samady, 2001). Samady (2001) raises strong doubt that any of these reforms actually had any effect more than proforma legislation.

6. These countries comprise Algeria, Cameroon, Malawi, Kenya, Tanzania.

these reforms.[7] Finally, we were able to identify compulsory schooling reforms for 48 countries that meet these criteria.[8] Figure 2.2.1 displays a map of the 48 countries with reforms in our data set. The map illustrates that the data largely cover Latin America, Europe, and Asia, but that the coverage of Africa and North America is rather scant.

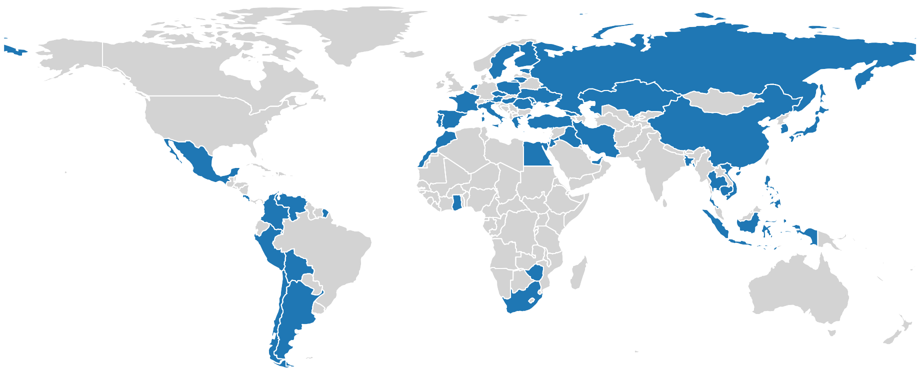


Figure 2.2.1. Countries included

Some countries passed and implemented more than one compulsory schooling reform within the respective time period, and we aimed at keeping all implemented reforms in our sample. Indonesia and the Netherlands, however, passed two compulsory schooling reforms within 10 and 12 years, respectively. In this case, we excluded one of the two reforms to avoid overlap of treatment from different reforms while keeping the number of observations sufficiently large. In the main specification, we include the reform that affects the larger number of observations.[9] We provide robustness checks for alternative codings in the Appendix. Moreover, seven countries in our sample were formerly part of the Soviet Union (Russia, Ukraine, Estonia, Georgia, Kazakhstan, Lithuania, Moldova) and thus subject to the same compulsory schooling reforms prior to 1990. Overall, our estimation sample comprises 50 reforms for 41 reform units in 48 countries. Appendix 2.C contains a detailed de-

7. For example, in Switzerland schooling laws are implemented at the level of the 24 cantons. In the United States, Australia, and Germany, schooling laws are implemented at the state level. In Canada, schooling laws are enacted at the province level. The United Kingdom enacted reforms separately for England/Wales, Scotland, and Northern Ireland. For the United Kingdom and Germany, the additional problem arises, that information about years of education is missing in the GPS data. Reforms of compulsory schooling for these two countries are available in Pischke and Wachter (2008), Gathmann, Jürges, and Reinhold (2015), and Fort, Schneeweis, and Winter-Ebmer (2016).

8. Note that some countries such as the Czech Republic and Egypt experienced increases as well as decreases in compulsory years of schooling. Here, we focus only on reforms that increase the compulsory years of schooling. See Appendix 2.C for details.

9. This means that for the Netherlands, we keep the reform of 1973 and exclude the reform of 1985. For Indonesia, we keep the 1984 reform and exclude the reform implemented in 1994. Below we present a robustness check that uses the reforms of 1985 and 1994, respectively.

scription of the reforms, including a detailed list of these reforms that is displayed in Table 2.C.1.

Figure 2.2.2 illustrates the timing of reforms as well as the level of compulsory schooling stipulated by the respective reform. In particular, the horizontal axis displays the year in which a reform was implemented, while the vertical axis shows the years of compulsory schooling it stipulates. The colors of the respective country markers indicate the world region in which the country is located. The figure shows that Japan's reform of 1947 is the first reform included in our sample. This is not because Japan was the first to introduce a compulsory schooling law, but rather because it affects the oldest cohort in our sample for which a sufficient number of individual observations is available in the GPS data both for the cohorts affected and for the cohorts not affected by the reform. In the case of Japan, sufficiently many observations of old cohorts are available to permit the analysis of a reform as early as 1947.

The figure illustrates that reforms came in geographically clustered reform waves. The majority of reforms in Europe that are covered by our sample were introduced between 1960 and 1970. These reforms typically raised compulsory schooling to a duration of eight to ten years. With the exception of Japan, all Asian countries included in our sample implemented a reform between the late 1970s and early 1990s. The level of compulsory years of schooling associated with these reforms is more heterogeneous than in Europe and ranges from five to nine years. South American countries introduced compulsory schooling quite early, and often before the relevant observation period (see Appendix 2.C). This implies that new reforms are not observed until the wave of substantive reforms in compulsory schooling laws that began in the 1990's. A similar pattern is observed for some European countries. Chile and Peru are the only countries in our sample that made upper secondary education compulsory during the observation period.

2.2.3 Empirical Strategy

In order to identify the effect of the compulsory schooling reforms on patience, we classify each individual in the estimation sample as being affected or unaffected by a reform. This classification is based on the year of birth, the country, and the pivotal cohort of a reform, i.e., the cohort that was the first to be potentially affected by the reform.[10] Accordingly, we construct a treatment indicator that takes the value 1 if the respondent is born in the same year as the pivotal birth cohort or in later years,

10. For 29 of the reforms, we were able to draw on research papers and/or laws that contained information on the implementation of the reform to construct the pivotal cohort. For the remaining reforms, we had to approximate the pivotal cohort. For these reforms, we assume that the reform applies to everyone still in compulsory schooling according to the law that was in place before the reform. The rationale behind this is that almost all laws that we recovered apply retroactively, but those who already left school or could leave school were rarely forced to return to school. In case our sources only state the year in which the law was enforced but not the first cohort for which it was

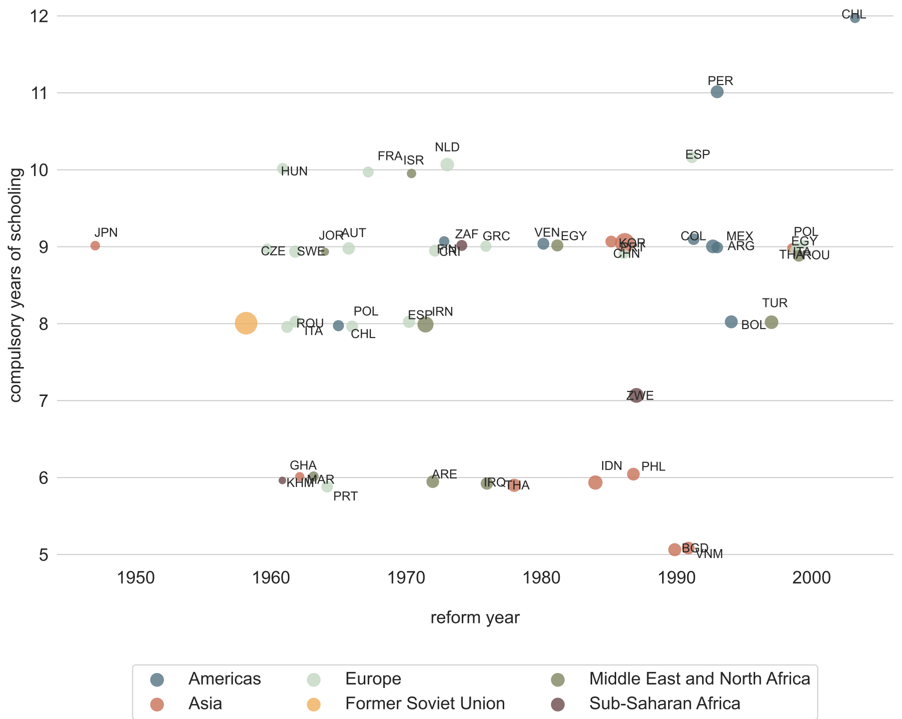


Figure 2.2.2. Distribution of Compulsory Schooling Across Countries and Time

Note: Data points are scaled by sample size in the *GPS* within an interval of +/- 10 years of the pivotal cohort. Labels correspond to the ISO code of the respective country.

and 0 otherwise. This treatment specification is based on the logic that the timing of the implementation of a compulsory schooling reform is governed by complex political processes that are unrelated to the particular birth cohorts. Hence, from the perspective of individuals of different birth cohorts, the implementation of a schooling reform, and hence the assignment of the treatment status, is exogenous.[11]

To make meaningful inference about the causal effect of the reform, individuals affected by the reform and individuals unaffected by the reform should be as comparable as possible in all relevant dimensions other than the reform. To avoid that

---

binding, we approximate the pivotal (birth-)cohort by using the following formula:

$$\begin{aligned}\text{pivotal cohort } &= \text{year of reform implementation}\\ &\quad - \text{primary school entrance age}\\ &\quad - \text{years of compulsory schooling before the reform}\end{aligned}$$

Table 2.C.1 indicates for which reforms we estimated the cohort according to this procedure. The third column of Table 2.C.1 in Appendix 2.C indicates the pivotal cohort of each reform.

11. While the implementation of schooling reforms might be related to other reforms or political processes, the particular cohort that is affected by the reform is largely random, since the implementation is typically the outcome of a lengthy political process with substantial uncertainty, see also Strittmatter and Sunde (2013) for an analogous discussion in the context of health reforms.

unobserved factors that change over time potentially confound the estimated reform effect, we would ideally only compare the youngest cohort that is just not affected by the reform to the oldest cohort that is just affected by the reform. However, due to the nature of the data, this would leave us with only a very small number of observations from which the reform effect would be identified. Therefore, we pursue a different strategy: We expand the comparison window to cohorts born within at most 10 years of the pivotal cohort of a reform.[12] This ensures a sufficiently large estimation sample. In order to control for unobserved factors that evolve over time and that might affect patience independently of the reform, we flexibly control for time trends as discussed below. Applying these restrictions leaves us with 19,680 observations for the analysis. Table 3.B.1 in the Appendix provides descriptive statistics for the resulting estimation sample.[13]

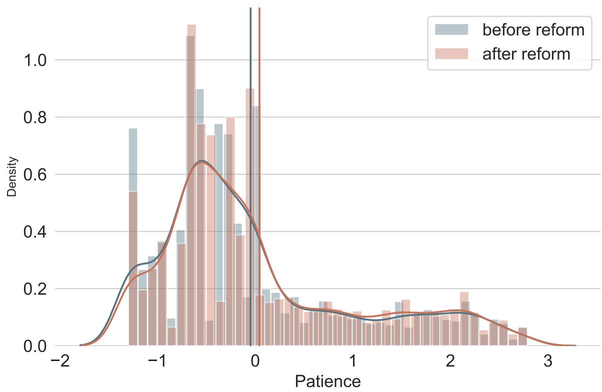
The sample is unbalanced within and across reforms (see Figure 2.2.2 and Table 2.C.1), for two obvious reasons: First, depending on the timing of the reform and the age distribution of a country, the data contains more or fewer individuals born within the 10-year window of the pivotal cohort. Second, the GPS contains more observations for China (2,574) and Russia (1,498) than the median sample size (1,000) for other countries. The resulting samples vary between less than 200 observations and more than 1,500 observations per reform. To avoid that the estimation results are driven by particular countries and/or reforms as a result of oversampling, we apply inverse probability weights to ensure that each reform receives equal weight (pre- and post-treatment) in the estimation.[14] As an additional advantage, this weighting procedure also facilitates the interpretation of the results. The estimated effect of the compulsory schooling reforms on patience can therefore be interpreted as the intention-to-treat effect of being randomly assigned to a reform. Moreover, with this weighting scheme in place, sample averages and distributions before and after reforms are more comparable because differences in the distributions are not affected by changes in composition of the sample. For instance, Figure 2.2.3a displays the weighted distribution of patience for individuals born before the pivotal cohort of a reform and individuals born after the pivotal cohort of a reform. Individuals born after the pivotal cohort, i.e., individuals that have been affected by the respective reform, are on average approximately 0.09 standard deviations more patient than

12. For reforms in Egypt and Thailand, this window is seven and nine years respectively, because otherwise the reforms would overlap within one country, so that some respondents would be the control group for one reform and the treatment group for another reform. In comparison to the overlapping reforms in Netherlands and Indonesia discussed above, these reforms are far enough apart to allow a comparably large number of observations per reform. See also Table 2.C.1.

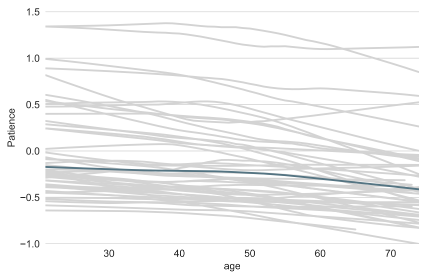
13. Figure 2.A.1 provides a histogram of the sample composition around the pivotal cohort.

14. The weights are constructed in the following way: Let N denote the full sample size, n_{1r} denotes the number of observations before the reform for reform r and n_{2r} denotes the number of observations after a reform. Then the non-normalized inverse probability weights correspond to $\hat{w}_{jr} = \frac{1}{n_{jr}}$ for $j = 1, 2$, and the normalized weights correspond to $w_{jr} = \frac{\hat{w}_{jr}}{\sum_r \sum_j \hat{w}_{jr}} \times N$.

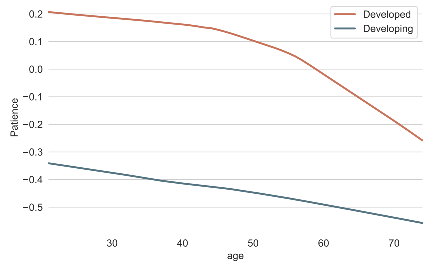
Figure 2.2.3. Distributions and Evolution of Patience



(a) Weighted Distribution of Patience



(b) Patience by Age



(c) Age Gradient by Level of Development

Note: Figure 2.2.3a displays weighted distributions of *Patience* within +/-10 years of reforms. Vertical lines correspond to the respective average. Figure 2.2.3b displays the evolution of *Patience* across age for all countries in the sample using locally smoothed regression (LOWESS). Blue line corresponds to the age gradient pooled across countries. Figure 2.2.3c shows the age gradient by level of development. Both figures exclude individuals older than 75 because the data gets very sparse for most countries making comparison across countries and groups difficult. 4% of the GPS sample restricted to the respective countries is above 75.

individuals born before pivotal cohort, i.e., individuals that were not affected by the respective reform. The difference between the distribution of pre-reform levels of patience and post-reform levels of patience is particularly marked at the lower end of the patience distribution.

This simple contrast across groups might conceal systematic differences across age groups or birth cohorts.[15] Figure 2.2.3b displays the age gradient in patience for each country separately (grey) and for the entire sample (blue). Patience decreases with age (during adulthood) across countries.[16] Despite this overall decrease in pa-

15. Recall that the GPS represents a cross-sectional data set. Hence, birth cohort and age are perfectly collinear. In the following, we will refer to age or cohort effects as age effects for simplicity.

16. Existing work suggests that the life cycle pattern of patience is hump-shaped, increasing during childhood/adolescence (see, e.g., Bettinger and Slonim, 2007).

tience, there seems to be country-specific variation in the strength and form of the decline. In line with findings in the existing literature, the form of the age gradient depends on the level of development (Falk, Becker, Dohmen, Enke, et al., 2018) and income (Burro, McDonald, Read, and Taj, 2022). Figure 2.2.3c shows that the age gradient is more concave for developed countries, while patience almost linearly decays in developing countries. Here and in what follows, we define a country to be developed if its Human Development Index (2020) is above or equal to 0.8.[17]

To account for systematic heterogeneity in age and other potentially relevant factors, we estimate the following baseline specification of the empirical model

$$PAT_{iarc} = \beta Treated_{arc} + \gamma_{rc} f_{rc} + \alpha_{rc} f_{arc} + \epsilon_{iarc} \,. \tag{2.2.1}$$

The dependent variable PAT_{iarc} denotes the level of patience of respondent i, as measured in the GPS data, of age a who is affected/not affected by reform r in country c. $Treated_{rac}$ is an age-reform-(country-)specific binary indicator that takes the value 1 if the respondent of age a is born in the year of birth of the pivotal cohort of reform r in country c or after, i.e., whether the respondent has been treated by the reform. To account for systematic variation across reforms implemented at different points in time, the specification includes separate fixed effects f_{rc} for each reform r in country c.[18] To account for systematic differences in age between individuals that were affected by the respective reform and individuals that were not, i.e., between the treated and the control group, the specification also controls for reform-country-specific age functions f_{arc}.

The baseline specification (2.2.1) implicitly assumes that (i) country-reform specific age effects in patience are well approximated by the respective age functions; and (ii) that all systematic deviations from the age trend post-reform can be attributed to the respective compulsory schooling reform. Since it is implicitly assumed in this specification that pre-reform and post-reform age trends are homogeneous, the country-specific age functions need to approximate these age trends well over a 20 year window. Figure 2.2.3b indicates that country-specific age effects in patience are well approximated by linear age effects over shorter horizons and quadratic age effects over longer horizons.

In additional analyses, we relax assumptions (i) and (ii) by estimating an extended version of the empirical specification that includes distinct pre-reform and

17. This corresponds to the sample median. According to this measure, countries such as Kazakhstan, Romania, Russia and Argentina are just above the cutoff, while countries that are classified as developing with a value just below this threshold include Iran, Turkey, and Costa Rica.

18. With most countries exhibiting only one reform during the observation period, the reform-country-specific effects collapse to reform-specific effects. The reform-specific and reform-country-specific effects mainly account for the systematic variation across the countries of the former Soviet Union, where multiple contemporary countries were affected by the same reform.

post-reform age effects,

$$PAT_{iarc} = \beta Treated_{arc} + \gamma_{rc} f_{rc} + \alpha_{rc} f_{arc} + \mu_{rc} Treated_{arc} \times f_{arc} + \epsilon_{iarc} \qquad (1')$$

where, as a baseline, we estimate a specification with linear pre-reform and post-reform age effects $Treated_{arc} \times f_{arc}$. This more flexible specification does not rely on the assumption of homogeneous pre-reform and post-reform trends. It instead entails that the effect of the reform is captured by the variation in the intercept.[19] Conceptually, this setting is similar to a regression discontinuity design with the birth cohort as the forcing variable, thus resembling a specification that has been widely used in the education literature (see, e.g., Brunello, Fort, and Weber, 2009; Fort, Schneeweis, and Winter-Ebmer, 2011; Brunello, Fort, Schneeweis, et al., 2016). In comparison to this literature, however, our setting includes more reforms but fewer observations per reform, which prevents the implementation of more flexible, non-parametric methods.

To further account for differences in life-cycle patterns of patience across developed and developing countries that might arise from systematic heterogeneity in economic living conditions, public health, or mortality, we also report estimation results for extended specifications that account for heterogeneity in age effects according to the level of development. We implement flexible controls for heterogeneous age effects across different levels of development by assigning individuals to five-year bins and interact the bins with a binary indicator that takes the value 1 if a country is developed and 0 otherwise.

Finally, we estimate a specification that addresses the potential concern that compulsory schooling reforms were not always introduced in every school or region at the same time. In fact, some reforms were phased-in over multiple years. This applies to the reforms in Argentina, Bolivia, Chile (2003), China, Finland, Israel, Soviet Union, and Sweden (for details see Appendix 2.C). For these reforms, it is not possible to determine a single pivotal cohort. The third column of Table 2.C.1, therefore provides a range of pivotal cohorts rather than a particular cohort.

Apart from incomplete coverage in the reform year, cut-off dates of school entry that fall in the same year can contaminate the treatment group, such that not every individual member of the pivotal cohort is necessarily treated. Reforms usually apply to school cohorts, but school cohorts do not perfectly overlap with birth cohorts. In many countries, the cut-off date for school entry does not coincide with the beginning of a calendar year, but rather a cut-off date during the year, so that school entry age applies to individuals born as of, e.g., August 1st or September 1st of a particular year. The lack of information about the month of birth makes it more difficult to construct cut-offs based on birth cohorts and to construct the pivotal cohort for

19. Some of the reforms in our data have less than 100 individual observations before or after the reform. Thus, using more flexible trend functions increases potential concerns of overfitting.

the treatment assignment of respondents. To deal with this issue, we construct a binary indicator variable $Partially\ NonTreated_{arc}$ that takes the value 1 if an individual is either born in the pivotal cohort of the respective reform in a country or during the period of the phase-in of the reform, and thus may not yet be affected by the reform. In the weighted sample, approximately 8% of the observations are partially non-treated (see Table 2.A.2). To account for potential misassignment to the pivotal cohort we estimate a specification of the empirical model that extends the baseline specification by including the control variable $Partially\ NonTreated_{ra}$:

$$\begin{aligned} PAT_{iarc} = &\beta Treated_{arc} + \gamma_{rc} f_{rc} + \alpha_{rc} f_{arc} \left(+\mu_{rc} Treated_{arc} \times f_{arc}\right) + \epsilon_{iarc} \\ &+ \eta Treated_{rac} \times Partially\ NonTreated_{arc}. \end{aligned} \tag{2.2.2}$$

Below, we present the results for this extended version of the empirical specifications along with the results of the baseline specification.[20]

2.3 Results

2.3.1 Main Results

Table 2.3.1 presents the estimated effects of compulsory schooling reforms on patience for the different specifications described in the previous section. Panel A of Table 2.3.1 reports the estimated effect of being treated by a compulsory schooling reform on patience for the baseline specification (2.2.1) of our empirical framework. In particular, the specifications in Columns (1) and (2) account for country-reform specific linear age effects, while the specifications in Columns (3) and (4) account for country-reform-specific quadratic age effects. The results in Columns (5)–(6) report results for the extended specification accounting for heterogeneous pre-reform and post-reform age effects as in the extended empirical model (1'). In particular, the specifications in Columns (5) and (6) allow for country-reform-specific linear pre-reform and post-reform age trends. The estimates across Columns (1) to (6) consistently imply that having been subject to a compulsory schooling implies an increase in patience of approximately 0.05 standard deviations.

The results reported in Panel A of Table 2.3.1 indicate that education has a positive effect on patience. Notably, the estimates are attenuated as some individuals might be erroneously classified as treated, as explained in the previous section. The estimates in Panel B of Table 2.3.1 address the potential problem of attenuation bias

20. Conceptually, this specification follows the logic of a fuzzy regression discontinuity design. The relatively small number of observations of individuals in the birth cohorts around the respective cut-off prevents the implementation of more demanding estimators along the lines of a fuzzy regression discontinuity design using non-parametric methods.

Table 2.3.1. Compulsory Schooling Reforms and Patience

	Patience					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A.						
Treated	0.058*	0.054*	0.054*	0.05	0.058*	0.045
	(0.032)	(0.032)	(0.033)	(0.031)	(0.032)	(0.03)
Adj. R^2	0.142	0.142	0.142	0.142	0.142	0.142
Panel B.						
Treated	0.103***	0.096***	0.095**	0.087**	0.099**	0.086**
	(0.036)	(0.033)	(0.039)	(0.034)	(0.039)	(0.035)
Partially Non-Treated	−0.088***	−0.079***	−0.078**	−0.069**	−0.085**	−0.081**
	(0.029)	(0.029)	(0.033)	(0.03)	(0.035)	(0.033)
Adj. R^2	0.142	0.142	0.142	0.142	0.142	0.142
Age trend	homogeneous linear	homogeneous linear	homogeneous quadratic	homogeneous quadratic	heterogeneous linear	heterogeneous linear
Age Bins × Developed		✓		✓		✓
N	19,680	19,680	19,680	19,680	19,680	19,680

Notes: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. OLS estimates of the effect of compulsory schooling reforms on patience. The dependent variable is the standardized measure of patience. Treated is defined as described in the text. All specifications include reform specific trends, country-reform specific fixed effects. Column (2), (4), (6) and (8) additionally include age bin fixed effects that are allowed to vary by level of development. Age bins are 5-year age intervals. Developed is an indicator that is 1 if a country can be classified as developed in terms of the Human Development Index, and 0 otherwise. Observations are weighted with the aforementioned weighting scheme. Standard errors are clustered on the country level.

caused by contamination of the treatment group, by showing estimates for the extended specification of the empirical framework (2.2.2) that accounts for potential heterogeneity of the effect of the reform for cohorts that were only partially treated. The results indicate that misclassification of the treatment group might indeed affect the estimates in Panel A, as the estimated effect of a compulsory schooling reform on patience increases to 0.1 standard deviations, but is significantly (and sizably) smaller for members of cohorts that are likely to be only partially treated.

In the following, we consider the specifications in Columns (5) and (6) of Panel B as the most preferred specifications, because they account for varying country-specific linear pre-reform and post-reform trends and, additionally, control for general age effects in Column (6). In this specification, the effect of being subject to a compulsory education reform is associated with an increase in individual patience by 0.099 and 0.086 standard deviations, respectively.

2.3.2 Robustness

Empirical Specification. The results are robust to several modifications of the empirical framework. In particular, we replicated the analysis for a specification with a narrower window of cohorts around the reform date, defining the comparison window to cohorts born within at most seven years of the pivotal cohort of a reform, instead of 10 years as in the baseline specification. The results are reported in Table 2.B.1 and deliver quantitatively larger effects of the reform on patience. In particular, for these results the estimated effect of a compulsory schooling reform on patience increases to almost 0.1 standard deviations in the baseline specification (1') and to 0.14 standard deviations for the extended specification (2.2.2). Again, the effect for members of partially treated cohorts is significantly smaller, but still positive in this narrower comparison window. At the same time, this specification implies that the reform effect is identified from a substantially smaller number of observations.

Alternatively, we estimated even more flexible specifications that allow for country-reform-specific quadratic pre-reform and post-reform age trends, with qualitatively similar results as in the baseline specification (see Table 2.B.2 in the Appendix). In fact, the respective estimates are even larger, at the order of 0.1 standard deviations for the baseline specification (1') and 0.15 standard deviations for the extended specification in (2.2.2). When interpreting these estimates one has to keep in mind, however, that this more flexible specification implies that another 96 parameters have to be estimated in comparison to the specifications with linear pre-/post-reform trends in Columns (5) and (6) of Table 2.3.1, and in a rather small sample. This might entail potential overfitting of the data due to exceedingly flexible specifications of quadratic pre-reform and post-reform cohort trends. As a consequence, not only do the estimates get larger, but so do the standard errors. Moreover, regardless of the specification, the estimates of the effect of being subject to a compulsory schooling reform are quantitatively somewhat smaller when including age-bin fixed effects that are allowed to vary by the level of development, in comparison to specifications without adjusting for age-bin fixed effects (i.e., when comparing odd-numbered to the corresponding even-numbered columns in the table). Hence, we take these findings as confirmation for the robustness of the main results from the preferred specification shown in Column (6) of Table 2.3.1.

The results are also similar when estimating clustered standard errors using the wild bootstrap procedure (see Table 2.B.3) and when clustering standard errors at the level of reforms r, rather than at the country level (see Table 2.B.4).

Alternative Specifications: Partially Non-Treated. The results are robust to accounting for partially non-treated cohorts in different ways. In particular, one may suspect that the effects for partially non-treated cohorts vary by reform or by time of implementation, or by both. For instance, in countries in which the reform was introduced in a staggered fashion, one would expect the effect to be lowest for those cohorts that are closest to the pivotal cohort (for which the implementation was rather imperfect), and higher for those cohorts farther away from the pivotal cohort (which are more distinct in terms of treatment status). Moreover, in some countries children of the pivotal cohort who are born after April 1st are treated, while in other countries this would only apply to children born after September 30th. To account for this, we replicated the analysis with flexible specifications that accommodate linear pre-reform and post-reform trends and allow for country-specific, time-specific or country-time-specific effects for the partially non-treated cohorts. The results, which are reported in Table 2.B.5 in the Appendix, reveal that, if anything, the effects of schooling on patience is even larger and remains significant at the 5% level.

Placebos: Alternative Implementation Dates. To account for potential concerns that the results are driven by a discontinuity in an age trend that is approximated imperfectly and that is picked up by the treatment indicator, or by country-specific shocks that affect treatment and control groups asymmetrically, we conducted a placebo analysis. This accounts for the possibility that the identifying variation does not stem from the reforms themselves, but instead from allowing for a discontinuity in age or cohort trends that are otherwise assumed to be continuous. To address this concern, we replicate the analysis for specifications in which the pivotal cohort is shifted by one to seven years backward in time (towards earlier implementation). We apply a cut-off of seven years because the artificial control group in some countries would be partly subject to an earlier schooling reform when shifting the placebo implementation date by more than seven years. We refrain from performing an analogous placebo shift forward in time (towards later implementation), because this could potentially pick up a stronger effect when pivotal cohorts are only very partially treated and reforms were introduced continuously over a several years. The results of the placebo estimates are reported in Figure 2.B.1. As expected, the estimates of the effect of schooling decrease towards zero when the pivotal cohort is artificially shifted away from the actual pivotal cohort, and eventually become insignificant and fluctuate around zero. The results of this placebo analysis therefore provide evidence that the identifying variation is indeed linked to the compulsory schooling reforms.

Sensitivity to Specific Reforms. To investigate the sensitivity of the results with respect to particular reforms or countries, we conducted three additional robustness checks. First, we replicated the analysis for alternative reform coding in countries with multiple reforms (as in Indonesia or the Netherlands). Instead of using the first reform in the data as in the baseline specification, re-estimating the model using the second reform delivers very similar results to the baseline (see Table 2.B.6 in the Appendix). Second, we investigated whether our estimates are driven by a particular reform by re-estimating the preferred specification, but for samples that each time leave out one reform. The corresponding estimates are displayed in Figure 2.B.2 in the Appendix. While there is some fluctuation in the estimates, the 95%-

confidence intervals of all estimates cover the 'global' estimate, and all estimates are significantly different from zero. Third, we divide the reforms by whether or not they have been shown in a causal or descriptive study to have effectively increased enrollment or years of schooling. The corresponding results are shown in Table 2.B.7. We find that reforms that were analyzed previously in a causal study have almost the same effect as those that have never been analyzed in any study. Reforms for which only descriptive studies regarding their effects on schooling exist do not display a significant effect on patience. This suggests that the effect on patience indeed reflects a causal effect of patience, whereas patience does not increase in contexts where there seems to be no compelling quantitative evidence for the effectiveness of the respective reform on school outcomes.

Controlling for Other Preference Measures. Due to its importance for intertemporal choice, and hence almost all economic decisions, the analysis so far has focused on patience. Previous evidence has documented a correlation between patience and other economically relevant traits, such as risk attitudes (Dohmen, Falk, Huffman, and Sunde, 2010). To investigate the robustness of the results with respect to other traits as potential confounds, we replicated the analysis by leveraging the availability of other preferences measures in the GPS. The results from estimating an extended specification that also includes controls for risk attitudes, positive reciprocity, negative reciprocity, trust and altruism deliver a qualitatively and quantitatively similar pattern of results. As consequence of the limited variation, the estimates exhibit larger p-values, particularly for the more demanding specifications, such that the estimates are not always statistically significant at conventional levels (see Table 2.B.8 in the Appendix).

2.3.3 Mediation Analysis: Years of Education

The results presented so far suggest that education indeed affects individual patience. The estimates can be viewed as reduced form (intention-to-treat) evidence that makes use of exogenous variation in compulsory schooling to identify the effect of interest. A more conventional, and potentially more intuitive, approach would be to relate patience to years of education. In fact, one might consider using compulsory schooling reforms as an instrument for years of education in order to estimate the causal effect of one additional year of education on patience. This is informative by documenting the relevance of the instrument for predicting actual years of education as reported in the survey. At the same time, the validity of this approach for estimating the effect of education years also requires assuming that compulsory schooling reforms affect patience *only* through their effect on years of education. A priori, it is not entirely clear whether this exclusion restriction is justified. For instance, compulsory schooling reforms might lead to a change in class sizes and class composition, school sizes and other factors that influence individual attitudes such as patience. In the following, we report the results of instrumental variable regressions to document the results, but we emphasize that their interpretation is subject to considerations about the validity of the exclusion restriction.

Table 2.3.2 contains the results from OLS and 2SLS regressions of patience on the years of education. The OLS results, shown in Panel A, suggest that more years of education are associated with a significantly higher level of patience.[21] However, this association might be affected by reverse causality (for instance, as more patient individuals are willing to spend more time in education to benefit from a higher return on the labor market) or simultaneity/omitted variables (since, e.g., respondents with highly educated parents might exhibit greater patience and higher education outcomes than respondents with less educated parents). Exogenous variation in compulsory schooling legislation offers the possibility to account for these confounds by applying an instrumentation strategy.

21. The estimates are based on the preferred empirical specification (1').

Table 2.3.2. Education and Patience – IV Estimates

	Patience		*Yrs. education*	
	(1)	(2)	(3)	(4)
Panel A. OLS				
Yrs. education	0.022***	0.022***		
	(0.003)	(0.003)		
Partially Non-Treated	−0.029	−0.034		
	(0.029)	(0.03)		
Panel B. IV				
Yrs. education	0.135*	0.111*		
	(0.07)	(0.062)		
Partially Non-Treated	−0.021	−0.028		
	(0.039)	(0.036)		
Treated			0.711***	0.728***
			(0.196)	(0.183)
Partially Non-Treated			−0.486**	−0.486***
			(0.192)	(0.184)
F			13.22	15.83
AR 95% CI	(0.017 0.378)	(0.007, 0.302)		
AR *p*-value	0.028	0.037		
Age trend	heterogeneous linear	heterogeneous linear	heterogeneous linear	heterogeneous linear
Age Bins × Developed		✓		✓
N	19,415	19,415	19,415	19,415

Notes: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$ of the respective *t*-test. Panel A: OLS estimates of Years of Education on Patience. Panel B: 2SLS estimates of Years of Education on Patience, using compulsory schooling reforms as an instrument. Panel B Columns (3) and (4): first stage estimates. *F* corresponds to the *F*-statistic of the excluded instrument. AR 95% CI is the Anderson-Rubin 95% confidence interval, AR *p*-value is the *p*-value of the Anderson-Rubin Test under the H_0 of no effect (Anderson and Rubin, 1949). All specifications include an indicator variable that is 1 if a cohort in a specific country is partially non-treated, country-reform fixed effects, country-reform specific linear pre- and post-reform trends. Specifications in Columns (2) and (4) additionally include age bin fixed effects that are allowed to vary by level of development. Age bins are 5-year age intervals. Developed is an indicator that is 1 if a country can be classified as developed in terms of the Human Development Index, and 0 otherwise. Individuals that indicate that they have more than 30 years of education are excluded. Observations are weighted with the aforementioned weighting scheme. Standard errors are clustered on the country-level.

The corresponding 2SLS results, shown in Panel B, reveal two sets of distinct findings. The first stage results in Columns (3) and (4) of Panel B document that being subject to a compulsory schooling reform in the sample implies an increase in the years of education by approximately 0.7 years. In light of the fact that some of the respondents already planned to attend secondary and tertiary schooling in the absence of the reform, as well as in view of potential error from misclassification, this implies a reasonably high compliance to the reform. The estimates are significant at the 1% level and the corresponding F-Statistic varies between 13 and 16. The second stage results in Columns (1) and (2) of Panel B show that the corresponding IV estimate of the effect of one additional year of education on patience is about 0.1 standard deviations. Based on a conventional t-test, the estimate is significantly different from zero at the 10% level. Given that the instrument is rather weak, however, inference by means of the conventional t-test might be incorrect. To account for this, we additionally report Anderson-Rubin confidence intervals as well as the p-value of the Anderson and Rubin (1949) test, which is robust to weak instruments. While the confidence intervals are rather wide (more than 0.3 standard deviations), they exclude zero, and the respective p-values suggest that the IV estimate is significantly different from zero at the 5% level. Hence, under the assumption that the exclusion restriction is satisfied, the results suggest that an additional year of education due to a compulsory schooling reform causally leads to greater patience. Whether the effect is larger or smaller than the OLS estimate is not clear, however, since the confidence intervals of the second-stage effect are fairly wide and include the OLS estimate in Panel A for all estimates.

2.3.4 Effect Heterogeneity

The analysis so far was based on the implicit assumption that all compulsory schooling reforms in the sample exert a homogeneous effect on patience. It appears plausible, however, that the effect of reforms on patience is heterogeneous along different dimensions related to the schooling reforms. First, the effect might depend on the stage of education and the corresponding content, i.e., whether the reform applied to grade 5 or grade 9. In fact, the literature on skill formation suggests that the age at which the education intervention occurs might matter (see, e.g., Cunha and Heckman, 2007). In addition, the content of education differs across the stages of education and across countries, whereas the content of the education program and the curriculum might be more important than just the length of education for the formation of patience. Also, the degree to which reforms effectively change school attainment and the acquisition of skills differs depending on whether the reform induces a minimum of six or ten years of schooling. These arguments imply that the average effect might conceal systematic effect heterogeneity according to the stage of education that the respective reforms target.

Table 2.3.3 reports estimation results regarding heterogeneity of the reform effect. To investigate the possibility of effect heterogeneity by education level, we divide our reforms with respect to whether they target primary education (operationalized as 8 years of education or less) or secondary education (more than 8 years of education). The estimates shown in Columns (1) and (5) indicate that the effect of compulsory schooling on patience is almost exclusively driven by reforms that target secondary education.[22]

Compulsory schooling reforms and their effects on patience may also differ across countries in terms of the circumstances in which they are introduced. The most salient difference relates to the level of economic development, which may itself influence the effect of a reform. In addition, we have documented above that developed countries are more likely to implement longer durations of compulsory schooling. In columns (2) and (6) of Table 2.3.3, we allow for effect heterogeneity by the level of development, using a binary classification to distinguish between developed and developing countries in terms of a median split of the sample. The estimates reveal that the effects are not subject to significant heterogeneity along this dimension. The respective coefficient of the interaction term is positive, suggesting that the effect of schooling reforms is larger in developed countries, but this difference is not significant at conventional levels.

Finally, we investigate whether the effect exhibits heterogeneity by the pre-reform distribution of patience. In particular, the distribution of the quantitative patience item exhibits substantial censoring at the lower end. This is also indicated by the asymmetry that is visible in the distribution of the compound measure of patience in Figure 2.2.3a. The degree of censoring differs across countries. When focusing on observations of individuals not affected by the reforms, approximately 48% of observations of the quantitative item are censored in countries that are coded as developed according to our classification. In contrast, about 64% of observations of the quantitative item are censored in developing countries. We explored the sensitivity of the results with respect to this issue by allowing for heterogeneity of the effect by the share of censored individuals that were not subject to the respective education reform. The corresponding results, shown in Columns (3) and (7) of Table 2.3.3, do not provide any evidence for heterogeneous effects by the share of censored individuals.

The inclusion of controls for all dimensions of heterogeneity does not affect the finding of a heterogeneous effect by the stage of education (primary vs. secondary), as illustrated by the results in Columns (4) and (8) of Table 2.3.3. Notably, the

22. Again, the estimates are based on a specification that includes an indicator variable that is 1 if a cohort in a specific is partially non-treated as in (1'). Allowing for heterogeneity among the partially treated cohorts delivers no evidence for heterogeneity among the partially treated, and virtually identical results for treatment heterogeneity in the effect of interest.

results corroborate that the effect of schooling reforms on patience is mainly driven by secondary schooling effects.[23]

Additional results show that there is no evidence for heterogeneity in the effects by the quality of the education system. In particular, we estimated extended specifications that additionally account for potential heterogeneity by different indexes of educational quality. The findings are robust to different specifications of the quality index (see Table 2.B.10 in the Appendix). Moreover, additional findings document that reforms affecting secondary schooling do not increase the years of education more than other reforms, suggesting that the heterogeneity result is also not driven by secondary schooling reforms being more binding (see Table 2.B.11 in the Appendix).

Another potential dimension of effect heterogeneity refers to gender. Existing evidence has documented systematic differences in patience between women and men and that these differences are related to the level of development and gender equality of a country (see, e.g., Falk and Hermle, 2018). Moreover, while compulsory schooling reforms have been shown to be effective in increasing female education and closing the education gap, the associated shifts in education have also been shown to have asymmetric effects in other domains, such as fertility or mortality (see, e.g., Gathmann, Jürges, and Reinhold, 2015; Fort, Schneeweis, and Winter-Ebmer, 2016). This raises the question about heterogeneity in the effect of education on patience between men and women. Replicating the analysis for an extended specification to explore this question documents level differences in patience, but shows little evidence for effect heterogeneity or for systematic differences in the gender-specific effect of education on patience across countries with different levels of development (see Table 2.B.12).

2.4 Conclusion

This paper has provided new evidence for a positive impact of education on patience. The causal effect of education is identified from 50 compulsory schooling reforms in 48 countries around the globe and using a regression design that isolates the heterogeneity in patience across treated and non-treated cohorts. The results show that compulsory schooling reforms are associated with an increase in patience by 0.1 standard deviations. Using compulsory schooling reforms as an instrumental variable for years of schooling, we find that one more year of schooling is associated with an increase in patience of approximately 0.1 standard deviations. This finding is robust to alternative specifications of the empirical framework, and is not sensi-

23. Additional results using GDP per capita at the time that the reform was implemented as proxy for the level of development instead of the Human Development Index deliver very similar results, see Table 2.B.9 in the Appendix.

Table 2.3.3. Effect Heterogeneity

	Dependent variable:							
	Patience							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated	0.014	0.053	0.083	−0.140	0.018	0.043	0.072	−0.122
	(0.043)	(0.049)	(0.102)	(0.097)	(0.040)	(0.047)	(0.087)	(0.101)
Treated × Secondary Edu. Reform	0.156***			0.152**	0.125**			0.119*
	(0.059)			(0.063)	(0.057)			(0.064)
Treated × Developed		0.089		0.078		0.084		0.079
		(0.063)		(0.066)		(0.059)		(0.063)
Treated × Share censored (pre reform)			0.027	0.197			0.024	0.175
			(0.150)	(0.131)			(0.131)	(0.137)
Age Trend	heterogeneous linear	heterogeneous linear	heterogeneous linear	heterogeneous linear	heterogeneous linear	heterogeneous linear	heterogeneous linear	heterogeneous linear
Age Bins × Developed					✓	✓	✓	✓
Observations	19,680	19,680	19,680	19,680	19,680	19,680	19,680	19,680
Adjusted R^2	0.142	0.142	0.142	0.142	0.143	0.142	0.142	0.143

Notes: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. OLS estimates. Secondary Edu. Reform is an indicator variable that is 1 if a reform targets years of schooling above 8 years. Edu. Quality Index is obtained from Altinok, Angrist, and Patrinos (2018). It corresponds to the standardized version *average harmonised learning outcome score* of a country in the available year that is closest to the reform's implementation year. It is unavailable for Bangladesh, Cambodia, Iraq. Thus, column (3) and (6) do not include these countries. All specifications include an indicator variable that is 1 if a cohort in a specific is partially non-treated, country-reform fixed effects, reform specefic linear pre- and post-reform trends. Column (4)-(6) additionally include 5-year age bin fixed effects that are allowed to vary by level of development. Observations are weighted with the aforementioned weighting scheme. Standard errors are clustered on the country level.

tive to particular reforms or countries. The effect does not depend on the level of development or the quality of the education system.

An analysis of effect heterogeneity suggests that the effect is mainly driven by variation in secondary schooling. This finding is surprising in light of earlier evidence that has suggested that personality is particularly malleable during the early years of childhood. Our results suggest that this malleability extends to years of secondary schooling. The findings also indicate that schooling may have important indirect effects on factor accumulation and productivity, by affecting personality and fostering future-oriented behavior. A promising direction for future research concerns the mechanisms underlying our reduced form results.

In sum, the findings presented in this paper contribute new evidence for school-based education as a potential determinant of patience, confirming conjectures in the literature and complementing scattered evidence from experiments and case studies with findings from compulsory schooling reforms across the globe. More work is needed to isolate the mechanisms behind the effect. For instance, an open question is whether other reforms that delay the entry in the labor market, such as changes in compulsory military service, also affect patience.

References

Åkerlund, David, Bart H.H. Golsteyn, Hans Grönqvist, and Lena Lindahl. 2016. "Time discounting and Criminal Behavior." *Proceedings of the National Academy of Sciences of the United States of America* 113 (22): 6160–65. [141]

Alan, Sule, and Seda Ertac. 2018. "Fostering Patience in the Classroom: Results from Randomized Educational Intervention." *Journal of Political Economy* 126 (5): 1865–911. [141]

Altinok, N., N. Angrist, and H.A. Patrinos. 2018. *Global Data Set on Education Quality (1965-2015).* Washington DC: World Bank. [162]

Anderson, Theodore W, and Herman Rubin. 1949. "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations." *Annals of Mathematical Statistics* 20 (1): 46–63. [158, 159]

Angrist, Joshua D, and Alan B Krueger. 1991. "Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics* 106 (4): 979–1014. [142]

Ashraf, Nava, Dean Karlan, and Wesley Yin. 2006. "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines." *Quarterly Journal of Economics* 121 (2): 635–72. [141]

Bajaj, Monisha, and Huma Kidwai. 2016. "Human Rights and Education Policy in South Asia." *Handbook of Global Education Policy*, 206–23. [144]

Bauer, Michal, and Julie Chytilová. 2010. "The Impact of Education on Subjective Discount Rate in Ugandan Villages." *Economic Development and Cultural Change* 58 (4): 643–69. [141]

Becker, Gary S, and Casey B Mulligan. 1997. "The Endogenous Determination of Time Preference." *Quarterly Journal of Economics* 112 (3): [141]

Benavot, Aaron, and Julia Resnik. 2006. "Lessons from the Past: A Comparative Socio-Historical Analysis of Primary and Secondary Education." *Global Educational Expansion Historical Legacies and Political Obstacles*, (April 2017): 1–90. [144]

Bettinger, Eric, and Robert Slonim. 2007. "Patience Among Children." *Journal of Public Economics* 91 (1-2): 343–63. [149]

Black, Sandra E, Paul J Devereux, and Kjell G Salvanes. 2008. "Staying in the Classroom and out of the Maternity Ward? The Effect of Compulsory Schooling Laws on Teenage Births." *Economic Journal* 118 (530): 1025–54. [142]

Bradford, W. David. 2010. "The Association Between Individual Time Preferences and Health Maintenance Habits." *Medical Decision Making* 30 (1): 99–112. [141]

Brunello, Giorgio, Margherita Fort, Nicole Schneeweis, and Rudolf Winter-Ebmer. 2016. "The Causal Effect of Education on Health: What is the Role of Health Behaviors?" *Health Economics* 25 (3): 314–36. [142, 151]

Brunello, Giorgio, Margherita Fort, and Guglielmo Weber. 2009. "Changes in Compulsory Schooling, Education and the Distribution of Wages in Europe." *Economic Journal* 119 (536): 516–39. [142, 151]

Bülbül, Ayse Eliüsük, and Gökhan Izgar. 2018. "Effects of the Patience Training Program on Patience and Well-Being Levels of University Students." *Journal of Education and Training Studies* 6 (1): 159–68. [141]

Burro, Giovanni, Rebecca McDonald, Daniel Read, and Umar Taj. 2022. "Patience decreases with age for the poor but not for the rich: an international comparison." *Journal of Economic Behavior & Organization* 193: 596–621. [150]

Castillo, Marco, Jeffrey L. Jordan, and Ragan Petrie. 2019. "Discount Rates of Children and High School Graduation." *Economic Journal* 129 (619): 1153–81. [141]

Cunha, Flavio, and James Heckman. 2007. "The Technology of Skill Formation." *American Economic Review* 97 (2): 31–47. [141, 159]

Cunha, Flavio, James J Heckman, and Susanne Schennach. 2010. "Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Econometrica* 78 (3): 883–931. [141]

Cutler, David M, and Adriana Lleras-Muney. 2006. "Education and Health: Evaluating Theories and Evidence." *NBER Working Paper 12352*, [141]

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde. 2010. "Are Risk Aversion and Impatience Related to Cognitive Ability?" *American Economic Review* 100 (3): 1238–60. [156]

Ejieh, Michael. 2009. "The Universal Basic Education as an Effective Strategy for Meeting the Millennium Development Goals in Nigeria." *Nebula* 6 (1): 112–21. [144]

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde. 2018. "Global Evidence on Economic Preferences." *Quarterly Journal of Economics* 133 (4): 1645–92. [140–143, 150]

Falk, Armin, Anke Becker, Thomas J. Dohmen, David Huffman, and Uwe Sunde. 2023. "The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences." *Management Science* 69 (4): 1935–50. [143]

Falk, Armin, and Johannes Hermle. 2018. "Relationship of gender differences in preferences to economic development and gender equality." *Science* 362 (6412): [161]

Finke, Michael S., and Sandra J. Huston. 2013. "Time preference and the importance of saving for retirement." *Journal of Economic Behavior and Organization* 89: 23–34. [141]

Fort, Margherita, Nicole Schneeweis, and Rudolf Winter-Ebmer. 2011. "More Schooling, More Children: Compulsory Schooling Reforms and Fertility in Europe." *IZA Discussion Paper 6015*, [142, 151]

Fort, Margherita, Nicole Schneeweis, and Rudolf Winter-Ebmer. 2016. "Is Education Always Reducing Fertility? Evidence from Compulsory Schooling Reforms." *Economic Journal* 126 (595): 1823–55. [145, 161]

Gathmann, Christina, Hendrik Jürges, and Steffen Reinhold. 2015. “Compulsory Schooling Reforms, Education and Mortality in Twentieth Century Europe.” *Social Science and Medicine* 127: 74–82. [142, 145, 161]

Godbardhan-Rambocus, Lila. 2015. “Surinam: The Development of Education.” In *Education in the Commonwealth Caribbean and Netherlands Antilles*. Edited by R Murray Thomas. Springer. Chapter 20, pp. 360. [144]

Golsteyn, Bart H.H., Hans Grönqvist, and Lena Lindahl. 2014. “Adolescent Time Preferences Predict Lifetime Outcomes.” *Economic Journal* 124 (580): F739–F761. [141]

Hampf, Franziska. 2019. “The Effect of Compulsory Schooling on Skills: Evidence from a Reform in Germany.” *ifo Working Papers*, (313): [142]

Hanushek, Eric A., Lavinia Kinne, Philipp Lergetporer, and Ludger Woessmann. 2022. “Patience, Risk-Taking, and Human Capital Investment across Countries.” *Economic Journal* 132 (646): 2290–307. [141]

Heckman, James, Jora Stixrud, and Sergio Urzua. 2006. “The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior.” *Journal of Labor Economics* 24 (3): 411–82. [141]

Isaacs, Shafika. 2007. “ICT in Education in Asia.” In *Survey of ICT and Education in Africa*. Washington D.C.: World Bank. [144]

Jaiyeoba, Adebola O. 2007. “Perceived Impact Of Universal Basic Education On National Development In Nigeria Adebola.” *International Journal of African and African American Studies* VI (1): 48–58. [144]

Jusot, Florence, and Myriam Khlat. 2013. “The Role of Time and Risk Preferences in Smoking Inequalities: A Population-Based Study.” *Addictive Behaviors* 38 (5): 2167–73. [141]

Khwaja, Ahmed, Dan Silverman, and Frank Sloan. 2007. “Time Preference, Time Discounting, and Smoking Decisions.” *Journal of Health Economics* 26 (5): 927–49. [141]

Kim, Younoh, and Vlad Radoias. 2016. “Education, Individual Time Preferences, and Asymptomatic Disease Detection.” *Social Science and Medicine* 150: 15–22. [141]

Kosse, Fabian, Thomas Deckers, Pia Pinger, Hannah Schildberg-Hoerisch, and Armin Falk. 2019. “The Formation of Prosociality: Causal Evidence on the Role of Social Environment.” *Journal of Political Economy*, [141]

Oreopoulos, Philip. 2007. “Do Dropouts Drop Out Too Soon? Wealth, Health, and Happiness from Compulsory Schooling.” *Journal of Public Economics* 91: 2213–29. [142]

Oreopoulos, Philip, and Kjell G. Salvanes. 2011. “Priceless: The nonpecuniary benefits of schooling.” *Journal of Economic Perspectives* 25 (1): 159–84. [141]

Perez-Arce, Francisco. 2017. “The Effect of Education on Time Preferences.” *Economics of Education Review* 56: 52–64. [141]

Pischke, Jörn Steffen, and Till von Wachter. 2008. “Zero Returns to Compulsory Schooling in Germany: Evidence and Interpretation.” *Review of Economics and Statistics* 90 (August): 592–98. [142, 145]

Plank, David N. 1990. “The Politics of Basic Education Reform in Brazil.” *Comparative Education Review* 34 (4): 538–59. [144]

Prou, Marc. 2009. “Attempts at Reforming Haiti’s Education System: The Challenges of Mending the Tapestry , 1979-2004.” *Journal of Haitian Studies* 15 (1/2): 29–69. [144]

Samady, Saif R. 2001. *Education and Afghan Society in the Twentieht Century*. Paris: United Nations Educational, Scientifc, and Cultural Organization. [144]

Smith, Patricia K., Barry Bogin, and David Bishai. 2005. "Are Time Preference and Body Mass Index Associated? Evidence from the National Longitudinal Survey of Youth." *Economics and Human Biology* 3 (2 SPEC. ISS.): 259–70. [141]

Strittmatter, Anthony, and Uwe Sunde. 2013. "Health and economic development – evidence from the introduction of public health care." *Journal of Population Economics* 26: 1549–84. [147]

Sunde, Uwe, Thomas Dohmen, Benjamin Enke, Armin Falk, David Huffman, and Gerrit Meyerheim. 2022. "Patience and Comparative Development." *Review of Economic Studies* 89 (5): 2806–40. [141, 143]

Tanaka, Tomomi, Colin F. Camerer, and Quang Nguyen. 2010. "Risk and time preferences: Linking experimental and household survey data from Vietnam." *American Economic Review* 100 (1): 557–71. [141]

Tomich, Vera. 1963. "Education in Yugoslavia and the New Reform: The Legal Basis, Organization, Adiminstration and Program of the Secondary Schools." In *Bulletin, No. 20*. U.S. Department of Health, Education, and Welfare: Office of Education. [144]

Ullah, Aman. 2013. "Right to Free and Compulsory Education in Pakistan after 18th Constitutional Amendment." *South Asian Studies* 28 (2): 329. [144]

Vahl, Monica Maciel. 2018. "A Pedagogy of Oppression: The Politics of Literace in Brazil, 1971-1989." Dissertation. Canterbury: University of Canterbury, New Zealand. [144]

Van Stipriaan, Alex. 1998. "Between State and Society: Education in Suriname, 1850-1950." In *Mediators Between State and Society*. Edited by Nico Randeraad. Verloren, 57–80. [144]

Weiner, Myron. 1991. "Child Labor and Compulsory-Education Policies." In *The Child and the State in India*. Princeton, New Jersey: Princeton University Press. Chapter 5, 77–108. [144]

Worldbank. last accessed: 2020-06-01. "Worldbank Data on Compulsory Education." https://data.worldbank.org/indicator/SE.COM.DURS. [144]

Appendix 2.A Measures and Descriptives

Table 2.A.1. Descriptives (unweighted)

Characteristic	Treated			p-value
	Overall, N = 19,680	**No**, N = 8,901	**Yes**, N = 10,779	
Age				<0.001
	45.42	49.21	42.29	
	(14.42)	(14.10)	(13.92)	
Female	0.57	0.56	0.58	0.069
Global region				<0.001
Americas	0.17	0.17	0.16	
Asia	0.23	0.24	0.22	
Europe	0.31	0.33	0.29	
Former Soviet Union	0.08	0.06	0.09	
Middle East and North Africa	0.17	0.16	0.18	
Sub-Saharan Africa	0.05	0.04	0.06	
Developed	0.46	0.47	0.46	0.13
Human Development Index				0.032
	0.79	0.79	0.79	
	(0.10)	(0.09)	(0.10)	
Yrs. comp. edu.				<0.001
	6.68	4.93	8.13	
	(2.81)	(3.03)	(1.46)	
Yrs. education				<0.001
	10.72	10.17	11.17	
	(4.48)	(4.74)	(4.21)	
Patience				<0.001
	0.00	−0.02	0.02	
	(1.01)	(1.01)	(1.01)	
Partially Non-Treated	0.09	0	0.16	<0.001

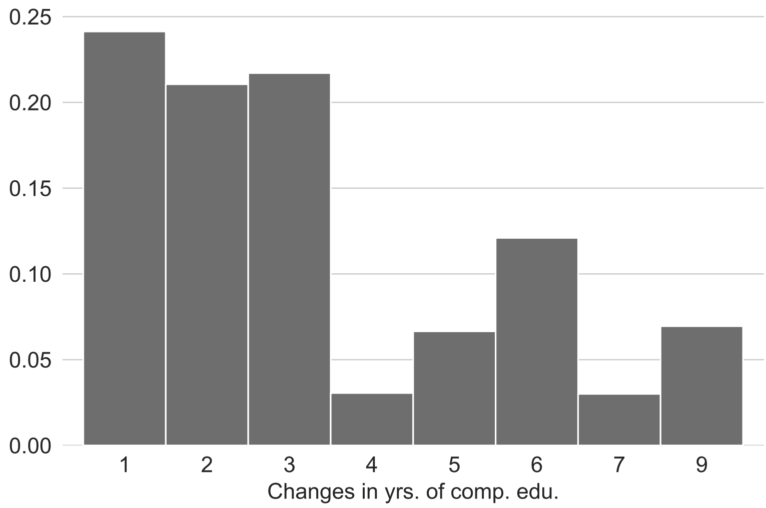
Note: The table displays the descriptives of the merged GPS-reform data set. The sample is restricted to observations that fall within 10 years of a compulsory schooling reform in their respective country. Pivotal cohorts are excluded. Means and standard deviation are shown for continuous variables and shares for dichotomous and categorical variables. p-values correspond to t-tests and χ^2-test of independence, respectively.

Table 2.A.2. Descriptives (weighted)

Characteristic	Treated **Overall**, N = 19,680	 **No**, N = 9,840	 **Yes**, N = 9,840	**p**-value
Age				<0.001
	45.67 (15.06)	50.57 (14.42)	40.76 (14.07)	
Female	0.57	0.56	0.57	0.085
Global region				>0.9
Americas	0.18	0.18	0.18	
Asia	0.20	0.20	0.20	
Europe	0.36	0.36	0.36	
Former Soviet Union	0.02	0.02	0.02	
Middle East and North Africa	0.18	0.18	0.18	
Sub-Saharan Africa	0.06	0.06	0.06	
Developed	0.51	0.51	0.51	0.9
Human Development Index				0.9
	0.79 (0.10)	0.79 (0.10)	0.79 (0.10)	
Yrs. comp. edu.				<0.001
	6.66 (2.89)	5.04 (3.02)	8.28 (1.54)	
Yrs. education				<0.001
	10.56 (4.68)	10.02 (4.92)	11.10 (4.37)	
Patience				<0.001
	0.00 (1.00)	−0.05 (0.99)	0.04 (1.01)	
Partially Non-Treated	0.08	0	0.16	

Note: Table displays the descriptives of the GPS data set, merged with the reform data set, weighted with inverse probability weights. Sample: Observations within 10 years of a compulsory schooling reform in their respective country. Pivotal cohorts excluded. Table includes mean (sd) for continuous variables and shares for dichotomous and categorical variables. P-values correspond to Wilcoxon rank-sum test for complex survey samples and χ^2-test with Rao & Scott's second-order correction, respectively.

Figure 2.A.1. Distribution of changes in compulsory years of education



Note: Sample consists of individuals born within +/-10 years of the pivotal cohort.

Table 2.B.1. Patience and compulsory schooling – 7-year window

	Patience					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A.						
Treated	0.096**	0.093**	0.096**	0.085**	0.101***	0.092**
	(0.037)	(0.038)	(0.038)	(0.038)	(0.037)	(0.038)
Adj. R^2	0.141	0.141	0.143	0.143	0.142	0.143
Panel B.						
Treated	0.154***	0.15***	0.151***	0.136***	0.152***	0.138***
	(0.044)	(0.044)	(0.05)	(0.048)	(0.049)	(0.048)
Partially Non-Treated	−0.093***	−0.091***	−0.091**	−0.082**	−0.091**	−0.08**
	(0.032)	(0.033)	(0.037)	(0.036)	(0.04)	(0.038)
Adj. R^2	0.141	0.141	0.143	0.143	0.143	0.143
Age trend	homogeneous linear	homogeneous linear	homogeneous quadratic	homogeneous quadratic	heterogeneous linear	heterogeneous linear
Age Bins × Developed		✓		✓		✓
N	14,689	14,689	14,689	14,689	14,689	14,689

Notes: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Table displays the effect of compulsory schooling reforms on patience allowing for a maximal seven year window around a reform. All specifications include country-reform specific trends, country-reform specific fixed effects, and an indicator variable indicating when a cohort is partially non-treated. Column (2), (4), and (6) additionally include age bin fixed effects that are allowed to vary by level of development. Age bins are 5-year age intervals. Developed is an indicator that is 1 if a country can be classified as developed in terms of the Human Development Index, and 0 otherwise. Observations are weighted with the aforementioned weighting scheme. Standard errors are clustered on the country level. Observations are weighted with the aforementioned weighting scheme.

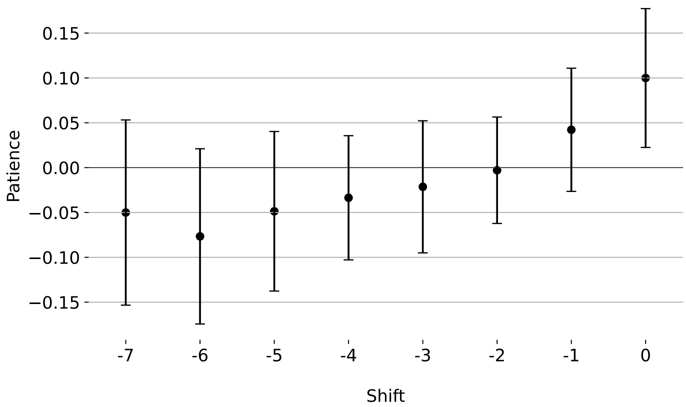
Appendix 2.B Additional Results

Table 2.B.2. Patience and compulsory: quadratic RDD-like specification

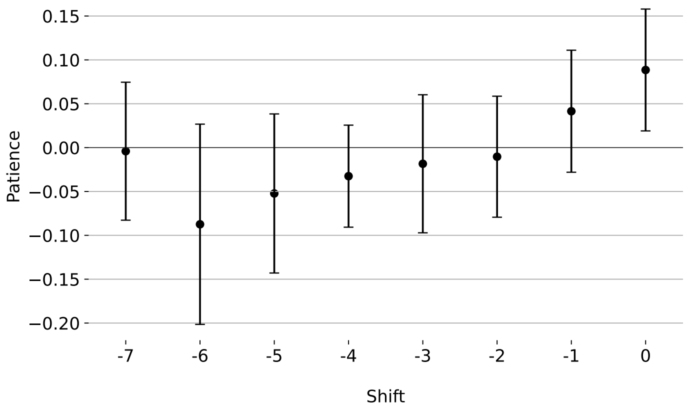
	Patience	
	(1)	(2)
Panel A.		
Treated	0.112**	0.103**
	(0.054)	(0.05)
Adj. R^2	0.144	0.144
Panel B.		
Treated	0.168**	0.153**
	(0.075)	(0.071)
Partially Non-Treated	−0.079	−0.069
	(0.052)	(0.052)
Adj. R^2	0.144	0.144
Age trend	heterogeneous quadratic	heterogeneous quadratic
Age Bins × Developed		✓
N	19,680	19,680

Notes: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Table displays the effect of compulsory schooling reforms on patience. All specifications include reform specific trends, country-reform specific fixed effects. Trend specification is *RDD: quadratic*, i.e. a quadratic pre- and post-reform trend. Column (2) additionally includes age bin fixed effects that are allowed to vary by level of development. Age bins are 5-year age intervals. Developed is an indicator that is 1 if a country can be classified as developed in terms of the Human Development Index, and 0 otherwise. Observations are weighted with the aforementioned weighting scheme. Standard errors are clustered on the country level.

Figure 2.B.1. Placebo Test: Shifting Reform Dates



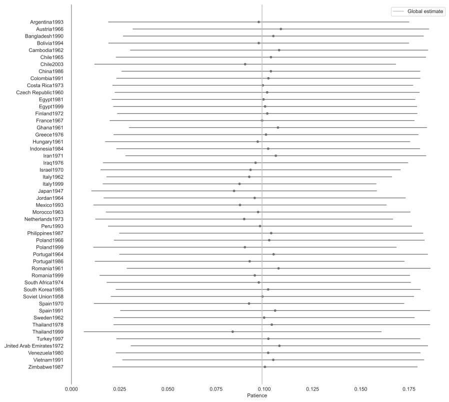
(a) Without Age Bin Fixed Effects



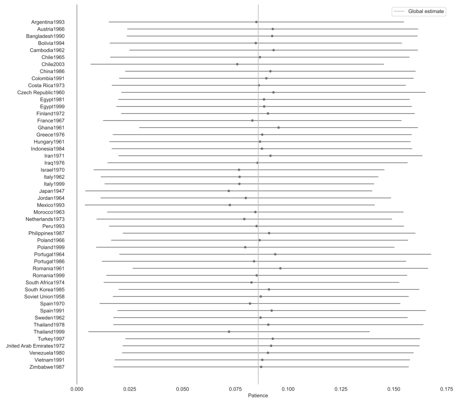
(b) With Age Bin Fixed Effects

Note: Displays estimates of our main specification when shifting treatment assignment one to seven years –the minimum distance before a reform in our sample would cover another reform – before the actual pivotal cohort. Trend specification is *heterogeneous: linear*, i.e. a linear pre- and post-reform trend. Lines correspond to 95% confidence intervals. All specification include an indicator variable indicating whether a cohort was partially non-treated. Figure (b) additionally includes 5-year Age-Bins interacted with in an indicator value to whether a country is developed. Observations are weighted with the aforementioned weighting scheme. Standard errors are clustered on the country level.

Figure 2.B.2. Leave one out estimates



(a) Without Age Bin Fixed Effects



(b) With Age Bin Fixed Effects

Note: Displays estimates when leaving out the respective reform. This is implemented by setting *Treated* and *Partial Non-Treated* to 0 for the respective reform-country. Lines represent 95% Confidence Intervals. Trend specification is *RDD: linear*, i.e. a linear pre- and post-reform trend. Lines correspond to 95% confidence intervals. All specification include an indicator variable indicating whether a cohort was partially non-treated. Figure (b) additionally includes 5-year Age-Bins interacted with in an indicator value to whether a country is developed. Observations are weighted with the aforementioned weighting scheme. Standard errors are clustered on the country level.

Table 2.B.3. Patience and compulsory schooling using wild cluster bootstrap

	Patience					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A.						
Treated	0.058*	0.054	0.054	0.05	0.058*	0.045
	(0.032)	(0.032)	(0.032)	(0.031)	(0.032)	(0.03)
Adj. R^2	0.142	0.142	0.142	0.142	0.142	0.142
Panel B.						
Treated	0.103***	0.096**	0.095**	0.087**	0.099**	0.086**
	(0.036)	(0.033)	(0.039)	(0.034)	(0.039)	(0.035)
Partially Non-Treated	−0.088***	−0.079**	−0.078**	−0.069**	−0.085**	−0.081**
	(0.028)	(0.029)	(0.032)	(0.03)	(0.035)	(0.033)
Adj. R^2	0.142	0.142	0.142	0.142	0.142	0.142
Age trend	homogeneous linear	homogeneous linear	homogeneous quadratic	homogeneous quadratic	heterogeneous linear	heterogeneous linear
Age Bins × Developed		✓		✓		✓
N	19,680	19,680	19,680	19,680	19,680	19,680

Notes: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Table displays the effect of compulsory schooling reforms on patience. All specifications include reform specific trends, country-reform specific fixed effects. Column (2), (4), and (6) additionally include age bin fixed effects that are allowed to vary by level of development. Age bins are 5-year age intervals. Developed is an indicator that is 1 if a country can be classified as developed in terms of the Human Development Index, and 0 otherwise. Observations are weighted with the aforementioned weighting scheme. Standard errors are obtained using wild cluster bootstrap using the *R* package *fwildclusterbootstrap* based on Roodman, Nielsen, MacKinnon, and Webb (2019). Clustering is performed on the country level.

Table 2.B.4. Patience and compulsory schooling clustering on the reform unit level instead of the country level

	Patience					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A.						
Treated	0.058*	0.054*	0.054*	0.05	0.058*	0.045
	(0.032)	(0.032)	(0.033)	(0.031)	(0.032)	(0.03)
Adj. R^2	0.142	0.142	0.142	0.142	0.142	0.142
Panel B.						
Treated	0.103***	0.096***	0.095**	0.087**	0.099**	0.086**
	(0.036)	(0.034)	(0.039)	(0.034)	(0.04)	(0.035)
Partially Non-Treated	−0.088***	−0.079***	−0.078**	−0.069**	−0.085**	−0.081**
	(0.029)	(0.029)	(0.033)	(0.03)	(0.035)	(0.033)
Adj. R^2	0.142	0.142	0.142	0.142	0.142	0.142
Age trend	homogeneous linear	homogeneous linear	homogeneous quadratic	homogeneous quadratic	heterogeneous linear	heterogeneous linear
Age Bins × Developed		✓		✓		✓
N	19,680	19,680	19,680	19,680	19,680	19,680

Notes: $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$. Table displays the effect of compulsory schooling reforms on patience. All specifications include country-greform specific trends, country-reform specific fixed effects, and an indicator variable indicating when a cohort is partially non-treated. Column (2), (4), and (6) additionally include age bin fixed effects that are allowed to vary by level of development. Age bins are 5-year age intervals. Developed is an indicator that is 1 if a country can be classified as developed in terms of the Human Development Index, and 0 otherwise. Observations are weighted with the aforementioned weighting scheme. Standard errors are clustered on the reform unit level, this deviates from the country level for countries of the former Soviet Union. Observations are weighted with the aforementioned weighting scheme.

Table 2.B.5. Patience and compulsory schooling using different specification for partially non-treated cohorts

	Dependent variable:					
	Patience					
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	0.102***	0.098***	0.096**	0.087**	0.095**	0.089**
	(0.040)	(0.035)	(0.040)	(0.035)	(0.040)	(0.036)
Age Trend	heterogeneous linear	heterogeneous linear	heterogeneous linear	heterogeneous linear	heterogeneous linear	heterogeneous linear
Age Bins × Developed		✓		✓		✓
Partial specification	Country Specific	Country Specific	Time Specific	Time Specific	C-T Specific	C-T Specific
Observations	19,680	19,680	19,680	19,680	19,680	19,680
Adjusted R^2	0.142	0.143	0.142	0.143	0.143	0.143

Notes: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Table displays the effect of compulsory schooling reforms on patience for different specifications of partially non-treated (not displayed). Column (1) and (2) allow the effect for the partially non-treated cohorts to vary with the reform, column (3) and (4) allow it to vary with time, column (5) and (6) allow it to vary with time and reform. All specification include country-reform specific fixed effects, country-reform specific age trends. Column (2), (4), (6) additionally include age bin fixed effects that are allowed to vary by level of development. Age bins are 5-year age intervals. Developed is an indicator that is 1 if a country can be classified as developed in terms of the Human Development Index, and 0 otherwise. Observations are weighted with the aforementioned weighting scheme. Observations are weighted with the aforementioned weighting scheme. Standard errors are clustered on the country level.

Table 2.B.6. Compulsory Schooling Reforms and Patience – Alternative Reform Coding Indonesia, Netherlands

	Patience					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A.						
Treated	0.055*	0.051*	0.048	0.04	0.055*	0.04
	(0.032)	(0.031)	(0.032)	(0.032)	(0.032)	(0.03)
Adj. R^2	0.146	0.146	0.146	0.147	0.147	0.147
Panel B.						
Treated	0.099***	0.091***	0.082**	0.071**	0.087**	0.072**
	(0.035)	(0.033)	(0.04)	(0.036)	(0.039)	(0.036)
Partially Non-Treated	−0.084***	−0.074***	−0.065*	−0.056*	−0.067*	−0.064*
	(0.028)	(0.028)	(0.034)	(0.031)	(0.037)	(0.036)
Adj. R^2	0.147	0.146	0.147	0.147	0.147	0.147
Age trend	homogeneous linear	homogeneous linear	homogeneous quadratic	homogeneous quadratic	heterogeneous linear	heterogeneous linear
Age Bins × Developed		✓		✓		✓
N	19,328	19,328	19,328	19,328	19,328	19,328

Notes: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. OLS estimates of the effect of compulsory schooling reforms on patience. The dependent variable is the standardized measure of patience. Treated is defined as described in the text. All specifications include reform specific trends, country-reform specific fixed effects. Column (2), (4), (6) additionally include age bin fixed effects that are allowed to vary by level of development. Age bins are 5-year age intervals. Developed is an indicator that is 1 if a country can be classified as developed in terms of the Human Development Index, and 0 otherwise. Observations are weighted with the aforementioned weighting scheme. Standard errors are clustered on the country level. In this specification, we use the reform of Indonesia from 1994 instead of the reform from 1984, and the reform of Netherlands from 1985 instead of the reform from 1973.

Table 2.B.7. The effect of the reform by whether there is evidence of its relevance from other studies

	Dependent variable:					
	Patience					
	(1)	(2)	(3)	(4)	(5)	(6)
Treated × causal study	0.102*	0.107**	0.088*	0.082	0.099*	0.090*
	(0.053)	(0.049)	(0.053)	(0.052)	(0.054)	(0.052)
Treated × descriptive study	−0.031	−0.011	−0.036	−0.018	−0.030	−0.010
	(0.094)	(0.100)	(0.092)	(0.098)	(0.095)	(0.101)
Treated × no study	0.127***	0.118**	0.124**	0.118**	0.121**	0.110**
	(0.048)	(0.051)	(0.053)	(0.050)	(0.051)	(0.050)
Age Trend	homogeneous linear	homogeneous linear	homogeneous quadratic	homogeneous quadratic	heterogeneous linear	heterogeneous linear
Age Bins × Developed		✓		✓		✓
Observations	19,680	19,680	19,680	19,680	19,680	19,680
Adjusted R^2	0.142	0.142	0.142	0.142	0.142	0.142

Notes: $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$. Table displays the effect of compulsory schooling reforms on patience, by whether the respective reform has been used in a causal study studying the reforms effect on education, a descriptive study indicating that the reform influenced years of schooling or enrollment into education, or has not been analyzed with respect to effectiveness, yet. All specifications include country-reform specific trends, country-reform specific fixed effects, and an indicator variable indicating when a cohort is partially non-treated. Column (2), (4), and (6) additionally include age bin fixed effects that are allowed to vary by level of development. Age bins are 5-year age intervals. Developed is an indicator that is 1 if a country can be classified as developed in terms of the Human Development Index, and 0 otherwise. Observations are weighted with the aforementioned weighting scheme. Observations are weighted with the aforementioned weighting scheme. Standard errors are clustered on the country level.

Table 2.B.8. Patience and compulsory schooling using controlling for other preference measures

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A.						
Treated	0.039	0.038	0.033	0.03	0.036	0.025
	(0.032)	(0.032)	(0.033)	(0.032)	(0.032)	(0.031)
Adj. R^2	0.19	0.191	0.191	0.191	0.191	0.191
Panel B.						
Treated	0.075^{**}	0.07^{**}	0.063	0.058	0.067^{*}	0.056
	(0.036)	(0.035)	(0.039)	(0.035)	(0.039)	(0.036)
Partially Non-Treated	-0.07^{**}	-0.062^{**}	-0.058^{*}	-0.051^{*}	-0.064^{*}	-0.062^{**}
	(0.028)	(0.029)	(0.03)	(0.029)	(0.033)	(0.032)
Adj. R^2	0.191	0.191	0.191	0.191	0.191	0.191
Other pref. measures	✓	✓	✓	✓	✓	✓
Age trend	homogeneous linear	homogeneous linear	homogeneous quadratic	homogeneous quadratic	heterogeneous linear	heterogeneous linear
Age Bins × Developed		✓		✓		✓
N	18,942	18,942	18,942	18,942	18,942	18,942

Notes: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. OLS estimates of the effect of compulsory schooling reforms on patience. The dependent variable is the standardized measure of patience. Treated is defined as described in the text. All specifications include reform specific trends, country-reform specific fixed effects. Column (2), (4), (6) additionally include age bin fixed effects that are allowed to vary by level of development. Age bins are 5-year age intervals. Developed is an indicator that is 1 if a country can be classified as developed in terms of the Human Development Index, and 0 otherwise. Observations are weighted with the aforementioned weighting scheme. Standard errors are clustered on the country level. *All specifications include the preference measures for: risktaking, pos. reciprocity, neg. reciprocity, trust, altruism*

Table 2.B.9. Patience: Heterogeneity of the Effect by Reform Type – different definition of developed

	Dependent variable:					
	Patience					
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	−0.021	−0.129	0.023	−0.118	0.111**	−0.074
	(0.065)	(0.115)	(0.047)	(0.096)	(0.056)	(0.104)
Treated × Rel. Developed (GDPPC)	0.147**	0.102				
	(0.069)	(0.070)				
Treated × Abs. Developed (rel. GDPPC)			0.106*	0.101*		
			(0.059)	(0.057)		
Treated: Abs. Developed (GDPPC, top tercile)					0.040	0.059
					(0.082)	(0.081)
Treated × Abs. Developing (GDPPC, bottom tercile)					−0.164**	−0.122*
					(0.072)	(0.063)
Treated × Secondary Edu. Reform		0.141**		0.131**		0.113**
		(0.070)		(0.062)		(0.057)
Treated × Share censored (pre reform)		0.105		0.126		0.177
		(0.136)		(0.129)		(0.136)
Age Trend	heterogeneous linear	heterogeneous linear	heterogeneous linear	heterogeneous linear	heterogeneous linear	heterogeneous linear
Age Bins × Developed	✓	✓	✓	✓	✓	✓
Observations	19,680	19,680	19,680	19,680	19,680	19,680
Adjusted R^2	0.143	0.143	0.143	0.143	0.143	0.143

Notes: $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$. Historical values of GDP per Capita are obtained from Bolt, Inklaar, Jong, and Zanden (2018). Abs. Developed (GDPPC) is an indicator that is 1 if a country's GDP per Capita relative to the U.S. GDPPC in the year the reform took place exceeded the median GDP per Capita (rel. to the U.S.) at reform in the sample. Top, and Bottom Tercile are defined accordingly. Rel. Developed (GDPPC) is an indicator that is 1 if the countries GDPPC at reform exceeded the median GDPPC in that year. Secondary Edu. Reform is an indicator variable that is 1 if a reform targets years of schooling above 8 years. Share censored (pre reform) is the share of respondents born before the pivotal cohort, that is left-censored in the quantitative patience measure and varies between 0 and 1. All specifications include an indicator variable that is 1 if a cohort in a specific is partially non-treated, country-reform fixed effects, linear pre- and post-reform trend, 5-year age bin fixed effects that are allowed to vary by level of development. Standard errors are clustered on the country level.

Table 2.B.10. Patience: Heterogeneity of the Effect by Reform Type – Controlling for Quality

	Dependent variable:					
	Patience					
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	−0.154	−0.130	−0.155	−0.173	−0.144	−0.214
	(0.142)	(0.146)	(0.151)	(0.160)	(0.160)	(0.161)
Treated × Secondary Edu. Reform	0.181***	0.174***	0.183***	0.187***	0.175**	0.193***
	(0.064)	(0.063)	(0.064)	(0.072)	(0.072)	(0.071)
Treated × Developed	0.063	0.012	0.074	0.085	0.014	0.093
	(0.081)	(0.107)	(0.074)	(0.075)	(0.099)	(0.072)
Treated × Share censored (pre reform)	0.247	0.251	0.233	0.210	0.235	0.182
	(0.205)	(0.188)	(0.198)	(0.209)	(0.189)	(0.204)
Treated × Edu. Quality Index	0.015			−0.006		
	(0.053)			(0.054)		
Treated × Edu. Quality Index_05		0.057			0.063	
		(0.058)			(0.059)	
Treated × Edu. Quality Index (Rank)			−0.004			0.089
			(0.131)			(0.141)
Age Trend	heterogeneous	heterogeneous	heterogeneous	heterogeneous	heterogeneous	heterogeneous
	linear 6					
Age Bins × Developed				✓	✓	✓
Observations	15,849	15,849	15,849	15,849	15,849	15,849
Adjusted R^2	0.155	0.155	0.155	0.155	0.155	0.155

Notes: $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$. Secondary Edu. Reform is an indicator variable that is 1 if a reform targets years of schooling above 8 years. Edu. Quality Index is obtained from Altinok, Angrist, and Patrinos (2018) who collected standardized test results across countries and harmonized these scores to make them comparable across countries. Part of their sample goes back to 1965. For each country, we use the score from the year of all available years that is closest to the year in which the reform was implemented. There is no score available for Bangladesh, Cambodia, and Iraq. The baseline measure corresponds to the standardized rank of *average harmonised learning outcome score* of a country from the year that is closest to the reform's implementation year in columns (1) and (4), from the year 2005 in columns (2) and (5). In columns (3) and (6) we use the yearly standardized rank of the *average harmonised learning outcome score* from the year closest to the respective reform's implementation year. The quality measure is unavailable for Bangladesh, Cambodia, Iraq. Developed is an indicator variable that indicates whether a country is developed according to the Human Development Index. All specifications include an indicator variable that is 1 if a cohort in a specific is partially non-treated, country-reform fixed effects, linear pre- and post-reform trend. Column (5)- (8) additionally include 5-year age bin fixed effects that are allowed to vary by level of development. Standard errors are clustered on the country level.

Table 2.B.11. Yrs. of Education: Heterogeneity of the Effect by Reform Type

	Dependent variable:							
	Yrs. education							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated	0.984***	0.877***	0.339	0.923	1.027***	1.035***	0.316	1.244*
	(0.322)	(0.304)	(0.408)	(0.866)	(0.277)	(0.270)	(0.321)	(0.695)
Treated × Secondary Edu. Reform	−0.501			−0.433	−0.546**			−0.411
	(0.314)			(0.303)	(0.277)			(0.284)
Treated × Developed		−0.322		−0.169		−0.599**		−0.492
		(0.337)		(0.412)		(0.273)		(0.347)
Treated × Share censored (pre reform)			0.628	0.186			0.699	−0.066
			(0.688)	(1.012)			(0.579)	(0.816)
Age Trend	heterogeneous linear	heterogeneous linear	heterogeneous linear	heterogeneous linear	heterogeneous linear	heterogeneous linear	heterogeneous linear	heterogeneous linear
Age Bins × Developed					✓	✓	✓	✓
Observations	19,415	19,415	19,415	19,415	19,415	19,415	19,415	19,415
Adjusted R^2	0.297	0.297	0.297	0.297	0.300	0.300	0.300	0.300

Notes: $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$. Secondary Edu. Reform is an indicator variable that is 1 if a reform targets years of schooling above 8 years. Share censored (pre reform) is the share of respondents born before the pivotal cohort, that is left-censored in the quantitative patience measure. This measure is defined between 0 and 1. All specifications include an indicator variable that is 1 if a cohort in a specific is partially non-treated, country-reform fixed effects, linear pre- and post-reform trend. Column (5)-(8) additionally include 5-year age bin fixed effects that are allowed to vary by level of development. Standard errors are clustered on the country level.

Table 2.B.12. Patience and compulsory schooling: Gender Differences

	Dependent variable:					
	Patience					
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	0.096**	0.081**	0.091**	0.077**	0.031	0.019
	(0.039)	(0.035)	(0.041)	(0.037)	(0.049)	(0.048)
Treated × Female			0.007	0.007	0.026	0.027
			(0.028)	(0.028)	(0.040)	(0.039)
Treated × Developed					0.121*	0.118*
					(0.068)	(0.066)
Treated × Female × Developed					−0.037	−0.040
					(0.056)	(0.055)
Female	−0.121***	−0.122***	−0.125***	−0.125***	−0.099***	−0.100***
	(0.022)	(0.022)	(0.025)	(0.025)	(0.034)	(0.033)
Female × Developed					−0.050	−0.050
					(0.050)	(0.050)
Age Trend	heterogeneous linear 12					
Age Bins × Developed		✓		✓		✓
Observations	19,680	19,680	19,680	19,680	19,680	19,680
Adjusted R^2	0.146	0.146	0.146	0.146	0.146	0.146

Notes: $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$. Table displays the effect of compulsory schooling reforms on patience. All specifications include an indicator variable that is 1 if a cohort in a specific is partially non-treated, country-reform fixed effects, linear pre- and post-reform trends. Column (2), (4), and (6) additionally include age bin fixed effects that are allowed to vary by level of development. Age bins are 5-year age intervals. Developed is an indicator that is 1 if a country can be classified as developed in terms of the Human Development Index, and 0 otherwise. Observations are weighted with the aforementioned weighting scheme. Standard errors are clustered on the country level. Observations are weighted with the aforementioned weighting scheme.

Appendix 2.C Compulsory Schooling Reforms

Table 2.C.1. Reforms

country	reform year	pivotal cohort	before/after	relevance	cohort info	time frame	no. obs. before	no. obs. after
Argentina	1993	1982 - 1986	7/9	causal study	precise	10	191	149
Austria	1966	1951	8/9	causal study	precise	10	184	227
Bangladesh	1990	1984	0/5	no study	estimate	10	231	217
Bolivia	1994	1982 - 1988	7/8	descriptive evidence	estimate	10	203	240
Cambodia	1962	1950	0/6	descriptive evidence	estimate	10	72	144
Chile	1965	1952	6/8	descriptive evidence	precise	10	121	194
	2003	1985 - 1988	8/12	causal study	precise	10	153	104
China	1986	1972 - 1977	0/9	causal study	precise	10	533	540
Colombia	1991	1981	5/9	no study	estimate	10	168	179
Costa Rica	1973	1961 - 1968	6/9	descriptive evidence	estimate	10	94	173
Czech Republic	1960	1947	8/9	causal study	precise	10	100	182
Egypt	1981	1970	6/9	no study	precise	7	204	170
	1999	1985	8/9	no study	estimate	7	187	165
Finland	1972	1961 - 1966	6/9	causal study	precise	10	215	166
France	1967	1953	8/10	causal study	precise	10	146	178
Ghana	1961	1954	0/6	descriptive evidence	estimate	10	51	78
Greece	1976	1963	6/9	causal study	precise	10	145	175
Hungary	1961	1946	8/10	no study	estimate	10	113	220
Indonesia	1984	1977	0/6	causal study	precise	10	249	298
Iran	1971	1961	6/8	no study	precise	10	238	473
Iraq	1976	1971	0/6	no study	precise	10	154	247
Israel	1970	1955 - 1958	8/10	causal study	precise	10	63	153
Italy	1962	1949	5/8	causal study	precise	10	182	185
	1999	1985	8/9	no study	estimate	10	122	72
Japan	1947	1934	6/9	no study	estimate	10	58	166
Jordan	1964	1952	6/9	no study	estimate	10	46	105
Mexico	1993	1980	6/9	causal study	precise	10	219	282
Morocco	1963	1957	0/6	no study	estimate	10	106	153
Netherlands	1973	1959	9/10	causal study	precise	10	237	252
Peru	1993	1982	6/11	causal study	precise	10	215	229
Philippines	1987	1981	0/6	no study	estimate	10	188	235
Poland	1966	1952	7/8	causal study	precise	10	157	216
	1999	1986	8/9	no study	precise	10	162	81
Portugal	1964	1956	4/6	causal study	precise	10	188	170
	1986	1980	6/9	no study	precise	10	152	118
Romania	1961	1947	7/8	no study	estimate	10	156	209
	1999	1987	8/9	no study	estimate	10	124	63
South Africa	1974	1967	0/9	no study	estimate	10	94	212
South Korea	1985	1972	6/9	no study	estimate	10	168	183
Soviet Union	1958	1945 - 1947	7/8	no study	precise	10	559	952
Spain	1970	1957	6/8	causal study	precise	10	174	218
	1991	1978	8/10	no study	precise	10	209	145
Sweden	1962	1950 - 1955	7/9	causal study	precise	10	183	218
Thailand	1978	1967	4/6	no study	estimate	9	223	240
	1999	1986	6/9	no study	estimate	9	151	80
Turkey	1997	1986	5/8	causal study	precise	10	309	186
United Arab Emirates	1972	1966	0/6	no study	estimate	10	101	319
Venezuela	1980	1968	6/9	causal study	precise	10	159	211
Vietnam	1991	1977	0/5	causal study	precise	10	236	224
Zimbabwe	1987	1981	0/7	no study	estimate	10	208	383

Notes: Reform year refers to the year in which the reform was implemented. Pivotal cohort corresponds to the first cohort potentially affected by the reform. Time frame refers to the number of periods before and after the reform that were included. Time frames are lower than 10 years if either another reform interferes with the respective reform, or it partly falls out of our age range. Relevance refers to whether a reform has been shown to effectively change enrollment or years of schooling. Cohort info indicates whether a cohort ha been estimated or whether it was obtained precisely.

Austria

Maria Theresia introduced six years of compulsory education in Austria already in 1774. More than one century later, in 1869, the duration of compulsory education was extended to 8 years by the introduction of the *Reichsvolksschulgesetz*.[24] Thus, Austria started in the 20th century with already 8 years of compulsory schooling. This duration of compulsory education was kept even beyond the break down of the Habsburger Monarchy in 1918. All through the 1st Republic (1919 - 1934), Austro-fascism (1934 -1938), and the annexation of Austria by the German Reich (1938 - 1945) the duration of compulsory education remained at 8 years (Rusinow, 1977; Engelbrecht, 1982).

After World War II, at the beginning of the 2nd Republic in 1945 the Education Law/Act of the 1st Republic was re-enacted (Rusinow, 1977; Loew, Markus, 2018). Even though, it was supposed to only be a provisional arrangement, it was not until 1962 that a new education act (*Bundesgesetz vom 25. Juli 1962 über die Schulpflicht (Schulpflichtgesetz)* 1962) was brought forward. It increased duration of compulsory education from 8 to 9 years. The law states that the change in the duration of compulsory education is to be introduced in 1966. The pivotal cohort is calculated based on this information and the information on the entry age (six years of age by september 1st). It, thus, deviates from, for instance, Brunello, Fort, and Weber (2009) who assumed the implementation to be in the year of 1962. Until today, the duration of compulsory education has remained 9 years.

Argentina

Argentina introduced compulsory schooling (seven years) as early as 1884.[25] This was expanded to cover nine years in 1993[26] and further increased to twelve years in 2006[27].

As the latter reform was phased in until 2011 and is, thus, at the very corner of our observation period, we use only variation that comes from the reform in 1993. This reform was phased in on the province level between 1996 and 2000. Never implemented in two provinces (City of Buenos Aires and Río Negro) (Crosta, 2007; Alzúa, Gasparini, and Haimovich, 2015). For a table of the years of implementation in each region see Alzúa, Gasparini, and Haimovich (2015). The phasing in was officially allowed between 1995 and 1999. In some cases the reform was gradually implemented within a province, in other provinces the reform was immediately fully implemented (Alzúa, Gasparini, and Haimovich, 2015). Individuals that are 14 years old at the time of the respective implementation are (potentially)[28] treated.

24. Note that this law did not only apply to what is contemporarily known as Austria but it covered Cisleithania, i.e., from todays perspective, the majority of territories of Czech Republic and Slovenia and parts of Poland, Montenegro, Italy and Croatia.

25. Ley Nº 1420, accessed: http://www.bnm.me.gov.ar/giga1/normas/5421.pdf

26. Ley Federal de Educaciòn, april 1993, accessed: www.fadu.uba.ar/application/post/download-filename/238

27. Ley Nº 26.206, accessed: http://www.me.gov.ar/doc_pdf/ley_de_educ_nac.pdf

28. Depending on whether the reform is gradually implemented, there is no grade repetition etc.

That is, in birth year terms, individuals born before 1982 were not treated, between 1982 and 1985 are treated depending on the province, and individuals born in and after 1986 are always treated.[29].

Bangladesh

Bangladesh made schooling (5 years) compulsory in 1990 (Ministry of Primary and Mass Education, 1990; Bajaj and Kidwai, 2016)

Bolivia

Bolivia first introduced compulsory education (four years) in the early 20th century.[30] With the education law of 1955, Bolivia increased this to seven years (Ströbele-Gregor, 2007) and further increased to eight and twelve years in 1994 and 2010 respectively.[31]

The 1994 reform was implemented over a horizon of 7 years, that is, between 1995 and 2002 (Contreras and Simoni, 2003). According to Contreras and Simoni (2003), it increased enrollment in primary education, the target stage of the compulsory schooling reform.

Cambodia

While there were attempts to introduce compulsory schooling for boys already in 1906 and 1916, these were never truly enforced (Bilodeau, Pathammavong, and Hông, 1955). Cambodia introduced compulsory primary education (6 years) in the school year of 1962 (Education, 1962). The typical primary entrance age was 6 years. Our reading of Education (1962) suggests that it applied to everyone in the population, regardless of their enrollment, so we estimate the pivotal cohort to 1962-6-6=1950.

Chile

Chile first introduced compulsory schooling (four years) in 1920.[32] This was increased to cover six years in 1929,[33] eight years in 1965,[34] and twelve years in

29. Ignoring the two exceptional provinces mentioned above
30. Ley de 6 de Febrero de 1900
31. Ley Nº 1565, Ley Nº 070
32. Ley de Instrucción Prmaria Obligatoria
33. Decreto 5.291
34. Decreto 27.952

2003.[35]. The reform of 1965 affected cohorts born in and after 1952 (Celhay and Gallegos, 2015). Results presented in Celhay and Gallegos (2015) are suggestive of the fact that the reform positively affected schooling. The reform of 2003 came into effect in 2004 and affected cohorts born in and after 1985 (see also Flores, Sanhueza, Atria, and Mayer, 2015). However, it is unclear that those that already left school, before entering secondary education, i.e. those that were born between 1985 and 1988 were only partially affected. Using a regression discontinuity design, (Flores et al., 2015) finds that the reform significantly increased years of schooling.

China

China introduced nine years of compulsory schooling in 1986. The law was implemented gradually across provinces. Fang, Eggleson, Rizzo, Rozelle, and Zeckhauser (2012) date the first potential cohort affected to 1972 and show that this significantly increased educational attainment.

Colombia

Colombia first introduced compulsory elementary schooling in 1927.[36] As the duration of elementary schooling varied between rural and urban areas, this also yielded two different lengths of compulsory years of schoooling in the beginning, namely three years and six years, respectively (Hanson, 1986, p.30). However, as there were no funds made available to implement this law, it was not really enforced at that time (Hanson, 1986, p.30).

Compulsory education was expanded to cover nine years of basic education (five years of primary schooling and four years of secondary education) in 1991.[37].

Costa Rica

The first law on compulsory and free education came into effect in 1862. It mandated school to be obilgatory for ages 8 through 12 (Furbay, 1946, p. 7). This provision was extended to cover ages 8 through 14 by the Common Education Law (Ley de Educaciòn Común) of 1886 (Furbay, 1946, p. 8). This coincided with the duration of primary education.

In 1971, Costa Rica put forth the National Educational Development Plan (NEDP). It proposed to extend compulsory education until the 9th grade. The

35. Ley Nº 19.876
36. Ley 56 de 1927
37. Constitucion Politica de 1991, Article 67

NEDP was put into law in 1973 (Solona and Olivera, 1974, p. 505). Nine years of schooling was not fully universalized immediately but instead there was phased in until 1980 (Solona and Olivera, 1974; Gill, 1980). Results in Gill (1980) are suggestive of increased enrollment due to the reform in the respective age groups.

Czech Republic

Compulsory schooling (8 years) was first introduced in 1869 (Garrouste, 2010). This was extended to 9 years in 1948 (Urban, 1972; Garrouste, 2010; Fort, Schneeweis, and Winter-Ebmer, 2011). This increase was reversed in 1953 and re-implemented in 1960 (Garrouste, 2010; Fort, Schneeweis, and Winter-Ebmer, 2011). In 1984, compulsory years of schooling was then increased to last 10 years (Baske, Benes, and Riedel, 1992; Kopp, 1992; Eurydice, 1997; Filer, Jurajda, and Plánovský, 1999). According to Kopp (1992), this increased schooling only for people that would have not stayed in school but decreased the time for individuals that would have stayed in school. As this implies a non-monotone effect of the reform, we will not use this reform. This increase was again reversed in 1990, after communism ended (Kopp, 1992; Eurydice, 1997; Fenoll and Kuehn, 2017).

Egypt

The Egyptian constitution of 1923 stipulated that “primary education shall be compulsory”. Primary education comprises six years of schooling (OECD, 2015). In 1981, this was increased to cover lower secondary education in 1981 (9 years). Assad, Aydemir, Dayioglu, and Kirdar (2016) suggests 1970 as the first cohort potentially affected. In 1988, this was decreased to 8 years by decreasing the length of primary schooling for one year. Assad et al. (2016) suggests 1978 as the pivotal cohort for this reform and shows that it effectively decreased years of schooling for the respective cohorts. This decision was reversed in 1999, increasing compulsory years of schooling back to nine years (National Center for Educational Research and Development, 2001).

Finland

Finland first introduced compulsory schooling (6 years) in 1921. In 1972, it started increasing it to 9 years. This reform was implemented on the municipality level within a five year time span. In line with (Brunello, Fort, and Weber, 2009), we date the first cohort to be potentially affected to 1961 and the first cohort that was fully affected to 1966 in line with Fort (2006). The reform was used as an instrument for years of schooling in (Brunello, Fort, and Weber, 2009).

France

France implemented its first compulsory schooling reform in 1882 (6 years). This was increased to 9 years in 1936 and again to 10 years in 1967. The pivotal cohort is taken from (Gathmann, Jürges, and Reinhold, 2015). The first reform was used as an instrument in (Gathmann, Jürges, and Reinhold, 2015), the second in (Brunello, Fort, and Weber, 2009; Fort, Schneeweis, and Winter-Ebmer, 2011; Gathmann, Jürges, and Reinhold, 2015; Brunello, Fort, Schneeweis, and Winter-Ebmer, 2016).

Ghana

Ghana introduced six years of compulsory schooling in 1961. This was extended to cover three years of junior secondary school in 1987. As the education system collapsed in 1981 due to the new military regime, we do not use this reform. (Kadingdi, 2006)

Greece

Greece introduced six years of compulsory schooling in 1927 (Garrouste, 2010). Before the military revolution, it tried to increase compulsory schooling to nine years, but this was immediately reversed (Murtin and Viarengo, 2011). The reform was instead implemented in 1976. The first potential cohort affected is dated to 1963 (Brunello, Fort, and Weber, 2009). The reform has been used as an instrument in (Brunello, Fort, and Weber, 2009; Fort, Schneeweis, and Winter-Ebmer, 2011; Brunello, Fort, Schneeweis, et al., 2016).

Hungary

As part of the Austro-Hungarian Empire, compulsory schooling (6 years) was introduced in Hungary 1868. It increased to eight years (6 - 14 years of age) in 1951[38], to nine years of age (6 - 15) in 1959[39], to ten (6 - 16) in 1961[40] (see also Ágoston,

38. Legislative Decree no. 15 of 1951 of the Presidential Council of the People's Republic on compulsory education and primary school

39. Legislative Decree no. 29 of 1959 of the Presidential Council of the People's Republic on compulsory education and on the amendment of Legislative Decree no. 15 of 1951 accessed from López-Falcón, Börsch-Supa, and Vieregg (2014)

40. Act 3 of 1961 on the educational system of the People's Republic of Hungary, accessed from López-Falcón, Börsch-Supa, and Vieregg (2014)

1980). The 1959 legislation, however only prescribes that children who do not finish the 8th grade by the age of 14 must stay in school for an additional year, while the 1961 legislation additionally introduces a training school that must be attended until 16 years of age unless children entered into secondary education, apprentice traineeship or major employment. For this reason, we code the de-facto increase in years of compulsory schooling to be in 1961.

It was further increased to last 12 (6-18) years in 1996[41] but implemented for children who begin primary school in 1998[42], i.e. children born in 1992. The last change was reversed in 2011.[43]. (see also Eurydice, 1997; Eurydice, 2013)

Japan

Japan first introduced compulsory primary education (4 years) in 1872, this was increased to cover 6 years in 1880. In postwar Japan, compulsory schooling was inreased to cover 3 years of lower secondary education by implementing the School Fundamental Law of 1947 (Self and Grabowski, 2003). Students enter primary school in the year in which they are 6 years of age in April. Thus, the potentially first cohort to be affected is the cohort born in mid to late 1934.

Jordan

Jordan's constitution of 1952 stipulates six years of compulsory primary schooling. This was extended to 9 years in 1964 (Education, 1964; Burke and al-Wakid, 1997; UNESCO-IBE, 2006). The typical entrance age in Jordan at that time was 6 years. We thus estimate the pivotal cohort to be 1952=1964-6-6.

Indonesia

Indonesia introduced compulsory primary education in 1984 (6 years) and increased this to cover lower secondary education in 1994 (Yeom, Acedo, Utomo, and Yeom, 2002; Purnastuti, Salim, and Joarder, 2015). Purnastuti, Salim, and Joarder (2015) dates the respective pivotal cohorts to 1977 and 1987 respectively and shows that they effectively increased educational attainment.

41. Act 62 of 1996 on the amendment of Act 79 of 1993 on public education
42. Act LXVIII of 1999 on the ammendment of Act 79 of 1993 on public education, section 6(5)
43. Act 190 of 2011 on national public education, accessed from López-Falcón, Börsch-Supa, and Vieregg (2014)

Iran

Iran increased its compulsory schooling from six to nine years in 1971. (Watson, 1976) states that the first cohort afftected by this reform entered secondary school in the school year of 1973/74 suggesting that the first pivotal cohort affected by this reform is the cohort of 1961. Note that the literature disagrees on changes on compulsory schooling after 1987. Also we cannot pin point exactly when the six years where introduced. The literature suggests that it was in or before 1943. (Watson, 1976; Wilber, 1981; UNESCO, 1987; George and Scatolini, 2015).

Iraq

Iraq already introduced compulsory education in 1940 but this law was never really enforced (Education, 1958; Ranjan and Jain, 2009). It reintroduced compulsory education for 6 years in 1976 (Ranjan and Jain, 2009).[44]. Article 10 of the law stipulates that the law shall be applied starting the school year 1978/1979. The compulsory schooling starts when children are six years at the beginning of the school year,.i.e cohorts born in late 1971 are the first one to be potentially affected.

Israel

Isreal first introduced compulsory education in 1949 (8 years of schooling and one year of kindergarten). In 1969, a law was passed that increased compulsory schooling to ten years. This was phased in and gradually implemented. The law first applied to 9th graders in 1970-73 and then to tenth graders in 1974-75. We take the respective pivotal cohorts from Krief (2009). The results of Krief (2009) suggest that these reforms effectively increased schooling, in particular for Asian and African Israelis and Non-jewish individuals.

Italy

Italy started into the 20th Century with four years of compulsory education. The next major reform of the compulsory education system followed under the Fascist Rule (1922 - 1943/45) with the Gentile Reform of 1923. It formally increased compulsory education to last until the age of 14 (8 years).[45] (Scarangello, 1964)

While the constitution of 1947 (implemented 1st of January, 1948) already stipulates that primary school shall be compulsory, free and shall last at least 8 years

44. Compulsory Education Law No. 118 of 1976

45. Garrouste (2010) is of the opinion that the Gentile Reform only stipulated 5 years of compulsory education.

(for an English translation of the relevant parts see, e.g., Scarangello 1964), the vast majority of literature (Shavit and Westerbeek, 1998; Brandolini and Cipollone, 2002; Brunello, Fort, and Weber, 2009; Garrouste, 2010; Murtin and Viarengo, 2011; Brunello, Fort, Schneeweis, et al., 2016) suggests that until the educational reform of 1962 compulsory schooling actually lasted only 5 years (the actual duration of primary school at the time) and was increased by the reform to 8 years by aboloshing the vocational track in lower secondary education (6th - 8th grade), and forcing all students to follow the then unitary (academic) junior high school (scuola media) (Scarangello, 1964). The law was enforced upon september 1963. Thus - in agreement with the literature, i.e. (Shavit and Westerbeek, 1998; Brandolini and Cipollone, 2002; Brunello, Fort, and Weber, 2009; Brunello, Fabbri, and Fort, 2013; Brunello, Fort, Schneeweis, et al., 2016) - the first cohort potentially affected by this reform was the cohort of 1949.

In 1999, the government increased the duration of compulsory education further to 9 years and then again to 10 years in 2007 (Eurydice, 1996; Eurydice, 2000; Eurydice, 2005; Eurydice, 2007; Eurydice, 2009; Murtin and Viarengo, 2011). According to Eurydice (2009) this law was introduced in the school year 2007/08. Thus, the cohort that is potentially first affected was born in 1991.

Mexico

The first mention of compulsory education we were able to recover can be dated to 1934 (six years).[46] This was extended to cover nine years in 1993.[47] Creighton and Park (2010) suggests that the first cohort to be affected is the cohort born in 1980. In 2012, Mexico increased its compulsory schooling to 11 years.[48]

Morocco

Morocco introduced compulsory education in 1963 for the ages 7 to 13 (6 years) (Diyen, 2004).[49]

Netherlands

Netherlands first introduced compulsory schooling in 1901 (six years). This was increased to 7 years in 1921, decreased again in 1924, increased again to seven in 1928, increased to 8 in 1948, decreased to seven years in 1942 (Levin and

46. Reforma de la Constitucion Politica de los Estados Unidos Mexicanos
47. Ley General de Educacion
48. Reforma de la Constitucion Politica de los Estados Unidos Mexicanos
49. Dahir No. 1-63-071 (1963)

Plug, 1999; Gathmann, Jürges, and Reinhold, 2015; Brunello, Fort, Schneeweis, et al., 2016). The cohorts, if available, are taken from Gathmann, Jürges, and Reinhold (2015) and Brunello, Fort, Schneeweis, et al. (2016). After this phase of implementations and reversals, compulsory schoooling monotonically increased from seven to nine in 1950, from nine to ten in 1973 and from ten to twelve in 1985. For the reforms in 1950 and 1973, we take the cohorts from Gathmann, Jürges, and Reinhold (2015) and Oosterbeek and Webbink (2004), respectively. The 1950 reform has been used by Gathmann, Jürges, and Reinhold (2015) as an instrument, the 1973 by Brunello, Fort, and Weber (2009) and Gathmann, Jürges, and Reinhold (2015). Fenoll and Kuehn (2017) suggest that the 1985 reform increased compulsory schooling for those born in 1973 or later to eleven years and for individuals born after 1980 to twelve years.

Peru

Peru introduced compulsory education in 1905 (six years). This was extended to cover eleven years in 1993 (Weitzman, 2017). Weitzman (2017) suggests that the pivotal cohort can be dated to 1982. Utilizing a regression discontinuity design, Weitzman (2017) finds that the reform significantly increased years of schooling by 0.32 (s.e. 0.1) years for females.

Philippines

The reform of the constitution in 1987 made elementary schooling compulsory (6 years) (Okabe, 2013).

Poland

As Poland was divided among Russia, Prussia, and Austria-Hungary until 1918, the first introduction of compulsory schooling depends on the part of the country. Austria-Hungary introduced compulsory schooling in Poland in 1895 (7 years) and Prussia in 1825 (8 years) (Baske, 1987). In 1932, Poland made his first own compulsory schooling law (Hessen, 1934) which was valid until the Nazis invaded Poland in 1939. During the war, Polish were only able be schooled in secret (Baske, 1987). After the war, the law of 1932 was reestablished and renewed in 1956 (Baske, 1987). It increased compulsory schooling to cover eight years in 1961 (Baske, 1987) and to cover nine years in 1999 (Eurydice, 2000; Jakubowski, 2015). For the calculation of the cohort of the 1961 reform, we rely on a speech of the Polish education minister, which states that the reform will be implemented in the school year of 1966/67 (see Baske, 1987, source 136, p. 454ff). This is broadly in line with the date (Mocan and Pogorelova, 2017) use. This reform has been used in an instrumental variable approach in Mocan and Pogorelova (2017). The Act of 8 January 1999 on the Implementation of the Education System Reform implies that the pivotal cohort of the 1999 reform is the cohort of 1986, this is also in line with Fenoll and Kuehn (2017).

Portugal

The first reform we were able to recover was implemented differentially for boys and girls. It increased the compulsory years of schooling from three to four years for boys in 1956 and girls in 1960 (Fort, 2006; Murtin and Viarengo, 2011). As we do not know when the three years were introduced we drop the respective cohorts born before the reforms. In 1964, Portugal increased the compulsory years of schooling to six years (Murtin and Viarengo, 2011; Brunello, Fabbri, and Fort, 2013; Mocan and Pogorelova, 2017) Cohort is taken from (Brunello, Fabbri, and Fort, 2013). This reform has been used as an instrument in Brunello, Fabbri, and Fort (2013) and Mocan and Pogorelova (2017). This was extended to cover nine years in 1986[50] (Murtin and Viarengo, 2011). The law applies first to cohorts entering primary school in the school year 1987/88, i.e. the inidividuals born late 1980. Portugal increased their compulsory schooling to twelve years in 2009.[51]. The implementation of the law states that the law is valid for all students starting the school year 2009/10 except for those already in eighth grade or higher.

Romania

We cannot exactly pin point the first introduction of compulsory schooling in Romania. The first reform we are aware of increased compulsory schooling from four to seven years in 1948(Antochi, 1981). It then increased compulsory schooling to eight years in 1961, and to ten years in 1968 (Education, 1962; Education, 1968; Miclescu, 1980).[52]. According to Miclescu (1980), the reform of 1968 the actual compulsory years of schooling mostly untouched because it left many exceptions. So we do not use this reform. Miclescu (1980) dates the de facto implementation of 10 (7-17 years of age) years of compulsory schooling to the school year 1974/75, i.e. the pivotal cohort affected to 1959. This change was reversed after the break down of the communist regime. It then increased it to nine years in 1999[53](Eurydice,

50. Law 46/36 of 14 October 1986
51. Law No. 85/2009, of August 27
52. Law 11 / 1968 on education in the Socialist Republic of Romania, accessed from López-Falcón, Börsch-Supa, and Vieregg (2014)
53. Law 151 of 30 July 1999

2005; UNESCO-IBE, 2010).

South Korea

South Korea implemented free compulsory primary education between 1954 and 1959 (Kim, 2011). It increased it to cover lower secondary edcuation (9-years) in 1985 in rural areas and in 2002 in the rest of the country. We use the first implementation of the reform (Kim, Kim, Kim, and Kim, 2006).

Spain

The first mention of compulsory schooling, we could recover, prescribes three years of mandatory schooling and was passed in 1857.[54]. This was increased to six years in 1945.[55] In 1970 this was further expanded to cover eight years (Garrouste, 2010; Murtin and Viarengo, 2011; Gathmann, Jürges, and Reinhold, 2015).[56]. The cohort is taken from Brunello, Fort, and Weber (2009) and Gathmann, Jürges, and Reinhold (2015) which use this reform in the instrumental variable estimation. In 1991, Spain further expanded their compulsory schooling to ten years (Eurydice, 1990; Garrouste, 2010; Murtin and Viarengo, 2011).[57]. The implementation decree suggests that the pivotal cohort was born in 1978.

Sweden

Sweden first introduced compulsory schooling (six years) in the 19th century[58]. In 1936, it increased it to seven years[59] and in 1962 to nine years.[60] The implementation of the last reform happened gradually on the municipality level with the vast majority covered in the first 5 years. The pivotal cohort is dated to 1950 in accordance with Brunello, Fort, and Weber (2009) and the first cohort that was

54. Ley de Instrucción Pública
55. Ley de Enseñanza Primaria
56. Law on Basic General Education
57. Ley Orgánica para la Mejora de la Calidad Educativa 1990, implementation in Royal Decree 986/1991, of 14 June
58. HM's statute (1842:19) regarding education among the general population, also known as the 1842 primary school code and HMs renewed statute (1882:8) regarding education among the general population, also known as the 1882 primary school code, accessed from López-Falcón, Börsch-Supa, and Vieregg (2014)
59. HM's ordinance regarding certain changes to the renewed statute of 26 September 1921 (604) regarding education among the general population, accessed from López-Falcón, Börsch-Supa, and Vieregg (2014)
60. Law regarding comprehensive schools (1962:319); The 1962 comprehensive school code (1962:439), accessed from López-Falcón, Börsch-Supa, and Vieregg (2014)

fully affected to 1955.

Thailand

Thailand introduced compulsory primary education (4 years) already in 1921. This was increased in 1978 to six years and in 1999 to nine years (OECD/UNESCO, 2016; Chankrajang and Muttarak, 2017). The typical entrance age into education was 7.

Turkey

Turkey first introduced compulsory years of schooling (5 years) in 1923. This was increased to cover eight years in 1997 (Güneş, 2015; Kirdar, Dayioglu, and Koç, 2016). The pivotal cohort is taken from Güneş (2015) which shows that this reform effectively increased educational attainment.

United Arab Emirates

United Arab Emirates introduced compulsory primary education (six years) in 1972 (Federal Law of 1972, No 1-M7 and No 11).

Vietnam

There were several early attempts to introduce compulsory schooling in Vietnam, but none of them were successful (Bilodeau, Pathammavong, and Hông, 1955). The first true introduction can be dated to 1991. In 1991 primary schooling (5 years) was made compulsory (Dang, 2017). Dang (2017) dates the first cohort to be affected by this reform to 1977 and finds that this significantly increased years of schoooling.

Venezuela

Venezuela first established compulsory education in 1870 (six years) (Hanson, 1986). This was not enforced at the time (Hanson, 1986). In 1980, this was extended to cover nine years of schooling (Patrinos, 2004). In line with Patrinos (2004) we take 1968 as the first potential cohort. Patrinos (2004) finds that the reform increased years of schooling by 1.363 (t-value=7.5) years.

South Africa

While South Africa had compulsory laws for certain racial groups in certain states already in place in the late 1940’s (UNESCO-IBE, 1949; UNESCO-IBE, 1950), the first nationwide legislation can be dated to 1974. It made schooling compulsory for whites/Africaans from the age of 7 to 16 (Johnson, 1982). Then in in 1984, education was made compulsory for all racial groups but at different levels:

- Whites: 7-16
- Black: 7th grade (7-)
- Asian and coloured: 7-15

However, this law was only weakly enforced for non-White students. With the school year of 1995, schooling was made compulsory from the age of 6 to 16 for all people (9 years). (Federal Research Division, 1997). We include only the reform of 1974, as the reform of 1984 did not bring changes for Whites and was only weakly enforced for the rest. We estimate the pivotal cohort to be 1974-7=1967.

Zimbabwe

Schooling was made compulsory for whites in 1931 and extended to the Asian community in 1938. This law was repelled in 1979 to abolish racial disparities in the education system. In 1987, primary education (7-years) was made compulsory for everyone. The entrance age to primary education is 6, so we estimate the pivotal cohort affected by this law to be 1981. (Lemon, 1995)

Soviet Union

Russia, Estionia (1940-1991), Lithuania (1940-1991), Moldova (1940-1991), Ukraine (1922-1991), Kazakhstan (1925-1991), Georgia (1936-1991)

Before the Russian Revolution in 1917, there was no compulsory schooling law in Russia. In 1918, Lenin tried to implement nine years of schooling, this failed and was reduced to seven years in 1921. This failed as well (Anweiler and Meyer, 1961; Koeder, 1977). Koeder (1977) dates the first successful introduction of compulsory schooling to 1930.[61] This prescribed seven years for children in industrial cities and four years for children in rural areas. This was increased to seven years for everyone in 1949 (Chabe, 1971), and increased to cover eight years in 1959[62]. The law stipulates that the eight years of schooling is supposed to be phased in between the school year 1959/60 and 1962. The Soviet Union then later introduced ‘universal

61. Resolution of the Central Committee of the Communist Party of the Soviet Union, 25th of July, 1930, German translation in Anweiler and Meyer (1961, p. 173)

62. Law on the Consolidation of the Relationship of School with Life and on the Development of the People’s School System in the Soviet Union, 24th of December, 1958, German translation in Anweiler and Meyer (1961, pp. 308)

secondary schooling' covering 10 years of schooling. The exact date of the implementation is difficult to pin down and there is some disagreement to what extent this was a legal increase of compulsory schooling. Chabe (1971) suggests that it was introduced in 1970; Anweiler, Kuebart, and Meyer (1976) and Anweiler (1985) suggest that there was no official increase in compulsory years of schooling; the statistical yearbooks of the UNESCO suggest that the official years of compulsory schooling were 10 years starting 1980 (UNESCO, 1963-1999). For that reason, we do not include this reform.

After the Soviet Union broke down Russia decreased its compulsory years of schooling to nine years, and increased again to eleven years in 2007 (UNESCO, 2011).

Estonia, late Soviet years, post-soviet

Toward the end of Soviet Union it was partially allowed to pursue its own education policies. Shortly before the collapse of the Soviet Union, it increased compulsory schooling to 12 years, this however was quite immediately reversed after the break down of the Soviet Union. (Saar, 2008)

Ukraine, post-soviet

Ukraine increased to twelve years in 1999 affecting cohorts born in and after 1995. [63]

Lithuania, Moldova, Kazakhstan, Georgia, post-soviet

We are not aware of any reforms in the post-soviet era in these countries.

63. Law on General Secondary Education 1999, translation obtained from the UNESCO. Law states that expansion of compulsory schooling (article 12) shall come into force for students entering primary school in 2001.

References

Ágoston, György. 1980. “Die Bildungsreform in Ungarn und die Perspektiven des Bildungswesens.” In *Bildungssysteme in Osteuropa - Reform oder Krise*. Edited by Oskar Anweiler and Friedrich Kuebart. Berlin: Berlin Verlag, 92–96. Osteuropaforschung Schriftenreiheder Deutschen Gesellschaft für Osteuropakunde Herausgegeben von Wolfgang Kasack Band 12. [189]

Altinok, N., N. Angrist, and H.A. Patrinos. 2018. *Global Data Set on Education Quality (1965-2015)*. Washington DC: World Bank. [181]

Alzúa, María Laura, Leonardo Gasparini, and Francisco Haimovich. 2015. “Education reform and labor market outcomes: The case of argentina’s ley federal de educaciÓn.” *Journal of Applied Economics* 18 (1): 21–43. [185]

Antochi, Iosif. 1981. “Die Entwicklung des Bildungswesens in Rumänien in den Jahren des sozialistischen Aufbaus.” In *Vergleichende Erziehungswissenschaften: Festschrift für Hermann Röhrs zum 65. Geburtstag*. Edited by Ulrich Baumann, Volker Lenhart, and Axel Zimmermann. Erziehungswissenschaftliche Reihe - Band 22. Wiesbaden: Akademische Verlagsgesellschaft, 133–46. [194]

Anweiler, Oskar. 1985. “Zentralismus und Förderalismus im sowjetischen Bildungswesen.” In *Sowjetsystem und Ostrecht: Festschrift für Boris Meissner zum 70. Geburtstag*. Edited by Georg Brunner. Berlin: Duncker & Humblot, 179–96. [198]

Anweiler, Oskar, Friedrich Kuebart, and Klaus Meyer. 1976. *Die sowjetische Bildungspolitik von 1958 bis 1973 - Dokumente und Texte*. Berlin: Osteuropa-Institut an der Freien Universität Berlin - Erziehungswissenschaftliche Veröffentlichung. [198]

Anweiler, Oskar, and Klaus Meyer. 1961. *Die sowjetische Bildungspolitik seit 1917: Dokumente und Texte*. Edited by Oskar Anweiler and Klaus Meyer. Heidelberg: Quelle & Meyer. [197]

Assad, Ragui, Abdurrahmann Aydemir, Meltem Dayioglu, and Murat Guray Kirdar. 2016. “Returns to Schooling in Egypt.” *Economic Research Forum: Working paper series* 1000: arXiv: arXiv: 1011.1669v3. [188]

Bajaj, Monisha, and Huma Kidwai. 2016. “Human Rights and Education Policy in South Asia.” *Handbook of Global Education Policy*, 206–23. [186]

Baske, Siegfried. 1987. *Bildungspolitik in der Volksrepublik Polen 1944 - 1986, Teil I*. Edited by Oskar Anweiler and Siegfried Baske. 1st edition. Berlin: Osteuropa-Institut an der Freien Universität Berlin - Erziehungswissenschaftliche Veröffentlichung Band 18. [193]

Baske, Siegfried, Milan Benes, and Rainer Riedel. 1992. *Der Übergang von der marxistisxhleninistischen zu einer freiheitlich-demokratischen Bildungspolitik in Polen, in der Tschechoslowakei und in Ungarn*. Berlin: Osteuropa-Institut an der Freien Universität Berlin - Erziehungswissenschaftliche Veröffentlichung. [188]

Bilodeau, Charles, Somlith Pathammavong, and Lê Quang Hông. 1955. *Compulsory Education in Cambodia, Laos and Viet-Nam*. Vol. XIV, Studies on Compulsory Education. IBE-Unesco, 1689–. arXiv: arXiv:1011.1669v3. [186, 196]

Bolt, Jutta, Robert Inklaar, Herman de Jong, and Jan Luiten van Zanden. 2018. “Rebasing ‘Maddison’: New Income Comparisons and the Shape of Long-Run Economic Development.” *Maddison Project Working paper*, (10): [180]

Brandolini, Andrea, and Piero Cipollone. 2002. *Return to Education in Italy: 1992 - 1997*. Rome: Bank of Italy. [192]

Brunello, Giorgio, Daniele Fabbri, and Margherita Fort. 2013. “The Causal Effect of Education on Body Mass: Evidence from Europe.” *Journal of Labor Economics* 31 (1): 195–223. [192, 194]

Brunello, Giorgio, Margherita Fort, Nicole Schneeweis, and Rudolf Winter-Ebmer. 2016. “The Causal Effect of Education on Health: What is the Role of Health Behaviors?” *Health Economics* 25 (3): 314–36. [189, 192, 193]

Brunello, Giorgio, Margherita Fort, and Guglielmo Weber. 2009. “Changes in Compulsory Schooling, Education and the Distribution of Wages in Europe.” *Economic Journal* 119 (536): 516–39. [185, 188, 189, 192, 193, 195]

“Bundesgesetz vom 25. Juli 1962 über die Schulpflicht (Schulpflichtgesetz).” 1962. “Bundesgesetz vom 25. Juli 1962 über die Schulpflicht (Schulpflichtgesetz).” [185]

Burke, D., and A. al-Wakid. 1997. “On the Threshold: Private Higher Education in Jordan.” *International Higher Education* 9: 2–4. [190]

Celhay, Pablo, and Sebastian Gallegos. 2015. “Persistence in the Transmission of Education: Evidence across Three Generations for Chile.” *Journal of Human Development and Capabilities* 16 (3): 420–51. [187]

Chabe, Alexander. 1971. “Soviet Educational Policies: Their Development, Administration, and Content.” *Educational Leadership*, 22–25. [197, 198]

Chankrajang, Thanyaporn, and Raya Muttarak. 2017. “Green Returns to Education: Does Schooling Contribute to Pro-Environmental Behaviours? Evidence from Thailand.” *Ecological Economics* 131: 434–48. [196]

Contreras, Manuel E., and Maria Luisa Talavera Simoni. 2003. “The Bolivian Education Reform 1992-2002 : Case Studies in Large-Scale Education Reform.” *Country Studies - Education Reform and Management Publication Series* II (2): 102. [186]

Creighton, Mathew, and Hyunjoon Park. 2010. “Closing the Gender Gap: Six Decades of Reform in Mexican Education.” *Comparative Education Review* 54 (4): 513–37. [192]

Crosta, Facundo. 2007. “Exploring the effects of the school levels reform on access and its quality: The Education Federal Law of Argentina.” *Well-Being and Social Policy Magazine* 3 (1): 97–122. [185]

Dang, Thang. 2017. “Education as Protection? The Effect of Schooling on Non-Wage Compensation in a Developing Country.” *MPRA Paper No. 79223*, arXiv: sureshgovindarajan. [196]

Diyen, Hayat. 2004. “Reform of Secondary Education in Morocco: Challenges and Prospects.” *Prospects* XXXIV (2): 211–22. [192]

Education, UNESCO / International Bureau of. 1958. *International Yearbook of Education*. Vol. 20, Paris, Geneva: UNESCO / International Bureau of Education. [191]

Education, UNESCO / International Bureau of. 1962. *International Yearbook of Education*. Vol. 18, Paris, Geneva: UNESCO / International Bureau of Education. [186, 194]

Education, UNESCO / International Bureau of. 1964. *International Yearbook of Education*. Vol. 26, Paris, Geneva: UNESCO / International Bureau of Education. [190]

Education, UNESCO / International Bureau of. 1968. *International Yearbook of Education*. Vol. 24, Paris, Geneva: UNESCO / International Bureau of Education. [194]

Engelbrecht, Helmut. 1982. *Geschichte des österreichischen Bildungswesens - Erziehung und Unterricht auf dem Boden Österreichs*. Vol. 5, Österreichischer Bundesverlag. [185]

Eurydice. 1990. *Major educational developments in the Member States of the European Community: January to September 1990*. Brussels: European Education, and Culture Executive Agency. [195]

Eurydice. 1996. *Key Data on Education in Europe 2002*. Brussels: European Education, and Culture Executive Agency. [192]

Eurydice. 1997. *Supplement to the Study in the Structures of Education and Initial Training Systems in the European Union - The Situation inn Bulgaria, the Czech Republic, Hungary, Poland,* *Romania and Slovakia*. Brussels: European Education, and Culture Executive Agency. [188, 190]

Eurydice. 2000. *Key Data on Education in Europe 2000*. Brussels: European Education, and Culture Executive Agency. [192, 193]

Eurydice. 2005. *Key Data on Education in Europe 2005*. Brussels: European Education, and Culture Executive Agency. [192, 194]

Eurydice. 2007. *Description of the structures of the education systems from pre-primary to higher education (ISCED 0 to 5) - School year 2007/08*. Brussels: European Education, and Culture Executive Agency. [192]

Eurydice. 2009. *Key Data on Education in Europe 2009*. Brussels: European Education, and Culture Executive Agency. [192]

Eurydice. 2013. *Compulsory Education in Europe 2013/14*. Brussels: European Education, and Culture Executive Agency. [190]

Fang, Hai, Karen N Eggleson, John A Rizzo, Scott Rozelle, and Richard J Zeckhauser. 2012. "The returns to Education in China: Evidence from the 1986 compulsory education law." *NBER Working Paper 1818*, 1–40. [187]

Federal Research Division, United States of America. 1997. *South Africa: A Country Study*. Edited by Rita M Byrnes. 3rd edition. Area Handbook Series. Library of Congress. [197]

Fenoll, Ainhoa Aparicio, and Zoë Kuehn. 2017. "Compulsory Schooling Laws and Migration Across European Countries." *Demography* 54: 2181–200. [188, 193, 194]

Filer, Randall K., Štěpán Jurajda, and Ján Plánovský. 1999. "Education and wages in the Czech and Slovak Republics during transition." *Labour Economics* 6 (4): 581–93. [188]

Flores, Ignacio, Claudia Sanhueza, Jorge Atria, and Ricardo Mayer. 2015. "Top Incomes in Chile: A Historical Perspective on Income Inequality, 1964–2017." *Review of Income and Wealth* 66 (4): 850–74. [187]

Fort, Margherita. 2006. "Educational reforms across Europe: A toolbox for empirical research." Working Paper. [188, 194]

Fort, Margherita, Nicole Schneeweis, and Rudolf Winter-Ebmer. 2011. "More Schooling, More Children: Compulsory Schooling Reforms and Fertility in Europe." *IZA Discussion Paper 6015*, [188, 189]

Furbay, John H. 1946. *Education in Costa Rica*. Bulletin, 1946, No. 4. US Office of Education, Federal Security Agency. [187]

Garrouste, Christelle. 2010. *100 Years of Educational Reforms in Europe: A Contextual Database*. Publications Office of the European Union, Luxembourg. [188, 189, 192, 195]

Gathmann, Christina, Hendrik Jürges, and Steffen Reinhold. 2015. "Compulsory Schooling Reforms, Education and Mortality in Twentieth Century Europe." *Social Science and Medicine* 127: 74–82. [189, 193, 195]

George, Milton A, and Sergio Saleem Scatolini. 2015. *Language, Culture and Education*. Oman: Euro-Khaleeji Research, and Publishing House. [191]

Gill, Clark C. 1980. *The Educational System of Costa Rica. Education Around the World.* Washington D.C.: Department of Education. [188]

Güneş, Pinar Mine. 2015. "The Role of Maternal Education in Child Health: Evidence From a Compulsory Schooling Law." *Economics of Education Review* 47: 1–16. [196]

Hanson, E. Mark. 1986. *Educational Reform and Administrative Development: The Cases of Colombia and Venezuela*. Education and Society. Stanford, California: Hoover Institution Press, Stanford University. [187, 196]

Hessen, Sergius. 1934. "Die polnische Schulreform und ihre Träger." In *Jahrbücher für Kultur und Geschichte der Slaven*. Vol. 10, Franz Steiner Verlag, 145–62. [193]

Jakubowski, Maciej. 2015. "Opening up Opportunities: Education Reforms in Poland." *IBS Policy Papers*, (January): 1–20. [193]

Johnson, Walton R. 1982. "Education: Keystone of Apartheid." *Anthropology & Education Quarterly*, 214–37. [197]

Kadingdi, Stanislaus. 2006. "Policy initiatives for change and innovation in basic education programmes in Ghana." *Educate*∼ 4 (2): 3–18. [189]

Kim, Gwang-Jo. 2011. "Education Policies and Reform in South Korea." *Pacific Focus* 26 (2): 260–86. [195]

Kim, Ee-gyeong, Kap-sung Kim, Do-ki Kim, and Eun-young Kim. 2006. "Improving School Leadership: Country Background for Korea." Report. Korean Educational Development Institute. [195]

Kirdar, Murat G., Meltemlu Dayioglu, and Ismet Koç. 2016. "Does longer compulsory education equalize schooling by gender and rural/urban residence?" *World Bank Economic Review* 30 (3): 549–79. [196]

Koeder, Kurt W. 1977. *Das Bildungswesen der UdSSR: von der Oktoberrevolution zum 25. Parteitag der KPdSU*. 1st edition. München: Ehrenwirth. [197]

Kopp, Botho von. 1992. "The Eastern European Revolution and Education in Czechoslovakia." *Comparative Education Review* 36 (1): 101–13. [188]

Krief, Tomer. 2009. "The compulsory education law in Israel and liquidity constraints." *Israel Economic Review* 7 (1): 73–112. [191]

Lemon, Anthony. 1995. "Education Africa: Post-apartheid lessons." *Comparative Education* 31 (1): 101–14. [197]

Levin, Jesse, and Erik J.S. Plug. 1999. "Instrumenting education and the returns to schooling in the Netherlands." *Labour Economics* 6 (4): 521–34. [192]

Loew, Markus. 2018. "Das Schulorganisationsgesetz 1962." [185]

López-Falcón, Diana, Axel Börsch-Supa, and Dirk Vieregg. 2014. "Social Policy Archive for SHARE (SPLASH)." https://splash-db.eu/. [189, 190, 194, 195]

Miclescu, Maria. 1980. "Probleme und Perspektiven der Bildungsreform in Rumänien." In *Bildungssysteme in Osteuropa - Reform oder Krise*. Edited by Oskar Anweiler and Friedrich Kuebart. Berlin: Berlin Verlag, 111–20. [194]

Ministry of Primary and Mass Education, Bangladesh. 1990. "Primary Education (Compulsory) Act, 1990." *Bangladesh Gazette, Extra*, [186]

Mocan, Naci, and Luiza Pogorelova. 2017. "Compulsory schooling laws and formation of beliefs: Education, religion and superstition." *Journal of Economic Behavior and Organization* 142: 509–39. [194]

Murtin, Fabrice, and Martina Viarengo. 2011. "The Expansion and Convergence of Compulsory Schooling in Western Europe, 1950-2000." *Economica* 78 (311): 501–22. [189, 192, 194, 195]

National Center for Educational Research and Development, Arab Republic of Egypt. 2001. "Education Development." Report. [188]

OECD. 2015. *Schools for Skills: A New Learning Agenda for Egypt*. Paris: OECD Publishing. [188]

OECD/UNESCO. 2016. *Education in Thailand: An OECD-UNESCO Perspective: Education in Thailand*. Reviews of National Policies for Education. Paris: OECD Publishing. [196]

Okabe, Masayoshi. 2013. "Where Does Philippine Education Go? The "K to 12" Program and Reform of Philippine Basic Education." *IDE DISCUSSION PAPER No. 425*, [193]

Oosterbeek, H.; and D. Webbink. 2004. “Wage effects of an extra year of lower vocational education: Evidence from a simultaneous change of compulsory school leaving age and program length.” *Scholar Working Paper Series No. 44/04*, 1–18. [193]

Patrinos, Harry Anthony. 2004. “Schooling and Labor Market Impacts of a Natural Policy Experiment.” *World Bank Policy Research Working Paper* No.3460: [196]

Purnastuti, Losina, Ruhul Salim, and Mohammad Abdul Munim Joarder. 2015. “The returns to education in Indonesia: Post reform estimates.” *Journal of Developing Areas* 49 (3): 183–204. [190]

Ranjan, Rakesh Kumar, and Prakash C Jain. 2009. “The decline of educational system in Iraq.” *Journal of Peace Studies* 16 (1-2): 1–12. [191]

Roodman, David, Morten Ørregaard Nielsen, James G MacKinnon, and Matthew D Webb. 2019. “Fast and Wild: Bootstrap Inference in Stata Using Boottest.” *Stata Journal* 19 (1): 4–60. [174]

Rusinow, Dennison I. 1977. “Educational Reforms in Austria and Yugoslvia - Part I: Five-Day School, All-Day School, Comprehensive School.” *Southeast Europe Series* XXII (3): 1–18. [185]

Saar, Ellu. 2008. “The Estonian educational system and the ISCED-97.” In *The International Standard Classification of Education (ISCED-97). An evaluation of content and criterion validity for 15 European Countries*. Edited by S. Schneider. Mannheimer Zentrum für Europäische Sozialforschung, 237–52. [198]

Scarangello, Anthony A. 1964. “Progress and Trends in Italian Education.” Report. U.S. Department of Health, Education, and Welfare. [191, 192]

Self, Sharmistha, and Richard Grabowski. 2003. “Education and long-run development in Japan.” *Journal of Asian Economics* 14 (4): 565–80. [190]

Shavit, Yassi, and Karin Westerbeek. 1998. “Educational Stratification in Italy - Reforms, Expansion, and Equality of Opportunity.” *European Sociological Review* 14 (September): 33–47. [192]

Solona, Uladisloa Gàmez, and Carlos E. Olivera. 1974. “Costa Rica: a national educational development plan.” *Prospects* IV (4): 503–11. [188]

Ströbele-Gregor, Juliana. 2007. “Bildungsreform und indianische Bewegung in Bolivien.” *Lateinamerika. Analysen-Daten-Dokumentation* 13 (31): 62–73. [186]

UNESCO. 1987. “Statistical Yearbook of Education.” Report. Geneva. [191]

UNESCO. 2011. *Russian Federation*. World Data on Education 2010/2011. Geneva: UNESCO-IBE. [198]

UNESCO. 1963-1999. *Statistical Yearbook of Education*. Vol. 1-36, Geneva: UNESCO. [198]

UNESCO-IBE. 1949. *International Yearbook of Education*. Vol. 11, Paris, Geneva: UNESCO / International Bureua of Education. [196]

UNESCO-IBE. 1950. *International Yearbook of Education*. Vol. 12, Paris, Geneva: UNESCO / International Bureua of Education. [196]

UNESCO-IBE. 2006. “Jordan.” In *World Data on Education*. 6th edition. Geneva: UNESCO-IBE, 1–31. [190]

UNESCO-IBE. 2010. *Romania*. World Data on Education 2010/2011. Geneva: UNESCO-IBE. [195]

Urban, Rudolf. 1972. *Die Entwicklung des tschechoslowakischen Schulwesens 1959 - 1970 - Ein dokumentarischer Bericht*. Edited by Oskar Anweiler and Siegfried Baske. Berlin: Quelle & Meyer Verlag. [188]

Watson, Keith. 1976. “The Shah ’ s White Revolution- Education and Reform in Iran.” *Comparative Education Review* 12 (1): 23–36. [191]

Weitzman, Abigail. 2017. “The effects of women’s education on maternal health: Evidence from Peru.” *Social Science and Medicine* 180: 1–9. [193]

Wilber, Donald N. 1981. *Iran, Past and Present: From Monarchy to Islamic Republic*. Princeton: Princeton University Press. [191]

Yeom, MinHo, Clementina Acedo, Erry Utomo, and MinHo Yeom. 2002. "The reform of secondary education in indonesia during the 1990s: Basic education expansion and quality improvement through curriculum decentralization." *Asia Pacific Education Review* 3 (1): 56–68. [190]

Chapter 3

Hours and income dynamics during the Covid-19 pandemic: The case of the Netherlands

Joint with Christian Zimpelmann, Hans-Martin von Gaudecker, Lena Janys, Bettina Siflinger

3.1 Introduction

Beginning in early 2020, the Covid-19 pandemic has strongly affected working lives around the world. A large number of studies have tracked the crisis' initial impact in the US and European countries on employment, hours worked, and income.[1] Along these dimensions, existing inequalities were generally exacerbated early in the crisis, although the degree varied widely across countries. The fact that inequalities went

* This a pre-print version of the study that is also published in Labour Economics, volume 73, 2021. The data collection was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2126/1 – 390838866, by the Dutch Research Council (NWO) under a Corona Fast track grant (440.20.043), and by the IZA – Institute of Labor Economics. Gaudecker and Zimpelmann are grateful for financial support by the German Research Foundation (DFG) through CRC-TR 224 (Project C01). This research would not have been possible without the help of many others at the CoViD-19 Impact Lab, a research group initiated in Bonn in Mid-March 2020. Special thanks to the team at CentERdata, who made the surveys underlying this research possible in record time. We would like to thank Egbert Jongen for very helpful comments.

1. Examples include Adams-Prassl, Boneva, Golin, and Rauh (2020), Alstadsæter, Bratsberg, Eielsen, Kopczuk, Markussen, et al. (2020), Béland, Brodeur, and Wright (2020), Bick and Blandin (2020), Brynjolfsson, Horton, Ozimek, Rock, Sharma, et al. (2020), Coibion, Gorodnichenko, and Weber (2020), Eurofound (2020), Farré, Fawaz, González, and Graves (2020), Meekes, Hassink, and Kalb (2020), von Gaudecker, Holler, Janys, Siflinger, and Zimpelmann (2020), and Crossley, Fisher, and Low (2021)

up is not surprising in light of the particularities of this pandemic-induced recession—e.g., social distancing behaviors, non-pharmaceutical interventions to reduce the virus' spread, or the huge increase in working from home. The first months of the pandemic were, however, also characterized by a substantial amount of uncertainty and by supply chain disruptions (e.g. Meier and Pinto, 2020). Neither is it well understood how employment, hours, and income developed throughout the first year of the pandemic; nor why variations across countries are so large.

We add to this understanding by providing an in-depth analysis of individual labor market trajectories throughout 2020 in the Netherlands, a stereotypical North-western European country along many core dimensions.[2] The Dutch government imposed a lockdown from March to May 2020, which was followed by re-opening most parts of the social and economic life over the summer. A second wave of the pandemic led to another lockdown in autumn and winter. Business closures were accompanied by labor hoarding schemes for the employed and various subsidies for the self-employed. Government restrictions and changes in consumer behavior directly affected firm demand; labor supply may be affected by fear of infection or childcare needs.

We make use of customized panel data collected for seven periods during the year 2020 in the LISS panel, a high-quality online survey based on a probability sample of the Dutch population. Doing so allows us to access a wealth of background characteristics from prior years in addition to contemporaneous measures of labor market outcomes and potential drivers thereof.

We document three stylized facts regarding the trends in employment, hours worked, and household income throughout the year 2020. First, the rates of unemployment and non-employment rose by 1.1 and 1.9 percentage points, respectively, between February and May. The unemployment rate slightly decreased thereafter while the rate of non-employment remained constant. Both of these patterns are consistent with administrative records, highlighting the quality of our data. The decrease in employment relationships is much smaller than in many other countries. For example, the US unemployment rate rose by 10 percentage points and labor force participation fell by 4 percentage points between February and April (Bick and Blandin, 2020) and in Canada employment fell by about 15 percentage points (Lemieux, Milligan, Schirle, and Skuterud, 2020).

Second, working hours declined strongly among those who were working just before the pandemic started to affect labor markets. Considering the extensive and intensive margin jointly, hours had dropped by 15 percent on average by April. They stayed roughly at this level for the rest of 2020—aggregate changes were within the realm of seasonal fluctuations. This pattern is very different when breaking down the evolution of working hours by socio-economic group, measured by education and

2. The Netherlands is fairly similar to countries such as Germany, Denmark, etc. in terms of the social safety net and labor protection laws; the reaction to the pandemic was broadly comparable.

personal income. Less educated or low-income individuals reduced working hours roughly twice as much as others. This socio-economic gradient becomes smaller during the summer when infection rates were low and social-distancing restrictions were more relaxed. Again, these facts are consistent with administrative microdata covering the first half of 2020 (Meekes, Hassink, and Kalb, 2020). The initial impact on aggregate working hours is only about half of what Lemieux et al. (2020) find for Canada, but the heterogeneity in the effect is comparable to their findings. During the second lockdown in December, the gradient becomes steeper again but stays below its spring levels. Throughout the year, the evolution of hours worked from home by socio-economic group tracks the differential evolution of total hours worked.

The third stylized fact is that the distribution of household income hardly changed throughout 2020. Relative to household income in the pre-pandemic months, the median of subsequent changes is zero. This is true across different socio-economic groups, whether these are measured by education, personal income, or long-run household income. Across these groups, the first and third quartiles of changes in household income are very similar and of limited magnitude. These patterns stand in contrast to the experiences of countries like the UK, where household earnings around the median decreased by 15 percentage points between February and May and poorer households were affected much stronger (Crossley, Fisher, and Low, 2021). Their earnings measure includes transfers made through the furlough-ing scheme; its dynamics should be similar for most parts of the income distribution to our comprehensive measure of income. Similarly, earnings decreased for almost 40 percent of the US population until April (Bick and Blandin, 2020) and vulnerable groups were hit much more strongly (Fazzari and Needler, 2021). Losses were, however, more than compensated by direct transfers from the unemployment insurance system which had a (temporary) replacement rate above pre-pandemic earnings for the lowest income groups (Cortes and Forsythe, 2020; Ganong, Noel, and Vavra, 2020). Other international comparisons are difficult to make due to different conventions of including transfer payments and different income measures or non-representative sampling. Overall, the picture that emerges is mixed, with some countries experiencing median income losses (e.g., Italy and Spain) and some countries having more stable household income dynamics (Germany, France, and Sweden, see, e.g., Bounie, Camara, Fize, Galbraith, Landais, et al., 2020; Clark, Ambrosio, and Lepinteur, 2021). Unlike our results, most countries surveyed in this literature experienced some form of heterogeneity in the income response, either by age or education (see, e.g., Belot, Choi, and Tripodi (2020) for China, Japan, Korea, Italy, UK, US; Osterrieder, Cuman, Pan-Ngum, Cheah, Cheah, et al. (2021) for Thailand, Malaysia, UK, Italy, Slovenia), but only some countries had regressive effects on the lowest income groups, such as France (Bounie et al., 2020)) and Italy (Belot, Choi, and Tripodi, 2020).

We then leverage our panel data and the tailor-made questionnaires to examine the drivers of these observed trends. During the initial lockdown, essential worker status and the fraction of work that can be done from home explain most of the socio-economic gradient in total hours worked.[3] The two characteristics interact strongly: telecommutability only plays a role for non-essential workers. In September—when infection rates were low and restrictions on social and economic life were few—these pandemic-specific mechanisms do not play a role and there hardly is a socio-economic gradient in hours worked. Their importance is large again in December, but weaker than in early spring. These patterns suggest that the best way to ameliorate the socio-economic gradient inherent in the pandemic's impact on labor markets is to keep infection rates low.

Finally, we relate changes in household income to employment transitions and hours changes using a set of quantile regressions. The median change for employees who remain employed throughout the year is very close to zero throughout. The first quartile of changes is between -7 and -13 percent, whereas the third quartile is between 13 and 17 percent. There is no relation with hours worked. By contrast, the first quartile of the distribution of household income innovations is a loss of about one quarter for the self-employed, for those who become unemployed, and for those who drop out of the labor force. The median is clearly negative for the three groups as well. For those who become unemployed, losses at the third quartile are still 14 percent.

Compared to other countries, separations to non-employment are very low in the Netherlands. The perfect insurance against changes in hours worked for employees that we just described is very rare. We thus run another set of quantile regressions of household income on employment transitions and whether employers' took up the wage subsidy scheme (NOW), which required to continue paying the full wage. Across quartiles, employer take-up of policies is unrelated to household income, suggesting that the combination of firing restrictions and large-scale support policies helped insure employees very well against the fallout of the crisis. The self-employed were hit much harder; the first quartile of those who benefited from any program targeting the self-employed saw their households' income drop by around 70%.

The next section describes the setting for our analysis and the data we collected. In Section 3.3, we distill the stylized facts on the evolution of employment, hours of work, and household income throughout the first year of the pandemic. We examine the drivers of the dynamics in working hours and household income in Section 3.4 before concluding in the last part.

3. Béland, Brodeur, and Wright (2020) show that early in the pandemic the ability to work from home and essential worker status mitigate labor market impacts in the US. We expand that analysis to a country where labor outcomes are mostly affected on the intensive margin and look at the relevance of these characteristics over different stages of the pandemic.

3.2 Context

The following section provides an overview of the development of the Covid-19 spread in the Netherlands and the social distancing policies. We moreover describe the key features of the Dutch labor market and economic support programs and present the data used in the empirical analysis.

3.2.1 Spread of Covid-19 and social distancing policies

Figure 3.2.1 displays the development of confirmed SARS-CoV-2 infections in the Netherlands on a logarithmic scale (left axis). By mid-March, when we collected our first wave of data, more than 10 new cases per million inhabitants were confirmed each day. This number reached 60 by the end of March and stayed roughly at that level for the first three weeks of April.[4] The incidence measure declined thereafter and reached 10 in mid-May, remaining at that level or somewhat below over the summer. In August, the infection numbers started rising again, reaching a temporary peak of 500 daily new cases per million inhabitants at the end of October. After falling below 300, confirmed infection numbers reached their 2020 peak at 700 new cases just before Christmas.[5]

Similar to other countries, the initial rise in infections prompted the Dutch government to impose restrictions on economic and social life to stop the spread of SARS-CoV-2. The Oxford Response Stringency Index measures the stringency of these policies (Hale, Petherick, Phillips, Webster, and Kira, 2020) and is shown in Figure 3.2.1 on the right axis. In mid-March, all schools and childcare facilities were closed along with restaurants, cafes, bars, and several other businesses involving personal contacts. People were advised to stay at home, to keep a distance of at least 1.5 meters to each other, and to avoid social contacts; the number of visitors at home was restricted to a maximum of three individuals. While most of the policy measures resembled those of other European countries, they did not involve a general curfew and some measures were more lenient. For instance, businesses such as stores for clothes, utilities, or coffee shops remained open as long as they could guarantee to maintain the social distancing rules. Public locations were accessible and traveling or the use of public transportation was possible throughout this lockdown period.

Beginning in May, the restrictions were gradually lifted. Daycare facilities and primary schools started opening in mid-May, businesses such as hairdressers and

4. The peak in daily cases was also between 60 and 70 in Germany, France, or the UK, although the plateau lasted shorter in Germany and France. It lasted much longer in the UK. During the March-April period, the peaks were substantially higher in Spain (160), Italy, and the US (both between 90 and 100).

5. These numbers include only confirmed cases. Since testing increased over time, the numbers are not directly comparable. The test positive rate peaked at 27% in late March but was about 5 % in September before increasing again to 16 % thereafter.

Figure 3.2.1. Daily new confirmed cases per million people and response stringency

Notes: The left axis (blue line) shows daily new cases as rolling 7-day average, based on (Roser, Ritchie, Ortiz-Ospina, Hasell, and Ritchie, 2020). The Oxford Response Stringency Index (right axis, orange line) measures the stringency of restrictions on economic and social life (Hale, Petherick, Phillips, Webster, and Kira, 2020). The vertical lines indicate the waves of data collection (see Section 3.2.3). They are located at our sample's median response dates for each wave: March 22, April 14, May 12, June 10, September 18, and December 17.

beauty salons were allowed to accept customers again. In early June, secondary schools started opening; restaurants, cafes, and cinemas could operate under restricted capacity. With the main exceptions of bans on larger (inside) gatherings, the requirement to wear masks in public transport, and the mandate to keep a distance of 1.5 meters to other people, social and economic life was largely back to what it was before.

In reaction to the increasing infection numbers during the fall, the Dutch government successively sharpened the restriction on September 30th, October 14th, and November 4th. The latter set of rules was similar to the one during the first lockdown in spring with the exception that schools were still open. Since the infection rate decreased in the first half of November, the Dutch government decided to lift the restrictions somewhat from November 18 but put an even stricter lockdown into place one month later. This implied that all sports locations, eating locations including room services in hotels, and shops, except supermarkets and essential services, had to close. Moreover, all schools switched to online teaching, and childcare facilities were closed.

3.2.2 Institutions and ad-hoc economic support measures

The Netherlands is a generic Western European welfare state. There is compulsory social insurance; unemployment insurance is obligatory for employees; and strong labor protection laws make firing employees without cause difficult for employers. To reduce the impact of the lockdown and behavioral reactions to the virus spread on the labor market, the Dutch government implemented several measures starting in mid-March 2020 for the period March to May. These programs were extended with minor adjustments and are in place until at least June 2021.

The first two emergency programs for the Dutch economy amount to about 30 billion Euros, which is about 3-4 percent of the Dutch GDP. The additional fiscal spending relative to GDP due to Covid-19 has been lower in the Netherlands than in other, larger economies such as Germany, UK, and the US; it has been similar to, for example, Sweden or Norway (IMF, 2021).

The most important policy measure targeting employees is the short-term allowance (Noodmaatregel Overbrugging voor Werkgelegenheid, NOW), which subsidizes labor hoarding. Internationally, job retention schemes can be classified into two different types (OECD, 2020): short term work schemes, as introduced in e.g. Germany, the UK or Japan, and wage subsidies as in e.g. Canada or Poland. NOW is classified as a hybrid scheme according to this definition, as employment subsidies were tied to employment guarantees. Under the NOW scheme, the Dutch government supports all businesses that expect a loss in gross revenues of at least 20% between March 2020 and July 2021 with advanced money for labor costs. The amount of advancement depends on the expected revenue loss. A business that expects a loss of 100% can request 90% of its labor costs from the government. The advancement is paid out at three points in time, with a first chunk being paid within 2-4 weeks after a positive decision on the request. Employers who get the advancement commit to paying full salaries to their employees and not fire employees due to reduced business activities. Only Denmark had a similar wage “top-up” requirement (OECD, 2020). Moreover, employers can revert dismissals that already have taken place. The advancement can also be requested for employees with fixed-term contracts or temporary workers. In contrast to labor hoarding arrangements in other countries, e.g. the UK or Germany, affected employees are not required to reduce working hours and their incomes remain the same by default.

The TOZO (Tijdelijke Overbruggingsregeling Zelfstandig Ondernemers, Temporary Bridging Measure for Self-employed Professionals) is the most relevant program for the self-employed. This income support measure was not means-tested in the first three months of existence. For the period June-December, a household-level income test was introduced. Another program for the self-employed is the TOGS (Tegemoetkoming Ondernemers Getroffen Sectoren Covid-19, Reimbursement for Entrepreneurs in Affected Sectors Covid-19), a one-time payment of 4000€ that is conditional on the sector being affected directly by the pandemic or pandemic-related measures between March and May. Further relief was provided through tax deferrals and loan guarantees for firms. We provide some more detail in Table 3.A.2 of the Online Appendix.

3.2.3 The LISS panel

To understand the behaviors and expectations of households during the different stages of the Covid-19 crisis, we designed a set of modules in the Longitudinal Internet Studies for the Social Sciences (LISS) panel. The LISS panel is based on a probability sample of individuals registered by Statistics Netherlands; it has been running since 2007 and consists of roughly 4,000 Dutch households comprising about 7,000 individuals. It is administered by CentERdata, a survey research institute affiliated with Tilburg University, the Netherlands, and has been used in several studies on individual and household behavior (e.g., Cherchye, De Rock, and Vermeulen, 2012; Noussair, Trautmann, and van de Kuilen, 2014; Cherchye, Demuynck, De Rock, and Vermeulen, 2017; Drerup, Enke, and von Gaudecker, 2017).

The first module of our questionnaire was fielded between March 20th and 31st 2020, a few days into the lockdown. Five more modules followed throughout April, May, June, September, and December. With roughly 80%, the response rate was at the top end of the span of usual response rates in the panel for all waves. Throughout this paper, we restrict our sample to respondents aged 18 to 66 years where the latter is the legal retirement age in the Netherlands in 2020. Whenever not stated otherwise, we furthermore restrict on all individuals working at least 10 hours before the pandemic. This leaves us with 17,314 observations over all waves. While the resulting panel is unbalanced, the distribution of demographic variables is very stable over time.[6]

Our questionnaires ask respondents about working hours at home and at the workplace during the last week. To assess the effect of the pandemic on labor supply in certain jobs, we elicit two job characteristics that are potentially important for labor supply during contact restrictions. First, we ask all subjects working before Covid-19 if their job qualifies as essential to the working of public life. Altogether, 35% of respondents work in an essential job. Second, in the May and December questionnaire, we ask about the fraction of usual work that can be done from home. In May, the question explicitly referred to the period before the pandemic. We find that the measure is very stable between May and December, both on the individual level and based on the aggregate distribution.[7] We, therefore, take the mean of the two elicitations. On average, 44% of all tasks can be done from home. The measure varies across the whole distribution; the first quartile is zero and the third quartile is

6. For brevity, we present descriptive statistics of our data in Section 3.B of the Online Appendix.

7. We would expect larger differences if we had also asked about telecommutability before the pandemic started. It is likely that many people only realized how much they could actually work from home in March/April.

90%.[8] Furthermore, we ask for household income every month during the pandemic. This allows us to examine how changes in working hours translate to the financial situation of households and how inequality is affected.

All questions are documented in von Gaudecker, Zimpelmann, Mendel, Siflinger, Janys, et al. (2021). Questionnaires of the LISS panel from 2019 and the first months of 2020 provide us with a rich set of additional background characteristics.

3.3 Work and income in 2020

To analyze the impact of the crisis on inequality within society, we document how changes in working hours and household income are related to the socio-economic status, measured by education, personal income, and household income.

3.3.1 Aggregate employment and working hours

While GDP contracted by 9.3% year-to-year in the second quarter of 2020, the non-employment rate and unemployment rate increased only slightly by roughly 1.1 and 1.9 percentage points each (more details in Section 3.C in the Online Appendix). The unemployment rate slightly decreased thereafter while the rate of non-employment stayed at this level.[9] These aggregate movements in the labor market are fairly similar to the movements experienced by countries such as Germany or the UK; they are less extreme than in Southern Europe or the US (see e.g. Anderton, Botelho, Consolo, Da Silva, Foroni, et al., 2020; Coibion, Gorodnichenko, and Weber, 2020; Crossley, Fisher, and Low, 2021).

To analyze the impact of the pandemic on the labor market, our main focus is on the dynamics of working hours. In a country like the Netherlands, with strong labor protection laws and comprehensive support policies implemented during the pandemic, focusing on job separations misses a large part of the effects of the crisis. As argued above, job separations were low even though aggregate output decreased substantially. To examine the extent and heterogeneity of productivity losses, it is, thus, vital to investigate the intensive margin, i.e. changes in working hours. Therefore, we analyze the dynamics of relative changes of unconditional working hours. This approach captures both the extensive (flow out of employment) and intensive margin of employment shocks.

8. The measure is with a correlation of 0.82 highly correlated between both points in time. For more information on the distribution and reliability of the measure, consult Appendix 3.B.3.

9. In official data by Statistics Netherlands, the level of un- and non-employment is somewhat lower, but the development over time overall lines up well with the numbers in our sample. We present a comparison to official data, visualizations of observed aggregate patterns, and robustness analyses of those patterns in Section 3.C in the Online Appendix. Robustness analyses include sample weights and an alternative before-Covid-19 measure that uses the time use and consumption survey conducted in November 2019.

From the workers' perspective, there are at least two reasons why reductions in working hours matter even if they do not loose their job. First, labor hoarding may not be sustainable in the medium term (the Dutch programs, for example, only ran for a few months and were renewed multiple times). A negative shock to working hours would then be an early indicator of future employment loss. This is certainly what respondents in our sample believe on average; working hours reductions are predictive of higher job loss expectations (Appendix 3.C.4). Second, working fewer hours might reduce the accumulation of human capital and delay future wage growth. This seems particularly plausible for recent job entrants.

The first row of Table 3.3.1 shows aggregate weekly unconditional working hours for each observed period. As we asked for the pre-Covid-19 working hours retrospectively, both, in March and April, the number of observations is higher for this period.[10] Working hours initially decreased by 4.3 hours or 12%. They bottomed out in May at a decrease of 7.7 weekly hours and rose thereafter by 2.5 hours until December. Based on the Dutch labor force survey (EBB), the drop in conditional working hours until April was 3 hours which is as expected slightly smaller than the changes in unconditional working hours in our sample (CBS, 2020). The EBB also shows that in the last years, working hours tended to be up to 3 hours larger in December than in May, June, and September. This might explain the increase in working hours despite increasing infections during the last wave of our data.

The most striking change in the labor market has been an unprecedented rise in the amount of work performed from home. Indeed, the second row of Table 3.3.1 shows a huge jump in March from 4 to over 15 hours until April. The share of hours worked from home increased from 11% to 50% in the aggregate. This fraction declined steadily to 31% in September before increasing again in December. The joint patterns of total hours and home office hours display the starting point of this paper: The pandemic led to both an increase in home office hours and a decrease in total working hours in March and April. The former quickly became much less important as infections dwindle and restrictions were lifted, while the overall amount of work stayed much lower than before the crisis.

3.3.2 Inequality in working hours and in working from home

Similar to studies for the US and UK, we find that the impact on hours is highly unequally distributed among socio-economic groups. The top row of Figure 3.3.1

10. A potential concern is that observed changes in working hours might be driven by the baseline being asked retrospectively. An alternative baseline measure is based on the time use and consumption survey that was in the field in November 2019. As participants are in this study also asked for their working hours in the last week, the elicitation method is closer to the one for our observations from March on. Appendix 3.C.2 shows that the distributions of both measures are closely aligned. Given that this alternative baseline was elicited longer before the pandemic and the joint sample is substantially lower, we rely on the retrospective measure from March/April 2020 for our analyses.

Table 3.3.1. Unconditional working hours over time

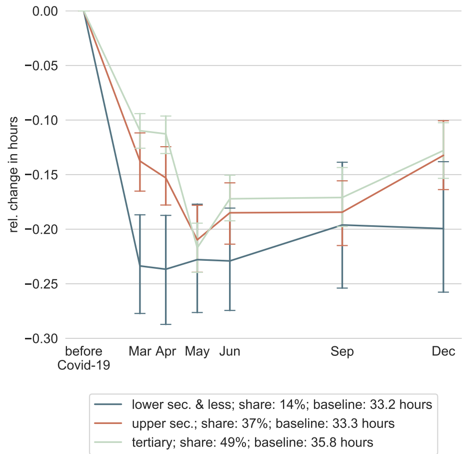
	before Covid-19	Mar	Apr	May	Jun	Sep	Dec
working hours	34.5	30.2	29.5	26.8	27.9	27.8	29.3
	(0.2)	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)
N	2962	2656	2634	2375	2518	2384	2298
hours worked from home	4.1	15.0	15.5	12.3	11.2	8.9	12.0
	(0.2)	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)
N	2962	2656	2634	2375	2518	2384	2298
share of hours worked from home	0.11	0.49	0.51	0.45	0.38	0.31	0.39
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
N	2962	2437	2408	2106	2317	2127	2052

Notes: The first two rows present unconditional total working hours and hours worked from home over time. All statistics are on respondents between ages 18 and 66 who worked for at least 10 hours in early March. The share of hours worked from home is only defined for individuals working in that period. Source: LISS.

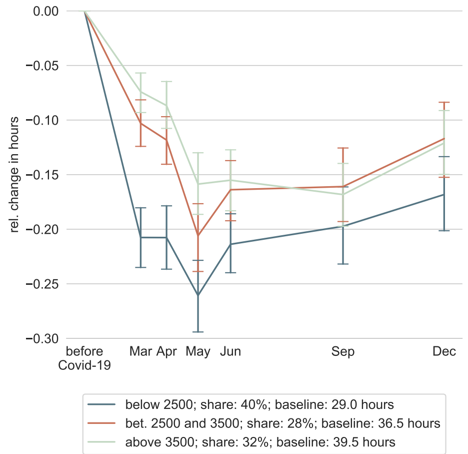
displays relative changes of total working hours, relative to early March 2020, by level of education (Figure 3.3.1a) and personal gross income (measured before the pandemic; Figure 3.3.1b). For individuals with lower secondary education or less, working hours fell by more than 22% on average in March and April. Better educated subjects reduced working hours significantly less: for those who completed tertiary education the reduction was just 11%. This difference becomes smaller in later months when restrictions were lifted before increasing again in December. Figure 3.3.1b shows that income is also predictive of changes in working hours: the group of individuals earning less than 2500 Euros reduced total working hours by more than 20% on average during March and April. This is roughly twice as much as individuals earning more. The difference to the highest-earning group decreases over time but is still roughly 3% in September and December.

The differences for hours worked from home by education (Figure 3.3.1c) are even stronger and more persistent over the full course of the pandemic. While the lowest educated group increased home office hours by less than 2.5 hours in all observed months, subjects with tertiary education did so by more than 15 hours during the first lockdown and still more than 7.5 hours in September. Figure 3.3.1d shows similar patterns for personal income: over the full course of the pandemic in 2020, better-earning individuals work consistently more from home although the level of working from home varies for all groups.

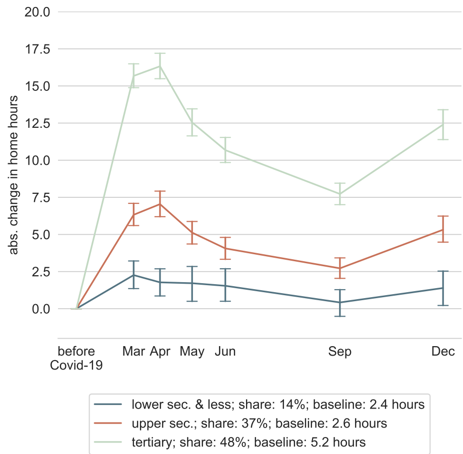
When splitting the sample by pre-crisis household income instead of personal income, the differential effects are substantially weaker indicating that personal characteristics are the main driver for the change in working hours (Figure 3.D.3 and Table 3.D.2 in the Online Appendix).



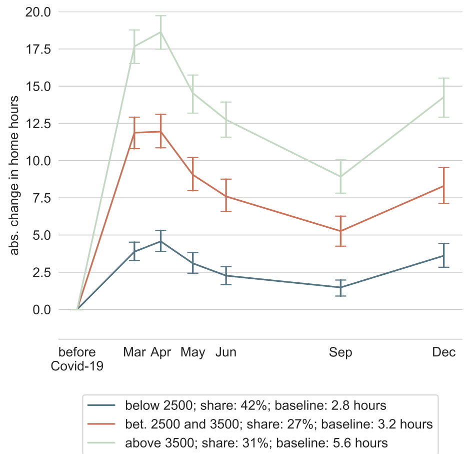
(a) Relative change in total working hours by education



(b) Relative change in total working hours by personal income



(c) Change in hours worked from home by education



(d) Change in hours worked from home by personal income

Figure 3.3.1. Mean changes in total working hours and hours worked from home, by socio-economic status

Notes: The top row shows mean relative changes in total hours worked by achieved education level (Figure 3.3.1a) and by personal gross income in three categories (Figure 3.3.1b). Figure 3.3.1c and Figure 3.3.1d display mean absolute changes in hours worked from home for the respective groups. Reference period is late February/early March. The legend displays hours and share of each group in early March. Vertical bars depict 95 %-confidence intervals. Sample: $18 \leq$ age ≤ 66; working hours of at least 10h in early March.

In summary, the impact of the pandemic on the amount and location of hours worked differed strongly by socio-economic status. More educated and better-paid individuals increased hours worked from home much more and decreased total working hours substantially less, the latter especially during the initial lockdown in March and April. We next examine whether these differences also translate into differences in household income during the pandemic.

3.3.3 Income Inequality

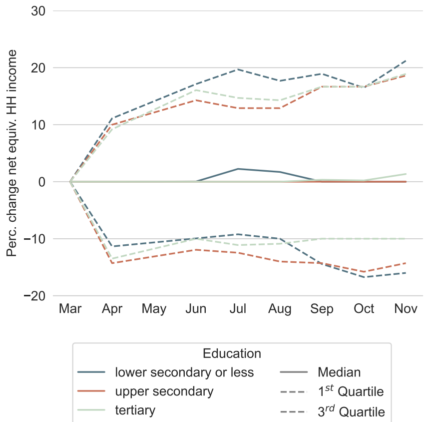
In April, June, September, and December, we asked individuals retrospectively about their household income in the previous months. Figure 3.3.2 depicts quantiles of changes in net equivalized household income relative to the average in January and February 2020, by socio-economic characteristics.[11] Median changes are close to zero in every month between March and November for all values of socio-economic variables that we condition on. Similar to our analysis of working hours, Figures 3.3.2a and 3.3.2b slice the data by education and individual gross income, respectively. Figure 3.3.2c conditions on pre-Covid household income—measured using LISS core questionnaires for the years 2018 and 2019—as a comprehensive measure of economic means. For all three measures of socio-economic status, the evolution of the first and the third quartile in changes is rather symmetric around zero. If anything, gains at the third quartile are slightly higher than losses at the first quartile. Again, there is no clear socio-economic gradient in any of the measures. Hence, we do not see an increase in income inequality in 2020 in the Netherlands. This is in stark contrast to, for example, the UK experience. Crossley, Fisher, and Low (2021) show that in May the earnings losses for the lowest quintile of the long-run income distribution were 60% at the first quartile and 13% at the median.[12] For the second-lowest quintile, the respective changes were −36% at the first quartile and −6% at the median.

3.4 Explanations and mechanisms

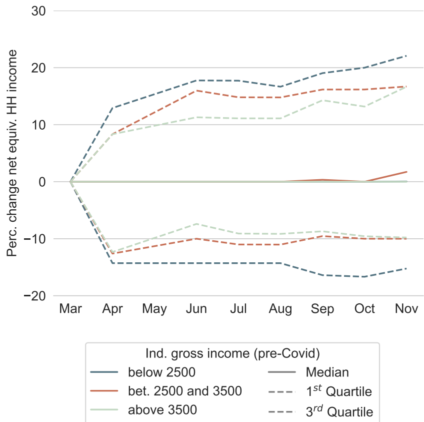
The previous section highlighted three important findings. First, the reduction in working hours is unequally distributed among socio-economic groups. Second, this seems to be particularly driven by an unequal substitution between working at the workplace and working from home. Third, despite the large and unequal decline in working hours, we do not observe a large and unequal decline in household income. In this section, we explore whether the dynamics in working hours are driven by pandemic-specific features. We then analyze the relation of working hour changes and changes in household income and examine why the socio-economic gradient for working hours changes does not carry over to household income.

11. We exclude the month of May because most employees receive a vacation payment mandated by law; the resulting jumps at all quantile make the graph very hard to read. See Figure 3.D.6 in the Online Appendix for the same graph as Figure 3.3.2 including the May data.

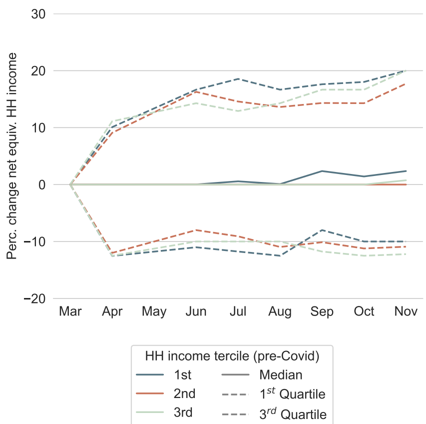
12. Earnings are defined as take-home pay and will thus include transfers made under the Job Retention Scheme via the employer.



(a) Relative changes in net equivalized household income by education



(b) Relative changes in net equivalized household income by pre-Covid individual gross income



(c) Relative changes in net equivalized household income by pre-Covid household income

Figure 3.3.2. Relative changes in net equivalized household income by socio-economic status

Notes: Relative change of net equivalized household income relative to the average of January and February 2020. Pre-Covid household income tercile calculated by using the terciles of the average household income of 2018 and 2019. Sample: $18 \leq \text{age} \leq 66$, working pre-Covid, report positive household income in either January or February (this excludes 170 individuals). We leave out May because the vacation bonus renders the graphs difficult to read; see Figure 3.D.6 in the Online Appendix for the same figure including the May numbers.

3.4.1 Working hours

Two job characteristics stand out that are potentially highly relevant during restrictions of economic activity: First, the ability to work from home. Doing so is the most natural way to continue working while keeping a distance from people outside the own household. Second, essential workers were exempted from most restrictions imposed on work lives. Table 3.4.1 shows the distribution of these job characteristics over socio-economic groups. The definition of essential workers was rather wide in the Netherlands and 35% of our sample state they are covered by this definition. This share does not vary strongly with the level of education but is negatively related to income: 40% of individuals earning less than 2500 Euros work in essential occupations while this is the case for only 27% of individuals earning more than 3500 Euros. By contrast, the ability to work from home is strongly positively related to both education and income. In the lowest education category, only 17 % of work can potentially be done from home, while this share is more than three times higher for individuals with tertiary education. These relations suggest that the strong gradient in realized home office hours described in the last section might be reflected in differing potentials to do so.

Table 3.4.1. Job characteristics by socio-economic status

	essential worker	frac. work doable from home
education: lower secondary and lower	0.37	0.17
education: upper secondary	0.40	0.31
education: tertiary	0.32	0.61
gross income: below 2500	0.41	0.29
gross income: bet. 2500 and 3500	0.39	0.45
gross income: above 3500	0.28	0.63

Notes: The table shows for different subsamples by socio-economic status (left side) the share of the sample that is an essential worker, and the average share of work that can be done from home. Sample: $18 \leq \text{age} \leq 66$; working hours of at least 10h in early March.

We next investigate whether pandemic-related job characteristics can explain the observed trajectory of aggregate working hours and especially the socio-economic gradient. We regress relative changes of working hours on socio-economic variables, essential worker status, telecommutability, and interaction of these two job characteristics. All regressions control for gender, work status before the pandemic (full-time employed, part-time employed, self-employed), and age. For conciseness, Table 3.4.2 focusses on the channels of particular interest and pools observations for the months March-June 2020 and for September/December 2020, respectively. Table 3.D.3 shows the full set of coefficients in a specification with disaggregated time effects and interactions.

Column 1 of Table 3.4.2 is the multivariate version of our analysis in the previous section. Not controlling for essential worker status and telecommutability, better educated and high-income individuals reduce their working hours less throughout. The disaggregated analysis (Table 3.D.3) shows that this relation is most pronounced in March/April and December, when the strongest restrictions were in place.

Column 2 of Table 3.4.2 adds job characteristics. Conditional on not being able to perform any tasks from home, essential workers' unconditional working hours are 13 percentage points higher than that of similar non-essential workers during the first four months of the pandemic. In September and December, the difference is even smaller and no longer statistically significant. For non-essential workers, moving the degree of telecommutability from zero to one increases average hours by 18 percentage points between March and June. The effect almost disappears during the second half of the year, where the specification hides the fact that it increases to 9 percentage points during the December lockdown (see Table 3.D.3). Importantly, the coefficient on the interaction of these two job characteristics implies that there is no effect of telecommutability for essential workers. Controlling for sector by month fixed effects in Column 3 does not change any of these coefficients in a meaningful way. Any potential spillover effects within sectors thus seem to be limited.

Table 3.4.2. Hours worked by individual and job characteristics

	change total working hours		
	(1)	(2)	(3)
march-june × education: upper sec.	0.05***	0.03*	0.03
	(0.02)	(0.02)	(0.02)
september/december × education: upper sec.	0.04	0.04	0.03
	(0.03)	(0.03)	(0.03)
march-june × education: tertiary	0.05***	0.01	0.01
	(0.02)	(0.02)	(0.02)
september/december × education: tertiary	0.06*	0.06*	0.06*
	(0.03)	(0.03)	(0.03)
march-june × income bet. 2500 and 3500	0.06***	0.05***	0.03**
	(0.02)	(0.02)	(0.02)
september/december × income bet. 2500 and 3500	0.05**	0.05**	0.03
	(0.02)	(0.02)	(0.02)
march-june × income above 3500	0.09***	0.06***	0.05***
	(0.02)	(0.02)	(0.02)
september/december × income above 3500	0.03	0.03	0.01
	(0.02)	(0.03)	(0.03)
march-june × essential worker		0.13***	0.12***
		(0.02)	(0.02)
september/december × essential worker		0.03	0.04
		(0.03)	(0.03)
march-june × frac. work doable from home		0.18***	0.18***
		(0.02)	(0.02)
september/december × frac. work doable from home		0.02	0.03
		(0.03)	(0.03)
march-june × essential × work doable from home		-0.16***	-0.15***
		(0.03)	(0.03)
september/december × essential × work doable from home		-0.07	-0.08*
		(0.04)	(0.04)
N	15738	15738	15133
R^2	0.151	0.163	0.168
demographic controls	Yes	Yes	Yes
month × sector FE	No	No	Yes

Notes: The table shows OLS regressions of relative changes in total (unconditional) working hours. Reference period = Early March. Further elements of the specifications include a full set of time dummies, gender, and pre-pandemic measures of part-time work and self-employment (all interacted with time dummies). Table 3.D.3 shows the full set of coefficients in a specification with disaggregated time effects and interactions. Standard errors are clustered on the individual level. The data are restricted to individuals who worked at least ten hours in early March. Notes: *p<0.1; **p<0.05; ***p<0.01.

Unsurprisingly, the relation of working hours reductions with socio-economic variables becomes somewhat weaker once we add essential worker status and telecommutability to the regression (Column 2). This indicates that the heterogeneous effects by income and education can be explained to some degree by pandemic-specific job characteristics. Low-educated individuals seem to reduce working hours more strongly due to their lacking ability to work from home in their current jobs. At the same time, however, the results show that for given job characteristics, higher-earning individuals were more successful in conserving their working hours. One explanation could be that they might have been better able to realize the potential to work from home while employees earning less might more often lack the technical support to do so. Furthermore, pre-pandemic earnings might proxy the robustness of firms towards the Covid-19 shock – especially for self-employed individuals.

In terms of other control variables, females see an extra loss of 4 to 6 percentage points in all months except June and September. These differences cannot be explained by job characteristics. We explore the gendered patterns of employment shocks and childcare in a separate paper, where we also discuss the nature of part-time work in greater detail (Holler, Janys, Zimpelmann, von Gaudecker, and Siflinger, 2021). The self-employed are hit very hard initially and see an additional average loss of 13 percentage points during the lockdown period compared to full-time employees. The difference in hours reductions falls to 3 percentage points and is no longer statistically significant in June. This pattern is consistent with many small businesses operating in industries that are hit particularly hard by the restrictions—bars and restaurants, hairdressers, etc.—as well as firms providing insurance to their employees (Guiso, Pistaferri, and Schivardi, 2005), potentially with the help of the government. Sectoral differences are large during the lockdown but become smaller in later months. All this is consistent with the broad line of our overall results, i.e., the specific features of a pandemic recession becoming less important in the months following the first lockdown.

A potential concern with our data is that pre-pandemic working hours are asked retrospectively for a few weeks earlier while working hours in all other periods are asked for the last week. We, therefore, make two robustness checks: First, we exclude subjects that took a day off out of turn, e.g. because of official holidays, vacation, or being sick. Second, we use the time use survey of November 2019, which also asked for working hours during the last week, as the reference period. Our results do not change substantially (Table 3.D.1 in the Online Appendix).

3.4.2 Household Income

To analyze why the relationship between employment shocks and socio-economic status does not translate into a similar gradient for changes in net equivalized household income, we regress relative changes in household income on relative changes in working hours and time fixed effects. We use quantile regressions and report results for the three inner quartiles. Compared to OLS regressions, quantile regressions allow us to study effects on household income at several points of the distribution. Furthermore, they are less affected by outliers. To distinguish between the extensive margin (movements out of employment) and the intensive margin (changes in working hours among employed and self-employed), we create multiple mutually exclusive indicator variables. In each period, an individual can either be employed, self-employed, unemployed, or out of the labor force (retired, student, homemaker, receiving social assistance). We consider the employed and self-employed separately if they kept their job. Conversely, we use groups for those who became unemployed and those who dropped out of the labor force, irrespective of whether they were employed or self-employed before the pandemic. If an individual was employed pre-Covid, she is classified as *employed (pre-Covid) ⇒ employed* if she is employed in the respective period; as *empl or self-empl (pre-Covid) ⇒ unemployed* if she is unemployed in the period; as *empl or self-empl (pre-Covid) ⇒ out of labor force* if she dropped out of the labor force. The definition for initially self-employed individuals is equivalent.[13] We leave out March because the working hours information refers to late March only, which will not be representative of the entire month.

The results are displayed in the first three columns of Table 3.4.3. The time dummies refer to individuals who remain in employment; for all three quartiles, they are very close to the unconditional quantiles in Figure 3.3.2 in April, but considerably narrower thereafter. Interestingly, changes in working hours do not affect the employed as is evident from the fifth row. Changes in hours refer to working hours in the respective month relative to working hours in late February/early March. All three coefficients are zero and precisely estimated. Unsurprisingly, the lower tail looks much worse for the self-employed, where the evolution of the first quartile implies an additional loss of 25% of pre-Covid household income relative to those who remain employed. At the median, the additional drop is 7%; it is smaller and insignificant for the third quartile. The point estimates for hours changes go in the opposite direction as the expected co-movement of hours and income, but these are estimated very imprecisely. The last two rows show that the magnitudes of changes in household income of individuals who transitioned from working to not working are similar to the self-employed who remain so. For those who become unemployed, point estimates are larger at the median and the third quartile. The effects of extensive margin adjustments on household income are likely similar to changes in household income of those who remain in self-employment because transitions out of work are more frequent for part-time workers. This leaves many households where one partner worked part-time the primary earner's income. Similarly, high

13. We drop respondents who transition from employment to self-employment and from self-employment to employment because of the small group size (maximized at 28 individuals in September).

Table 3.4.3. Relationship between labor market outcomes, support policies, and household income

	Dependent variable: Rel. change in net equ. HH income					
	Hours worked			Support policies		
	p25	p50	p75	p25	p50	p75
April	-12.5***	0.00	13***	-10***	0.41	13.28***
	(1.01)	(0.01)	(1.03)	(1.37)	(0.54)	(1.45)
May	-4.05***	7.14***	44.44***	-2.17	7.31***	44.87***
	(1.23)	(0.95)	(2.13)	(1.33)	(1.05)	(2.32)
June	-7.41***	0.09	15.79***	-6.25***	0.41	15.89***
	(1.04)	(0.48)	(1.03)	(1.08)	(0.63)	(1.14)
September	-8.56***	1.35*	16.76***	-7.94***	1.54**	16.73***
	(0.97)	(0.71)	(1.32)	(1.16)	(0.73)	(1.38)
rel. change in work. hours × employed (pre-Covid) ⇒ employed	0.07	0.00	-0.01			
	(0.58)	(0.21)	(1.67)			
Policy: Yes × employed (pre-Covid) ⇒ employed				0.16	-0.41	-2.16
				(1.54)	(0.59)	(2.43)
Policy: I don't know × employed (pre-Covid) ⇒ employed				-4.58***	-0.41	1.13
				(1.59)	(0.58)	(2.05)
self-empl (pre-Covid)⇒ self-empl	-25.82***	-7.14**	-3.2	-19.76***	-5.92**	-3.05
	(3.34)	(3.17)	(4.81)	(3.11)	(2.76)	(4.33)
rel. change in work. hours × self-empl (pre-Covid) ⇒ self-empl	-2.06	-2.94	-4.15			
	(3.16)	(15.35)	(13.23)			
Policy: Yes × self-empl (pre-Covid) ⇒ self-empl				-51.49***	-10.48	4.06
				(14.87)	(9.05)	(11.07)
empl or self-empl (pre-Covid) ⇒ unemployed	-26.52***	-16.04***	-14.44**	-29.08***	-19.28***	-14.73**
	(7.14)	(5.81)	(6.92)	(7.79)	(5.55)	(6)
empl or self-empl (pre-Covid) ⇒ out of labor force	-24.77***	-7.14**	-4.76	-25.1***	-7.31***	-4.73
	(4.87)	(2.82)	(6.37)	(4.59)	(2.06)	(5.74)
N		8,595			8,564	

Notes: Quantile regressions with relative changes in net equalized household income (relative to the average of January and February 2020) as the dependent variable. Standard errors are clustered on the household level using the wild bootstrap procedure proposed by Hagemann (2017) and implemented in the R package quantreg. Sample: $18 \leq \text{age} \leq 66$; employed or self-employed while working at least 10 hours pre-Covid (early March); positive household income either in January or February 2020 (this excludes 170 individuals). Reference group: employed (pre-Covid) ⇒ employed. Policy: Yes = respondent's employer/respondent applied for policy support and was not rejected; "I don't know" = respondent does not know whether employer applied for support policies. For employed only the NOW policy was considered. For self-employed, all potential policies were considered.

replacement rates from unemployment insurance or pensions will often be higher for part-time workers with relatively low incomes.

In the second set of columns of Table 3.4.3, we replace changes in working hours with an indicator of whether individuals received any policy in case they continue to work. For individuals who become unemployed or drop out of the labor force, we do not make a distinction whether they benefitted from any policy before.[14] Unsurprisingly, their coefficients look very similar to those in columns 1-3; so do the coefficients on the time dummies. The most interesting results are those for the employed, where we only consider the NOW (labor hoarding) program. There are no significant differences in the innovations to household income conditional on policy receipt or not, except for a small drop at the first quartile for individuals who do not know whether their employer applied for the NOW. Although we lack a precise counterfactual for what would have happened in absence of this policy, the experience in other countries suggests that incomes would likely have dropped with hours reductions for employees.[15] For the self-employed, we see much larger reductions in household income if they made use of any support policy. This is an indicator that the programs seem reasonably well-targeted. Altogether, the results from the regressions including support policies suggest that the NOW achieved its goal of near-perfect insurance against changes along the intensive margin for employees. Given the low numbers of separations into non-work relative to many other countries, they are likely to have helped in limiting these transitions, too.

3.5 Conclusion

This study has analyzed how the Covid-19 pandemic affected the Dutch labor market over the entire year 2020. Compared to countries like the US (Bick and Blandin, 2020), much fewer job separations occurred, but working hours were substantially affected. We show that subjects with lower socio-economic status faced the strongest decreases in working hours. At the same time, their hours worked from home increased only slightly. This heterogeneous effect did not translate to a socio-economic gradient in household income changes.

Examining the drivers of these patterns, we find that pandemic-specific job characteristics (telecommutability and essential worker status) are highly predictive of working hours changes while social distancing restrictions are in place. We stress the interaction of those two job characteristics: home office capability only mattered for changes in working hours of non-essential workers. When case numbers are low and economic restrictions are widely abolished, these job characteristics

14. Remember from Section 3.2.2 that in total, both rows contain less than 3% of individuals at any point in time.

15. Figure 3.D.5 in the Online Appendix shows that policy take-up was strongly related to reductions in working hours for both employees and the self-employed.

hardly influence hours worked. As a consequence, the socio-economic gradient in employment outcomes was low during the summer albeit working hours were still substantially lower than before the pandemic.

Household income did not decrease in the medium term and was decoupled from employment shocks for individuals who remained employed. This stands in stark contrast to the UK, where the pandemic led to a large negative shock on earnings inclusive of transfers made through the Job Retention Scheme (Crossley, Fisher, and Low, 2021). The finding is also very different from the impact of the Great Recession in the Netherlands. Income declined by 13% in 2009 while movements out of employment were similar (van den Berge, Erken, de Graaf-Zijl, and Van Loon, 2014). It seems likely that the government support programs are responsible for these differences: the NOW program not only aims at job retention but also at full wage insurance for workers. This was not the case for the job retention scheme during the Great Recession in the Netherlands (Hijzen and Venn, 2011). Our explanation is supported by the finding that the take-up of NOW is unrelated to changes in household income. Thus, we provide suggestive evidence that inducing full wage stability through job retention schemes might counteract medium-term regressivities in income better than other work retention schemes. Household income of self-employed subjects was hit particularly hard and could only be partly cushioned by support policies. This likely reflects the fact that it is much harder to design incentive-compatible support measures for the self-employed. It thus is crucial to continue supporting the self-employed during the pandemic and help them to get back to business when infection numbers allow it.

Future research may shed more light on the effects of support policies by comparing household income dynamics to institutionally more similar countries with different job retention schemes not targeting full wages such as Germany. We are not aware of any study that analyzes household income dynamics in 2020 in any other Northwestern European country.

References

Adams-Prassl, Abi, Teodora Boneva, Marta Golin, and Christopher Rauh. 2020. “Inequality in the Impact of the Coronavirus Shock: Evidence from Real Time Surveys.” *Journal of Public Economics*, (7): [205]

Alstadsæter, Annette, Bernt Bratsberg, Gaute Eielsen, Wojciech Kopczuk, Simen Markussen, Oddbjorn Raaum, and Knut Røed. 2020. “The First Weeks of the Coronavirus Crisis: Who Got Hit, When and Why? Evidence from Norway.” Working Paper. [205]

Anderton, Robert, Vasco Botelho, Agostino Consolo, António Dias Da Silva, Claudia Foroni, Matthias Mohr, and Lara Vivian. 2020. “The Impact of the COVID-19 Pandemic on the Euro Area Labour Market.” *Economic Bulletin Articles* 8: [213]

Béland, Louis-Philippe, Abel Brodeur, and Taylor Wright. 2020. “Covid-19, Stay-at-Home Orders and Employment: Evidence from CPS Data.” Working Paper. [205, 208]

Belot, Michèle, Syngjoo Choi, and Egon Tripodi. 2020. "Unequal Consequences of COVID-19 across Age and Income: Representative Evidence from Six Countries," 23. [207]

Bick, Alexander, and Adam Blandin. 2020. "Real-Time Labor Market Estimates During the 2020 Coronavirus Outbreak." Working Paper. [205–207, 225]

Bounie, David, Youssouf Camara, Etienne Fize, John Galbraith, Camille Landais, Chloe Lavest, Tatiana Pazem, and Baptiste Savatier. 2020. "Consumption Dynamics in the Covid Crisis: Real Time Insights from French Transaction Bank Data." *Covid Economics* 59: 1–39. [207]

Brynjolfsson, Erik, John J Horton, Adam Ozimek, Daniel Rock, Garima Sharma, and Hong-Yi TuYe. 2020. "COVID-19 and Remote Work: An Early Look at US Data." Working Paper. [205]

CBS. 2020. "Maandcijfers gewerkte uren EBB." nl-NL. https://www.cbs.nl/nl-nl/maatwerk/2020/23/maandcijfers-gewerkte-uren-ebb. [214]

Cherchye, Laurens, Bram De Rock, and Frederic Vermeulen. 2012. "Married with Children: A Collective Labor Supply Model with Detailed Time Use and Intrahousehold Expenditure Information." *American Economic Review* 102 (7): 3377–405. [212]

Cherchye, Laurens, Thomas Demuynck, Bram De Rock, and Frederic Vermeulen. 2017. "Household Consumption When the Marriage Is Stable." *American Economic Review* 107 (6): 1507–34. [212]

Clark, Andrew, Conchita Ambrosio, and Anthony Lepinteur. 2021. "The Fall in Income Inequality during COVID-19 in Five European Countries." Working Paper. [207]

Coibion, Olivier, Yuriy Gorodnichenko, and Michael Weber. 2020. "Labor Markets During the COVID-19 Crisis: A Preliminary View." Working Paper. [205, 213]

Cortes, Guido Matias, and Eliza Forsythe. 2020. "Impacts of the Covid-19 Pandemic and the CARES Act on Earnings and Inequality." Working Paper. [207]

Crossley, Thomas F., Paul Fisher, and Hamish Low. 2021. "The Heterogeneous and Regressive Consequences of COVID-19: Evidence from High Quality Panel Data." *Journal of Public Economics* 193 (1): 104334. [205, 207, 213, 217, 226]

Drerup, Tilman, Benjamin Enke, and Hans-Martin von Gaudecker. 2017. "The Precision of Subjective Data and the Explanatory Power of Economic Models." *Journal of Econometrics* 200 (2): 378–89. [212]

Eurofound. 2020. "Living, Working and COVID-19 - First Findings – April 2020." Working Paper. [205]

Farré, Lídia, Yarine Fawaz, Libertad González, and Jennifer Graves. 2020. "How the COVID-19 Lockdown Affected Gender Inequality in Paid and Unpaid Work in Spain." Working Paper. [205]

Fazzari, Steven, and Ella Needler. 2021. "US Employment Inequality in the Great Recession and the COVID-19 Pandemic." Working Paper. [207]

Ganong, Peter, Pascal Noel, and Joseph Vavra. 2020. "US Unemployment Insurance Replacement Rates during the Pandemic." *Journal of Public Economics* 191 (11): 104273. [207]

Guiso, Luigi, Luigi Pistaferri, and Fabiano Schivardi. 2005. "Insurance within the Firm." *Journal of Political Economy* 113 (5): 1054–87. [222]

Hagemann, Andreas. 2017. "Cluster-Robust Bootstrap Inference in Quantile Regression Models." *Journal of the American Statistical Association* 112 (517): 446–56. [224]

Hale, Thomas, Anna Petherick, Toby Phillips, Samuel Webster, and Beatriz Kira. 2020. "Variation in Government Responses to COVID-19." Working Paper. [209, 210]

Hijzen, Alexander, and Danielle Venn. 2011. "The Role of Short-Time Work Schemes during the 2008-09 Recession." Working Paper. [226]

Holler, Radost, Lena Janys, Christian Zimpelmann, Hans-Martin von Gaudecker, and Bettina Siflinger. 2021. "The Early Impact of the Covid-19 Pandemic on the Gender Division of Market and Household Work." Unpublished Manuscript. [222]

IMF, Fiscal Affairs Department. 2021. "Fiscal Monitor Database of Country Fiscal Measures in Response to the COVID-19 Pandemic." https://www.imf.org/en/Topics/imf-and-covid19/Fiscal-Policies-Database-in-Response-to-COVID-19. [211]

Lemieux, Thomas, Kevin Milligan, Tammy Schirle, and Mikal Skuterud. 2020. "Initial Impacts of the COVID-19 Pandemic on the Canadian Labour Market." *Canadian Public Policy* 46 (S1): S55–S65. [206, 207]

Meekes, Jordy, Wolter Hassink, and Guyonne Kalb. 2020. "Essential Work and Emergency Childcare: Identifying Gender Differences in COVID-19 Effects on Labour Demand and Supply." Working Paper. [205, 207]

Meier, Matthias, and Eugenio Pinto. 2020. "Covid-19 Supply Chain Disruptions." Working Paper. [206]

Noussair, Charles N., Stefan T. Trautmann, and Gijs van de Kuilen. 2014. "Higher Order Risk Attitudes, Demographics, and Financial Decisions." *Review of Economic Studies* 81 (1): 325–55. [212]

OECD. 2020. *Job Retention Schemes during the COVID-19 Lockdown and Beyond*. Paris: OECD Publishing. [211]

Osterrieder, Anne, Giulia Cuman, Wirichada Pan-Ngum, Phaik Kin Cheah, Phee-Kheng Cheah, Pimnara Peerawaranun, Margherita Silan, Miha Orazem, Ksenija Perkovic, Urh Groselj, Mira Leonie Schneiders, Tassawan Poomchaichote, Naomi Waithira, Supa-at Asarath, Bhensri Naemiratch, Supanat Ruangkajorn, Lenart Skof, Natinee Kulpijit, Constance R S Mackworth-Young, Darlene Ongkili, Rita Chanviriyavuth, Mavuto Mukaka, and Phaik Yeong Cheah. 2021. "Economic and Social Impacts of COVID-19 and Public Health Measures: Results from an Anonymous Online Survey in Thailand, Malaysia, the UK, Italy and Slovenia." *BMJ Open* 11 (7): [207]

Roser, Max, Hannah Ritchie, Esteban Ortiz-Ospina, Joe Hasell, and Hannah Ritchie. 2020. "Our World in Data: Coronavirus Pandemic (COVID-19)." https://ourworldindata.org/coronavirus. [210]

van den Berge, Wiljan, Hugo Erken, Marloes de Graaf-Zijl, and Eric Van Loon. 2014. "CPB Background Document." In *The Dutch Labour Market during the Great Recession*. Den Haag: CPB Netherlands Bureau for Economic Policy Analysis. [226]

von Gaudecker, Hans-Martin, Radost Holler, Lena Janys, Bettina M. Siflinger, and Christian Zimpelmann. 2020. "Labour Supply in the Early Stages of the COVID-19 Pandemic: Empirical Evidence on Hours, Home Office, and Expectations." eng. Working Paper. [205]

von Gaudecker, Hans-Martin, Christian Zimpelmann, Moritz Mendel, Bettina Siflinger, Lena Janys, Jürgen Maurer, Egbert Jongen, Radost Holler, Renata Abikeyeva, Felipe Augusto Azuero Mutis, Annica Gehlen, and Eva Lucia Kleifgen. 2021. "CoVID-19 Impact Lab Questionnaire Documentation." eng. https://doi.org/10.5281/zenodo.4338730. [213]

Appendix 3.A Context

3.A.1 Policies

Table 3.A.1. Overview government support program to fight the Corona crisis

program & period	type	eligilibility & content	target group
noodpakket 1.0 March-May 2020	NOW 1.0	• company with at least 20% expected loss in gross revenues relative to actual loss in gross revenue for a 3-month period can request up to 90% of labor costs; maximum labor cost compensation/employee is set to €9,538 which is 2x the maximum “dagloon” (fiscal number to determine social security benefits) • obligation: employer pays 100% of wages to employees; no lay offs for business related reasons • consequence lay offs: fine of 50% of requested subsidy, thus 150% of subsidy has to be paid back • applies to employees with permanent and fixed term contracts • number of working hours is set by an agreement between employer and employee • advance money: 80% of requested subsidy; actual loss in gross revenues is evaluated afterwards and corrected retrospectively (employer either has to pay back or receives additional subsidies); large requests require auditor’s report • reference period: expected gross revenues are compared to revenue from January-December 2019 divided by four (different for companies not existing on Jan 1, 2019). • a compensation of labor costs of 30% has been chosen for all cases (not sure here)	all companies
	TOZO 1.0	• income support program for self-employed; lump sum payments up to social minimum (see https://www.uwv.nl/particulieren/bedragen/detail/sociaal-minimum) • eligibile: businesses founded before March 17, 2020; business was founded before January 1, 2019: minimum number of hours worked is 1,225 hrs/a; founded after January 1, 2019: at least 23.5 hours/wk • TOZO 1.0: income of partner was not taken into account • self-employed can request loan on business capital (berijfskapitaal); maximum loan: €10,517 at reduced interest rate to solve liquidity problems	self employed
	TOGS	• direct lump sum payment of €4,000 to employers particularly affected by the social distancing regulations to fight the Corona crisis	self-employed directly affected by social distancing regulations

Table 3.A.2. Overview government support program to fight the Corona crisis, cont.

program & period	type	eligilibility & content	target group
noodpakket 2.0 June-September 2020	NOW 2.0	• very similar to NOW 1.0, few main differences – expected loss in gross revenues for 4 months; reference period for calculation: March 2020 – compensation for labor costs increases from 30% to 40% – fine for lay offs due to business related reasons is abolished; subsidy is reduced by 5% if companies with 20 and more employees does not request lay off of employees in time (law WMCO) during subsidy period – employer encourages employee to participate in on-the-job-training programs (extra budget) – no pay out of bonuses to management or profits to shareholders, buy back own shares	all companies
	TOZO 2.0	• similar to TOZO 1.0 • main difference: partner income is also taken into account; amount of income support based on social minimum is now calculated on household income rather than individual income	self employed
	TVL (replaces TOGS)	• Compensation for fixed costs from €1,000 up to €50,000 if loss in gross revenues is more than 30%; minimumn fixed costs: €4,000 • maximum of fixed costs subsidized is 50%; Minimum subsidy per company: €1,000; maximum subsidy: €50,000 • compensation period: 4 months	applies to micro, small, medium sized companies (MKB). Medium sized companies have less than 250 employees, less than €50 Mio gross annual revenues, a maximum of €43 Mio annual balance

Appendix 3.B Data

In this part of the appendix, we describe and examine additional aspects of our data and the variables we use.

3.B.1 Descriptive Statistics

The first row of Panel A of Table 3.B.1 shows that just over half of our sample is female. Thirteen percent left school with a primary or lower secondary degree (bo/vmbo), 37% have completed upper secondary education (havo/vwo/mbo), just under one half of the workforce has some form of tertiary education (wo/hbo). Before the Covid-19 crisis started, just over a quarter of the sample were employed part-time, defined as working no more than 30 hours per week; 62% were in full-time employment while one in ten individuals was self-employed. Individuals' gross monthly income before the crisis was 3,710€ on average; median income is at 2,870€. We also make use of long-run household income which allows us to examine the impact on inequality. It is measured as the average monthly net household income in 2018 and 2019 and equivalized by the number of household members.

Table 3.B.1. Descriptive statistics main sample

	N	mean	std. dev.	$q_{0.25}$	$q_{0.5}$	$q_{0.75}$
female	2962	0.52				
age	2962	44.24	12.33	34	45	55
education: lower sec. and below	2962	0.14				
education: upper secondary	2962	0.37				
education: tertiary	2962	0.49				
net hh income 18/19 (equiv)	2468	2.39	3.38	1.67	2.18	2.82
full time employed pre-Covid	2962	0.62				
part time employed pre-Covid	2962	0.28				
self-employed pre-Covid	2962	0.10				
gross income	2781	3.71	31.53	1.94	2.87	3.91
essential worker	2962	0.35				
frac. work doable from home	2634	0.44	0.41	0	0.38	0.9
affected by policy: yes	2962	0.16				
affected by policy: no	2962	0.33				
affected by policy: don't know	2962	0.26				

Notes: Source LISS. Household income in thousands. All statistics are on respondents between ages 18 and 66 who worked for at least 10 hours in at least one of the 6 periods.

In the questionnaires of May and September, we asked all subjects that were employed or self-employed, for which support policies their employer or they themselves – if they were self-employed – applied and were not rejected. Among the self-employed, the policies with the most frequent take-up was the TOZO (26% in May; 14% in September). Tax deferrals and TOGS were the second most frequent in May (17%), followed by the NOW program (11% in May, 6% in September). Employees are targeted through the NOW program. 13% (11%) of employees indicate that their employer applied for the NOW program in May (September). A large fraction of employees indicates that they don't know whether their employer applied for NOW (27% in May, 30% in September). According to official statistics roughly 24% of employees were affected by NOW between March-May.[16] This indicates that a lot of employees are not aware of the policy take-up of their employer. We code every respondent who indicated that their employer applied and was not rejected by NOW in May or September as being affected by a support program.[17] For self-employed we consider all policies and code them as being affected by policy if they applied to any policy between March-September. We do not distinguish between take-up between March and May and June and September because the number of people affected only by the second round of policies is very small.

As additional control variable, we also use the sector an individual works in. This information is elicited in the work and schooling questionnaire in April 2020. When this information is not available, we use the answer from April 2019.

16. Absolute numbers can be found here: https://www.nowinzicht.nl/factsheet
17. Rejection rates are very low see https://www.nowinzicht.nl/factsheet.

3.B.2 Essential worker status

The Dutch government has identified a number of areas of the economy that are exempt from the restrictions on public life. Facilities in these areas remain open and parents working in these occupations are eligible for emergency daycare and after school care. A non-exhaustive list of occupations and industries includes care, youth aid and social support, including transportation and production of medicine and medical devices; teachers and school staff, required for online learning, exams and childcare; public transportation; food production and distribution, such as supermarkets, food production and food transportation, farmers, farmworkers and so forth; transportation of fuel, coal, diesel and so forth; transportation of waste and garbage; daycare; media and communications; emergency services such as fire department, ambulance, regional medical organizations; necessary administrative services on the provincial and municipality level. In addition, about 100 companies have been identified as necessary to sustain public life, operating in sectors such as gas and fuel production, distribution and transportation, communication and online services, water supply, securities trading, infrastructure, etc..

We asked the respondents directly for their essential worker status in April, but also obtain an indirect measure in March from a question about compliance to a potential curfew. The answering options were "yes", "no" or "I work in a critical profession". Whenever available we make use of the direct measure. Overall, 35% of individuals indicate that they work in an essential occupation (Table 3.B.1). The level and the distribution over sectors lines up well with estimates based on the 2019 Labor force survey (LFS) of Statistics Netherlands.[18] In the fourth quarter of 2019, about 34% of respondents worked in an occupation later to be declared essential.

18. For details see https://www.cbs.nl/nl-nl/faq/corona/economie/hoeveel-mensen-werken-er-in-cruciale-beroepen-.

3.B.3 Ability to work from home

In May 2020, we ask individuals "What percentage of your normal work prior to the coronavirus outbreak can you do while working from home?". Subjects could answer a number between 0 and 100. In December, we repeated this question about their current job by asking "What percentage of your normal work can you do with working from home?". We recode this measure to range from 0 to 1, instead. Table 3.B.2 displays number of observations, mean, standard deviation, as well as quantiles of the responses. Comparing the distribution of the measures of May and of December does not reveal large differences. 2,177 subjects answered the question in May and December. For those subjects, we can directly compare the answers, to investigate the stability of the measure. The measure may vary because (1) individuals change jobs or tasks at jobs or (2) measurement error. The correlation between the measure in May and the measure in December is 0.82. That is, the measure is fairly stable. It is with 0.63 lower for those individuals that changed employment status at some point between May and December (N=215). The average difference between May and September is 0.01 and approximately half of subjects do not change their answer at all. This stability in the measure indicates that measurement error is not substantial even though the question is asked retrospectively in May.

Table 3.B.2. Distribution of work from home capability in December and May

	count	mean	std	min	25%	50%	75%	max
May	2746	0.45	0.42	0.0	0.0	0.40	0.90	1.0
Dec.	2671	0.44	0.43	0.0	0.0	0.30	0.90	1.0
dev. in meas.	2177	0.01	0.25	-1.0	0.0	0.00	0.02	1.0
abs. dev. in meas.	2177	0.13	0.22	0.0	0.0	0.01	0.18	1.0

Notes: First (second) row displays the distribution of work from home capability in May (December). Third row displays the distribution of the intra-subject changes in answers between May and December. Deviations are calculated by subtracting the May answer from the December answer of subjects. The fourth row displays the distirbution of the absolute value of deviations.

Given the high stability of the measure and the low labor market turnover in our sample, we use the mean between the answers in May and in December in our analysis to measure the work from home capability.

3.B.4 Sample attrition

Tables 3.B.3 displays summary statistics of respondents in all waves. Table 3.B.4 shows the same measures for our main sample, i.e. all individuals working at least 10 hours in the pre-pandemic period.

Except the increasing age of our sample, the only variable with a significant difference over time is essential worker status. We elicit essential worker status twice and measure a slightly higher share of essential workers in the April wave than in the March wave. Since the question in April is more precisely asked, we take the April measure as default and make use of the March measure whenever the former is missing. This leads to the combined measure being 4-5 % higher in April than in the other waves which doesn't seem to influence our main results.

Altogether, the characteristics of respondents are very stable over the waves which suggests that sample attrition does not introduce a bias in any direction.

Table 3.B.3. Characteristics of respondents in each survey wave – full sample

	before Covid-19	march 2020	april 2020	may 2020	june 2020	september 2020	december 2020
age	44.806	45.226	45.470	45.442	45.218	45.668	45.875
	(0.215)	(0.226)	(0.226)	(0.234)	(0.225)	(0.230)	(0.237)
female	0.560	0.553	0.560	0.550	0.557	0.557	0.547
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
education: lower sec. and below	0.191	0.194	0.195	0.196	0.196	0.194	0.197
	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	(0.007)	(0.007)
education: upper secondary	0.387	0.388	0.384	0.389	0.384	0.384	0.391
	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
education: tertiary	0.422	0.418	0.421	0.415	0.420	0.423	0.412
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
net hh income 18/19 (equiv)	2233.167	2202.449	2250.052	2216.935	2212.931	2213.192	2258.957
	(66.630)	(61.474)	(73.906)	(64.824)	(60.151)	(64.310)	(79.815)
gross income: below 2500	0.538	0.536	0.540	0.538	0.537	0.534	0.535
	(0.008)	(0.008)	(0.008)	(0.009)	(0.008)	(0.009)	(0.009)
gross income: bet. 2500 and 3500	0.224	0.225	0.220	0.227	0.224	0.229	0.225
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
gross income: above 3500	0.238	0.239	0.240	0.235	0.239	0.236	0.240
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
full time employed pre-Covid	0.426	0.426	0.422	0.420	0.424	0.424	0.430
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)
part time employed pre-Covid	0.221	0.217	0.220	0.216	0.214	0.216	0.213
	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
self-employed pre-Covid	0.076	0.076	0.074	0.074	0.073	0.075	0.072
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
has partner	0.693	0.694	0.696	0.699	0.696	0.694	0.700
	(0.007)	(0.007)	(0.007)	(0.008)	(0.007)	(0.008)	(0.008)
married	0.487	0.491	0.496	0.493	0.493	0.492	0.497
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
no. children below 12	0.363	0.359	0.341	0.337	0.341	0.344	0.332
	(0.012)	(0.013)	(0.012)	(0.013)	(0.013)	(0.013)	(0.013)
frac. work doable from home	0.427	0.423	0.423	0.429	0.428	0.428	0.426
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
essential worker	0.354	0.351	0.398	0.364	0.356	0.356	0.358
	(0.009)	(0.009)	(0.009)	(0.010)	(0.009)	(0.009)	(0.010)
affected by policy: yes	0.212	0.211	0.208	0.210	0.210	0.201	0.205
	(0.008)	(0.009)	(0.009)	(0.008)	(0.009)	(0.008)	(0.009)
affected by policy: no	0.423	0.432	0.430	0.439	0.433	0.441	0.438
	(0.010)	(0.011)	(0.011)	(0.010)	(0.010)	(0.010)	(0.011)
affected by policy: don't know	0.365	0.357	0.361	0.350	0.357	0.359	0.357
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
N	4283	3850	3844	3631	3895	3641	3494

Notes: Sample: $18 \leq \text{age} \leq 66$. Not all variables are non-missing for each observation.

Table 3.B.4. Characteristics of respondents in each survey wave – working sample

	before Covid-19	march 2020	april 2020	may 2020	june 2020	september 2020	december 2020
age	44.238	44.579	44.847	44.941	45.041	45.240	45.365
	(0.227)	(0.238)	(0.239)	(0.252)	(0.243)	(0.249)	(0.254)
female	0.524	0.518	0.522	0.519	0.519	0.518	0.505
	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
education: lower sec. and below	0.135	0.137	0.137	0.138	0.133	0.136	0.137
	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
education: upper secondary	0.372	0.373	0.370	0.376	0.376	0.369	0.381
	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)
education: tertiary	0.492	0.489	0.493	0.486	0.491	0.496	0.481
	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
net hh income 18/19 (equiv)	2391.263	2334.973	2411.652	2353.283	2359.641	2359.043	2432.614
	(67.975)	(46.616)	(75.945)	(51.495)	(48.508)	(51.150)	(85.101)
gross income: below 2500	0.397	0.393	0.397	0.392	0.386	0.387	0.386
	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
gross income: bet. 2500 and 3500	0.282	0.284	0.277	0.287	0.284	0.290	0.284
	(0.009)	(0.009)	(0.009)	(0.010)	(0.009)	(0.010)	(0.010)
gross income: above 3500	0.321	0.323	0.326	0.320	0.330	0.324	0.330
	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)
full time employed pre-Covid	0.616	0.618	0.615	0.618	0.622	0.618	0.629
	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)
part time employed pre-Covid	0.279	0.276	0.282	0.280	0.277	0.277	0.271
	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
self-employed pre-Covid	0.105	0.105	0.103	0.103	0.102	0.105	0.100
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
has partner	0.713	0.714	0.719	0.724	0.718	0.714	0.723
	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
married	0.504	0.505	0.515	0.515	0.519	0.508	0.516
	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
no. children below 12	0.425	0.419	0.406	0.405	0.404	0.407	0.396
	(0.015)	(0.016)	(0.016)	(0.017)	(0.016)	(0.017)	(0.017)
frac. work doable from home	0.440	0.437	0.435	0.440	0.440	0.439	0.437
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.009)
essential worker	0.353	0.349	0.397	0.371	0.363	0.365	0.370
	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
affected by policy: yes	0.216	0.216	0.212	0.211	0.211	0.203	0.207
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
affected by policy: no	0.437	0.445	0.444	0.461	0.453	0.461	0.456
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
affected by policy: don't know	0.347	0.339	0.345	0.328	0.335	0.336	0.337
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)
N	2962	2656	2634	2375	2518	2384	2298

Notes: Sample: $18 \leq \text{age} \leq 66$; working hours of at least 10h in early March. Not all variables are non-missing for each observation.

Appendix 3.C Aggregate Trends

3.C.1 Labor force and unemployment over time

The first row of Table 3.C.1 shows the dynamics of the labor force for all respondents between the ages of 18 and 66. The share of respondents that are out of the labor force, i.e., neither working nor unemployed, but e.g., in education, retired or a home maker, increases from 24.4% before the onset of the crisis to 26.2% in May. Thereafter, it remains roughly at this level until December. Next, we focus on those individuals in the labor force and look at the unemployment rate. The second row of Table 3.3.1 reveals that before the Covid-19 crisis, we estimate the unemployment rate to be 4.5%. Until May, it gradually rises by 1.1 percentage points and decreases slightly thereafter.

Table 3.C.1. Labor force status and working hours over time

	before Covid-19	Mar	Apr	May	Jun	Sep	Dec
out of laborforce (perc.)	24.4	24.7	25.1	26.2	25.8	26.2	26.9
	(0.7)	(0.7)	(0.7)	(0.7)	(0.7)	(0.7)	(0.7)
N	4285	3866	3863	3645	3910	3656	3509
unemployed (perc.)	4.5	4.9	5.5	5.6	5.1	5.6	5.2
	(0.4)	(0.4)	(0.4)	(0.4)	(0.4)	(0.4)	(0.4)
N	3241	2912	2892	2689	2902	2698	2566

Notes: Source LISS. All statistics are on respondents between ages 18 and 66. For the unemployment rate, only individuals in the labor force are considered.

We next compare these trends to official data of Statistics Netherlands (CBS)[19]. We focus on the group of individuals aged 25-44 years since official records are not available specifically for the age range used in our analysis. Table 3.C.2 reports the rates of unemployment and non-employment in our sample and in the official records. The trajectory are overall very similar. Until April, the rate of non-employed individuals increases by 0.8 percentage points in our sample and by 0.5 in official data. Until December, it falls even slightly below the pre-pandemic level. The level of the unemployment rate is about 1 percentage point larger in our sample compared to official records. The maximal raise in the unemployment rate and the small increase until December (0.3 and 0.2 percentage points) are fairly similar, but the timing of this pattern is different: In official data, the increase starts only in June while we measure increasing unemployment in our sample already in the months before. The deviation could be partly caused by the fact that we didn't ask for employment

19. See https://opendata.cbs.nl/statline/#/CBS/en/dataset/80590ENG/table?ts=1620213584059

status explicitely in March and April, but infer those from reported working hours and qualitative follow-up questions.

Table 3.C.2. Labor force status and working hours over time (age 25-44)

	before Covid-19	Mar	Apr	May	Jun	Sep	Dec
out of laborforce (perc.)	11.1	11.6	11.9	11.7	11.6	10.6	10.3
	(0.8)	(0.9)	(0.9)	(0.9)	(0.9)	(0.9)	(0.9)
N	1560	1384	1341	1251	1372	1261	1180
unemployed (perc.)	3.6	4.3	4.5	4.8	3.9	4.0	4.0
	(0.5)	(0.6)	(0.6)	(0.6)	(0.6)	(0.6)	(0.6)
N	1387	1223	1182	1105	1213	1127	1059
out of laborf CBS	11.6	11.6	12.1	12.0	11.9	11.4	11.2
unemployed CBS	3.0	3.0	3.1	3.0	3.5	3.5	3.2

Notes: Source LISS. The last two rows report the numbers based on official records by CBS (Statistics Netherlands). All statistics are on respondents between ages 25 and 44. For the unemployment rate, only individuals in the labor force are considered.

The official data is also available for a larger sample of individuals between 15 and 75 years. For this sample, the observed differences to our sample are similar. We, however, observe a higher level of non-employment and an increase of this rate over time. This is likely associated with older individuals having a higher response rate. Overall, the comparison in this section reveals that the most important changes over time visible in official records are replicated in our sample. The observed differences are unlikely to bias the result of our main analyses which is based on unconditional working hours.

3.C.2 Robustness for aggregate trends

Our main baseline measure of working hours before the onset of the pandemic are the working hours of early March 2020. Those are asked retrospectively in late March and April. Conversely, for the working hour measures in all other periods, we ask for the working hours in the last seven days. A potential concern is that observed changes in labor supply might be driven by the different ways working hours are elicited. An alternative baseline measure is based on the time use and consumption survey that was in the field in November 2019. As participants are in this study also asked for their working hours in the last week, the elicitation method is closer to the one for our observations from March on. On the other hand, this data was elicited longer before the pandemic and the joint sample is substantially lower.

Table 3.C.3 compares the distributions of the two measures. Based on the time use survey, mean total working hours are about one hour larger. The third row reveals that mean deviation on the individual level is below 0.2 which shows that the mean of the two measures are very similar. The absolute deviation is 7 hours on

average with a median of 3 hours. The correlation between the measures is 0.51 which indicates that none of the samples seem to be strongly biased in any direction. Because of the larger sample size, we make use of the February data in the main body of the paper and use the time use data for robustness analyses.

Table 3.C.3. Pre-Covid working hours based on Covid survey and time use survey

	N	mean	std. dev.	min	$q_{0.25}$	$q_{0.5}$	$q_{0.75}$	max
hours early March 2020 (retrospective)	3112	33.23	12.51	0	25	36	40	80
hours November 2019 (time use survey)	1827	34.34	13.58	0	28	36	40	80
dev. in measures	1827	0.19	12.68	-60	0	0	4	63
abs. dev. in measures	1827	6.96	10.60	0	0	3	8	63

Notes: First row displays the distribution of working hours in early March 2020 while the second row shows the respective distribution for the measure based on the time use survey in November 2019. Third row displays the distribution of the intra-subject differences between November 2019 and March/April 2020. The fourth row displays the distribution of the absolute value of deviations.

Table 3.C.4 replicates Table 3.3.1 for a different sample which includes all individuals that work at least 10 hours in any of the seven periods. Importantly, we include individuals in this sample that were not working shortly before Covid-19 hit the economy, but do so afterwards. We hence avoid a mechanical drop in average unconditional working hours.

As expected, unconditional working hours are smaller for this sample. Furthermore, reductions in aggregate working hours are smaller which implies that Table 3.3.1 overestimates those, especially in later months. For our analyses, we nevertheless prefer the restriction on individuals working before the pandemic for two reasons: First, it allows to look at relative changes in working hours. Second, we only have complete information on essential worker status and ability to work from home for these individuals.

Table 3.C.5 shows aggregate trends making use of sample weights. The weights are based on age, sex, and marital status of the respondents.

Table 3.C.4. Working hours over time for subjects working at least 10 hours in any period

	before Covid-19	Mar	Apr	May	Jun	Sep	Dec
working hours	32.2	28.2	27.7	26.3	27.1	27.4	29.0
	(0.2)	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)
N	3182	2857	2832	2658	2869	2693	2580
hours worked from home	3.8	14.0	14.6	12.2	10.8	8.7	12.0
	(0.2)	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)
N	3182	2857	2832	2658	2869	2693	2580
share of hours worked from home	0.11	0.49	0.51	0.45	0.38	0.31	0.39
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
N	2962	2437	2408	2106	2317	2127	2052

Notes: Source LISS. Household income in thousands. All statistics are on respondents between ages 18 and 66 who worked for at least 10 hours in at least one of the 7 periods.

Table 3.C.5. Labor force status and working hours over time (weighted)

	before Covid-19	Mar	Apr	May	Jun	Sep	Dec
out of laborforce (perc.)	23.0	23.0	23.2	24.3	24.1	24.0	24.3
	(0.6)	(0.7)	(0.7)	(0.7)	(0.7)	(0.7)	(0.7)
N	4285	3866	3851	3645	3910	3656	3509
unemployed (perc.)	4.3	4.8	5.4	5.4	5.1	5.5	4.9
	(0.4)	(0.4)	(0.4)	(0.5)	(0.4)	(0.5)	(0.4)
N	3241	2912	2883	2689	2902	2698	2566
working hours	35.0	30.8	30.0	27.1	28.2	28.1	29.8
	(0.3)	(0.4)	(0.4)	(0.4)	(0.4)	(0.4)	(0.4)
N	2962	2656	2634	2375	2518	2384	2298
hours worked from home	4.1	15.4	15.9	12.4	11.4	9.0	12.4
	(0.2)	(0.4)	(0.4)	(0.3)	(0.3)	(0.3)	(0.4)
N	2962	2656	2634	2375	2518	2384	2298
share of hours worked from home	0.11	0.49	0.52	0.45	0.39	0.31	0.40
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
N	2962	2437	2408	2106	2317	2127	2052

Notes: Source LISS. All statistics are on respondents between ages 18 and 66. The sample for unemployment includes all individuals in the labor force. The sample for hours include individuals who worked for at least 10 hours in any one of the 5 periods. Observations are weighted based on age, sex, and marital status.

3.C.3 Figures for trends over time

This subsection presents visualizations of the trajectories of labor force participation, unemployment, and total working hours.

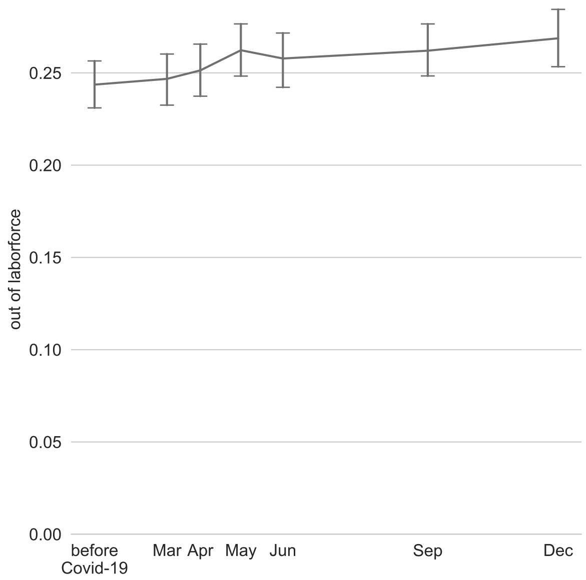


Figure 3.C.1. Non-participation rate

The figure shows the rate of respondents in our sample over that are neither employed nore self-employed over time. Vertical bars depict 95 %-confidence intervals. Sample: Age $\leq$ 65.

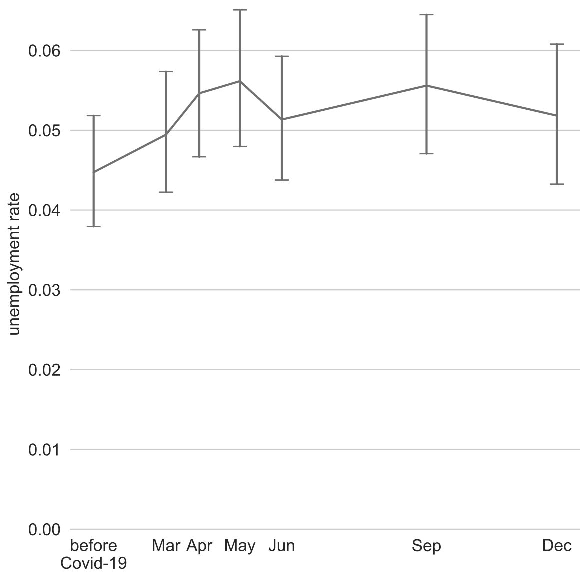


Figure 3.C.2. Unemployment rate

The figure shows the unemployment rate in our sample over time. Vertical bars depict 95 %-confidence intervals. Sample: $18 \leq \text{age} \leq 66$; being employed, self-employed or unemployed in the respective month.

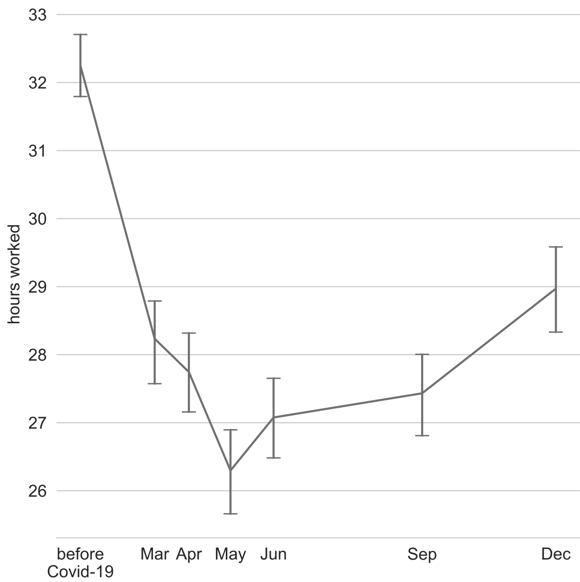


Figure 3.C.3. Working hours

Notes: The figure shows total hours worked over time. Vertical bars depict 95 %-confidence intervals. Sample: $18 \leq \text{age} \leq 66$; working hours of at least 10h in at least one period.

3.C.4 Working hours reductions and expected job loss

Working less while still earning the same might be for many individuals not a bad thing per se. However, they are likely a good proxy of who will loose their job in case the pandemic continues and economic support measures run out. Even if people who reduce working hours are going to keep their job later, they might face increased mental stress with respect to job security. Table 3.C.6 shows that a reduction in working hours in March by 10 hours is associated with a 1.2 higher expected probability to loose one's job within the next two months (column (2)). This relation is not mainly driven by individuals that lost their job already (column (3)). Furthermore, it relates to an increase of self-reported job worries by 0.12 std (column (1)).

Appendix 3.D Predictors of working hours and household income

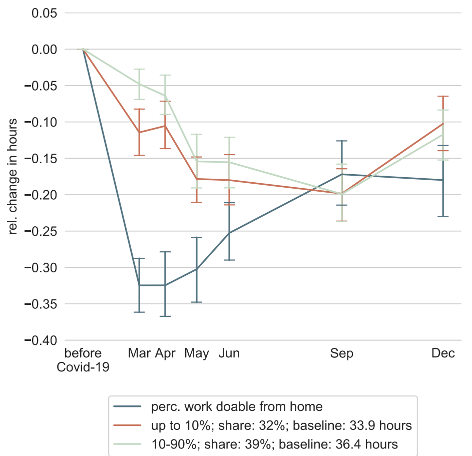
3.D.1 Working hours changes by characteristics

The top row of Figure 3.D.1 shows total working hours by the degree of telecommutability in three categories: For the subset of non-essential workers (Figure 3.D.1a), roughly 3 in 10 individuals can work up to 10 % of their work from home and the same share can do so for more than 90 % of their work. This leaves 40 % of non-essential workers in the middle category. For workers who are not classified as essential, the relevance of telecommutability during the first lockdown is enormous. The fifth of the workforce that is not classified as essential worker and has very little possibility to work from home lost one third of pre-pandemic working hours, compared to 11 and 5 percentage point for intermediate and high degrees of telecommutability. These gaps have narrowed considerably to 10 percentage points or less by June and are slightly reversed in September. Until December, working hours for individuals with high or medium capability to work from home go up again, but stagnate for low telecommutability jobs.

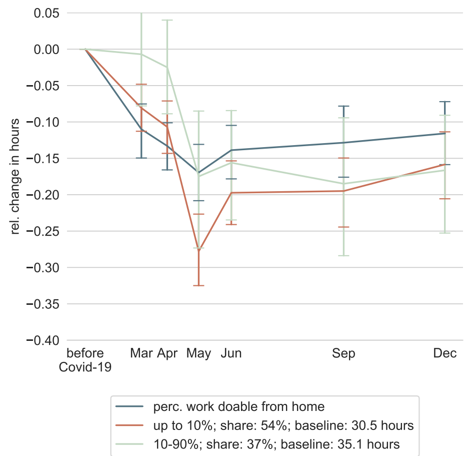
In stark contrast to this, the ability to work from home does not have salient effects on the overall quantity of work for essential workers. Figure 3.D.1b shows that initially, reductions are only slightly stronger for workers without the ability to work from home. Starting from May, there is an additional 15 percentage point decrease for the group of essential workers with intermediate degrees of telecommutability. The relation between telecommutability and hours changes is generally not monotone for essential workers, whereas it is for non-essential workers.

Figure 3.D.1c suggests that substituting workplace hours by home office hours is driving many of these patterns. For non-essential workers with more than 90% capability to work from home, home office hours are up by more than 20 hours in March and April. For subjects in jobs with medium degrees of telecommutability, hours worked from home increase by more than 15 hours during the first months of the pandemic. As restrictions are gradually lifted, home office hours decrease again in these two groups, both in terms of absolute numbers and the share of total working hours. In December, home office hours increase strongly again although not quite to the levels during the first lockdown. Conversely, in jobs in which almost all work has to be done at the workplace, the change in home office is very close to zero over the full observed period. for essential workers (Figure 3.D.1d), changes in hours worked from home are very similar to non-essential workers, for a given level of telecommutability.

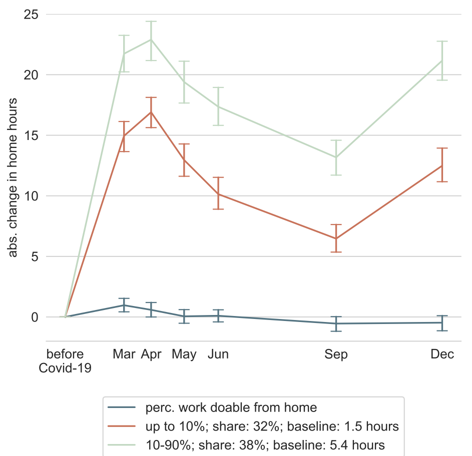
Figure 3.D.2 displays absolute changes in working hours for socio-economic groups. Especially for the income groups, baseline working hours differ strongly between the groups. Therefore, absolute changes are harder to interpret as relative changes which we use in the main part of the paper.



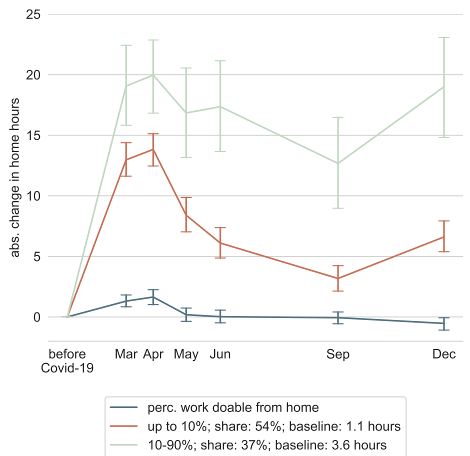
(a) Non-essential workers: Relative change in total working hours



(b) Essential workers: Relative change in total working hours



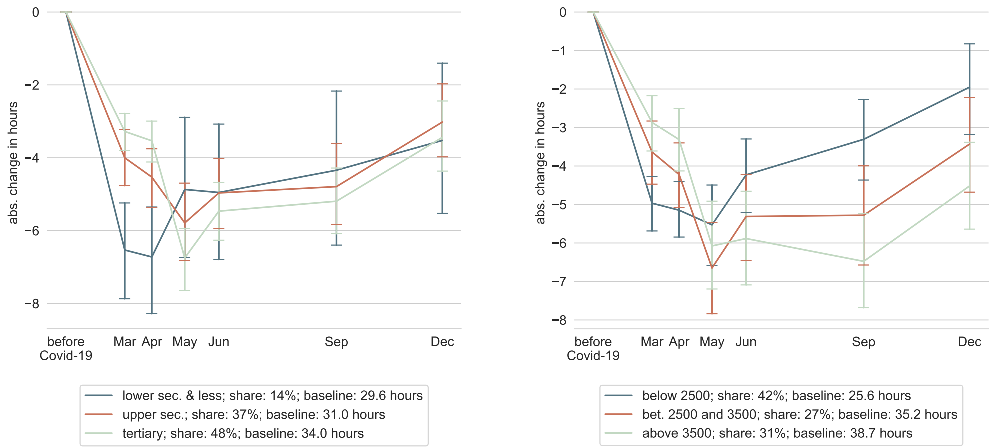
(c) Non-essential workers: Change in hours worked from home



(d) Essential workers: Change in hours worked from home

Figure 3.D.1. Changes in total working hours and hours worked from home, by essential worker status and the percentage of work that can be done from home

Notes: The figure shows mean relative changes in total hours worked (top row) and mean absolute changes in hours worked from home (Panel b) over time by percentage of work that can be done from home (in three categories). The sample in the first column is restricted on non-essential workers while the second column considers only essential workers. Reference period is late February/early March. Vertical bars depict 95 %-confidence intervals. Sample: $18 \leq \text{age} \leq 66$; working hours of at least 10h in early March. The legend displays hours and share of each group in early March.



(a) Absolute change in total working hours by education

(b) Absolute change in total working hours by personal income

Figure 3.D.2. Absolute changes in total working hours, by socio-economic status

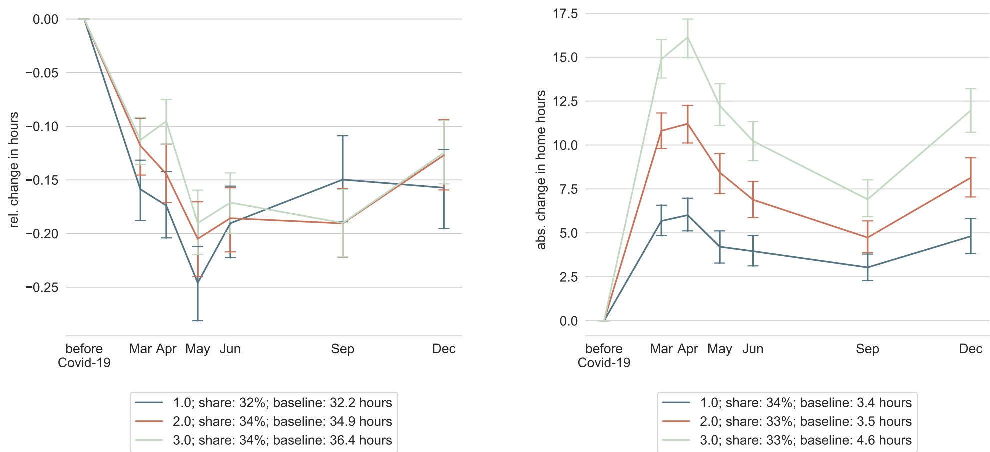
Notes: The figure shows mean absolute changes in total hours worked by level of education (Panel a) and personal gross income (Panel b) over time. Reference period is late February/early March. The legend displays hours and share of each group in early March. Vertical bars depict 95 %-confidence intervals. Sample: $18 \leq \text{age} \leq 66$; working hours of at least 10h in early March.

Table 3.C.6. Working hours reductions in March and job loss expectations

	concerned about job	expected job loss prob.	
	(1)	(2)	(3)
change hours March	-0.013***	-0.123***	-0.095***
	(0.002)	(0.030)	(0.026)
female	-0.039	-1.165**	-0.913
	(0.044)	(0.581)	(0.556)
N	2485	2487	2470
R^2	0.128	0.033	0.027
mean dependent variable	0.034	4.464	4.304
Subset: didn't loose job	No	No	Yes
Demographic controls	Yes	Yes	Yes

Notes: Source LISS. Job concerns are measued by a 5-point Likert scale and standardized. Sample: 18 $\leq$ age $\leq$ 66; working hours of at least 10h in early March. For the first three columns the sample is additionally restricted to individuals working pre-Covid. *p<0.1; **p<0.05; ***p<0.01.

Figure 3.D.3 show changes in workings hours over time by long-run household income. Figure 3.D.4 does so for the employed and self-employed.

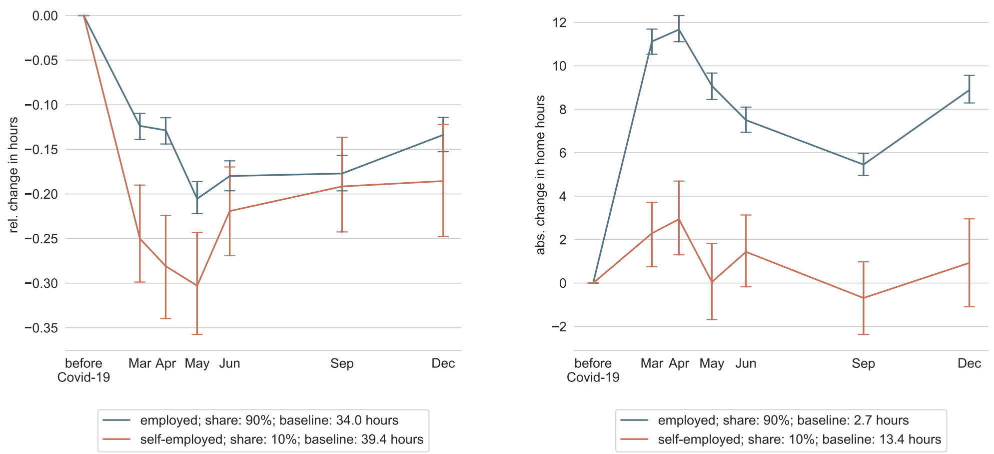


(a) Relative change in total working hours by household income tercile

(b) Change in hours worked from home by household income tercile

Figure 3.D.3. Changes in total working hours and hours worked from home, by long-run household income before Covid-19

Notes: The figure shows mean relative changes in total hours worked (Panel a) and mean absolute changes in hours worked from home (Panel b) over time by long-run household income tercile (equivalized). Reference period is late February/early March. The legend displays hours and share of each group in early March. Vertical bars depict 95 %-confidence intervals. Sample: 18 $\leq$ age $\leq$ 66; working hours of at least 10h in early March.

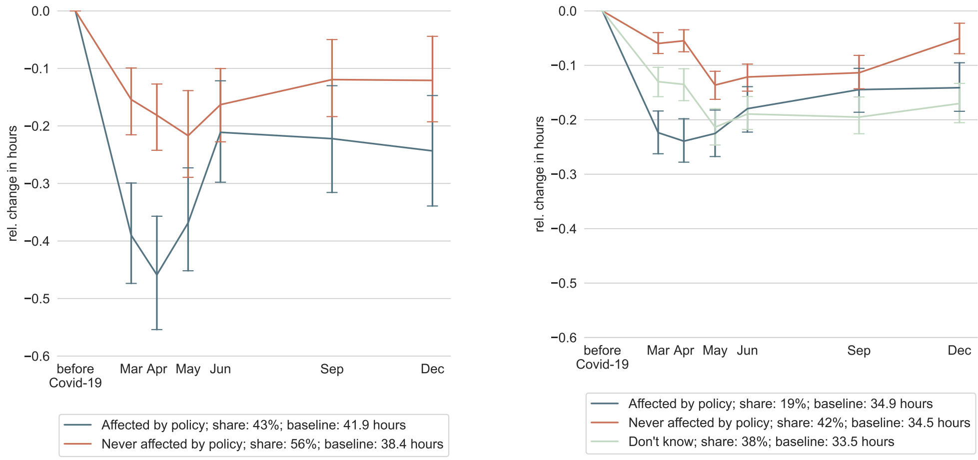


(a) Relative hange in total working hours

(b) Change in hours worked from home

Figure 3.D.4. Changes in total working hours and hours worked from home, by type of employment

Notes: The figure shows mean relative changes in total hours worked (Panel a) and mean absolute changes in hours worked from home (Panel b) over time for self-employed and employees. Reference period is late February/early March. The legend displays hours and share of each group in early March. Vertical bars depict 95 %-confidence intervals. Sample: $18 \leq \text{age} \leq 66$; working hours of at least 10h in early March.



(a) Initially self-employed **(b)** Initially employee

Figure 3.D.5. Total working hours and hours worked from home, by being affected by any support measure as elicited between March and September

Notes: The figure shows mean relative changes in total hours worked by being affected by any support measure sometime between March and September for initially self-employed (Panel a) and initially employed (Panel b) over time. Reference period is late February/early March. The legend displays hours and share of each group in early March. Vertical bars depict 95 %-confidence intervals. Sample: $18 \leq \text{age} \leq 66$; working hours of at least 10h in early March.

Figure 3.D.5 shows that those self-employed that applied for government support decreased their working hours substantially in March/April. This is reassuring, as TOGS and TOZO – while not explicitly restricting working hours – targeted those who were directly affected by the social distancing regulations and those whose income fell below the social minimum. Employees affected by a policy reduced their working hours on average much less than the self-employed, however, they still reduced working hours quite substantially by more than 20 %. Further, they weakly increase their working hours between May and December.

While these results cannot tell us anything about the counterfactual scenario, they indicate that on average policies did not overcompensate the productivity loss of firms. Even though there was no formal requirement of decreasing working hours under the NOW policy, workers still worked on average substantially less hours during the policy receipt as right before the pandemic.

3.D.2 Predictors of changes in working hours

A potential issue with our data is that pre-pandemic working hours are asked retrospectively for a few weeks earlier while working hours in all other periods are asked for the last week. Table 3.D.1 shows robustness analyses for the regressions in Table 3.4.2. In the first three columns all individuals are excluded who report that they took a day off out of turn, e.g. because of official holidays, vacation, or being sick. March and June observations are dropped since we don't have this information for these months. In the last three columns, pre-pandemic working hours are based on the time use survey conducted in November 2020 that also asks for working hours during the last seven days (see Section 3.C.2). Standard errors are larger due to the lower sample size, but observed patterns are very similar to Table 3.4.2 indicating that the different elicitation method does not drive our results.

Table 3.D.1. Hours worked by individual and job characteristics (Robustness)

	change total working hours					
	subset: no day taken off		baseline: time use survey			
	(1)	(2)	(3)	(4)	(5)	(6)
march/april × education: upper sec.	0.06**	0.03	0.02	0.03	0.01	0.00
	(0.02)	(0.02)	(0.02)	(0.05)	(0.05)	(0.05)
may × education: upper sec.	0.03	0.01	0.00	-0.04	-0.05	-0.06
	(0.05)	(0.05)	(0.05)	(0.06)	(0.05)	(0.06)
june × education: upper sec.				-0.06	-0.06	-0.07
				(0.06)	(0.06)	(0.07)
september × education: upper sec.	-0.01	0.00	0.01	-0.02	0.00	-0.03
	(0.04)	(0.04)	(0.04)	(0.06)	(0.06)	(0.07)
december × education: upper sec.	0.02	0.02	0.01	0.02	0.03	-0.00
	(0.04)	(0.04)	(0.04)	(0.06)	(0.05)	(0.06)
march/april × education: tertiary	0.07***	0.01	0.00	0.06	0.00	0.00
	(0.02)	(0.02)	(0.03)	(0.05)	(0.05)	(0.06)
may × education: tertiary	-0.02	-0.07	-0.04	-0.02	-0.04	-0.06
	(0.05)	(0.05)	(0.06)	(0.06)	(0.07)	(0.07)
june × education: tertiary				-0.01	-0.01	-0.06
				(0.06)	(0.08)	(0.08)
september × education: tertiary	-0.01	0.01	0.03	0.08	0.12	0.07
	(0.04)	(0.04)	(0.04)	(0.07)	(0.09)	(0.10)
december × education: tertiary	0.05	0.04	0.03	0.07	0.08	0.03
	(0.04)	(0.04)	(0.04)	(0.06)	(0.07)	(0.07)
march/april × income bet. 2500 and 3500	0.07***	0.06***	0.04**	-0.01	-0.02	-0.04
	(0.02)	(0.02)	(0.02)	(0.04)	(0.04)	(0.04)
may × income bet. 2500 and 3500	0.01	-0.00	-0.04	0.03	0.03	0.00
	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.06)
june × income bet. 2500 and 3500				0.02	0.02	-0.00
				(0.04)	(0.05)	(0.05)
september × income bet. 2500 and 3500	0.03	0.03	0.03	-0.01	-0.00	-0.02
	(0.03)	(0.03)	(0.03)	(0.06)	(0.06)	(0.06)
december × income bet. 2500 and 3500	0.00	-0.00	-0.03	0.00	0.00	-0.03
	(0.03)	(0.03)	(0.03)	(0.06)	(0.06)	(0.06)
march/april × income above 3500	0.09***	0.06***	0.04*	0.04	0.01	-0.01
	(0.02)	(0.02)	(0.02)	(0.06)	(0.07)	(0.07)
may × income above 3500	0.06	0.04	0.01	0.13	0.11	0.08
	(0.05)	(0.05)	(0.04)	(0.09)	(0.10)	(0.10)
june × income above 3500				0.07	0.07	0.06
				(0.09)	(0.11)	(0.11)
september × income above 3500	0.04	0.05	0.04	-0.06	-0.04	-0.06
	(0.03)	(0.03)	(0.03)	(0.10)	(0.11)	(0.12)
december × income above 3500	0.00	-0.01	-0.02	0.02	0.02	0.00
	(0.03)	(0.03)	(0.03)	(0.09)	(0.10)	(0.11)
march/april × essential worker		0.18***	0.16***		0.16***	0.13
		(0.02)	(0.03)		(0.05)	(0.09)
may × essential worker		0.14***	0.14**		0.06	-0.03
		(0.05)	(0.06)		(0.07)	(0.14)
june × essential worker					0.02	-0.06
					(0.08)	(0.16)
september × essential worker		0.01	0.02		-0.04	-0.12
		(0.03)	(0.04)		(0.09)	(0.16)
december × essential worker		0.04	0.06*		-0.07	-0.16
		(0.03)	(0.03)		(0.08)	(0.15)
march/april × frac. work doable from home		0.23***	0.22***		0.21***	0.21***
		(0.02)	(0.02)		(0.07)	(0.07)
may × frac. work doable from home		0.24***	0.17***		0.11	0.12
		(0.05)	(0.06)		(0.11)	(0.09)
june × frac. work doable from home					0.00	0.04
					(0.12)	(0.11)
september × frac. work doable from home		-0.03	-0.02		-0.09	-0.06
		(0.03)	(0.03)		(0.13)	(0.12)
december × frac. work doable from home		0.07**	0.06*		-0.04	-0.03
		(0.03)	(0.04)		(0.11)	(0.10)
march/april × essential × work doable from home		-0.15***	-0.12***		-0.18**	-0.14
		(0.03)	(0.04)		(0.08)	(0.09)
may × essential × work doable from home		-0.29***	-0.21**		-0.20**	-0.14
		(0.09)	(0.09)		(0.09)	(0.11)
june × essential × work doable from home					-0.07	-0.02
					(0.09)	(0.12)
september × essential × work doable from home		-0.07	-0.07		-0.08	-0.05
		(0.06)	(0.06)		(0.10)	(0.13)
december × essential × work doable from home		-0.09*	-0.09*		-0.02	0.05
		(0.05)	(0.05)		(0.09)	(0.12)
N	8161	8161	7872	10529	10529	10356
R^2	0.054	0.082	0.101	0.009	0.011	0.016
demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
month × sector FE	No	No	Yes	No	No	Yes

Notes: The table shows robustness analyses for the regressions in Table 3.4.2. In the first three columns all individuals are excluded who report that they took a day off because of a vacation, an official holiday, being sick, or another exceptional reason. Since we don't have this information in June, we don't make use of these observations. For the last three columns, the baseline is based on the time use and consumption survey conducted in November 2019. Further elements of the specifications include a full set of time dummies, gender, a self-employed dummy and a part-time dummy. Standard errors are clustered on the individual level. Notes: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table 3.D.2. Hours worked by long-run household income

	change total working hours
	(1)
march/april	-0.20***
	(0.03)
may	-0.19***
	(0.04)
june	-0.09**
	(0.04)
september	0.02
	(0.05)
december	-0.02
	(0.05)
march/april × working hours pre-Covid	0.00
	(0.00)
may × working hours pre-Covid	-0.00**
	(0.00)
june × working hours pre-Covid	-0.00***
	(0.00)
september × working hours pre-Covid	-0.01***
	(0.00)
december × working hours pre-Covid	-0.00***
	(0.00)
march/april × net hh income 18/19 Q2	0.03
	(0.02)
may × net hh income 18/19 Q2	0.07**
	(0.03)
june × net hh income 18/19 Q2	0.04
	(0.02)
september × net hh income 18/19 Q2	-0.01
	(0.03)
december × net hh income 18/19 Q2	0.04
	(0.03)
march/april × net hh income 18/19 Q3	0.05***
	(0.02)
may × net hh income 18/19 Q3	0.07***
	(0.03)
june × net hh income 18/19 Q3	0.03
	(0.02)
september × net hh income 18/19 Q3	-0.03
	(0.03)
december × net hh income 18/19 Q3	0.05*
	(0.03)
N	14938
R^2	0.144

Notes: The table shows regressions of relative changes in working hours relative to pre-corona levels. Independent variables are the long-run net household income in quintiles and baseline working hours. The former is measured as the average monthly net household income in 2018 and 2019. This variable is equivalized by the number of household members. All variables are fully interacted with month-dummies. Standard errors are clustered on the individual level. Notes: *p<0.1; **p<0.05; ***p<0.01.

Table 3.D.3. Hours worked and not working by individual and job characteristics

	change total working hours			no job		
	(1)	(2)	(3)	(4)	(5)	(6)
march/april	-0.22***	-0.32***	-0.51***	0.014**	0.019***	0.011
	(0.03)	(0.03)	(0.07)	(0.006)	(0.006)	(0.010)
may	-0.28***	-0.33***	-0.47***	0.077***	0.093***	0.080**
	(0.04)	(0.04)	(0.07)	(0.019)	(0.020)	(0.037)
june	-0.35***	-0.40***	-0.38***	0.069***	0.076***	0.019
	(0.04)	(0.04)	(0.07)	(0.020)	(0.021)	(0.027)
september	-0.31***	-0.31***	-0.25**	0.108***	0.118***	0.133***
	(0.04)	(0.05)	(0.11)	(0.023)	(0.024)	(0.048)
december	-0.27***	-0.29***	-0.31***	0.101***	0.100***	0.117**
	(0.04)	(0.05)	(0.10)	(0.024)	(0.024)	(0.046)
march/april × female	-0.04***	-0.07***	-0.06***	0.000	0.001	0.001
	(0.02)	(0.02)	(0.02)	(0.004)	(0.004)	(0.004)
may × female	-0.05**	-0.05**	-0.05**	-0.013	-0.009	0.000
	(0.02)	(0.02)	(0.02)	(0.011)	(0.011)	(0.010)
june × female	-0.01	-0.02	-0.03	-0.019*	-0.016	-0.007
	(0.02)	(0.02)	(0.02)	(0.011)	(0.011)	(0.011)
september × female	-0.03	-0.03	-0.04	-0.022*	-0.019	-0.017
	(0.02)	(0.02)	(0.02)	(0.012)	(0.013)	(0.013)
december × female	-0.05**	-0.05**	-0.05*	-0.015	-0.015	-0.012
	(0.02)	(0.02)	(0.02)	(0.013)	(0.013)	(0.013)
march/april × education: upper sec.	0.06***	0.04	0.03	0.004	0.004	0.006
	(0.02)	(0.02)	(0.02)	(0.004)	(0.005)	(0.004)
may × education: upper sec.	0.03	0.01	0.01	-0.012	-0.012	0.000
	(0.03)	(0.03)	(0.03)	(0.016)	(0.016)	(0.015)
june × education: upper sec.	0.05*	0.04	0.03	-0.012	-0.012	0.004
	(0.03)	(0.03)	(0.03)	(0.016)	(0.016)	(0.015)
september × education: upper sec.	0.01	0.01	0.01	-0.021	-0.021	-0.013
	(0.04)	(0.04)	(0.04)	(0.019)	(0.019)	(0.019)
december × education: upper sec.	0.06*	0.06	0.05	-0.032	-0.032	-0.019
	(0.04)	(0.04)	(0.04)	(0.020)	(0.020)	(0.020)
march/april × education: tertiary	0.07***	0.01	0.01	0.005	0.005	0.007
	(0.02)	(0.02)	(0.03)	(0.005)	(0.006)	(0.005)
may × education: tertiary	0.00	-0.03	-0.02	-0.016	-0.018	0.002
	(0.03)	(0.03)	(0.03)	(0.017)	(0.017)	(0.017)
june × education: tertiary	0.07**	0.05	0.03	-0.018	-0.022	-0.000
	(0.03)	(0.03)	(0.03)	(0.016)	(0.017)	(0.016)
september × education: tertiary	0.04	0.06	0.06	-0.032*	-0.034*	-0.026
	(0.04)	(0.04)	(0.04)	(0.019)	(0.020)	(0.020)
december × education: tertiary	0.08**	0.06*	0.05	-0.030	-0.033	-0.018
	(0.04)	(0.04)	(0.04)	(0.021)	(0.022)	(0.022)
march/april × income bet. 2500 and 3500	0.07***	0.05***	0.04**	-0.008*	-0.008*	-0.005
	(0.02)	(0.02)	(0.02)	(0.005)	(0.004)	(0.004)
may × income bet. 2500 and 3500	0.05*	0.04	0.01	-0.032***	-0.032***	-0.017*
	(0.02)	(0.02)	(0.02)	(0.011)	(0.011)	(0.010)
june × income bet. 2500 and 3500	0.06**	0.05**	0.04	-0.013	-0.013	-0.001
	(0.02)	(0.02)	(0.02)	(0.011)	(0.011)	(0.011)
september × income bet. 2500 and 3500	0.04*	0.05*	0.03	-0.030**	-0.029**	-0.015
	(0.03)	(0.03)	(0.03)	(0.012)	(0.012)	(0.012)
december × income bet. 2500 and 3500	0.06**	0.05**	0.03	-0.028**	-0.029**	-0.016
	(0.03)	(0.03)	(0.03)	(0.014)	(0.014)	(0.014)
march/april × income above 3500	0.11***	0.07***	0.06***	-0.010**	-0.010**	-0.007
	(0.02)	(0.02)	(0.02)	(0.005)	(0.005)	(0.004)
may × income above 3500	0.09***	0.07**	0.05*	-0.022*	-0.024*	-0.010
	(0.03)	(0.03)	(0.03)	(0.012)	(0.012)	(0.012)
june × income above 3500	0.04	0.03	0.03	-0.008	-0.011	-0.000
	(0.03)	(0.03)	(0.03)	(0.012)	(0.012)	(0.012)
september × income above 3500	0.01	0.02	0.00	-0.006	-0.006	0.010
	(0.03)	(0.03)	(0.03)	(0.014)	(0.014)	(0.014)
december × income above 3500	0.04	0.03	0.02	-0.021	-0.023	-0.011
	(0.03)	(0.03)	(0.03)	(0.015)	(0.015)	(0.015)
march/april × part time pre-Covid	0.02	0.03	0.03	0.006	0.007	0.007
	(0.02)	(0.02)	(0.02)	(0.005)	(0.005)	(0.005)
may × part time pre-Covid	0.04	0.04	0.03	0.042***	0.046***	0.052***
	(0.03)	(0.03)	(0.03)	(0.014)	(0.015)	(0.014)
june × part time pre-Covid	0.05*	0.05*	0.05*	0.039***	0.042***	0.041***

Continued on next page

Table 3.D.3. Hours worked and not working by individual and job characteristics

	change total working hours			no job		
	(1)	(2)	(3)	(4)	(5)	(6)
	(0.03)	(0.03)	(0.03)	(0.014)	(0.014)	(0.014)
september × part time pre-Covid	0.07**	0.06**	0.07**	0.058***	0.061***	0.063***
	(0.03)	(0.03)	(0.03)	(0.016)	(0.016)	(0.017)
december × part time pre-Covid	0.10***	0.11***	0.10***	0.048***	0.049***	0.054***
	(0.03)	(0.03)	(0.04)	(0.017)	(0.017)	(0.017)
march/april × self-employed pre-Covid	-0.13***	-0.11***	-0.11***	-0.008***	-0.009***	-0.011***
	(0.03)	(0.02)	(0.03)	(0.002)	(0.003)	(0.003)
may × self-employed pre-Covid	-0.09**	-0.08**	-0.10***	0.010	0.003	0.011
	(0.03)	(0.03)	(0.03)	(0.013)	(0.014)	(0.014)
june × self-employed pre-Covid	-0.03	-0.03	-0.02	0.001	-0.003	-0.000
	(0.03)	(0.03)	(0.03)	(0.012)	(0.012)	(0.013)
september × self-employed pre-Covid	-0.00	-0.00	-0.00	0.006	0.002	-0.001
	(0.03)	(0.03)	(0.04)	(0.015)	(0.015)	(0.016)
december × self-employed pre-Covid	-0.03	-0.03	-0.04	0.038*	0.037*	0.041*
	(0.03)	(0.03)	(0.03)	(0.020)	(0.020)	(0.021)
march/april × age: between 36 and 55	0.01	0.02	0.01	-0.009**	-0.009**	-0.009**
	(0.02)	(0.02)	(0.02)	(0.004)	(0.004)	(0.004)
may × age: between 36 and 55	0.06**	0.07***	0.06**	-0.031***	-0.031***	-0.028***
	(0.02)	(0.02)	(0.03)	(0.011)	(0.011)	(0.010)
june × age: between 36 and 55	0.12***	0.12***	0.12***	-0.033***	-0.033***	-0.030***
	(0.02)	(0.02)	(0.02)	(0.010)	(0.010)	(0.010)
september × age: between 36 and 55	0.14***	0.14***	0.13***	-0.057***	-0.057***	-0.051***
	(0.03)	(0.03)	(0.03)	(0.012)	(0.012)	(0.013)
december × age: between 36 and 55	0.09***	0.09***	0.07***	-0.038***	-0.038***	-0.033***
	(0.03)	(0.03)	(0.03)	(0.012)	(0.012)	(0.012)
march/april × age: above 55	-0.04*	-0.03	-0.03	-0.003	-0.003	-0.002
	(0.02)	(0.02)	(0.02)	(0.005)	(0.005)	(0.006)
may × age: above 55	0.03	0.03	0.04	-0.006	-0.005	0.002
	(0.03)	(0.03)	(0.03)	(0.014)	(0.014)	(0.014)
june × age: above 55	0.10***	0.10***	0.09***	0.006	0.007	0.013
	(0.03)	(0.03)	(0.03)	(0.014)	(0.014)	(0.014)
september × age: above 55	0.06**	0.06*	0.05*	0.004	0.004	0.012
	(0.03)	(0.03)	(0.03)	(0.017)	(0.017)	(0.017)
december × age: above 55	0.00	0.00	-0.00	0.036**	0.036**	0.042**
	(0.03)	(0.03)	(0.03)	(0.018)	(0.018)	(0.018)
march/april × essential worker		0.17***	0.15***		-0.013**	-0.015**
		(0.02)	(0.03)		(0.005)	(0.007)
may × essential worker		0.09***	0.08**		-0.048***	-0.031**
		(0.03)	(0.03)		(0.013)	(0.013)
june × essential worker		0.10***	0.10***		-0.027**	-0.020
		(0.03)	(0.03)		(0.012)	(0.012)
september × essential worker		0.03	0.03		-0.031**	-0.023
		(0.03)	(0.04)		(0.016)	(0.017)
december × essential worker		0.03	0.04		-0.002	-0.002
		(0.03)	(0.03)		(0.017)	(0.018)
march/april × frac. work doable from home		0.24***	0.22***		-0.004	-0.006
		(0.02)	(0.02)		(0.007)	(0.007)
may × frac. work doable from home		0.15***	0.15***		-0.008	-0.022
		(0.03)	(0.03)		(0.016)	(0.018)
june × frac. work doable from home		0.10***	0.13***		0.006	-0.014
		(0.03)	(0.03)		(0.016)	(0.018)
september × frac. work doable from home		-0.04	-0.02		-0.004	-0.015
		(0.03)	(0.03)		(0.018)	(0.019)
december × frac. work doable from home		0.07**	0.09**		0.010	-0.008
		(0.03)	(0.04)		(0.019)	(0.021)
march/april × essential × work doable from home		-0.15***	-0.12***		0.013	0.017*
		(0.03)	(0.04)		(0.008)	(0.009)
may × essential × work doable from home		-0.19***	-0.16***		0.033*	0.036*
		(0.05)	(0.05)		(0.019)	(0.019)
june × essential × work doable from home		-0.16***	-0.19***		0.006	0.015
		(0.05)	(0.05)		(0.019)	(0.019)
september × essential × work doable from home		-0.05	-0.06		0.012	0.005
		(0.06)	(0.06)		(0.026)	(0.025)
december × essential × work doable from home		-0.09*	-0.09*		-0.013	-0.010

Continued on next page

Table 3.D.3. Hours worked and not working by individual and job characteristics

	change total working hours			no job		
	(1)	(2)	(3)	(4)	(5)	(6)
		(0.05)	(0.05)		(0.028)	(0.028)
march/april × sector: construction			0.32***			0.009
			(0.07)			(0.012)
may × sector: construction			0.29***			0.005
			(0.08)			(0.044)
june × sector: construction			0.05			0.030
			(0.08)			(0.028)
september × sector: construction			0.03			-0.048
			(0.11)			(0.047)
december × sector: construction			0.09			-0.055
			(0.10)			(0.048)
march/april × sector: education			0.18**			0.001
			(0.07)			(0.011)
may × sector: education			0.05			-0.041
			(0.08)			(0.038)
june × sector: education			0.08			0.012
			(0.08)			(0.022)
september × sector: education			-0.05			-0.024
			(0.10)			(0.046)
december × sector: education			0.01			-0.033
			(0.09)			(0.047)
march/april × sector: env., culture, recr.			0.09			0.010
			(0.08)			(0.017)
may × sector: env., culture, recr.			0.09			-0.018
			(0.08)			(0.043)
june × sector: env., culture, recr.			-0.15*			0.043
			(0.09)			(0.032)
september × sector: env., culture, recr.			-0.07			0.015
			(0.11)			(0.056)
december × sector: env., culture, recr.			-0.04			-0.048
			(0.11)			(0.050)
march/april × sector: financial & business services			0.25***			0.006
			(0.07)			(0.009)
may × sector: financial & business services			0.19**			-0.010
			(0.08)			(0.039)
june × sector: financial & business services			-0.02			0.043*
			(0.08)			(0.025)
september × sector: financial & business services			-0.06			-0.023
			(0.10)			(0.045)
december × sector: financial & business services			0.04			-0.024
			(0.09)			(0.046)
march/april × sector: healthcare & welfare			0.25***			0.010
			(0.07)			(0.010)
may × sector: healthcare & welfare			0.21***			-0.051
			(0.07)			(0.037)
june × sector: healthcare & welfare			0.02			0.008
			(0.08)			(0.021)
september × sector: healthcare & welfare			-0.03			-0.052
			(0.10)			(0.044)
december × sector: healthcare & welfare			0.02			-0.049
			(0.09)			(0.046)
march/april × sector: industry			0.25***			0.001
			(0.07)			(0.009)
may × sector: industry			0.17**			-0.026
			(0.07)			(0.038)
june × sector: industry			0.04			0.019
			(0.08)			(0.023)
september × sector: industry			-0.02			-0.056
			(0.10)			(0.044)
december × sector: industry			0.06			-0.059
			(0.09)			(0.045)
march/april × sector: other			0.25***			0.006
			(0.07)			(0.010)
may × sector: other			0.16**			-0.025

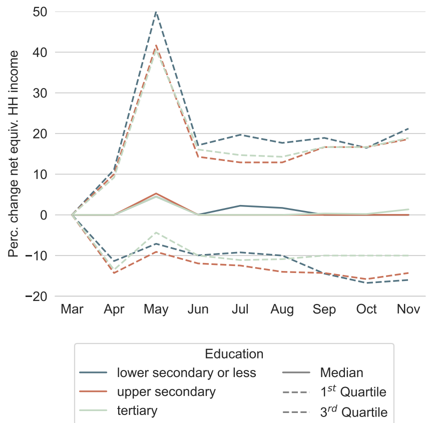
Continued on next page

Table 3.D.3. Hours worked and not working by individual and job characteristics

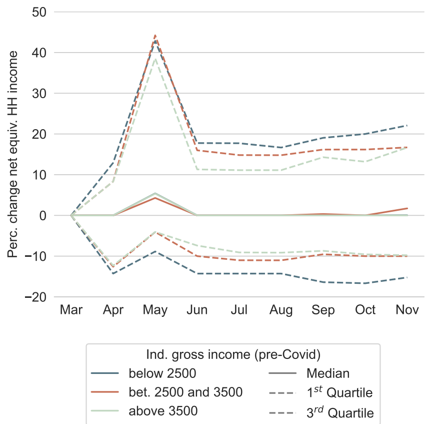
	change total working hours			no job		
	(1)	(2)	(3)	(4)	(5)	(6)
			(0.07)			(0.038)
june × sector: other			-0.00			0.033
			(0.08)			(0.024)
september × sector: other			-0.05			-0.034
			(0.10)			(0.045)
december × sector: other			0.06			-0.040
			(0.09)			(0.046)
march/april × sector: public services			0.24***			-0.002
			(0.07)			(0.009)
may × sector: public services			0.14*			-0.021
			(0.07)			(0.039)
june × sector: public services			-0.04			0.034
			(0.08)			(0.025)
september × sector: public services			-0.04			-0.045
			(0.10)			(0.045)
december × sector: public services			0.02			-0.034
			(0.09)			(0.046)
march/april × sector: retail			0.22***			0.007
			(0.07)			(0.010)
may × sector: retail			0.18**			0.000
			(0.07)			(0.040)
june × sector: retail			0.03			0.047*
			(0.08)			(0.027)
september × sector: retail			-0.08			-0.037
			(0.10)			(0.046)
december × sector: retail			0.09			-0.043
			(0.10)			(0.046)
march/april × sector: transport, communication, & utilities			0.22***			0.000
			(0.08)			(0.009)
may × sector: transport, communication, & utilities			0.15*			-0.012
			(0.08)			(0.040)
june × sector: transport, communication, & utilities			-0.10			0.062**
			(0.08)			(0.031)
september × sector: transport, communication, & utilities			-0.09			-0.035
			(0.11)			(0.047)
december × sector: transport, communication, & utilities			0.01			-0.010
			(0.10)			(0.050)
N	15738	15738	15133	15796	15796	15181
R^2	0.159	0.173	0.182	0.073	0.077	0.077

Dependent variable in the first columns are unconditional working hours. This part of the table shows the full set of covariates for the regressions shown in Table 3.4.2. The dependent variable in the last three columns is a dummy variable if the individual is either out of the laborforce or unemployed. Standard errors are clustered on the individual level. The data are an unbalanced panel restricted to individuals who worked more than ten hours in early March. Reference period = Early March. Notes: *p<0.1; **p<0.05; ***p<0.01.

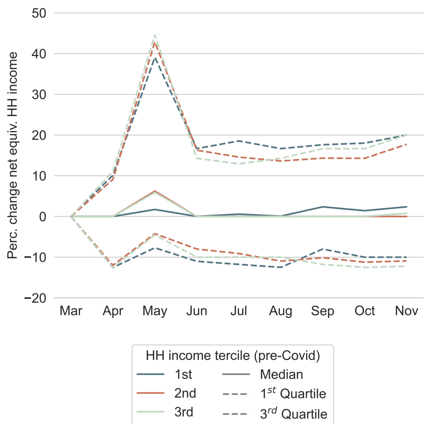
3.D.3 Predictors of household income



(a) Relative changes in net equivalized household income by education



(b) Relative changes in net equivalized household income by pre-Covid individual gross income



(c) Relative changes in net equivalized household income by pre-Covid household income

Figure 3.D.6. Relative changes in net equivalized household income by socio-economic status

Notes: Relative change of net equivalized household income relative to the average of January and February 2020. Pre-Covid household income tercile calculated by using the terciles of the average household income of 2018 and 2019. Sample: $18 \leq \text{age} \leq 66$, working pre-Covid, report positive household income in either January or February. In May, a vacation bonus is paid out, which is prescribed by law to be at least 8% of the yearly gross income. See https://wetten.overheid.nl/BWBR0002638/2017-01-01#HoofdstukIII for more information.

Table 3.D.4. Net equivalized household income by characteristics

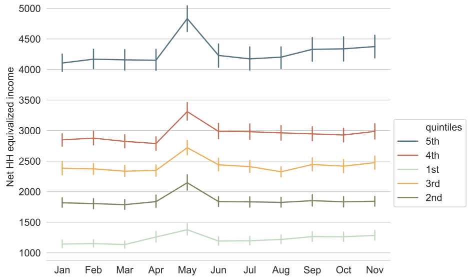
	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov
All	2395	2406	2381	2421	2792	2454	2435	2425	2489	2482	2519
Employment status pre-CoVid											
employed	2727	2750	2737	2773	3261	2816	2793	2789	2833	2822	2877
self-employed	2787	2821	2597	2491	2756	2922	2857	2745	3022	2959	2973
not working	1603	1591	1586	1714	1928	1644	1653	1643	1698	1706	1716
Initial employment shock											
decreased at least 20h	2404	2394	2159	2151	2609	2313	2265	2350	2486	2431	2511
decreased less than 20h	2545	2585	2560	2517	2915	2641	2607	2549	2602	2582	2638
did not decrease	2363	2372	2366	2442	2828	2451	2441	2433	2490	2486	2516
Policy Take-up											
Affected by policy, March-Sept	2655	2678	2525	2498	2893	2705	2698	2699	2601	2575	2610
Affected by policy, March-May	2512	2567	2485	2380	2772	2707	2637	2653	2820	2732	2753
Affected by policy, June-Sept	2567	2564	2504	2498	3005	2563	2539	2496	2767	2773	2770
Never affect by policy	2835	2877	2862	2848	3362	2907	2895	2881	2954	2950	3009
Reason for reduction											
closure	2354	2360	2209	2161	2527	2273	2278	2297	2463	2432	2462
less business	2456	2469	2388	2383	2683	2472	2443	2460	2481	2438	2478
care	2894	3004	2853	2816	3373	3194	2986	3104	3275	3263	3280
other	2617	2670	2692	2737	3299	2764	2724	2633	2669	2644	2772
no reduction	2356	2363	2359	2428	2811	2447	2436	2425	2479	2478	2506
Income quintile pre-Covid											
1st	1143	1151	1135	1259	1381	1195	1200	1221	1269	1267	1289
2nd	1826	1813	1796	1847	2156	1841	1839	1827	1877	1859	1868
3rd	2382	2371	2332	2343	2720	2443	2411	2330	2428	2401	2458
4th	2849	2876	2823	2788	3307	2987	2978	2960	2945	2926	2985
5th	4111	4173	4162	4154	4835	4229	4177	4208	4335	4343	4380

Notes: Average monthly net equivalized household income by characteristics. Long run income quintile calculated by using the quintiles of the average household income of 2018 and 2019. Sample: $18 \leq \text{age} \leq 66$.

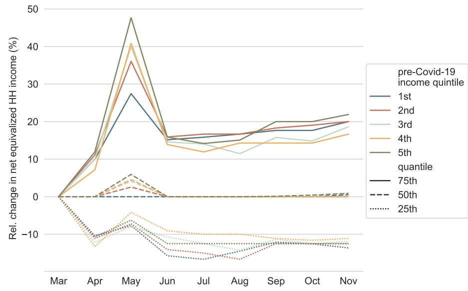
Table 3.D.5. Relative change in equivalized household income by characteristics

month	Mar			Apr			May			June			July			Aug			Sep			Oct			Nov		
quantile	p25	p50	p75	p25	p50	p75	p25	p50	p75	p25	p50	p75	p25	p50	p75	p25	p50	p75	p25	p50	p75	p25	p50	p75	p25	p50	p75
All	0	0	0	-12	0	10	-7	2	40	-13	0	14	-14	0	14	-14	0	15	-13	0	17	-13	0	17	-13	0	20
Employment status pre-CoVid																											
employed	0	0	0	-11	0	10	-4	6	46	-9	0	15	-10	0	14	-10	0	14	-11	0	17	-11	0	17	-11	1	18
self-employed	-20	0	0	-33	-7	11	-29	0	21	-25	0	22	-29	0	20	-29	0	20	-29	0	25	-31	0	25	-31	0	25
not working	0	0	0	-12	0	12	-10	0	33	-19	0	12	-20	0	12	-20	0	14	-14	0	18	-14	0	20	-14	0	20
Initial employment shock																											
decreased at least 20h	-8	0	0	-29	0	7	-20	0	40	-20	-2	11	-20	0	12	-20	0	12	-19	0	21	-20	0	20	-20	0	24
decreased less than 20h	0	0	0	-16	0	11	-11	4	43	-16	0	17	-18	0	17	-19	0	17	-17	0	17	-19	0	17	-17	0	20
did not decrease	0	0	0	-12	0	12	-7	4	40	-12	0	14	-12	0	14	-13	0	14	-12	0	17	-12	0	17	-12	0	19
Policy Take-up																											
Affected by policy, March-Sept	0	0	0	-16	0	7	-12	0	37	-10	0	23	-12	0	21	-11	0	18	-12	0	14	-14	0	14	-16	0	20
Affected by policy, March-May	0	0	0	-16	0	9	-12	0	29	-17	0	17	-17	0	16	-18	0	20	-22	0	20	-24	0	18	-24	0	20
Affected by policy, June-Sept	0	0	0	-11	0	2	-8	2	49	-9	0	17	-8	0	17	-9	0	17	-9	0	22	-9	0	21	-9	0	21
Never affect by policy	0	0	0	-12	0	11	-4	7	48	-10	0	16	-10	0	15	-10	0	15	-10	1	17	-10	1	17	-10	2	20
Reason for reduction																											
closure	0	0	0	-22	0	11	-17	2	46	-20	0	17	-19	0	18	-17	0	19	-17	0	24	-19	0	25	-17	1	25
less business	0	0	0	-17	0	10	-12	0	30	-17	0	14	-17	0	14	-18	0	14	-17	0	14	-21	0	14	-20	0	15
care	0	0	0	-19	0	4	-12	5	40	-18	0	13	-20	-2	12	-15	-1	12	-9	0	17	-9	0	17	-9	0	17
other	0	0	0	-11	0	12	-5	10	56	-14	0	16	-16	0	14	-18	0	14	-17	0	19	-17	0	19	-17	1	23
no reduction	0	0	0	-12	0	11	-7	4	40	-12	0	14	-12	0	14	-13	0	14	-12	0	17	-12	0	17	-12	0	19
Income quintile pre-Covid																											
1st	0	0	0	-10	0	10	-8	0	27	-16	0	15	-17	0	16	-15	0	17	-12	0	18	-13	0	18	-14	0	20
2nd	0	0	0	-11	0	11	-7	3	36	-14	0	16	-15	0	17	-17	0	17	-13	0	18	-13	0	19	-12	0	20
3rd	0	0	0	-12	0	10	-8	4	40	-10	0	15	-12	0	14	-14	0	11	-11	0	16	-12	0	15	-12	0	18
4th	0	0	0	-13	0	7	-4	5	40	-9	0	14	-10	0	11	-10	0	14	-11	0	14	-11	0	14	-11	0	17
5th	0	0	0	-11	0	12	-6	6	48	-12	0	16	-12	0	14	-12	0	15	-12	0	20	-12	0	20	-12	1	22

Notes: Quartiles of the relative changes in net equivalized household income by characteristics. Long run income quintile calculated by using the quintiles of the average household income of 2018 and 2019. Sample: $18 \leq \text{age} \leq 66$ and household income positive in January or February 2020.



(a) Net equivalized household income by long-run income quintile pre-Covid



(b) Relative et equivalized household income by long-run income quintile pre-Covid

Figure 3.D.7. Evolution of net equivalized household income by pre-Covid income quintile.

Notes: Net equivalized household income by long run income quintile. Long run income quintile calculated by using the quintiles of the average household income of 2018 and 2019. Sample: $18 \leq$ age ≤ 66

Table 3.D.6. Quantile regression: household income and pre-Covid income quintiles

	Rel. change net equiv. HH inc. (%)		
	p25	p50	p75
Apr	-16.48***	0	21.07***
	(4.43)	(0.49)	(4.65)
May	-11.66**	11.89***	44.3***
	(4.65)	(3.37)	(6.49)
Jun	-14.93***	0	26.5***
	(4.66)	(1.01)	(3.89)
Sep	-14***	3.78	24.95***
	(4.89)	(2.52)	(4.52)
Apr × 2nd income quintile	-0.3	0	0.73
	(4.26)	(0.55)	(4.57)
Apr × 3rd income quintile	2.84	0	-0.07
	(3.56)	(0.5)	(4.77)
Apr × 4th income quintile	-0.53	0	-3.67
	(3.56)	(0.49)	(3.99)
Apr × 5th income quintile	2.32	0	0.29
	(3.77)	(0.5)	(4.1)
May × 2nd income quintile	1.24	-0.53	1.23
	(4.82)	(3.17)	(7.56)
May × 3rd income quintile	2.86	2.54	9.13
	(4.61)	(3.29)	(8.7)
May × 4th income quintile	6.89	1.87	9.13
	(4.3)	(3.02)	(6.7)
May × 5th income quintile	3.12	3.43	9.15
	(4.05)	(3.66)	(7.56)
Jun × 2nd income quintile	4.38	2.82**	0.06
	(4.78)	(1.35)	(4.21)
Jun × 3rd income quintile	3.46	0	-0.42
	(4.57)	(0.76)	(4.31)
Jun × 4th income quintile	4.36	0	-3.26
	(4.19)	(0.7)	(4.08)
Jun × 5th income quintile	-0.81	0	-2.61
	(4.85)	(0.6)	(4.06)
Sep × 2nd income quintile	-5.61	-3.92*	-4.79
	(5.99)	(2.2)	(4.85)
Sep × 3rd income quintile	-3.5	-4.65**	-5.54
	(4.96)	(1.88)	(4.82)
Sep × 4th income quintile	-6	-4.65**	-8.89**
	(4.6)	(2.01)	(3.76)
Sep × 5th income quintile	-8.29	-4.79**	-3.99
	(5.17)	(2.01)	(4.27)
Apr × work. hours (pre-Covid)	0.02	0	-0.21**
	(0.11)	(0)	(0.09)
May × work. hours (pre-Covid)	0.05	-0.21***	-0.22
	(0.11)	(0.07)	(0.2)
Jun × work. hours (pre-Covid)	0.09	0	-0.27***
	(0.09)	(0.02)	(0.08)
Sep × work. hours (pre-Covid)	0.25***	0.04	-0.06
	(0.07)	(0.05)	(0.09)
N	9030	9030	9030

Notes: Quantile regression of relative changes in net equalized household income on pre-Covid income quintiles. Standard errors clustered on the household level using wild bootstrapped procedure as proposed by Hagemann (2017) and implemented in the R package quantreg. Sample: $18 \leq \text{age} \leq 66$; employed or self-employed pre-Covid (early March) and working hours of at least 10h in early March; positive household income either in January or February 2020.

References

Hagemann, Andreas. 2017. “Cluster-Robust Bootstrap Inference in Quantile Regression Models.” *Journal of the American Statistical Association* 112 (517): 446–56. [261]

Chapter 4

Shift to remote work and the parental division of labor

Joint with Hans-Martin von Gaudecker, Lenard Simon, Christian Zimpelmann

4.1 Introduction

Despite some progress towards gender convergence in the division of labor within households in recent decades, in many countries, mothers still tend to assume a disproportionate share of childcare and domestic responsibilities, while fathers work outside the home. This pattern is at least partially driven by the need for at least one parent's job to be compatible with childcare needs. Most parents must be able to step in at short notice when children are unable to attend school or daycare due to illness or other reasons. These responsibilities are often taken on by mothers, who may choose jobs with fewer hours or greater flexibility in order to accommodate them. Fathers, on the other hand, typically specialise in market work – potentially driven by non-linear returns to working hours (Gicheva, 2013; Bick, Blandin, and Rogerson, 2022) which make it less attractive for parents to share market and non-market work equally.

* The data collection was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2126/1 – 390838866, by the Dutch Research Council (NWO) under a Corona Fast track grant (440.20.043), and by the IZA – Institute of Labor Economics. Gaudecker and Zimpelmann are grateful for financial support by the German Research Foundation (DFG) through CRC-TR 224 (Project A05). This research would not have been possible without the help of many others at the CoViD-19 Impact Lab, a research group initiated in Bonn in Mid-March 2020. Special thanks to the team at CentERdata, who made the surveys underlying this research possible in record time. We would like to thank Egbert Jongen for very helpful comments.

One promising approach to mitigating the gendered division of labor is, thus, to ensure that both parents' jobs are compatible with childcare responsibilities, without altering other factors such as remuneration. A potential avenue to achieve this is by increasing the amount of work that can be done from home, provided that employers do not use this as a means of selecting employees. Working from home inherently involves an increase in time spent at home, as well as a typically higher degree of schedule flexibility and a reduction in commuting time and associated frictions.

In this paper, we exploit the way the CoViD-19 pandemic has evolved in the Netherlands in order to quantify this channel. We do so by using representative survey data from the LISS Panel, an online survey based on a true probability sample of the Dutch population, combined with administrative labor market records from CBS Netherlands. We argue that among the multitude of effects that the pandemic had on family lives, we can isolate the effects of working from home for several reasons. First, schools and daycare were open in the Netherlands except for two (primary schools and daycare) to three (secondary schools) months in the spring of 2020. Consequently, total hours spent on childcare did not change in the months of November of 2020 or 2021 relative to 2019. Second, generous wage-support schemes were in place, which left income unchanged for most households and helped that the unemployment rate did not move much in general and actually decreased for parents. Third, we show that the potential for working from home has little explanatory power for hours worked from home just before the pandemic. This drastically changed with the onset of the pandemic and the government's advice to work from home. Put differently, the potential to work from home was there before the pandemic, but it was realized to a large extent only after March 2020.

We start out by showing that the gains in job flexibility through the shift to remote work are asymmetrically distributed among parents. On average, fathers gained more flexibility than mothers. This asymmetry is driven by two factors. First, fathers tend to work in jobs with a higher degree of remote work potential. Second, they work more hours, which is more important quantitatively.

Relying on time use data from the LISS Panel, we find that fathers as well as mothers use their newly gained job flexibility for childcare provision. Given the asymmetric changes in job flexibility, the gender gap in childcare provision decreased substantially. Before the pandemic, mothers provided 12.5 more hours of care to their children than their partners. In late 2021, this gap had shrunk to 9 hours. Two thirds of the decline can be attributed to families where fathers' remote working potential was high.

To investigate the effect of the shift to remote work on labor supply, we use labor market information on the full-population of Dutch parents contained in the Dutch administrative data provided by the Centraal Bureau voor de Statistiek (CBS). The larger and longer panel compared to the time use data of the LISS Panel enables us to detect more subtle changes in the labor supply as well as to implement a more sophisticated identification strategy.

Using the administrative data, we first show that a pre-existing trend of increasing full-time work among mothers strongly accelerated during the pandemic. We then aim to identify whether this acceleration is driven by fathers gaining more job flexibility using an identification strategy that resembles an Event-Study combined with a Difference-in-Differences approach with continuous treatment. That is, we compare the relationship of partners' remote work potential and own working hours over the 2018–2021 period with the same relationship between 2013 and 2016. We find that mothers and fathers indeed increase their labor supply in response to their partners' newly gained job flexibility. Given that fathers gained asymmetrically more job flexibility, this in turn means that mothers disproportionately increased their labor supply. We find no evidence that this result is driven by reduced commuting time. Instead, it seems that the increased availability of parents at home drives the results.

Our results, thus, suggest that increased possibilities to work from home allowed couples to choose a more balanced distribution between market and non-market work. More generally, it highlights that policies which make it easier to combine career ambitions and childcare time can be effective in reducing gender inequality within households.

Our results are related to several strands of the literature. First, women and in particular mothers have preferences for jobs with higher employee-side flexibility and tend to work in more flexible jobs than men. Mas and Pallais (2017) find that in the U.S., mothers of younger children have a higher willingness to pay for remote work, as well as to avoid employer scheduling discretion. Consistent with that, U.S. women have a higher willingness to pay for flexible work arrangements as measured by the option to do part-time work and for job stability (Wiswall and Zafar, 2018). In Germany, Felfe (2012) finds suggestive evidence that women who change their job after child birth choose jobs with more schedule flexibility. Magda and Lipowska (2021), studying the distribution of job flexibility all over Europe, find that the likelihood of mothers working in positions with schedule flexibility does not differ strongly from that of fathers, with the exception of Anglo-Saxon countries where mothers are more likely to have schedule flexibility. However, across all countries, women are less likely to work in positions with a high degree of employer scheduling discretion.

Furthermore, even within the same jobs, women choose more flexible working schedules which are more aligned with childcare needs. For example, Houghton (2020) analyzes a wealth of publicly available records of workers' coding activity on GitHub. Examining the impact of unexpected, weather-related public school closures, she finds that women starkly reduce their work activities in response to childcare shocks, while men do not react at all. Similarly, Adams-Prassl (2021) analyzes gender differences in crowdwork on Amazon mechanical turk. She finds that women who do crowdwork are more likely than men who do crowdwork to interrupt their tasks, which leads them to earn 20% lower wages on average. These effects are concentrated among mothers with children at home. Such patterns, however, need not persist everywhere. The overall picture emanating from the literature is that women do take direct wage hits in order to be able to provide childcare.

Second, a set of papers examines the relation of job flexibility and gendered labor market outcomes. Le Barbanchon, Rathelot, and Roulet (2021) find that in France, women search for jobs within a smaller commuting radius than men, which leads to a subsequent wage penalty in outcomes. Meekes and Hassink (2022) find a similar result for the Netherlands among individuals displaced because of firm bankruptcies; women's working hours in their subsequent job are differentially lower than men's, too. Constructing an occupation-level measure of flexibility, Bang (2021) shows that the flexibility of both partners in the year before child birth affects the child penalty. Mothers whose partners are working in flexible jobs experience smaller drops in earnings and wages in the medium run. Pointing to the role of other care options, Cortés and Pan (2019) use inflows of low-wage migrants as an exogenous change in the supply of housework, which leads women to move to occupations with higher returns to long working hours. Goldin (2014) shows that differences in flexibility of work arrangements across genders is strongly related to the remaining gender wage gap.

Our paper complements the aforementioned literature on gendered patterns of market work and non-market work and their relationship to job flexibility in two ways. First, these studies typically focus only on the relationship between job flexibility and labor supply while implicitly assuming that the effects operate through childcare provision. We make this explicit, by investigating childcare provision directly. Second and more importantly, we provide the first causal evidence for the effect of job flexibility on labor supply. We take advantage of "windfall" gains in job flexibility induced by the CoVid-19 pandemic and, thereby, circumvent the typically encountered problem that job characteristics and labor supply are jointly determined.

Furthermore, a wide-range of studies analyzes the effect of the pandemic on the intra-household allocation of labor. A wealth of papers looks at how couples share the increased childcare burden early in the pandemic while childcare facilities and schools were closed in a wide-range of countries.[1] The evidence is mixed, sometimes even within the same country, but in most cases the childcare gap increased in absolute but decreased in relative terms. Alon, Coskun, Doepke, Koll, and Tertilt (2022) look at the effect on the labor market and find that the pandemic let to a 'shecession' in many countries—however, interestingly not so in the Netherlands. This is consis-

1. A non-exhaustive list encompasses data collections in the UK (Andrew, Cattan, Costa Dias, Farquharson, Kraftman, et al., 2020; Sevilla and Smith, 2020), Italy (Del Boca, Oggero, Profeta, and Rossi, 2020; Mangiavacchi, Piccoli, and Pieroni, 2021), Spain (Farré, Fawaz, González, and Graves, 2020), Germany (Hank and Steinbach, 2020; Jessen, Spiess, Waights, and Wrohlich, 2022), and the US (Zamarro and Prados, 2021; Pabilonia and Vernon, 2022).

tent with Meekes, Hassink, and Kalb (2020) who find the same (small) negative effects on average for men and women and no differential effect for parents in couples. For the US, Heggeness and Suri (2021) find negative labor supply effects for mothers compared to fathers and compared to women without children in a period in which the closure of childcare facilities and schools was frequent in the U.S.. For the first 9 months of the pandemic, they find that negative labor supply shocks were slightly larger for mothers in remote work jobs. Their interpretation is that parents in onsite occupations were not exposed to the same level of intense simultaneous multitasking of increased childcare duties and working. We contribute to this literature in two ways. First, we extend the time horizon to more than one and a half years, thus focusing on the medium term effects of the pandemic. Second, by studying an institutional setting in which childcare facilities and school closures played only a minor role in the medium term, we can isolate the effect of the acceleration in remote work on both labor supply and childcare provision.

Our paper is structured as follows. We describe our data and the basic socioeconomic characteristics in the next section. Subsequently, we present the setting of our analysis: The way the pandemic evolved in the Netherlands, background on trends in parents' labor supply and childcare division, and our measures of job flexibility. In Section 4.4, we present our results on the effects of the pandemic on parents' childcare division and labor supply. We conclude in Section 4.5.

4.2 Data sources, sample selection, and basic demographics

Our study is based on customized survey data from the LISS panel, population-wide administrative records from Statistics Netherlands, and both datasets linked at the individual level. We describe both data sources in the following subsections. The last subsection describes the basic socio-demographic characteristics of our sample.

4.2.1 Customized survey data from the LISS Panel

In our study, we use the Longitudinal Internet Studies for the Social Sciences (LISS) panel. The LISS panel is based on a probability sample of individuals registered by Statistics Netherlands; it has been running since 2007 and comprises about 7000 individuals in 4000 households. The LISS panel is administered by CentERdata, a survey research institute affiliated with Tilburg University, the Netherlands. Each year, the LISS panel runs ten core surveys, which cover a wide range of topics, including health, education, work, and family. Taken together, these data are comparable in scope to popular surveys like the Panel Study of Income Dynamics (U.S.), Understanding Society (U.K.), or the Socio-Economic Panel (Germany).

On top of that, the LISS panel allows researchers to run their own questionnaires. In this paper, we make use of two sets of surveys that we ran ourselves or helped design.

First, in the period between mid-March and December 2020, we fielded six questionnaires on the impact that the CoViD-19 pandemic had on peoples' lives. From those surveys, we employ information on remote work potential and on working hours at the point in time just before the pandemic started affecting working lives. The documentation of the entire questionnaires can be found in von Gaudecker, Zimpelmann, Mendel, Siflinger, Janys, et al. (2021). In May 2020, we ask participants "What percentage of your normal work *prior to the coronavirus outbreak* can you do while working from home?". We repeated this question in December 2020, but inquired about the share of tasks at the current job that can be done from home instead of the pre-pandemic situation.[2] The resulting answers measure the remote work potential, abstracting from any changes in task content that happened during the period of social distancing. The fact that we ask this when the pandemic was already in full swing allows individuals to better assess the *potential* for remote work – it would not have occurred to many people that essentially all meetings could be held in virtual formats. The correlation between the measure in May and the measure in December is 0.82.[3] Given the high stability of this measure, we take the mean across these two dates for each individual for which we observe both, and the one

2. The question in December 2020 reads: "What percentage of your normal work can you do while working from home?"

3. In the Appendix of Zimpelmann, Gaudecker, Holler, Janys, and Siflinger (2021), we discuss the correlation between the answers in May and December as well as the distributions in great details.

that is available for those who we observe only once. This maximizes the number of observations for which we have this measure available while at the same time reducing potential measurement error if one believes that the variable is approximately stable across time.

Second, we employ time use information collected in comparable questionnaires in November 2019, April 2020, November 2020, and November 2021. In these surveys, people are asked to distribute the hours of the past week over different activities. We use the information on time spent working (beginning with the April 2020 wave, these hours are recorded separately by whether work was done at the usual workplace or at home), commuting, and on childcare. See van Soest, Been, Pinger, von Gaudecker, and Centerdata (2019), von Gaudecker and Centerdata (2020a), von Gaudecker and Centerdata (2020b), and Been and Centerdata (2021), respectively, for the documentation of the four questionnaires.

4.2.2 Population-wide administrative data, Working Conditions Survey

We access detailed administrative microdata from Statistics Netherlands (CBS) via a secure online environment which we use in our analyses in two ways: First, we obtain more precise measures of household composition and labor supply for our survey sample. Second and more importantly, we can greatly expand our sample for the analysis of labor supply and consider the full population of the Netherlands. We make use of gender, household composition, education, labor force status (dependent work in full time or part time, self-employment, unemployment, and being outside the labor force), sector, commuting distance, and working hours.

The labor market information is recorded monthly for each individual. To harmonize the CBS data with the LISS data (and for computational feasibility), we extract the labor market information in November of each year from 2013 to 2021. We use actual working hours, which are recorded at the spell level. Spells cover one month in case an employee works the whole month, and shorter than a month in case he or she does not work the whole month. We convert them to weekly hours throughout.

The administrative data does not contain direct information on remote work. We thus impute remote work ability based on the National Working Condition Survey (NEA). Using survey information on actual remote work in the fall of 2020 from 35,000 individuals, we calculate the average share of remote work by sector and education. We then use this information to impute a measure of potential remote work for all individuals in our data. See Appendix 4.A.1 for a more detailed description of the imputation procedure.

Finally, we are able to link our LISS survey data to the CBS data at the individual level. Doing so is possible for all panel members who gave their consent to the linkage, which holds true for around 90 % of individuals in our sample. For these respondents, we are able to update information on working hours and household composition, which is particularly useful when these individuals did not participate in one or more waves of the survey.

4.2.3 Statistics on socio-demographic variables

Throughout our analysis, we consider heterosexual couples where both partners are between 18 and 55 years of age and who have at least one child below the age of 16 in the household. For some of our analyses below, we require information that is missing for a subset of individuals. For example, we can calculate our measure of potential remote working hours in the LISS panel only if people were working just before the pandemic; in the CBS data we do not observe working hours for self-employed individuals. Where applicable, we exclude individuals with data that is missing by construction from our sample. If doing so affects the descriptive statistics shown here, we comment on it.

Table 4.2.1 display the socio-demographic characteristics in the two samples, we mostly rely on in our analysis pooled across time. It reveals that most socio-demographic statistics line up well between the LISS sample and the population data. Mothers are somewhat younger than fathers, families comprise slightly more than 2 children on average and the age of the youngest child falls just below the middle of the age interval we require.

The one exception is that respondents in the LISS panel are better educated. In particular, 3% of parents do not have a secondary degree. This compares to 10% in the CBS data and it is a well-known bias in surveys. The composition of our LISS sample changes somewhat over time. In particular, the average age of the youngest child is lower for mothers who respond in 2021 compared to 2019 and April 2020 (6.3 years vs 7.2 years, see Table 4.A.2a). This will affect the analysis of childcare hours below, where it will be important to control for the age of the youngest child.

Table 4.2.1. Socio-demographic variables by data source and gender pooled over time

	LISS		CBS	
	Fathers	Mothers	Fathers	Mothers
Age	42.56	40.27	41.41	39.0
	(6.51)	(6.42)	(6.99)	(6.72)
Age youngest child	6.75	6.85	6.66	6.7
	(4.69)	(4.74)	(4.83)	(4.86)
Number of children	2.08	2.03	1.96	1.94
	(0.78)	(0.8)	(0.8)	(0.78)
Education: High	0.46	0.47	0.39	0.44
	(0.5)	(0.5)	(0.49)	(0.5)
Education: Middle	0.26	0.27	0.29	0.31
	(0.44)	(0.44)	(0.45)	(0.46)
Education: Low	0.04	0.03	0.08	0.07
	(0.21)	(0.18)	(0.27)	(0.25)
Education: Unknown	0.24	0.23	0.24	0.18
	(0.42)	(0.42)	(0.42)	(0.39)
Observations	1,044	1,190	3,304,273	3,322,747

Notes: The first column displays basic demographic characteristics of the LISS sample by gender pooled over all months. The age variable is taken directly from the LISS survey. The values for the variables age of youngest child and number of children are taken from the administrative records for all linked individuals and from the LISS survey for all those who are not linked. The education variable is taken from the administrative records and therefore only available for linked individuals (note that even for linked individuals it is possible that the education is unknown). The second column displays basic demographic characteristics of the of all working-age (18-55 years old) who were employed some time in 2018 and 2019 parents with a child below 16 years old by gender pooled over November 2018 - November 2021. The education variable is unknown if there is no available administrative record on the education for the individual. See Table 4.A.2 for the numbers over time.

4.3 Setting

In this section, we describe the broader environment for our analysis along with stylized features emanating from our data. First, we illustrate the policy environment during the first two years of the CoViD-19 pandemic. We then highlight some key features of the parental division of labor regarding market and non-market work before and during the pandemic. Finally, we go through our measures of remote work—both the potential for doing so and its realizations—over the period of our analysis.

Taken together, from the contents of this section it becomes clear why we deem it plausible that we can isolate the effect of remote work ability on parents' outcomes during the time period of our analysis.

4.3.1 The CoViD-19 pandemic in the Netherlands

From March 2020 until the end of our data collection in November 2021, a set of measures were in place to slow the spread of the SARS-CoV-2 virus in the Netherlands. We will highlight the policy environment that was effective in and around the months of our data collection: November 2019, April 2020, November 2020, and November 2021. In general, measures were more lenient than in many other countries; in particular, no general curfew or stay-at-home mandate was in place at any point in time. We describe some key features relevant for our analysis; see Zimpelmann et al. (2021) for a more detailed description with a focus on labor market issues during the first year of the pandemic.

Figure 4.3.1. Timeline of relevant government policy measures at the points in time of our data collection.

Notes: The policy measures are obtained from the official government recommendations, which can be found on https://www.government.nl/latest/news. The unemployment rates are taken from the official statistics from CBS Netherlands.

Figure 4.3.1 shows the timeline of relevant government policy measures at the points in time for which we have data. In November 2019, the world lived in blissful ignorance of SARS-CoV-2's existence. In mid-March 2020, limits on social gatherings were imposed and many businesses involving personal contacts were closed, such as restaurants, bars, and hairdressers. However, others like stores for clothes or utilities remained open as long as they were able to to maintain the social distancing rules. Public locations were accessible and the use of public transportation was possible.

Many of these restrictions were lifted over the summer of 2020. The majority, however, were in place again during November 2020. After the winter, they were eased again and only much milder measures came back in the subsequent fall/winter.

With the onset of the initial restrictions, schools and childcare facilities were closed for a period of two (daycare, primary schools) to three (secondary schools) months. In the late spring and summer of 2020, policy makers made it very clear that schools and childcare facilities would be the last institutions to close again in case of renewed tightening of restrictions. Except for slightly prolonged vacations around Christmas 2020, this promise was kept. Actual closures were thus very limited in comparison to many other countries.

A comprehensive set of economic support measures accompanied the social distancing restriction. The largest and most influential policy was the short-term allowance (Noodmaatregel Overbrugging voor Werkgelegenheid, NOW), which subsidized labor hoarding with a 100% wage replacements rate. In turn, rather few people became unemployed or dropped out of the labor force; dependent employees did not see their incomes drop (Zimpelmann et al., 2021). Figure 4.3.1 shows that the unemployment rate was low throughout the 2019–2021 period. While it was slightly higher in November 2020 than in November 2019, it fell well below these levels in November 2021.

Throughout the part of our sample period that coincides with the CoViD-19 pandemic, the government strongly encouraged remote work.

4.3.2 Market and non-market work

Parents' labor force participation was high before the pandemic and increased further during 2020 and 2021. The distribution over different categories of employment (full-time employed, part-time employed, self-employed) or lack thereof (unemployment, out of the labor force) varies considerably with gender.

Table 4.3.1 contains the labor market status for our sample of parents for the months of November in the 2016–2021 period. The first two columns in the upper panel show that the share of mothers who are not working decreased considerably over those years. To be precise, the fraction outside the labor force went from 23% to 19%; the unemployed share decreased from 3% to 0.9%.[4] In the bottom panel, we also see the same trend for fathers, albeit at lower levels. The fraction outside the

4. Note that unemployment is measured as receipt of unemployment benefits, so by ordinary economic definitions, we might be putting some individuals into the wrong category of inactivity.

labor force went from 10.5% to 9.3%; the fraction of unemployed fathers decreased from 2.5% to 0.7%.

Table 4.3.1. Labor market status over time

		Out of the labor force	Unemployed	Self-employed	Part-time employed	Full-time employed
Mothers	2016	0.232	0.030	0.093	0.555	0.092
		(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
	2017	0.226	0.018	0.097	0.568	0.092
		(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
	2018	0.214	0.013	0.101	0.576	0.097
		(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
	2019	0.206	0.012	0.107	0.576	0.099
		(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
	2020	0.203	0.013	0.113	0.568	0.106
		(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
	2021	0.189	0.009	0.118	0.568	0.117
		(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
Fathers	2016	0.103	0.025	0.151	0.092	0.631
		(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
	2017	0.106	0.014	0.154	0.100	0.628
		(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
	2018	0.102	0.010	0.160	0.104	0.626
		(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
	2019	0.099	0.009	0.168	0.108	0.617
		(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
	2020	0.099	0.011	0.173	0.108	0.610
		(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
	2021	0.093	0.007	0.177	0.111	0.613
		(0.001)	(0.000)	(0.001)	(0.001)	(0.001)

Notes: The table shows the labor market participation for a 20 % sample of all working-age (18-55 years old) parents with a child below 16 years of age by month and gender. Individuals are classified as unemployed when they are receiving unemployment benefits and classified as out of the labor force when there are no working hours, no self-employment status, and no unemployment benefits recorded in the administrative data. Consistent with the official definition of CBS Netherlands, we classify individuals to be working part-time if they work less than 35 hours per week.

Hence, when it comes to the extensive margin, trends of increasing employment of parents continued or even accelerated during pandemic's first two years. Comparing the numbers with aggregate employment trends, parents experienced only a negligible uptick in unemployment/inactivity during 2020; they were thus less affected than the rest of the population. This may partly be explained by the type of jobs (e.g., relatively few parents work in the catering sector). Importantly for our purposes—and in stark contrast to countries where schools and daycare facilities were closed for prolonged periods of time (e.g. Heggeness and Suri, 2021)—there is no evidence that parents dropped out of the labor force to take care of their children.

The share of self-employed parents was high and increased over our observation period. For mothers, it started at 9.3% and gradually went up to 11.8% in late 2021. Among fathers, it increased from 15.1% to 17.7%. For mothers, the rise accounts for half of the decrease in the share of mothers who are not working. For fathers, the rise accounts for the entire decrease.

Mothers' part time employment went up from 55.5% in 2016 to peak at 57.6% in 2019, and then decreased again to 56.8%. In 2016, 9.2% of mothers were employed full-time – i.e. worked 35 hours or more. The share went up by 0.7 percentage points between 2016 and 2019 and increased by another 1.8 percentage points between 2019 and 2021. Hence, there was a strong acceleration in the increase of mothers' full-time employment during the first two years of the pandemic. As a result, 11.7% of mothers were employed full-time in 2021 as opposed to 9.9% in 2019 and 9.2% in 2016.

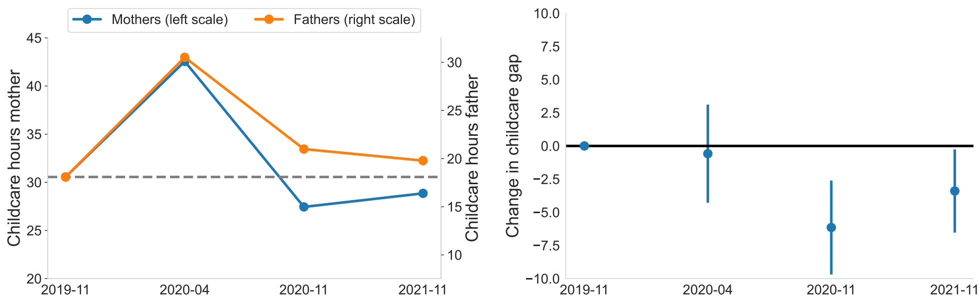
Fathers see a slight decrease in full-time employment and an increase in part-time employment over the observation period. In 2016, 63.1% of fathers worked 35 hours or more, while 9.2% worked less than 35 hours. The share of fathers in full-time employment decreased by 1.4 percentage points between 2016 and 2019, while the share of fathers in part-time employment increased by 1.6 percentage points to 10.8% in 2019. During the pandemic full-time employment dropped by another 0.7 percentage points until November 2020, but recovered again by 0.3 percentage points by November 2021. Hence, decreases in fathers' full-time employment over the entire pandemic period are similar to their pre-pandemic trends. Similarly, part-time employment of fathers increased only by additional 0.3 percentage points over the two years of the pandemic.

The trends described in the previous paragraph hold up when looking at working hours of dependent employees directly instead of categories. In particular, average working hours of mothers increased from 25.3 in 2016 to 26.1 in 2019. This trend accelerated during the pandemic and by 2021, mothers worked 27 hours on average (all numbers referred to in this paragraph are in Appendix Table 4.A.5). Among fathers, average working hours declined slightly from 38.6 in 2016 to 38.4 in 2019 and further to 38.3 at the end of our sample period.

Women take on a much larger share of childcare work than men. Figure 4.3.2a displays the evolution of childcare provision by fathers and mothers controlling for changes in household composition over time.[5] It shows that before the pandemic, mothers on average spent 29.6 hours per week providing care for their children. Fathers' childcare hours, with units depicted on the right axis, were well below that at 18.1 hours. The location of both lines is normalized so they visually start at the same level. This normalization makes the evolution during the pandemic salient.

During the period of closed schools and daycare facilities, combined childcare hours went up by about 25. This number is plausible given typical times spent at school/daycare and the fact that emergency childcare was available for parents

5. As previously mentioned we have a slight unbalancedness across time in the age of the youngest child in the LISS sample with parents in newer waves having on average younger children than those in older waves. To take out this variation we present results controlling for the (standardized) age of the youngest child and the number of children. Raw averages can be found in Table 4.A.10.



(a) Childcare hours **(b)** Change in the childcare gap

Figure 4.3.2. Evolution of the childcare gap 2019–2021.

Notes: Figure 4.3.2a shows the development of childcare hours by mothers and fathers in the LISS time use data. Figure 4.3.2b shows the development of differences in childcare provision between fathers and mothers. Both are based on a regression of childcare hours on the interaction of time dummies and gender, including as additional controls the number of children, and the standardized age of the youngest child interacted with gender. Standard errors clustered on the household level. The regression coefficients underlying the Figure are listed in Column (1) of Table 4.4.2.

working in essential occupations.[6] The large increase in April 2020 was distributed about equally among both genders.

When school and childcare facilities were reopened, total childcare provision recovered to pre-pandemic levels but now with mothers doing approximately 3 hours less, and fathers doing about 3 hours more childcare in November 2020 as compared to November 2019. Towards November 2021, mothers increase their childcare provision again by one and a half hours while fathers decrease their childcare provision by approximately 1.3 hours as opposed to November 2020.

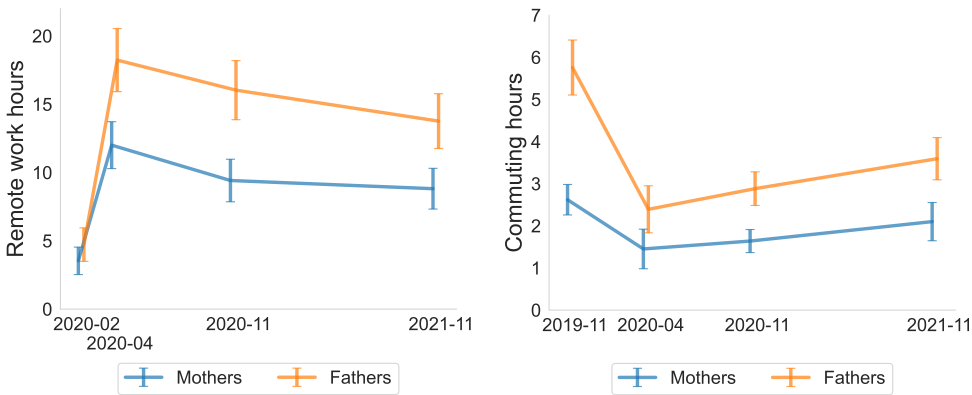
A different way to look at it is to consider the gap between genders directly. Figure 4.3.2b visualizes the result of this exercise, showing that the gender differences we described in the previous paragraphs are very robust in statistical terms. Normalizing the difference between mothers' and fathers' childcare provision to its pre-pandemic level, there virtually was no change in April 2020. Subsequently, the difference shrank by 3–6 hours. When accounting for statistical uncertainty, a range from 1 to 9 hours seems possible.

We will argue below that the change in the gender care gap can be explained by increased flexibility of parents when it comes to their work schedule and location. Next, we thus describe how remote work and commuting evolved over our period of study.

6. This was the most relevant difference for essential workers, hence we do not mention them elsewhere. See Zimpelmann et al. (2021) for a more detailed analysis of essential worker status.

4.3.3 Remote work and commuting

As early as 2016, the Netherlands introduced a law aimed at facilitating flexible work (Wet flexibel werken). This law defines processes and rights for employees to request adjustments to their working hours, their work schedules, or their work location. Before the CoViD-19 crisis, however, the effects were limited. E.g., ten Hoeve, Talman, van Mierlo, and Engelen (2021) find that 16% of employees made a request regarding flexible work along *any* of the three dimensions between 2016 and mid-March 2020. Consistent with those findings, our data shows that while 32% of individuals reported to have performed *some* work from home (see Appendix Table 4.A.6), the hours are very limited. Figure 4.3.3a reveals that on average, they are below five for mothers and fathers with fathers' remote hours being about 50 % higher than mothers'. Fathers on average spent almost 6 hours per week commuting to work; mothers spent about 2.5 (Figure 4.3.3b).



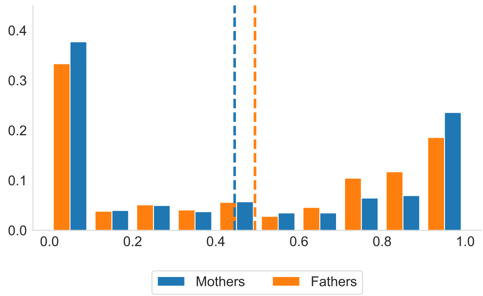
(a) Remote working hours

(b) Commuting hours

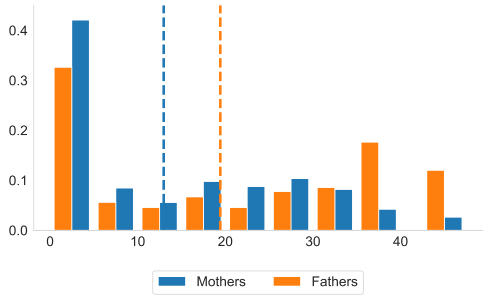
Figure 4.3.3. Realized work from home and commuting over time

Notes: Figure 4.3.3a displays average remote working hours in the LISS sample over time and by gender. Figure 4.3.3b displays average commuting hours in the LISS sample over time and by gender. For underlying numbers see Tables 4.A.8 and 4.A.9. Additionally, Table 4.A.6 contains the evolution over time and by gender of a variable measuring *any* remote work and Table 4.A.7 contains the evolution over time and by gender of the share of remote work. In the pre-pandemic period, remote working hours are measured in February 2020 and commuting hours in November 2019.

With the onset of the pandemic, these numbers changed dramatically for parents of both genders. In April 2020, weekly hours worked from home increased to 12 among mothers and 18 among fathers. Put differently, about fifty percent of actual hours were done from home. Commuting time dropped to 1.5 hours for mothers and 2.4 hours for fathers. Even as the pandemic progressed, all these numbers remained closer to the values they took during the initial lockdown than to their prior levels.



(a) Share of tasks that can be done from home



(b) Potential remote working hours

Figure 4.3.4. Remote working potential by gender

Notes: Figure 4.3.4a displays the distribution of the variable “share of tasks that can be done from home” by gender in the LISS sample. The vertical dashed lines display the mean by gender. Figure 4.3.4b shows the distribution of the variable potential remote work hours by gender in the LISS sample. Potential remote working hours are calculated by multiplying the share of tasks that can be done from home with the pre-covid (November 2019) working hours of an individual. The vertical dashed lines display the mean by gender. Samples conditional on working before the pandemic. Similar graphs for the CBS and NEA data are relegated to the Appendix, Figure 4.A.4.

These large differences prompt us to investigate the *potential* remote working hours. We argue that the capacity to work from home is roughly constant over the 2019-2021 period, but the extent to which this potential was realized changed due to the pandemic. Figure 4.3.4a shows the density of the reported share of tasks that can be done from home by gender in the LISS data, using the variable described in Section 4.2.1. The distribution is very polarized with the largest share of jobs admitting no remote work at all (37.7 % among mothers and 33.3 % among fathers). Fathers are also more likely to work in jobs where more than 50 % of tasks can be done from home (48.1 % vs. 43.9 %). Overall, the mean for mothers is 44.8 % and for fathers 48.3 %.

The overall potential gain in flexibility because of remote work is even larger for men because they work more hours. Figure 4.3.4b shows the distribution of potential remote working hours, which are obtained by multiplying the share of potential remote work from Figure 4.3.4a with the working hours just before the pandemic. More than 30 % of fathers can work at least 30 hours from home, while only 15 % of mothers can do so. The averages are 18.9 and 13.1 weekly hours, respectively.

The explanatory power of potential remote working hours for actual remote working hours is high and it increased dramatically during the pandemic. Column (1) of Table 4.3.2 shows that time dummies and family characteristics alone can explain about 8% of the variation in realized remote working hours. The R^2 increases to 0.46 when adding the remote work ability interacted with the time dummies. The coefficient on potential remote working hours increases from 0.2 before the

Table 4.3.2. Predictive power of potential remote working hours for realized hours worked from home and commuting time

	Remote working hours		Commuting hours	
	(1)	(2)	(3)	(4)
Constant	5.15***	1.16	4.64***	4.05***
	(1.16)	(0.76)	(0.27)	(0.37)
2020-04	11.81***	2.71***	−2.66***	−0.92*
	(0.87)	(0.87)	(0.28)	(0.49)
2020-11	9.54***	0.01	−2.26***	−0.82**
	(0.82)	(0.72)	(0.22)	(0.33)
2021-11	7.09***	0.43	−1.6***	−0.17
	(0.74)	(0.74)	(0.26)	(0.46)
Pot. hours remote work		0.21***		0.04**
		(0.04)		(0.02)
Pot. hours remote work × 2020-04		0.61***		−0.11***
		(0.06)		(0.02)
Pot. hours remote work × 2020-11		0.60***		−0.09***
		(0.05)		(0.02)
Pot. hours remote work × 2021-11		0.41***		−0.09***
		(0.05)		(0.02)
Observations	1,876	1,876	1,876	1,876
R^2	0.078	0.462	0.069	0.11

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors obtained by clustering on the household level. Sample conditional on working pre-CoVid. Baseline commuting hours based on LISS Time Use Survey from November 2019. Baseline remote work hours obtained from LISS Covid-19 Survey and based on February 2020. Sample restricted on parents who work in November 2019. For the full table see Table 4.B.1 in the Appendix. For the the interaction by gender see Table 4.B.2 in the appendix.

pandemic to 0.8 during its first year. That is, before the pandemic an hour of remote work potential translated into 12 minutes of actual remote work. During the initial lockdown, individuals worked more than 45 minutes remotely for every hour they could potentially do so. In late 2021, when overall remote work was slightly lower and more individuals may have changed jobs, the coefficient drops somewhat but remains high at 0.6 (i.e., 35 minutes for every potential hour).

Columns (3) and (4) of Table 4.3.2 reveal a similar pattern for realized commuting time as the dependent variable. Adding the interaction of potential remote working hours and time dummies leads to an increase in the R^2 by 60 %. Prior to the pandemic, a 40 hour job with the potential to do all tasks at home was associated 1.6 hours more time spent commuting compared to a job that would not admit any remote work. After the pandemic's onset, the relationship was reversed and commuting time was about 4 hours *less* for a person who works full-time and can do all his tasks from home.

Table 4.3.2 contains a very simple specification, not differentiating by gender. Adding the full set of interactions with gender in Table 4.B.2 does not reveal any gender differences in the take-up of remote work given equal remote work potential.

Actual remote work in the LISS data is consistent with the corresponding numbers from the much-larger working conditions survey (NEA, see Section 4.2.2). The NEA data also reveal a stark increase in remote work during the pandemic, from approximately 2.7 hours in late 2019 to 16 hours in late 2020. Further, investigating the remote work share by sectors (as a proxy for remote work potential), we find that in the pre-pandemic period, sectors only mildly predict the remote work share of their workers, while in late 2020 the share of hours a worker works remotely strongly depends on the sector he or she works in.[7] This supports our previous point that during the pandemic, remote work potential becomes much more important for its take-up, while take-up is more idiosyncratic before the pandemic.

Summing up, we find that remote working hours have strongly increased during the pandemic years. Before the pandemic, take up of remote work was low and rather idiosyncratic. Because of the pandemic, it became intimately tied to job characteristics. The potential hours that can be worked remotely strongly vary across genders. These hours are closely related to increases in actual remote work during the pandemic and to decreases in time spent commuting.

7. Details are in Appendix Section 4.A.1.

4.4 Results

Our main results establish that the trend towards a more equal division of childcare during the pandemic was entirely driven by households who gained flexibility because their potential to work remotely was realized. Similarly, we show in Section 4.4.2 that the same households are driving the acceleration of the trend towards mothers working longer hours.

4.4.1 Childcare

We first establish that the potential to work remotely had no effect on the hours spent on childcare before the pandemic and that this relationship changed dramatically with its onset. Beginning in early 2020, the potential to work remotely is closely associated with more time spent on childcare. We then show that remote work potential largely explains the decrease in the childcare gap between mothers and fathers, established in Section 4.3.2.

Table 4.4.1 illuminates the relationship between hours of childcare among parents pooled across gender. We include non-working parents by setting their potential hours of remote work to zero.[8] Column (1) displays the results when pooling across all working hours (including non-working). It shows a significant negative relationship between the potential hours of remote work and hours of childcare provision in November 2019.

Columns (2), (3) and (4) of Table 4.4.1 contain results of the same regression when splitting up the sample by hours of work prior to the pandemic. We use three bins: Full-time work (35 working hours or more), part-time work (between 20 and 34 working hours), and short hours (less than 20 working hours) or no work at all. Within these bins, potential remote working hours show only a slightly negative and statistically insignificant relationship to hours of childcare provision prior to the pandemic. Hence, the ability to work from home was unrelated to how much childcare a parent took over for a given level of working hours. This is not surprising as only a tiny fraction of remote work potential was realized.

Mirroring the effect of potential remote work hours on commuting and actual hours of remote work, this drastically changes with the onset of the pandemic. In April 2020, during the first lockdown in which childcare facilities and schools were closed, the relationship turns positive. One hour of potential remote work translates into almost half an hour of childcare for full-time working parents, one fifth of an hour for part-time working parents, and more than an hour for parents who only work little before the pandemic. In November 2020, when childcare facilities

8. We report results only including parents who worked before the pandemic in Appendix Table 4.B.4, results do not change. We prefer the sample in Table 4.4.1 because when we condition on a parent working before the pandemic, we disproportionately drop mothers, leaving fathers in single-earner households in the sample. Conceptually, we prefer to avoid this imbalance.

Table 4.4.1. Childcare hours and potential remote working hours before and during the CoVid-19 Pandemic

	Childcare hours			
	(1)	(2)	(3)	(4)
Constant	28.22***	19.56***	27.34***	37.79***
	(1.43)	(1.69)	(2.72)	(2.88)
2020-04	6.20***	6.69***	10.93***	0.38
	(1.40)	(2.14)	(2.22)	(2.35)
2020-11	−4.78***	1.52	−3.84*	−13.24***
	(1.30)	(1.80)	(2.05)	(2.34)
2021-11	−3.19***	−0.51	0.06	−10.1***
	(1.21)	(1.75)	(2.07)	(2.12)
Pot. hours remote work	−0.26***	−0.04	−0.13	−0.07
	(0.05)	(0.05)	(0.09)	(0.49)
Pot. hours remote work × 2020-04	0.49***	0.46***	0.35**	1.78***
	(0.07)	(0.09)	(0.14)	(0.44)
Pot. hours remote work × 2020-11	0.38***	0.21**	0.28**	1.14*
	(0.07)	(0.08)	(0.13)	(0.62)
Pot. hours remote work × 2021-11	0.26***	0.19***	0.12	0.74
	(0.06)	(0.07)	(0.13)	(0.52)
Working hours (2019-11)	All	>34	20-34	<20
Observations	2,234	1,010	706	518
R^2	0.273	0.245	0.375	0.359

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors obtained by clustering on the household level. Baseline month is November 2019. All specifications control for age of the youngest child interacted standardized by subtracting the pooled sample mean (6.8) divided by the standard deviation (4.7) interacted with gender, as well as indicator variables indicating number of children. Column (2) restricts the sample to parents that work 35 hours or more, column (3) restricts the sample to parents working between 20 and 34 hours, and column (4) restricts the sample to parents working less than 20 hours pre-CoVid. The latter includes non-working which are assigned 0 hours of remote work potential. Results of the same regressions when restricting the sample to working parents are included in Appendix Table 4.B.4. The full table is available in Appendix Table 4.B.3

and schools were open again, the relationship becomes somewhat weaker but stays strong for full-time working parents. Every hour of remote work potential translates into approximately a quarter of an hour of childcare. This effect remains stable until November 2021. For parents working part-time before the pandemic, the effect becomes weaker and statistically indistinguishable from its pre-pandemic value in November 2021. One reason for this might be that some parents with low working hours and high remote work potential increased their working hours during the pandemic, an effect we shall investigate in the next section.

Table 4.4.2 brings together the changes in the childcare gap between mothers and fathers and the shift to remote work. Column (1) repeats the numbers underlying Figure 4.3.2b, which plotted the coefficients on the indicator variables for mother by time period during the pandemic. The absolute difference in childcare provision between parents did not change in April 2020, when childcare facilities and schools were closed, because both parents increased their childcare provision with similar magnitudes. There is a sharp decline (six hours) in the gender gap in childcare in November 2020 accompanied by decrease in childcare provision by mothers and an increase by fathers. The decrease carries over to November 2021 but only at half the size.

Our key specification is column (2), which adds the potential hours of remote work. This yields a difference-in-differences design with a continuous treatment variable. The basic assumption is that in the absence of the pandemic, childcare hours would have evolved independently from remote work ability. While this assumption might be too strong, we would likely err in a direction that attenuates our coefficients of interest. In particular, we just established that before the pandemic, potential hours of remote work are negatively related with childcare hours or unrelated, depending on whether we control for working hour bins or not. This makes sense, as jobs with high remote work potential tend to yield relatively high earnings, so—to the extent that income effects dominate—c.p., fathers are more likely to work longer hours and mothers are more likely to return to work early and work longer hours. Note that none of our results is driven by the fact that we use potential remote working *hours* as a measure for the flexibility gains; everything also goes through if we use the share instead.

In column (2) of Table 4.4.2, we standardize the potential hours of remote work so that we can compare the evolution of the gender care gap across specifications. The coefficients on the mother by time period interactions measure the gender care gap, evaluated at the sample mean of potential hours of remote work.[9] Including the standardized potential hours of remote work in the regression diminishes the changes in the gender care gap in November 2020 and November 2021 by cutting coefficients in half, rendering them statistically indistinguishable from zero or

9. Defined as potential hrs remote work (std) $= \frac{(\text{potential hrs remote work}-\mu)}{\sigma}$ with the sample mean $\mu = 12,1$ and standard deviation $\sigma = 14.7$.

Table 4.4.2. The effect of potential remote working hours on the evolution of the gender care gap

	Childcare hours		
	(1)	(2)	(3)
Constant	18.09***	18.84***	17.88***
	(1.17)	(1.23)	(1.41)
2020-04	12.43***	10.33***	10.36***
	(1.46)	(1.42)	(1.43)
2020-11	2.90**	1.20	1.72
	(1.28)	(1.29)	(1.27)
2021-11	1.70	0.39	0.53
	(1.13)	(1.15)	(1.18)
Mother	12.46***	11.11***	12.51***
	(1.31)	(1.34)	(2.09)
Mother × 2020-04	−0.43	3.16	3.05
	(1.94)	(1.92)	(1.93)
Mother × 2020-11	−6.01***	−3.24*	−3.33*
	(1.84)	(1.86)	(1.85)
Mother × 2021-11	−3.4**	−1.41	−1.38
	(1.63)	(1.66)	(1.65)
Pot. hours remote work (std)		−2.34***	−1.49**
		(0.68)	(0.73)
Pot. hours remote work (std) × 2020-04		7.27***	7.56***
		(1.04)	(1.31)
Pot. hours remote work (std) × 2020-11		5.06***	3.57***
		(1.01)	(1.18)
Pot. hours remote work (std) × 2021-11		3.56***	3.16***
		(0.88)	(0.99)
Pot. hours remote work (std) × Mother × 2020-04			−0.66
			(2.02)
Pot. hours remote work (std) × Mother × 2020-11			3.73**
			(1.86)
Pot. hours remote work (std) × Mother × 2021-11			1.04
			(1.89)
Observations	2,234	2,234	2,234
R^2	0.326	0.347	0.349

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors obtained by clustering on the household level. The potential hours of remote work are standardized to mean zero and unit standard deviation to facilitate comparison of coefficients across columns. All specifications control for the (demeaned) age of the youngest child interacted with gender, as well as indicator variables indicating number of children, the left-out category is a single child. In column (5), we additionally interact the number of children with gender, so that the model is fully satiated. Potential remote work hours are set to zero if the individual did not work before the pandemic. The full set of coefficients can be found in Appendix Table 4.B.5. Appendix Table 4.B.6 shows results for the same specifications restricting the sample to individuals who were working before the pandemic.

marginally so. This indicates that the changes in the gender care gap can be largely explained by the shift to remote work which made, in particular, fathers more available in many families.

Column (3) additionally includes an interaction between standardized potential hours of remote work and the mother dummy. This does not change the previous results. Further, it shows that mothers tend to be more inclined to use their potential hours of remote work for childcare. The effect is, however, only statistically significant for November 2020.

4.4.2 Labor Supply

The result that remote work induced a decrease in the childcare gap gives rise to the question whether these changes also translate to effects on labor supply. In particular, mothers whose partners are now taking over a larger share of childcare duties might be willing to increase the time spent on market work. In Section 4.3.2, we saw that the trend of increasing full-time work of mothers accelerated over the 2020–2021 period. In this section, we analyze whether partners' remote work induces individuals to work more and to which extent this effect operates through a direct effect of increased remote working hours and to which extent through reduced commuting.

The mechanisms at play are thus more subtle and likely to operate with some time lag. In April 2020, there was an immediate need for childcare and parental involvement had to be adjusted instantly. In contrast, changing one's (paid) hours of work requires at least some preparation and potentially negotiations with the employer as well as within the household. Hence, we would expect changes in working hours to lag behind changes in childcare hours. Because effects are rather small, we cannot expect to find much in the LISS data. Hence, we recur to the CBS data, where we have information on hours worked as well.

Because we have a longer pre-CoViD time series, we can apply a more sophisticated design to estimate the labor supply response to the shift to remote work. In particular, we aim to alleviate a measurement problem in the treatment intensity. Given that the calculation of number of potential hours remote work relies on the pre-pandemic labor supply of parents, this is generically on average lower for mothers of younger children than for mothers of older children.[10] That is, for mothers with a youngest child of age 5 in 2019, treatment intensity is defined when the child is 5 years old, while for mothers with a youngest child of age 5 in 2020, treatment intensity is defined when the child is 4 years old. Thus, the partner's effect of remote work hours on labor supply is potentially confounded by mean reversion in working hours. One solution to this problem is to condition on the age of the youngest child in 2019 instead of controlling for age of the youngest child in each period. While this approach solves the measurement problem of the remote work potential, it can

10. Note that the same pattern does not apply to the same extent for fathers.

potentially lead to another problem. If, for instance, more educated mothers generally return faster to work or start working with more hours than less educated mothers, then our measure for potential remote work is correlated with transition rates within child birth cohorts.

To alleviate these problems, we, thus, opt for a more complex design that relies on a triple difference: we calculate the difference in working hours over time between parents with high partners' potential remote work hours, and those with lower partners' potential remote work hours, for individuals in the time period 2018-2021, when the Covid shock materialized, and for individuals in the time period 2013-2016, when no Covid shock took place. We then take these two separate Diff-in-Diff estimates and analyze how the difference between the two estimates evolves over time. Doing so, we always compare parents having a youngest child of the same age in period 0. This approach avoids the aforementioned measurement problem by always comparing parents with a youngest child of the same age, while at the same time removing generic differences in transitions rates between mothers after childbirth related to remote work capability.

The main identifying assumption is that—conditional on composition and demographic characteristics of households—the relation between working hours and remote work potential would have been the same across both time periods, absent the pandemic. As the mechanism operates through households' total time budgets, we include the remote work potential of both partners. This is also possible because of the large sample size; which further allows us to run separate estimations by gender. The regression equation becomes:

$$
\begin{aligned}
\text{Working Hours}_{i,t} &= \alpha + \chi \text{ Pot. hrs remote work}_i + \phi \text{ Pot. hrs remote work partner}_i \\
&+ \sum_{t=-1}^{2} (\beta_t \text{ Pot. hrs remote work}_i + \delta_t \text{ Pot. hrs remote work partner}_i) \\
&\times \mathbb{1}(Year = t) \\
&+ \sum_{t=-1}^{2} (\gamma_t \text{ Pot. hrs remote work}_i + \omega_t \text{ Pot. hrs remote work partner}_i) \\
&\times \mathbb{1}(Year = t) \times \text{Pandemic}_i \\
&+ \sum_{t=-1}^{2} \mu_t \mathbb{1}(Year = t) + \sum_{t=-1}^{2} \sigma_t \mathbb{1}(Year = t) \times \text{Pandemic}_i \\
&+ \pi \text{ Pandemic}_i + \rho \text{ Age youngest child}_{i,0} + \eta \text{ Number children}_{i,t} \\
&+ \iota \text{ Age}_{i,t} + \xi \text{ Age partner}_{i,t} + \epsilon_{i,t}
\end{aligned}
$$

The time index t measures time relative to just before the (placebo) pandemic – it is 0 in 2019 and 2014, respectively. Our dependent variable are unconditional working hours. Factually, this applies only to $t \in \{1, 2\}$, because we individuals to be working in the earlier years to obtain a measure of remote work ability in the administrative data (see Section 4.2.2 and, for all details, Appendix 4.A.1).

Of course, we also rely on the assumption that no other attributes of the CoVid-19 pandemic are correlated with potential remote work hours of my partner as well as my subsequent labor supply conditionally on own remote work potential, number of children, age and partners age. The most likely confounder is that the economic shock induced by the CoVid-19 pandemic differentially affected workers depending on their ability to work remotely. Zimpelmann et al. (2021) show that individuals who have a higher remote work capability decrease their working hours less, early in the pandemic in the Netherlands. That is, they experience less of a labor market shock and work more than those with a lower remote work capability which means they are less available for childcare. This, however, should dampen the effect of partner's remote work capability on own labor supply of parents which implies that, if anything, we may underestimate the effect.

Our coefficients of interest are ω_1 and ω_2, which measure how the effect of the partners' potential hours of remote work on working hours has evolved differently across the two time periods for members of households with identical characteristics.

Figures 4.4.1a and 4.4.1b show our estimates, namely the difference in the effect of partners' potential hours of remote work between CoVid-19 and Placebo sample relative to baseline period 0, over time. We estimate the effects separately for mothers and for fathers. The blue coefficients show the estimates when setting all β_t and γ_t to zero, i.e. when leaving own potential hours remote work out of the regression equation; the orange coefficients additionally control for own potential hours of remote work. Without controlling for own remote work ability, we see a slight negative pre-trend. When controlling for own potential remote work hours, this effect disappears and the pre-trend is insignificant and virtually zero numerical terms. For mothers we find that in both specifications, the DiD estimate is around 0.01, while for fathers the estimates range between 0.01 and 0.02. For illustration, a coefficient of 0.01 implies that for a level of partners' potential hours of remote work of 12, which is around the average of the fathers' potential hours of remote work in our sample, the shift towards remote work leads, on average, to around 0.12 working hours more.

To be able to disentangle the effect of the partners' remote work into a direct effect of increased potential hours of remote work and an indirect effect through potential commuting gains, we further include own and partner potential commuting gains into the regression.[11] Potential commuting gains are calculated by multiplying the imputed remote work capability with the commuting distance (in km) for the periods -1 and 0 and then averaging it. This implicitly assumes that all individuals

11. The exact regression equation can be found in Appendix 4.B.3.

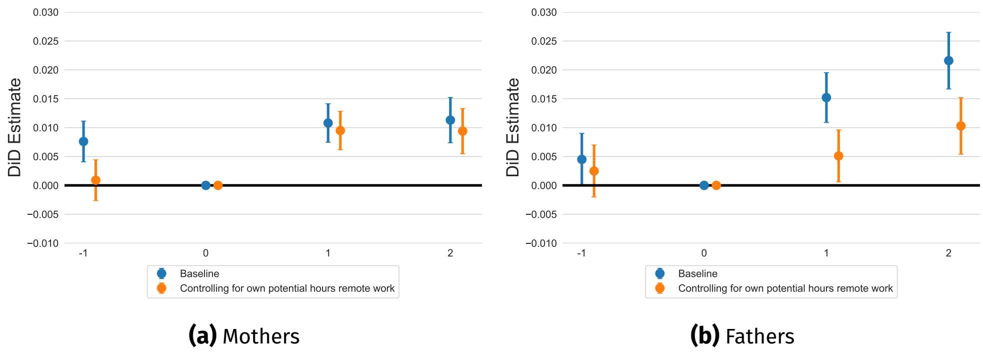


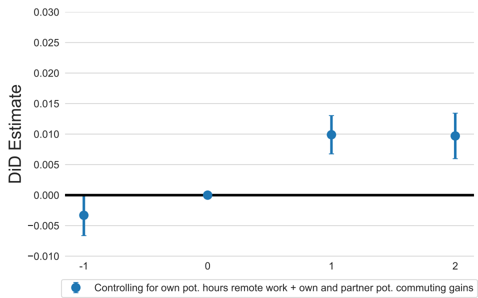
Figure 4.4.1. Effect of potential hours of remote work of the partner on own working hours

Notes: The figure displays the event-study DiD estimates for the effect of the potential hours of remote work of the partner on own working hours relative to the year of the Covid/Placebo shock. Results are reported separately for all mothers and fathers. The baseline specification only includes potential hours of remote work of the partner, while the controlling for own potential hours of remote work specification also controls for own potential hours of remote work. potential hours of remote work are calculated by multiplying the imputed remote work capability with actual working hours for the periods -1 and 0 and then taking the mean of it. All specifications include controls for own and partner age and fixed effects for the age of the youngest child and the number of children. Standard errors are obtained by clustering on the individual level. Complete regression results can be found for mothers in table 4.B.7 and for fathers in table 4.B.9. The baseline specification can be found in column (1), while the controlling for own potential hours of remote work specification can be found in column (2).

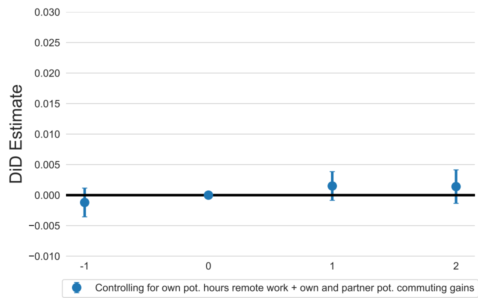
commute the same number of working days per week in the absence of the shift to remote work.

We separate the effect of potential hours of remote work into a direct effect in Figures 4.4.2a and an indirect effect through potential commuting gains in Figure 4.4.2b for mothers. Figure 4.4.2b shows that potential commuting gains do not seem to have explanatory power and that the results come from the direct effect of potential hours of remote work as shown in Figure 4.4.2a.

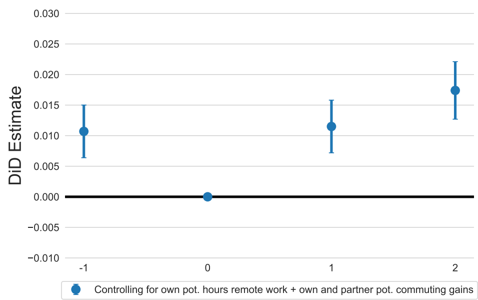
Figure 4.4.2c and Figure 4.4.2d show the equivalent for fathers. Figure 4.4.2d shows that potential commuting gains do not seem to have explanatory power. However, for the direct effect of potential hours of remote work as shown in Figure 4.4.2c the parallel trend assumption seems to be potentially violated, so that one should not overinterpret the effect size.



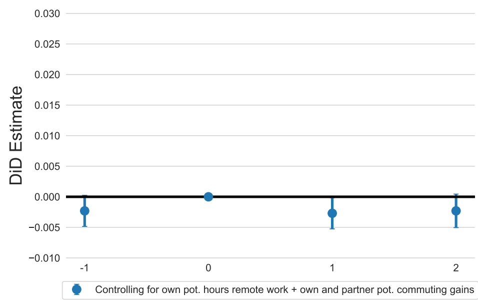
(a) Effect of fathers' potential hours of remote work on mothers' working hours



(b) Effect of fathers' potential commuting gains on mothers' working hours



(c) Effect of mothers' potential hours of remote work on fathers' working hours



(d) Effect of mothers' potential commuting gains on fathers' working hours

Figure 4.4.2. Direct and indirect effect of potential hours of remote work of the partner on own working hours

Notes: The figure separates the event-study DiD estimate for the effect of potential hours of remote work of the partner on own working hours in a direct effect through potential hours of remote work and an indirect effect through potential commuting gains. Figures 4.4.2a and 4.4.2c show the event-study DiD estimates for the direct effect of potential hours of remote work of the partner on own working hours relative to the year of the Covid/Placebo shock. Figures 4.4.2b and 4.4.2d show the event-study DiD estimates for the effect through potential commuting gains of the partner on own working hours relative to the year of the Covid/Placebo shock. Results are reported separately for all mothers and fathers. The specification includes own and partner potential hours of remote work and own and partner potential commuting gains into the regression. Potential hours of remote work are calculated by multiplying the imputed remote work capability with actual working hours for the periods -1 and 0 and then taking the mean of it. Potential commuting gains are calculated by multiplying the imputed remote work capability by the commuting distance (in km), assuming that all individuals commute the same number of working days per week, for the periods -1 and 0 and then averaging it. All specifications include controls for own and partner age and fixed effects, for the age of the youngest child and the number of children. Standard errors are obtained by clustering on the individual level. Complete regression results for mothers can be found in column (3) of table 4.B.7. Complete regression results for fathers can be found in column (3) of table 4.B.9.

4.5 Conclusion

We have investigated how the acceleration in the shift towards remote work during the CoViD-19 pandemic has impacted the division of childcare duties and working hours. The way the pandemic has been handled in the Netherlands—the most important feature being relative short school and daycare closures—has allowed us to isolate this effect. Our analysis has relied on self-collected survey data and population-wide administrative data.

We find that the average gap between mothers' and fathers' childcare provision shrinks by 3.4 hours or 27 % in the period from November 2019 to November 2021. Most of this decline can be attributed to households where the remote work potential was high. The partner's potential remote work also helps to explain the trend towards mothers working longer hours, which was accelerated during the pandemic.

Our results show that remote work can help many households to find a division of labor that is more equal across genders. It is likely that more working from home will remain very common in the future, so employers will be less able to condition wages and career progression on it than they were before the pandemic. This also means that a convenient excuse for some parents, in particular fathers, for not being available for childcare duties is gone on some days.

In other institutional environments, the effects we found might take longer to materialize. The infrastructure for remote work and childcare is well-developed and reliable in the Netherlands. Mothers had a high labor force participation rate—albeit with low hours—already before the pandemic, while fathers' weekly hours were low in international comparison (Bick, Brüggemann, and Fuchs-Schündeln, 2019). Finally, of course, in many countries the pandemic had a differentially larger direct effect on the labor market outcomes of women (Alon et al., 2022).

Overall, our results have shown that working from home might have a bright side in bringing about more gender equality within households.

References

Adams-Prassl, Abi. 2021. "The Gender Wage Gap in an Online Labour Market: The Cost of Interruptions." *Oxford Department of Economics Discussion Paper Series* 944: 1–38. [265]

Alon, Titan, Sena Coskun, Matthias Doepke, David Koll, and Michèle Tertilt. 2022. "From Mancession to Shecession: Women's Employment in Regular and Pandemic Recessions." *NBER Macroeconomics Annual 2021* 36 (1): 83–151. [266, 290]

Andrew, Alison, Sarah Cattan, Monica Costa Dias, Christine Farquharson, Lucy Kraftman, Sonya Krutikova, Angus Phimister, and Almudena Sevilla. 2020. "The Gendered Division of Paid and Domestic Work under Lockdown." IZA Discussion Paper 13500. Institute of Labor Economics (IZA). [266]

Bang, Minji. 2021. "Job Flexibility and Household Labor Supply: Understanding Gender Gaps and the Child Wage Penalty." Job Market Paper. [266]

Been, J., and Centerdata. 2021. *LISS Panel - Time Use and Consumption - Part 9*. DANS/KNAW. [269]

Bick, Alexander, Adam Blandin, and Richard Rogerson. 2022. "Hours and Wages." *The Quarterly Journal of Economics* 137 (3): 1901–62. [263]

Bick, Alexander, Bettina Brüggemann, and Nicola Fuchs-Schündeln. 2019. "Hours Worked in Europe and the United States: New Data, New Answers." *Scandinavian Journal of Economics* 121 (4): 1381–416. [290]

Cortés, Patricia, and Jessica Pan. 2019. "When Time Binds: Substitutes for Household Production, Returns to Working Long Hours, and the Skilled Gender Wage Gap." *Journal of Labor Economics* 37 (2): 351–98. [266]

Del Boca, Daniela, Noemi Oggero, Paola Profeta, and Mariacristina Rossi. 2020. "Women's and Men's Work, Housework and Childcare, before and during COVID-19." *Review of Economics of the Household* 18 (4): 1001–17. [266]

Farré, Lídia, Yarine Fawaz, Libertad González, and Jennifer Graves. 2020. "How the COVID-19 Lockdown Affected Gender Inequality in Paid and Unpaid Work in Spain." Working Paper. [266]

Felfe, Christina. 1, 2012. "The Motherhood Wage Gap: What about Job Amenities?" *Labour Economics* 19 (1): 59–67. [265]

Gicheva, Dora. 2013. "Working Long Hours and Early Career Outcomes in the High-End Labor Market." *Journal of Labor Economics* 31 (4): 785–824. [263]

Goldin, Claudia. 2014. "A Grand Gender Convergence: Its Last Chapter." *American Economic Review* 104 (4): 1091–119. [266]

Hank, Karsten, and Anja Steinbach. 2020. "The Virus Changed Everything , Didn't It ? Couples' Division of Housework and Childcare before and during the Corona Crisis." *Journal of Family Research*, 1–16. [266]

Heggeness, Misty, and Palak Suri. 2021. "Telework, Childcare, and Mothers' Labor Supply." [267, 274]

Houghton, Kendall. 2020. "Childcare and the New Part-Time: Gender Gaps in Long-Hour Professions." Job Market Paper. [265]

Jessen, Jonas, C. Katharina Spiess, Sevrin Waights, and Katharina Wrohlich. 2022. "The Gender Division of Unpaid Care Work throughout the COVID-19 Pandemic in Germany." *German Economic Review*, 641–67. [266]

Le Barbanchon, Thomas, Roland Rathelot, and Alexandra Roulet. 2021. "Gender Differences in Job Search: Trading off Commute against Wage." *The Quarterly Journal of Economics* 136 (1): 381–426. [266]

Magda, Iga, and Katarzyna Lipowska. 2021. "Flexibility of Working Time Arrangements and Female Labor Market Outcome." *IZA Discussion Paper Series No. 14812*, 20. [265]

Mangiavacchi, Lucia, Luca Piccoli, and Luca Pieroni. 2021. "Fathers Matter: Intrahousehold Responsibilities and Children's Wellbeing during the COVID-19 Lockdown in Italy." *Economics & Human Biology* 42: 1–19. [266]

Mas, Alexandre, and Amanda Pallais. 2017. "Valuing Alternative Work Arrangements." *American Economic Review* 107 (12): 3722–59. [265]

Meekes, Jordy, Wolter Hassink, and Guyonne Kalb. 2020. "Essential Work and Emergency Childcare: Identifying Gender Differences in COVID-19 Effects on Labour Demand and Supply." Working Paper. [267]

Meekes, Jordy, and Wolter H. J. Hassink. 1, 2022. "Gender Differences in Job Flexibility: Commutes and Working Hours after Job Loss." *Journal of Urban Economics* 129: 1–15. [266]

Pabilonia, Sabrina, and Victoria Vernon. 2022. "Who Is Doing the Chores and Childcare in Dual-earner Couples during the COVID-19 Era of Working from Home?" Working Paper. [266]

Sevilla, Almudena, and Sarah Smith. 2020. "Baby Steps: The Gender Division of Childcare during the COVID-19 Pandemic." *Oxford Review of Economic Policy* 36: (Supplement_1), S169–S186. [266]

Ten Hoeve, Youri, Jildou Talman, Jorrit van Mierlo, and Mirjam Engelen. 2021. "Evaluatie Wet Flexibel Werken." Onderzoek in opdracht van het ministerie van Sociale Zaken en Werkgelegenheid. [277]

Van Soest, Arthur, Jim Been, Pia Pinger, Hans-Martin von Gaudecker, and Centerdata. 2019. *LISS Panel - Time Use and Consumption - Part 6*. DANS/KNAW. [269]

von Gaudecker, Hans-Martin, Christian Zimpelmann, Moritz Mendel, Bettina Siflinger, Lena Janys, Jürgen Maurer, Egbert Jongen, Radost Holler, Renata Abikeyeva, Felipe Augusto Azuero Mutis, Annica Gehlen, and Eva Lucia Kleifgen. 2021. "CoVID-19 Impact Lab Questionnaire Documentation." eng. https://doi.org/10.5281/zenodo.4338730. [268]

Von Gaudecker, Hans-Martin, and Centerdata. 2020a. *LISS Panel - Time Use and Consumption - Part 7 (Additional Coronavirus Part)*. DANS/KNAW. [269]

Von Gaudecker, Hans-Martin, and Centerdata. 2020b. *LISS Panel - Time Use and Consumption - Part 8*. DANS/KNAW. [269]

Wiswall, Matthew, and Basit Zafar. 2018. "Preference for the Workplace, Investment in Human Capital, and Gender." *Quarterly Journal of Economics* 133 (1): 457–507. [265]

Zamarro, Gema, and María J. Prados. 1, 2021. "Gender Differences in Couples' Division of Childcare, Work and Mental Health during COVID-19." *Review of Economics of the Household* 19 (1): 11–40. [266]

Zimpelmann, Christian, Hans-Martin von Gaudecker, Radost Holler, Lena Janys, and Bettina Siflinger. 2021. "Hours and Income Dynamics during the Covid-19 Pandemic: The Case of the Netherlands." *Labour Economics* 73: 1–29. [268, 272, 273, 276, 287]

Appendix 4.A Details on data sets and descriptives

4.A.1 Imputation of remote work potential in the administrative data

For the imputation of the remote work capability in the administrative records, we make use of the National Working Condition Survey (NEA). It is currently available until 2020, i.e. the wave of 2021 is not yet published. Its goal is to gather information on the topics of working conditions, occupational accidents, work content, employment relationships and employment conditions of employees. The NEA is carried out yearly since 2005 by Statistics Netherlands and TNO, in collaboration with the Ministry of Social Affairs and Employment. Its target population are all employees aged 15 to 74 who work in the Netherlands, from whom a sample is surveyed during the period of 1st of October to 31st of December of each year.[12]

12. The documentation of the survey and all questionnaires are available at https://www.cbs.nl/nl-nl/onze-diensten/methoden/onderzoeksomschrijvingen/korte-onderzoeksbeschrijvingen/nationale-enquete-arbeidsomstandigheden--nea--.

Around 50,000 individuals answer the survey each year and around 30,000 to 35,000 of those respondents answered the questions on remote work, which we use for our imputation. In particular, we use the following variables for calculating a remote work share:

- Remote Work Hours (Afl_AantUurTW): "On average, how many hours a week do you work from home for your employer?"
- Remote Work Dummy (Afl_Telewerk): "Teleworker (works at least half a day a week outside the company location with access to the company's IT system)"
- Working Hours (Afl_Uren): "Working hours in hours per week in current job"

We calculate a remote work share for each individual by dividing the remote work hours by total working hours. For individuals for whom we do not observe information on the remote work hours, but for whom we observe the remote work dummy being 0, we impute a remote work share of 0.

Figure 4.A.4a displays the distribution of the remote work share by gender in the NEA in the year 2020. Dashed vertical lines indicate the mean for each gender. The figure shows that the remote work share in the NEA exhibits a similar distribution like the share of tasks that can be done from home variable in the LISS Sample (see Figure 4.3.4a). The distribution is bi-modal and men have, on average, a higher remote work share than women.

To be able to impute the remote work capability for each individual in the administrative records, we have to find highly predictive characteristics along which we can make the imputation. Table 4.A.1 displays the regression results from regressing the remote work share in 2020 on education, gender and sector. The table shows that education and sector are highly predictive for the remote work share, while gender is not predictive. We therefore perform the imputation with the help of education and sector.

We only consider the remote work share in the year 2020, since before this year there were no large differences in the remote work share across sectors and remote work shares were generally on a low level. Figures 4.A.1, 4.A.2, and 4.A.3 illustrate this by showing the mean remote work share by sector for the years 2018, 2019 and 2020.

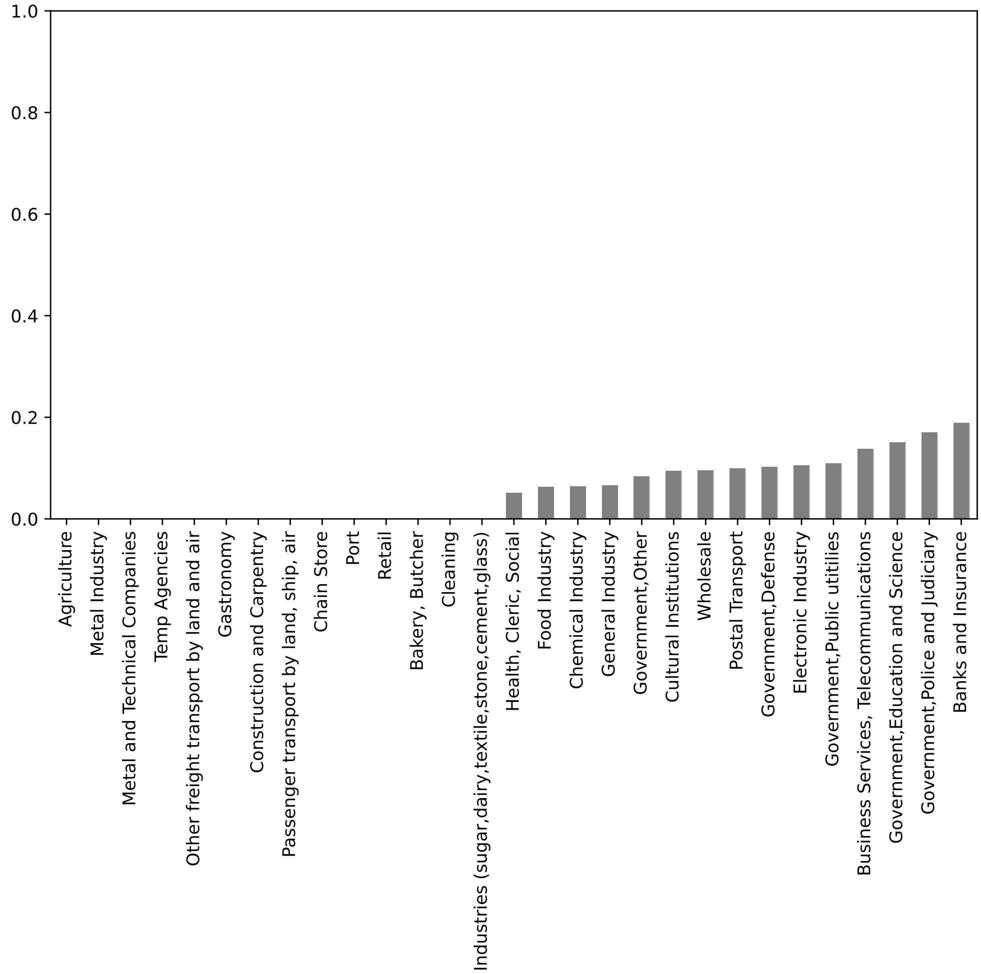


Figure 4.A.1. Share of Remote Work by Sector 2018

Notes: This figure displays the mean remote work share in the year 2018 aggregated by sector. The population are the participants of the National Working Conditions Survey (NEA) in the year 2018. The figure shows that the remote work potential does not vary substantially between sectors and is generally at a low level.

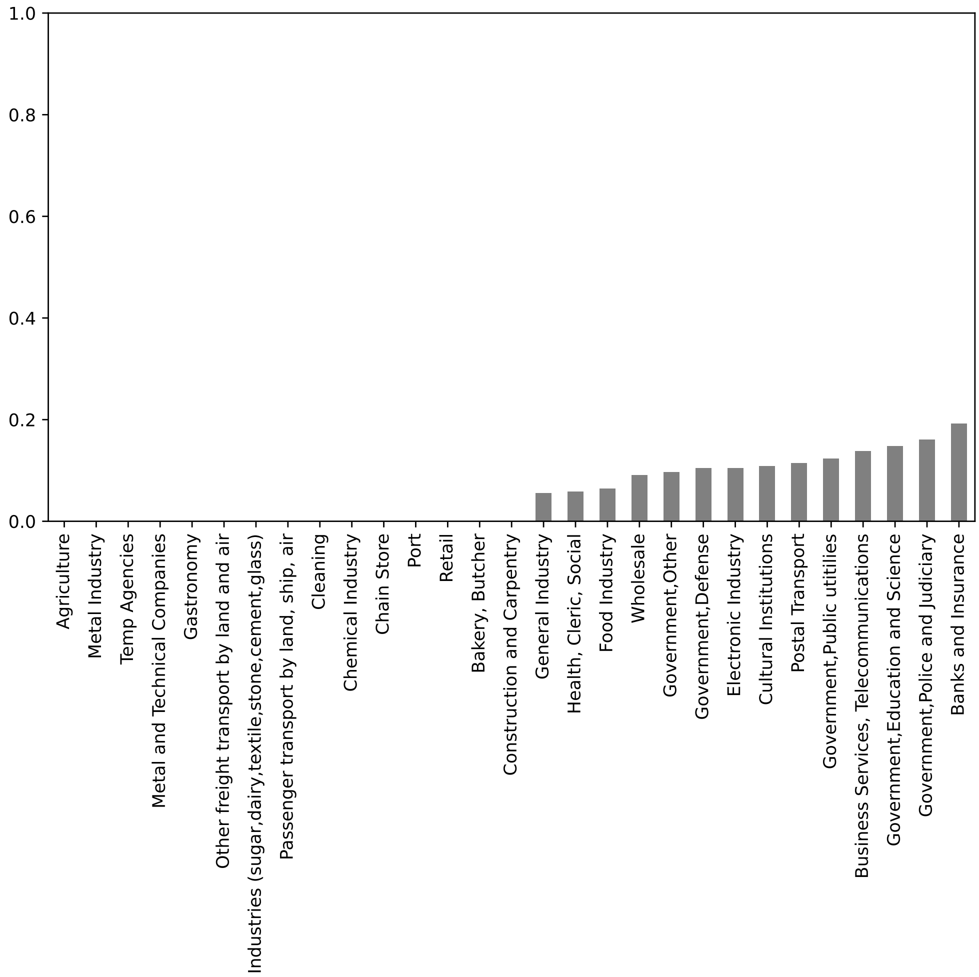


Figure 4.A.2. Share of Remote Work by Sector 2019

Notes: This figure displays the mean remote work share in the year 2019 aggregated by sector. The population are the participants of the National Working Conditions Survey (NEA) in the year 2019. The figure shows that the remote work potential does not vary substantially between sectors and is generally at a low level.

Table 4.A.1. Determinants remote work share in Fall 2020

	Share remote hours
Constant	0.44***
	(0.0086)
Education: Low	−0.29***
	(0.0079)
Education: Middle	−0.19***
	(0.0046)
Education: Unknown	−0.16***
	(0.0043)
Male	−0.0044
	(0.0038)
Agriculture	−0.15***
	(0.018)
Bakery, Butcher	−0.21***
	(0.024)
Banks and Insurance	0.41***
	(0.014)
Business Services	0.23***
	(0.0088)
Chain Store	−0.16***
	(0.015)
Chemical Industry	−0.029
	(0.019)
Cleaning	−0.21***
	(0.019)
Construction and Carpentry	−0.1***
	(0.013)
Cultural Institutions	−0.0071
	(0.023)
Electronic Industry	0.18***
	(0.021)
Food Industry	−0.014
	(0.02)
Gastronomy	−0.22***
	(0.016)
General Industry	−0.019
	(0.015)
Government, Defense	0.028
	(0.023)
Government, Education	−0.14***
	(0.0087)
Government, Other	0.099***
	(0.014)
Government, Police	0.24***
	(0.013)
Government, Public utitilies	0.27***
	(0.012)
Health, Cleric, Social	−0.15***
	(0.0089)
Industries (sugar, dairy, textile, stone, cement, glass)	−0.076***
	(0.021)
Metal Industry	−0.03**
	(0.014)
Metal and technical companies	−0.15***
	(0.011)
Other freight transport	−0.16***
	(0.015)
Passenger transport	−0.093***
	(0.016)
Port	0.022
	(0.021)
Postal Transport	−0.23***
	(0.029)
Retail	−0.17***
	(0.013)
Temp Agencies	−0.068***
	(0.014)
Wholesale	0.016
	(0.01)
N children	−0.0043***
	(0.0016)
Observations	37295
R^2	0.27

Notes: The table displays the regression results from regressing the remote work share in 2020 on education, gender and sector. The population are all individuals in the NEA sample for whom we have information on actual remote work. The table shows that education and sector are highly predictive for the remote work share, while gender is not predictive. We therefore use education and sector for our imputation.

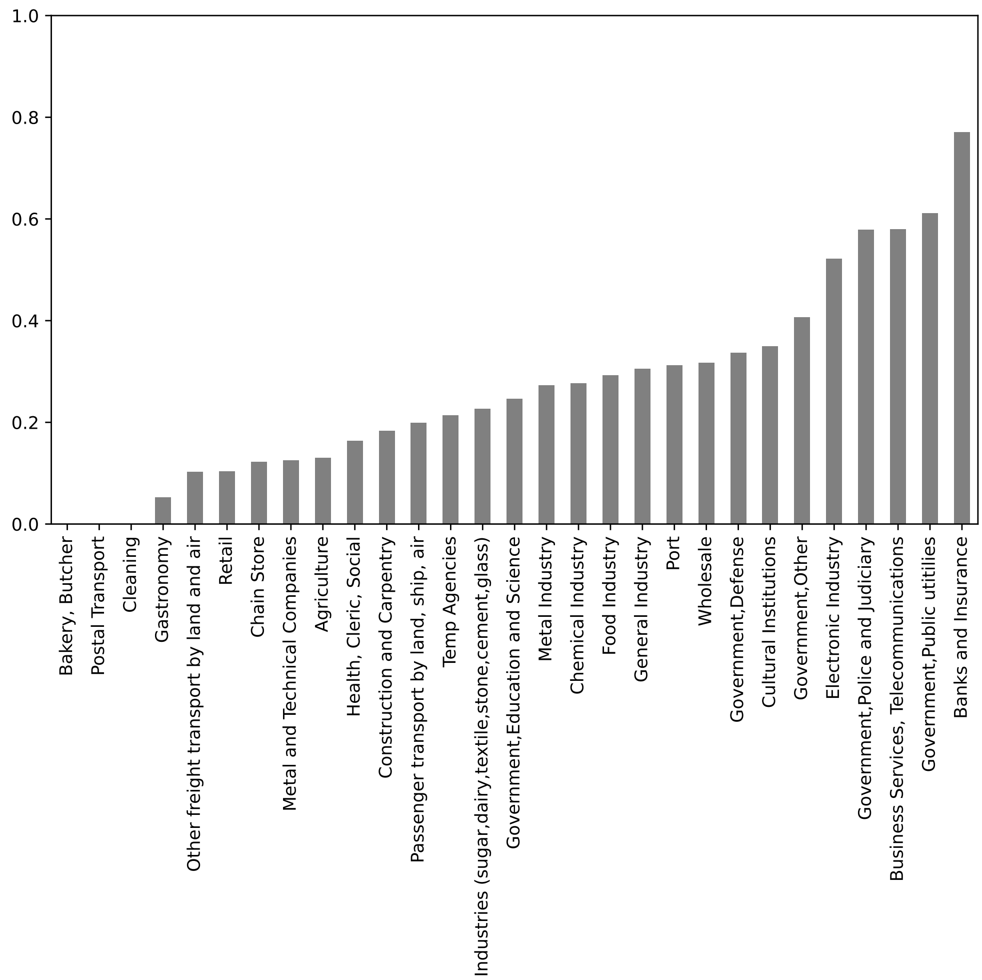
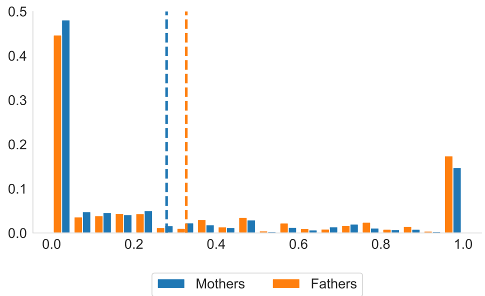
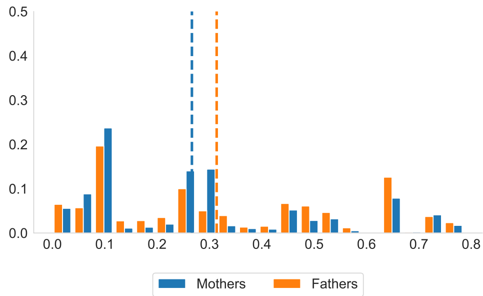


Figure 4.A.3. Share of Remote Work by Sector 2020

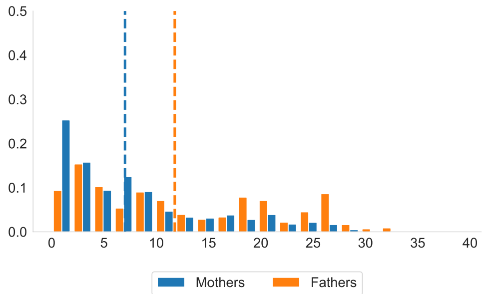
Notes: This figure displays the mean remote work share in the year 2020 aggregated by sector. The population are the participants of the National Working Conditions Survey (NEA) in the year 2020. The figure shows that the remote work potential varies substantially between sectors.



(a) Remote work capability (NEA)



(b) Imputed remote work capability (CBS)



(c) Potential hours of remote work (CBS)

Figure 4.A.4. Potential remote work by gender: in NEA and CBS

Notes: Figure 4.A.4a displays the distribution of remote work capability by gender in the National Working Conditions Survey (NEA) in the year 2020. Dashed vertical lines indicate the mean for each gender. Remote work capability is calculated by dividing the hours of remote work by total working hours. The figure shows that the remote work capability in the NEA exhibits a similar distribution like the share of tasks that can be done from home variable in the LISS Sample. The distribution is bi-modal and men work in jobs with an, on average, higher remote work capability than women. Figure 4.A.4b shows the distribution of the imputed remote work capability by gender in the CBS. Dashed vertical lines indicate the mean for each gender. For the imputation we calculate the average remote work capability by sector and education in the NEA and impute the remote work capability in the CBS with the help of these two variables. The figure shows that after the imputation, the distribution naturally looks less bi-modal, since averaging across sector and education reduces the occurrence of extreme values of the remote work capability. Figure 4.A.4c shows the distribution of the potential hours of remote work by gender in the CBS. Dashed vertical lines indicate the mean for each gender. For the imputation we calculate the average remote work capability by sector and education in the NEA and impute the remote work capability in the CBS with the help of those two variables. The imputed remote work capability is then multiplied with the pre-covid (November 2019) working hours to obtain the potential hours of remote work.

Table 4.A.2. Basic demographics by data source and gender over time

(a) LISS

		Age	Age youngest child	Number of children	Education: High	Education: Middle	Education: Low	Education: Unknown	Observations
Mothers	2019-11	40.55	7.17	2.05	0.45	0.26	0.03	0.26	260
		(0.38)	(0.28)	(0.05)	(0.03)	(0.03)	(0.01)	(0.03)	
	2020-04	40.90	7.18	2.05	0.45	0.25	0.05	0.25	280
		(0.37)	(0.29)	(0.05)	(0.03)	(0.03)	(0.01)	(0.03)	
	2020-11	40.09	6.85	2.03	0.44	0.26	0.04	0.26	339
		(0.35)	(0.26)	(0.04)	(0.03)	(0.02)	(0.01)	(0.02)	
	2021-11	39.68	6.29	2.01	0.54	0.30	0.02	0.14	311
		(0.38)	(0.27)	(0.05)	(0.03)	(0.03)	(0.01)	(0.02)	
Fathers	2019-11	42.63	6.94	2.10	0.48	0.25	0.03	0.24	229
		(0.41)	(0.30)	(0.05)	(0.03)	(0.03)	(0.01)	(0.03)	
	2020-04	42.96	6.80	2.07	0.42	0.27	0.05	0.25	257
		(0.41)	(0.30)	(0.05)	(0.03)	(0.03)	(0.01)	(0.03)	
	2020-11	42.33	6.69	2.07	0.43	0.25	0.05	0.27	283
		(0.39)	(0.28)	(0.05)	(0.03)	(0.03)	(0.01)	(0.03)	
	2021-11	42.35	6.60	2.08	0.51	0.27	0.04	0.18	275
		(0.40)	(0.29)	(0.05)	(0.03)	(0.03)	(0.01)	(0.02)	

(b) CBS

		Education: High	Education: Middle	Education: Low	Education: Unknown	Observations
Mothers	2016-11	0.36	0.28	0.11	0.25	268113
		(0.00)	(0.00)	(0.00)	(0.00)	
	2017-11	0.37	0.29	0.10	0.24	264208
		(0.00)	(0.00)	(0.00)	(0.00)	
	2018-11	0.39	0.29	0.10	0.23	261210
		(0.00)	(0.00)	(0.00)	(0.00)	
	2019-11	0.40	0.29	0.10	0.21	258413
		(0.00)	(0.00)	(0.00)	(0.00)	
	2020-11	0.41	0.30	0.09	0.20	254761
		(0.00)	(0.00)	(0.00)	(0.00)	
	2021-11	0.42	0.30	0.09	0.19	254539
		(0.00)	(0.00)	(0.00)	(0.00)	
Fathers	2016-11	0.34	0.26	0.10	0.30	268113
		(0.00)	(0.00)	(0.00)	(0.00)	
	2017-11	0.35	0.26	0.10	0.29	264208
		(0.00)	(0.00)	(0.00)	(0.00)	
	2018-11	0.36	0.27	0.10	0.27	261210
		(0.00)	(0.00)	(0.00)	(0.00)	
	2019-11	0.36	0.28	0.10	0.26	258413
		(0.00)	(0.00)	(0.00)	(0.00)	
	2020-11	0.37	0.29	0.10	0.24	254761
		(0.00)	(0.00)	(0.00)	(0.00)	
	2021-11	0.38	0.30	0.09	0.23	254539
		(0.00)	(0.00)	(0.00)	(0.00)	

Notes: Table 4.A.2a displays means and standard errors of basic demographic characteristics of the LISS sample by month and gender. The age variable is taken directly from the LISS survey. The values for the variables age of youngest child and number of children are taken from the administrative records for all linked individuals and from the LISS survey for all those who are not linked. The education variable is taken from the administrative records and therefore only available for linked individuals (note that even for linked individuals it is possible that the education is unknown). Table 4.A.2a displays means and standard errors of basic demographic characteristics of the 20 % sample of all working-age (18-55 years old) parents with a child below 16 years old by month and gender. The education variable is unknown if there is no available administrative record on the education for the individual.

4.A.2 Additional descriptive statistics

Table 4.A.3. Basic demographics, CBS sample for the analysis in Section 4.4.2

	Mothers		Fathers	
	Control: 2013-2016	Treated: 2018-2021	Control: 2013-2016	Treated: 2018-2021
Age	39.07 (6.61)	39.0 (6.72)	41.47 (6.84)	41.41 (6.99)
Age youngest child	6.76 (4.84)	6.7 (4.86)	6.76 (4.84)	6.66 (4.83)
Education: High	0.38 (0.48)	0.44 (0.5)	0.35 (0.48)	0.39 (0.49)
Education: Low	0.08 (0.27)	0.07 (0.25)	0.08 (0.28)	0.08 (0.27)
Education: Middle	0.28 (0.45)	0.31 (0.46)	0.25 (0.43)	0.29 (0.45)
Education: Unknown	0.27 (0.44)	0.18 (0.39)	0.32 (0.46)	0.24 (0.42)
Full-time	0.11 (0.32)	0.14 (0.35)	0.83 (0.38)	0.83 (0.38)
Number of children	1.96 (0.77)	1.94 (0.78)	1.98 (0.8)	1.96 (0.8)
Out of labor force	0.05 (0.23)	0.04 (0.2)	0.03 (0.17)	0.02 (0.15)
Part-time	0.83 (0.38)	0.81 (0.4)	0.14 (0.34)	0.14 (0.35)
Potential commuting gains	6.02 (12.36)	6.11 (11.93)	9.94 (16.04)	10.15 (16.13)
Potential hours remote work	7.01 (6.87)	7.75 (7.13)	11.77 (8.72)	12.37 (8.89)
Remote work capability (imputed)	26.65 (21.73)	27.56 (21.52)	31.38 (22.93)	32.35 (23.02)
Unemployed	0.01 (0.09)	0.01 (0.11)	0.01 (0.08)	0.01 (0.1)
Working hours	24.55 (8.37)	26.4 (7.99)	38.33 (5.78)	38.59 (5.6)
Working hours (unconditional)	23.09 (9.95)	25.08 (9.69)	37.01 (8.97)	37.52 (8.43)
N	3,427,090	3,322,747	3,549,700	3,304,273

Notes: The table displays means and standard deviations of the pooled event-study DiD sample for parents with a youngest child below 16. All variables are reported separately for the treatment and control group and for mothers and fathers. The difference between working hours and unconditional working hours is that the former excludes working hours of 0, while the latter is not conditional on working and therefore includes working hours of 0. Individuals are classified as unemployed when they are receiving unemployment benefits and classified as out of the labor force when there are no working hours, no self-employment status, and no unemployment benefits recorded in the administrative data. Consistent with the official definition of CBS Netherlands, we classify individuals to be working part-time if they work less than 35 hours per week. Imputed remote work capability is calculated with the procedure in Section 4.A.1. Potential hours of remote work are calculated by multiplying the imputed remote work capability with actual working hours for the periods -1 and 0 and then taking the mean of it. Potential commuting gains are calculated by multiplying the imputed remote work capability with the commuting distance for the periods -1 and 0, assuming that all individuals commute on five working days per week, and then taking the mean of it.

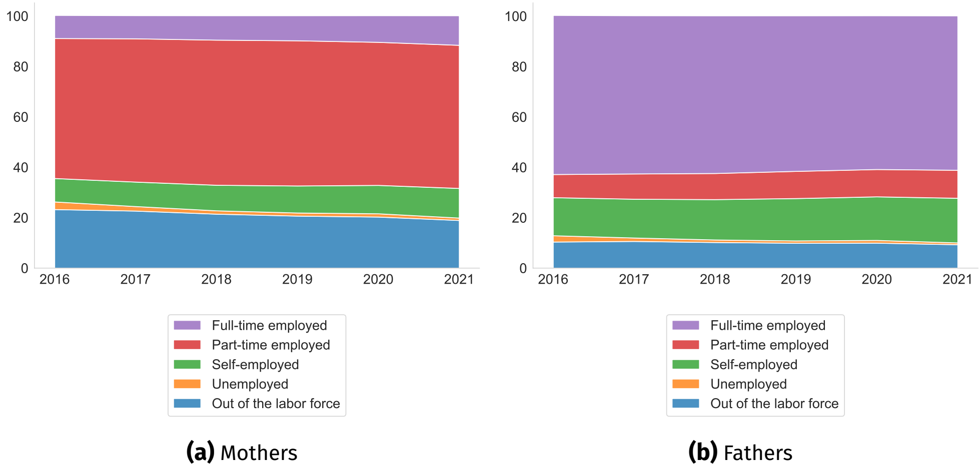


Figure 4.A.5. Labor force participation and hours categories over time

Notes: The figure provides an illustration of in Table 4.3.1. The data source is a 20 % sample of all working-age (18-55 years old) households with a child below 16 years of age by month and gender. Individuals are classified as unemployed when they are receiving unemployment benefits and classified as out of the labor force when there are no working hours, no self-employment status, and no unemployment benefits recorded in the administrative data. Consistent with the official definition of CBS Netherlands, we classify individuals to be working part-time if they work less than 35 hours per week.

Table 4.A.4. Labor market status over time in the LISS data

		Out of the labor force	Unemployed	Self-employed	Part-time employed	Full-time employed
Mothers	2019-11	0.188	0.008	0.131	0.581	0.092
		(0.024)	(0.005)	(0.021)	(0.031)	(0.018)
	2020-04	0.146	0.004	0.132	0.593	0.125
		(0.021)	(0.004)	(0.020)	(0.029)	(0.020)
	2020-11	0.156	0.003	0.103	0.605	0.133
		(0.020)	(0.003)	(0.017)	(0.027)	(0.018)
	2021-11	0.138	0.006	0.122	0.621	0.113
		(0.020)	(0.005)	(0.019)	(0.028)	(0.018)
Fathers	2019-11	0.026	0	0.074	0.170	0.729
		(0.011)	(0)	(0.017)	(0.025)	(0.029)
	2020-04	0.016	0.004	0.066	0.171	0.743
		(0.008)	(0.004)	(0.016)	(0.024)	(0.027)
	2020-11	0.025	0.004	0.074	0.152	0.746
		(0.009)	(0.004)	(0.016)	(0.021)	(0.026)
	2021-11	0.018	0.004	0.076	0.164	0.738
		(0.008)	(0.004)	(0.016)	(0.022)	(0.027)

Notes: The table shows the labor market participation by month and gender for the LISS sample. For all variables means and standard errors are reported. Individuals are classified as unemployed when they are receiving unemployment benefits and classified as out of the labor force when there are no working hours, no self-employment status, and no unemployment benefits recorded in the administrative data. Consistent with the official definition of CBS Netherlands, we classify individuals to be working part-time if they work less than 35 hours per week.

Table 4.A.5. Total working hours over time

	CBS			LISS		
	All	Fathers	Mothers	All	Fathers	Mothers
2016-11	32.3	38.6	25.3			
	(0.0162)	(0.0134)	(0.0204)			
2017-11	32.3	38.5	25.5			
	(0.016)	(0.0135)	(0.0202)			
2018-11	32.4	38.5	25.9			
	(0.0158)	(0.0136)	(0.0199)			
2019-11	32.4	38.4	26.1	28.7	37.3	21
	(0.0156)	(0.0136)	(0.0198)	(0.607)	(0.577)	(0.775)
2020-04				28.8	36.6	21.6
				(0.58)	(0.572)	(0.764)
2020-11	32.5	38.3	26.5	29.1	36.9	22.5
	(0.0154)	(0.0134)	(0.0198)	(0.539)	(0.536)	(0.711)
2021-11	32.8	38.3	27	29.6	37.6	22.6
	(0.0151)	(0.0133)	(0.0195)	(0.541)	(0.499)	(0.716)

Notes: The table displays mean and standard errors for the variable working hours by month and gender. The first column shows the average working hours for a 20 % sample of all working-age (18-55 years old) parents with a child below 16 years old. The second column shows the average working hours for the LISS sample. For all individuals in the LISS sample, which can be linked to the administrative records, we take the actual working hours from the administrative records. For the individuals which cannot be linked, we take the information on working hours from the LISS survey.

Table 4.A.6. Remote work dummy over time

	LISS		
	All	Fathers	Mothers
2020-02	0.32	0.36	0.28
	(0.02)	(0.031)	(0.027)
2020-04	0.58	0.63	0.54
	(0.021)	(0.03)	(0.03)
2020-11	0.48	0.56	0.42
	(0.02)	(0.029)	(0.027)
2021-11	0.49	0.56	0.42
	(0.02)	(0.03)	(0.028)

Notes: The table shows the mean and standard errors of the variable remote work dummy by month and gender for the LISS sample. We construct the remote work dummy ourselves such that it measures whether an individual did any remote work or none at all.

Table 4.A.7. Remote work share over time

	LISS		
	All	Fathers	Mothers
2020-02	0.12	0.12	0.13
	(0.012)	(0.015)	(0.018)
2020-04	0.47	0.49	0.44
	(0.022)	(0.031)	(0.031)
2020-11	0.37	0.43	0.33
	(0.021)	(0.029)	(0.03)
2021-11	0.34	0.35	0.33
	(0.021)	(0.026)	(0.031)

Notes: The table shows mean and standard errors of the the variable remote work share by month and gender for the LISS sample. The remote work share is calculated as hours worked from home divided by total working hours.

Table 4.A.8. Remote working hours over time

	LISS		
	All	Fathers	Mothers
2020-02	4.1	4.7	3.5
	(0.4)	(0.63)	(0.51)
2020-04	15	18	12
	(0.74)	(1.2)	(0.88)
2020-11	12	16	9.4
	(0.68)	(1.1)	(0.79)
2021-11	11	14	8.8
	(0.63)	(1)	(0.76)

Notes: The table shows mean and standard errors for the variable remote work hours by month and gender for the LISS sample.

Table 4.A.9. Commuting hours over time

	LISS		
	All	Fathers	Mothers
2019-11	4.1	5.8	2.6
	(0.2)	(0.33)	(0.18)
2020-04	1.9	2.4	1.5
	(0.19)	(0.29)	(0.24)
2020-11	2.2	2.9	1.6
	(0.12)	(0.2)	(0.14)
2021-11	2.8	3.6	2.1
	(0.17)	(0.26)	(0.23)

Notes: The table shows mean and standard errors for the variable remote work hours by month and gender for the LISS sample.

Table 4.A.10. Childcare hours over time

	LISS		
	All	Fathers	Mothers
2019-11	23.1	16.8	28.7
	(0.854)	(0.926)	(1.3)
2020-04	35.1	29.6	40.3
	(0.955)	(1.32)	(1.31)
2020-11	23.4	20.1	26.1
	(0.776)	(1.07)	(1.09)
2021-11	24.6	19.3	29.3
	(0.799)	(0.968)	(1.18)

Notes: The table shows mean and standard errors for the variable childcare hours by month and gender for the LISS sample.

Appendix 4.B Results

4.B.1 Remote work and Commuting

Table 4.B.1. Predictive power of potential remote working hours for realized hours worked from home and commuting time

	Remote working hours		Commuting hours	
	(1)	(2)	(3)	(4)
Constant	5.15***	1.16	4.64***	4.05***
	(1.16)	(0.76)	(0.27)	(0.37)
2020-04	11.81***	2.71***	−2.66***	−0.92*
	(0.87)	(0.87)	(0.28)	(0.49)
2020-11	9.54***	0.01	−2.26***	−0.82**
	(0.82)	(0.72)	(0.22)	(0.33)
2021-11	7.09***	0.43	−1.6***	−0.17
	(0.74)	(0.74)	(0.26)	(0.46)
Pot. hours remote work		0.21***		0.04**
		(0.04)		(0.02)
Pot. hours remote work × 2020-04		0.61***		−0.11***
		(0.06)		(0.02)
Pot. hours remote work × 2020-11		0.60***		−0.09***
		(0.05)		(0.02)
Pot. hours remote work × 2021-11		0.41***		−0.09***
		(0.05)		(0.02)
N children == 2	−0.06	−0.15	−0.32	−0.29
	(1.45)	(0.84)	(0.23)	(0.22)
N children == 3	−2.03	−0.38	0.01	−0.07
	(1.70)	(1.00)	(0.34)	(0.32)
N children == 4	−0.32	0.28	1.36	1.31
	(2.52)	(1.82)	(0.94)	(0.89)
N children >4	−4.06	0.94	1.38	0.98
	(2.68)	(0.88)	(1.08)	(0.95)
Age youngest child (std)	−0.75	0.48	0.21**	0.13
	(0.56)	(0.37)	(0.11)	(0.10)
Observations	1,876	1,876	1,876	1,876
R^2	0.078	0.462	0.069	0.11

Notes: Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors obtained by clustering on the household level. The table displays the relationship between commuting hours and remote work hours and remote work potential. All specification control for age of the youngest child standardized by subtracting the pooled sample mean (6.8) divided by the standard deviation (4.7), as well as indicator variables indicating number of children.. Sample restricted to parents working before the pandemic (Nov 2019).

Table 4.B.2. Predictive power of potential remote working hours for realized hours worked from home and commuting time by gender

	Hrs remote work (1)	Hrs commuting (2)
Constant	0.93	6.14***
	(0.88)	(0.67)
2020-04	3.64**	−2.34***
	(1.42)	(0.78)
2020-11	1.38	−1.71***
	(1.26)	(0.64)
2021-11	0.67	−1.01
	(1.25)	(0.76)
Pot. hours remote work	0.21***	0.00
	(0.04)	(0.02)
Pot. hours remote work × 2020-04	0.60***	−0.09***
	(0.08)	(0.03)
Pot. hours remote work × 2020-11	0.56***	−0.08***
	(0.07)	(0.02)
Pot. hours remote work × 2021-11	0.41***	−0.07***
	(0.06)	(0.03)
Pot. hours remote work × Female	0.02	0.04
	(0.08)	(0.03)
Pot. hours remote work × Female × 2020-04	0.01	−0.02
	(0.11)	(0.04)
Pot. hours remote work × Female × 2020-11	0.06	−0.0
	(0.11)	(0.03)
Pot. hours remote work × Female × 2021-11	−0.02	−0.02
	(0.10)	(0.04)
Youngest child age	0-15	0-15
Observations	1,876	1,876
R^2	0.165	0.463

Notes: Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors obtained by clustering on the household level. The table displays the relationship between commuting hours and remote work hours and remote work potential interacted with gender. All specification control for age of the youngest child standardized by subtracting the pooled sample mean (6.8) divided by the standard deviation (4.7) interacted with gender, as well as indicator variables indicating number of children. Sample restricted to parents working before the pandemic (Nov 2019).

4.B.2 Childcare

Table 4.B.3. Hours childcare and potential hours of remote work before and during the CoVid-19 Pandemic – full table

	Childcare hours			
	(1)	(2)	(3)	(4)
Constant	28.22***	19.56***	27.34***	37.79***
	(1.43)	(1.69)	(2.72)	(2.88)
2020-04	6.20***	6.69***	10.93***	0.38
	(1.40)	(2.14)	(2.22)	(2.35)
2020-11	−4.78***	1.52	−3.84*	−13.24***
	(1.30)	(1.80)	(2.05)	(2.34)
2021-11	−3.19***	−0.51	0.06	−10.1***
	(1.21)	(1.75)	(2.07)	(2.12)
Pot. hours remote work	−0.26***	−0.04	−0.13	−0.07
	(0.05)	(0.05)	(0.09)	(0.49)
Pot. hours remote work × 2020-04	0.49***	0.46***	0.35**	1.78***
	(0.07)	(0.09)	(0.14)	(0.44)
Pot. hours remote work × 2020-11	0.38***	0.21**	0.28**	1.14*
	(0.07)	(0.08)	(0.13)	(0.62)
Pot. hours remote work × 2021-11	0.26***	0.19***	0.12	0.74
	(0.06)	(0.07)	(0.13)	(0.52)
N children == 2	−0.82	−0.81	−1.68	0.80
	(1.20)	(1.65)	(2.25)	(2.53)
N children == 3	−1.63	−1.7	−2.28	−2.25
	(1.45)	(2.12)	(2.70)	(2.67)
N children == 4	−1.25	−1.08	−2.94	0.18
	(2.59)	(3.48)	(7.34)	(4.55)
N children >4	3.46*	3.24	0.60	1.11
	(2.06)	(3.86)	(2.33)	(4.33)
Age youngest child (std)	−9.22***	−6.66***	−11.07***	−11.48***
	(0.48)	(0.70)	(0.85)	(0.90)
Working hours (2019-11)	All	>34	20-34	<20
Observations	2,234	1,010	706	518
R^2	0.273	0.245	0.375	0.359

Notes: Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors obtained by clustering on the household level. This table displays for the LISS Sample, how the relationship between remote work ability and childcare changes over the course of the pandemic. All specification control for age of the youngest child standardized by subtracting the pooled sample mean (6.8) divided by the standard deviation (4.7), as well as indicator variables indicating number of children. Potential hours of remote work is set to zero for parents who do not work before the pandemic. Column (2) restricts the sample to parents who work 35 hours or more before the pandemic, column (3) restricts the sample to parents working between 20 and 34 hours, and column (4) restricts the sample to parents working less than 20 hours.

Table 4.B.4. Hours spent on childcare and potential hours of remote work before and during the CoVid-19 Pandemic, conditional on working in November 2019

	Childcare Hours			
	(1)	(2)	(3)	(4)
Constant	25.84***	20.42***	25.58***	30.17***
	(1.46)	(1.83)	(2.53)	(3.03)
2020-04	5.04***	3.11	10.47***	−0.6
	(1.62)	(2.27)	(2.41)	(3.53)
2020-11	−6.56***	−3.3*	−4.47**	−15.31***
	(1.37)	(1.75)	(2.17)	(3.41)
2021-11	−5.45***	−4.94***	−0.9	−13.17***
	(1.34)	(1.68)	(2.30)	(2.92)
Pot. hours remote work	−0.19***	−0.07	−0.02	0.18
	(0.05)	(0.05)	(0.10)	(0.44)
Pot. hours remote work × 2020-04	0.54***	0.58***	0.36***	1.89***
	(0.07)	(0.09)	(0.14)	(0.46)
Pot. hours remote work × 2020-11	0.47***	0.40***	0.29**	0.88
	(0.07)	(0.08)	(0.13)	(0.60)
Pot. hours remote work × 2021-11	0.38***	0.37***	0.18	0.96*
	(0.06)	(0.07)	(0.13)	(0.50)
N children == 2	−0.59	−0.96	−1.55	6.71***
	(1.23)	(1.61)	(2.25)	(2.57)
N children == 3	−0.53	−1.57	−1.54	4.51
	(1.52)	(2.03)	(2.64)	(3.08)
N children == 4	−0.87	−0.37	−4.57	10.36
	(3.00)	(3.43)	(6.96)	(6.39)
N children >4	5.84***	2.76	2.69	10.11***
	(2.14)	(3.70)	(2.12)	(2.34)
Age youngest child (std)	−8.54***	−6.5***	−10.57***	−11.65***
	(0.53)	(0.70)	(0.90)	(1.33)
Working hours (2019-11)	All	>34	20-34	<20
Observations	1,876	979	681	216
R^2	0.293	0.282	0.38	0.483

Notes: Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors obtained by clustering on the household level. This table displays for the LISS Sample, how the relationship between remote work ability and childcare changes over the course of the pandemic. All specification control for age of the youngest child standardized by subtracting the pooled sample mean (6.8) divided by the standard deviation (4.7), as well as indicator variables indicating number of children. Sample restricted to parents working before the pandemic (Nov 2019). Column (2) restricts the sample to parents that work 35 hours or more, column (3) restricts the sample to parents working between 20 and 34 hours, and column (4) restricts the sample to parents working less than 20 hours.

Table 4.B.5. Evolution of the gender care gap and potential hours of remote work – full table

	Hrs childcare		
	(1)	(2)	(3)
Constant	18.09***	18.84***	17.88***
	(1.17)	(1.23)	(1.41)
2020-04	12.43***	10.33***	10.36***
	(1.46)	(1.42)	(1.43)
2020-11	2.90**	1.20	1.72
	(1.28)	(1.29)	(1.27)
2021-11	1.70	0.39	0.53
	(1.13)	(1.15)	(1.18)
Mother	12.46***	11.11***	12.51***
	(1.31)	(1.34)	(2.09)
Mother × 2020-04	−0.43	3.16	3.05
	(1.94)	(1.92)	(1.93)
Mother × 2020-11	−6.01***	−3.24*	−3.33*
	(1.84)	(1.86)	(1.85)
Mother × 2021-11	−3.4**	−1.41	−1.38
	(1.63)	(1.66)	(1.65)
Pot. hours remote work (std)		−2.34***	−1.49**
		(0.68)	(0.73)
Pot. hours remote work (std) × 2020-04		7.27***	7.56***
		(1.04)	(1.31)
Pot. hours remote work (std) × 2020-11		5.06***	3.57***
		(1.01)	(1.18)
Pot. hours remote work (std) × 2021-11		3.56***	3.16***
		(0.88)	(0.99)
Pot. hours remote work (std) × Mother × 2020-04			−0.66
			(2.02)
Pot. hours remote work (std) × Mother × 2020-11			3.73**
			(1.86)
Pot. hours remote work (std) × Mother × 2021-11			1.04
			(1.89)
N children == 2	−0.33	−0.29	0.48
	(1.17)	(1.15)	(1.52)
N children == 3	−1.04	−0.81	0.05
	(1.41)	(1.38)	(1.93)
N children == 4	−0.3	0.13	1.12
	(2.64)	(2.60)	(3.69)
N children >4	1.15	2.00	6.67*
	(2.48)	(2.39)	(3.87)
Age youngest child (std)	−6.6***	−6.4***	−6.4***
	(0.63)	(0.63)	(0.64)
Age youngest child (std) × Mother	−5.25***	−5.16***	−5.19***
	(0.81)	(0.81)	(0.85)
Observations	2,234	2,234	2,234
R^2	0.326	0.347	0.349

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors obtained by clustering on the household level.All specification control for age of the youngest child interacted standardized by subtracting the pooled sample mean (6.8) divided by the standard deviation (4.7) interacted with gender, as well as indicator variables indicating number of children. In column (5), we additionally interact the number of children with gender, so that the model is fully satiated. Potential remote work hours are set to zero if the individual did not work before the pandemic. Table 4.B.6 replicates this table conditional on working before the pandemic.

Table 4.B.6. Evolution of the gender care gap and potential hours of remote work – conditional on working in November 2019

	Hrs childcare		
	(1)	(2)	(3)
Constant	16.97***	17.13***	16.58***
	(1.19)	(1.16)	(1.33)
2020-04	13.07***	12.3***	12.37***
	(1.47)	(1.39)	(1.38)
2020-11	3.66***	2.43**	2.80**
	(1.29)	(1.24)	(1.20)
2021-11	2.04*	1.05	1.29
	(1.12)	(1.08)	(1.06)
Mother	11.15***	10.39***	11.09***
	(1.33)	(1.37)	(2.17)
Mother × 2020-04	1.20	3.95**	3.90*
	(2.06)	(1.97)	(2.01)
Mother × 2020-11	−5.49***	−2.65	−2.35
	(1.95)	(1.91)	(1.92)
Mother × 2021-11	−2.94*	−0.76	−0.62
	(1.74)	(1.74)	(1.76)
Pot. hours remote work (std)		−1.75**	−1.02
		(0.70)	(0.70)
Pot. hours remote work (std) × 2020-04		8.24***	8.14***
		(1.12)	(1.33)
Pot. hours remote work (std) × 2020-11		6.79***	4.80***
		(1.04)	(1.18)
Pot. hours remote work (std) × 2021-11		5.49***	4.17***
		(0.97)	(1.02)
Pot. hours remote work (std) × Mother			−1.87
			(1.38)
Pot. hours remote work (std) × Mother × 2020-04			0.28
			(2.13)
Pot. hours remote work (std) × Mother × 2020-11			4.96**
			(1.96)
Pot. hours remote work (std) × Mother × 2021-11			3.21
			(1.99)
Observations	1,876	1,876	1,876
R^2	0.319	0.368	0.37

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors obtained by clustering on the household level.All specification control for age of the youngest child interacted standardized by subtracting the pooled sample mean (6.8) divided by the standard deviation (4.7) interacted with gender, as well as indicator variables indicating number of children. Sample restricted to parents working before the pandemic (Nov 2019).

4.B.3 Labor supply: Regression equation with pot. commuting gains

$$
\begin{aligned}
\text{Working Hours}_{i,t} = &\ \alpha + \chi \text{ Pot. hrs remote work}_i + \phi \text{ Pot. hrs remote work partner}_i \\
&+ \upsilon \text{ Pot. commuting gains}_i + \nu \text{ Pot. commuting gains partner}_i \\
&+ \sum_{t=-1}^{2} (\beta_t \text{ Pot. hrs remote work}_i + \delta_t \text{ Pot. hrs remote work partner}_i \\
&+ \lambda_t \text{ Pot. commuting gains}_i + \phi_t \text{ Pot. commuting gains partner}_i) \\
&\times \mathbb{1}(\mathit{Year} = t) \\
&+ \sum_{t=-1}^{2} (\gamma_t \text{ Pot. hrs remote work}_i + \omega_t \text{ Pot. hrs remote work partner}_i \\
&+ \theta_t \text{ Pot. commuting gains}_i + \kappa_t \text{ Pot. commuting gains partner}_i) \\
&\times \mathbb{1}(\mathit{Year} = t) \times \text{Pandemic}_i \\
&+ \sum_{t=-1}^{2} \mu_t \mathbb{1}(\mathit{Year} = t) + \sum_{t=-1}^{2} \sigma_t \mathbb{1}(\mathit{Year} = t) \times \text{Pandemic}_i \\
&+ \pi \text{ Pandemic}_i + \rho \text{ Age youngest child}_{i,0} + \eta \text{ Number children}_{i,t} \\
&+ \iota \text{ Age}_{i,t} + \xi \text{ Age partner}_{i,t} + \epsilon_{i,t}
\end{aligned}
$$

4.B.4 Labor supply

Table 4.B.7. Event study: Women with children 0-15

	Working hours		
	(1)	(2)	(3)
Constant	16.86***	18.87***	20.4***
	(0.07)	(0.06)	(0.06)
Partner: Pot. hrs remote work × Time rel. treat. == -1 × Pandemic	0.01***	0.00	−0.0*
	(0.00)	(0.00)	(0.00)
Partner: Pot. hrs remote work × Time rel. treat. == 1 × Pandemic	0.01***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)
Partner: Pot. hrs remote work × Time rel. treat. == 2 × Pandemic	0.01***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)
Partner: Pot. commuting gains × Time rel. treat. == -1 × Pandemic			−0.0
			(0.00)
Partner: Pot. commuting gains × Time rel. treat. == 1 × Pandemic			0.00
			(0.00)
Partner: Pot. commuting gains × Time rel. treat. == 2 × Pandemic			0.00
			(0.00)
Pot. hrs remote work × Time rel. treat. == -1 × Pandemic		0.03***	0.02***

		(0.00)	(0.00)
Pot. hrs remote work × Time rel. treat. == 1 × Pandemic		0.01***	0.01***
		(0.00)	(0.00)
Pot. hrs remote work × Time rel. treat. == 2 × Pandemic		0.02***	0.02***
		(0.00)	(0.00)
Pot. commuting gains × Time rel. treat. == -1 × Pandemic			0.02***
			(0.00)
Pot. commuting gains × Time rel. treat. == 1 × Pandemic			0.01***
			(0.00)
Pot. commuting gains × Time rel. treat. == 2 × Pandemic			0.02***
			(0.00)
Time rel. treat. == -1	0.23***	−0.04*	−0.87***
	(0.02)	(0.02)	(0.02)
Time rel. treat. == 1	−0.14***	0.25***	0.00
	(0.02)	(0.02)	(0.02)
Time rel. treat. == 2	0.07***	0.79***	0.45***
	(0.02)	(0.02)	(0.02)
Pandemic	2.38***	2.45***	1.47***
	(0.02)	(0.03)	(0.02)
Time rel. treat. == -1 × Pandemic	−0.68***	−0.83***	−0.0
	(0.03)	(0.03)	(0.03)
Time rel. treat. == 1 × Pandemic	−0.53***	−0.59***	−0.32***
	(0.03)	(0.03)	(0.03)
Time rel. treat. == 2 × Pandemic	−0.43***	−0.52***	−0.15***
	(0.03)	(0.03)	(0.03)
Pot. hrs remote work		0.66***	0.76***
		(0.00)	(0.00)
Pot. hrs remote work × Pandemic		−0.07***	−0.07***
		(0.00)	(0.00)
Pot. hrs remote work × Time rel. treat. == -1		0.02***	0.03***
		(0.00)	(0.00)
Pot. hrs remote work × Time rel. treat. == 1		−0.04***	−0.06***
		(0.00)	(0.00)
Pot. hrs remote work × Time rel. treat. == 2		−0.08***	−0.1***
		(0.00)	(0.00)
Partner: Pot. hrs remote work	0.07***	−0.05***	−0.04***
	(0.00)	(0.00)	(0.00)
Partner: Pot. hrs remote work × Pandemic	−0.01***	−0.0	0.00
	(0.00)	(0.00)	(0.00)
Partner: Pot. hrs remote work × Time rel. treat. == -1	0.00	−0.0	0.00***
	(0.00)	(0.00)	(0.00)
Partner: Pot. hrs remote work × Time rel. treat. == 1	0.00***	0.01***	0.01***

	(0.00)	(0.00)	(0.00)
Partner: Pot. hrs remote work × Time rel. treat. == 2	0.00***	0.02***	0.01***
	(0.00)	(0.00)	(0.00)
Partner: Pot. commuting gains			−0.06***
			(0.00)
Partner: Pot. commuting gains × Pandemic			0.00**
			(0.00)
Partner: Pot. commuting gains × Time rel. treat. == -1			−0.0
			(0.00)
Partner: Pot. commuting gains × Time rel. treat. == 1			0.01***
			(0.00)
Partner: Pot. commuting gains × Time rel. treat. == 2			0.02***
			(0.00)
Pot. commuting gains			−0.06***
			(0.00)
Pot. commuting gains × Pandemic			−0.0**
			(0.00)
Pot. commuting gains × Time rel. treat. == -1			−0.02***
			(0.00)
Pot. commuting gains × Time rel. treat. == 1			0.00***
			(0.00)
Pot. commuting gains × Time rel. treat. == 2			0.00***
			(0.00)
N children == 2	−1.58***	−1.17***	−1.2***
	(0.02)	(0.01)	(0.01)
N children == 3	−3.32***	−2.19***	−2.17***
	(0.02)	(0.02)	(0.02)
N children == 4	−5.82***	−3.64***	−3.55***
	(0.06)	(0.05)	(0.05)
N children >4	−8.89***	−5.63***	−5.51***
	(0.15)	(0.13)	(0.13)
Age	0.18***	0.00**	−0.01**
	(0.00)	(0.00)	(0.00)
Partner: Age	0.04***	0.02***	0.02***
	(0.00)	(0.00)	(0.00)
Youngest child age	0-15	0-15	0-15
Observations	5,160,628	5,160,628	4,934,264
R^2	0.041	0.216	0.233

Notes: The table displays the coefficients of the event-study DiD pooled for mothers with a youngest child below 16. The first column includes only the partner remote work capability. The second column includes both the own and the partner remote work capability. The third column includes both the own and the partner remote work capability and the own and partner potential commuting gains. Remote work capability is calculated by multiplying the imputed remote work capability with actual working hours for the periods -1 and 0 and then taking the mean of it. Potential commuting gains are calculated by multiplying the imputed remote work share with the commuting distance for the periods -1 and 0, assuming that all individuals commute on five working days per week, and then taking the mean of it. The exact procedure of the variable generation and pooling is described in Section 4.4.2. All specifications include controls for own and partner age and fixed effects for the age of the youngest child and the number of children. Standard errors are obtained by clustering on the individual level.

Table 4.B.8. Event study: Women with children 0-5

	Working hours		
	(1)	(2)	(3)
Constant	13.65***	16.76***	18.68***
	(0.09)	(0.08)	(0.07)
Partner: Pot. hrs remote work × Time rel. treat. == -1 × Pandemic	0.01**	−0.0	−0.01**
	(0.00)	(0.00)	(0.00)
Partner: Pot. hrs remote work × Time rel. treat. == 1 × Pandemic	0.01***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)
Partner: Pot. hrs remote work × Time rel. treat. == 2 × Pandemic	0.01***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)
Partner: Pot. commuting gains × Time rel. treat. == -1 × Pandemic			−0.0
			(0.00)
Partner: Pot. commuting gains × Time rel. treat. == 1 × Pandemic			0.00
			(0.00)
Partner: Pot. commuting gains × Time rel. treat. == 2 × Pandemic			0.00
			(0.00)
Pot. hrs remote work × Time rel. treat. == -1 × Pandemic		0.03***	0.02***
		(0.00)	(0.00)
Pot. hrs remote work × Time rel. treat. == 1 × Pandemic		0.01***	0.01
		(0.00)	(0.00)
Pot. hrs remote work × Time rel. treat. == 2 × Pandemic		0.02***	0.00
		(0.00)	(0.00)
Pot. commuting gains × Time rel. treat. == -1 × Pandemic			0.02***
			(0.00)
Pot. commuting gains × Time rel. treat. == 1 × Pandemic			0.01***
			(0.00)
Pot. commuting gains × Time rel. treat. == 2 × Pandemic			0.02***
			(0.00)
Time rel. treat. == -1	0.48***	0.24***	−0.85***
	(0.03)	(0.03)	(0.03)
Time rel. treat. == 1	−0.36***	0.01	−0.37***
	(0.03)	(0.03)	(0.03)
Time rel. treat. == 2	−0.32***	0.32***	−0.21***
	(0.03)	(0.03)	(0.03)
Pandemic	1.89***	2.03***	0.72***
	(0.04)	(0.04)	(0.04)
Time rel. treat. == -1 × Pandemic	−0.73***	−0.88***	0.21***
	(0.04)	(0.05)	(0.04)
Time rel. treat. == 1 × Pandemic	−0.52***	−0.58***	−0.17***
	(0.04)	(0.04)	(0.04)
Time rel. treat. == 2 × Pandemic	−0.38***	−0.47***	0.07*

	(0.04)	(0.05)	(0.04)
Pot. hrs remote work		0.62***	0.69***
		(0.00)	(0.00)
Pot. hrs remote work × Pandemic		−0.05***	−0.04***
		(0.00)	(0.00)
Pot. hrs remote work × Time rel. treat. == -1		0.02***	0.04***
		(0.00)	(0.00)
Pot. hrs remote work × Time rel. treat. == 1		−0.04***	−0.06***
		(0.00)	(0.00)
Pot. hrs remote work × Time rel. treat. == 2		−0.06***	−0.08***
		(0.00)	(0.00)
Partner: Pot. hrs remote work	0.08***	−0.04***	−0.04***
	(0.00)	(0.00)	(0.00)
Partner: Pot. hrs remote work × Pandemic	−0.01***	−0.0	0.01***
	(0.00)	(0.00)	(0.00)
Partner: Pot. hrs remote work × Time rel. treat. == -1	0.00	0.00	0.01***
	(0.00)	(0.00)	(0.00)
Partner: Pot. hrs remote work × Time rel. treat. == 1	0.00	0.01***	0.01***
	(0.00)	(0.00)	(0.00)
Partner: Pot. hrs remote work × Time rel. treat. == 2	−0.0	0.01***	0.01***
	(0.00)	(0.00)	(0.00)
Partner: Pot. commuting gains			−0.05***
			(0.00)
Partner: Pot. commuting gains × Pandemic			0.00
			(0.00)
Partner: Pot. commuting gains × Time rel. treat. == -1			−0.0
			(0.00)
Partner: Pot. commuting gains × Time rel. treat. == 1			0.01***
			(0.00)
Partner: Pot. commuting gains × Time rel. treat. == 2			0.01***
			(0.00)
Pot. commuting gains			−0.06***
			(0.00)
Pot. commuting gains × Pandemic			−0.01***
			(0.00)
Pot. commuting gains × Time rel. treat. == -1			−0.02***
			(0.00)
Pot. commuting gains × Time rel. treat. == 1			0.00
			(0.00)
Pot. commuting gains × Time rel. treat. == 2			0.00
			(0.00)
N children == 2	−1.98***	−1.42***	−1.45***

	(0.02)	(0.02)	(0.02)
N children == 3	−4.43***	−3.07***	−3.02***
	(0.03)	(0.03)	(0.03)
N children == 4	−7.6***	−5.05***	−4.9***
	(0.07)	(0.06)	(0.06)
N children >4	−10.97***	−7.33***	−7.17***
	(0.18)	(0.16)	(0.17)
Age	0.31***	0.11***	0.09***
	(0.00)	(0.00)	(0.00)
Partner: Age	0.02***	0.00**	0.01***
	(0.00)	(0.00)	(0.00)
Youngest child age	0-15	0-15	0-15
Observations	2,653,084	2,653,084	2,524,400
R^2	0.058	0.215	0.231

Notes: The table displays the coefficients of the event-study DiD pooled for mothers with a youngest child below 6. The first column includes only the partner remote work capability. The second column includes both the own and the partner remote work capability. The third column includes both the own and the partner remote work capability and the own and partner potential commuting gains. Remote work capability is calculated by multiplying the imputed remote work capability with actual working hours for the periods -1 and 0 and then taking the mean of it. Potential commuting gains are calculated by multiplying the imputed remote work share with the commuting distance for the periods -1 and 0, assuming that all individuals commute on five working days per week, and then taking the mean of it. The exact procedure of the variable generation and pooling is described in Section 4.4.2. All specifications include controls for own and partner age and fixed effects for the age of the youngest child and the number of children. Standard errors are obtained by clustering on the individual level.

Table 4.B.9. Event study: Men with children 0-15

	Working hours		
	(1)	(2)	(3)
Constant	38.97***	38.59***	39.75***
	(0.05)	(0.06)	(0.05)
Partner: Pot. hrs remote work × Time rel. treat. == -1 × Pandemic	0.00*	0.00	0.01***
	(0.00)	(0.00)	(0.00)
Partner: Pot. hrs remote work × Time rel. treat. == 1 × Pandemic	0.02***	0.01**	0.01***
	(0.00)	(0.00)	(0.00)
Partner: Pot. hrs remote work × Time rel. treat. == 2 × Pandemic	0.02***	0.01***	0.02***
	(0.00)	(0.00)	(0.00)
Partner: Pot. commuting gains × Time rel. treat. == -1 × Pandemic			−0.0*
			(0.00)
Partner: Pot. commuting gains × Time rel. treat. == 1 × Pandemic			−0.0**
			(0.00)
Partner: Pot. commuting gains × Time rel. treat. == 2 × Pandemic			−0.0
			(0.00)
Pot. hrs remote work × Time rel. treat. == -1 × Pandemic		0.01***	−0.02***
		(0.00)	(0.00)
Pot. hrs remote work × Time rel. treat. == 1 × Pandemic		0.03***	0.02***
		(0.00)	(0.00)
Pot. hrs remote work × Time rel. treat. == 2 × Pandemic		0.04***	0.02***
		(0.00)	(0.00)
Pot. commuting gains × Time rel. treat. == -1 × Pandemic			0.01***
			(0.00)
Pot. commuting gains × Time rel. treat. == 1 × Pandemic			0.00**
			(0.00)
Pot. commuting gains × Time rel. treat. == 2 × Pandemic			0.00***
			(0.00)
Time rel. treat. == -1	0.69***	0.67***	−0.2***
	(0.02)	(0.02)	(0.02)
Time rel. treat. == 1	0.60***	0.90***	0.48***
	(0.02)	(0.02)	(0.02)
Time rel. treat. == 2	1.36***	1.83***	1.29***
	(0.02)	(0.03)	(0.02)
Pandemic	1.52***	1.82***	0.70***
	(0.02)	(0.03)	(0.02)
Time rel. treat. == -1 × Pandemic	−0.54***	−0.59***	0.26***
	(0.02)	(0.03)	(0.03)
Time rel. treat. == 1 × Pandemic	−1.05***	−1.35***	−0.93***
	(0.02)	(0.03)	(0.03)
Time rel. treat. == 2 × Pandemic	−1.56***	−1.9***	−1.34***

	(0.02)	(0.03)	(0.03)
Pot. hrs remote work		0.12***	0.12***
		(0.00)	(0.00)
Pot. hrs remote work × Pandemic		−0.03***	−0.0***
		(0.00)	(0.00)
Pot. hrs remote work × Time rel. treat. == -1		−0.0	0.02***
		(0.00)	(0.00)
Pot. hrs remote work × Time rel. treat. == 1		−0.03***	−0.03***
		(0.00)	(0.00)
Pot. hrs remote work × Time rel. treat. == 2		−0.04***	−0.04***
		(0.00)	(0.00)
Partner: Pot. hrs remote work	−0.07***	−0.11***	−0.1***
	(0.00)	(0.00)	(0.00)
Partner: Pot. hrs remote work × Pandemic	−0.01***	−0.0*	−0.01***
	(0.00)	(0.00)	(0.00)
Partner: Pot. hrs remote work × Time rel. treat. == -1	−0.01***	−0.01***	−0.02***
	(0.00)	(0.00)	(0.00)
Partner: Pot. hrs remote work × Time rel. treat. == 1	−0.01***	0.00	−0.01***
	(0.00)	(0.00)	(0.00)
Partner: Pot. hrs remote work × Time rel. treat. == 2	−0.01***	0.00	−0.01***
	(0.00)	(0.00)	(0.00)
Partner: Pot. commuting gains			−0.02***
			(0.00)
Partner: Pot. commuting gains × Pandemic			−0.0***
			(0.00)
Partner: Pot. commuting gains × Time rel. treat. == -1			0.00
			(0.00)
Partner: Pot. commuting gains × Time rel. treat. == 1			0.01***
			(0.00)
Partner: Pot. commuting gains × Time rel. treat. == 2			0.01***
			(0.00)
Pot. commuting gains			−0.02***
			(0.00)
Pot. commuting gains × Pandemic			−0.0
			(0.00)
Pot. commuting gains × Time rel. treat. == -1			−0.01***
			(0.00)
Pot. commuting gains × Time rel. treat. == 1			0.00***
			(0.00)
Pot. commuting gains × Time rel. treat. == 2			0.01***
			(0.00)
N children == 2	0.34***	0.27***	0.21***

	(0.01)	(0.01)	(0.01)
N children == 3	0.27***	0.16***	0.13***
	(0.02)	(0.02)	(0.02)
N children == 4	0.08*	0.06	0.11***
	(0.04)	(0.04)	(0.04)
N children >4	0.17	0.29**	0.32***
	(0.12)	(0.11)	(0.11)
Age	−0.07***	−0.06***	−0.06***
	(0.00)	(0.00)	(0.00)
Partner: Age	−0.01***	−0.03***	−0.04***
	(0.00)	(0.00)	(0.00)
Youngest child age	0-15	0-15	0-15
Observations	5,070,314	5,070,314	4,854,396
R^2	0.013	0.022	0.023

Notes: The table displays the coefficients of the event-study DiD pooled for fathers with a youngest child below 16. The first column includes only the partner remote work capability. The second column includes both the own and the partner remote work capability. The third column includes both the own and the partner remote work capability and the own and partner potential commuting gains. Remote work capability is calculated by multiplying the imputed remote work capability with actual working hours for the periods -1 and 0 and then taking the mean of it. Potential commuting gains are calculated by multiplying the imputed remote work share with the commuting distance for the periods -1 and 0, assuming that all individuals commute on five working days per week, and then taking the mean of it. The exact procedure of the variable generation and pooling is described in Section 4.4.2. All specifications include controls for own and partner age and fixed effects for the age of the youngest child and the number of children. Standard errors are obtained by clustering on the individual level.

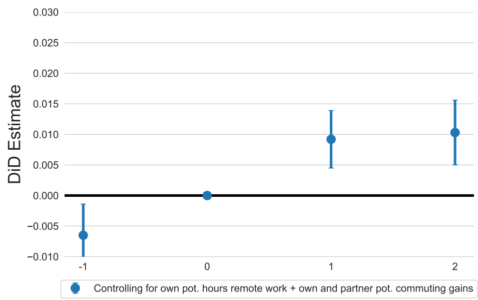
Table 4.B.10. Event study: Men with children 0-5

	Working hours		
	(1)	(2)	(3)
Constant	38.64***	38.24***	39.48***
	(0.07)	(0.07)	(0.07)
Partner: Pot. hrs remote work × Time rel. treat. == -1 × Pandemic	0.00	0.00	0.01***
	(0.00)	(0.00)	(0.00)
Partner: Pot. hrs remote work × Time rel. treat. == 1 × Pandemic	0.02***	0.01***	0.02***
	(0.00)	(0.00)	(0.00)
Partner: Pot. hrs remote work × Time rel. treat. == 2 × Pandemic	0.02***	0.01***	0.02***
	(0.00)	(0.00)	(0.00)
Partner: Pot. commuting gains × Time rel. treat. == -1 × Pandemic			−0.0
			(0.00)
Partner: Pot. commuting gains × Time rel. treat. == 1 × Pandemic			−0.0
			(0.00)
Partner: Pot. commuting gains × Time rel. treat. == 2 × Pandemic			−0.0
			(0.00)
Pot. hrs remote work × Time rel. treat. == -1 × Pandemic		0.01**	−0.02***
		(0.00)	(0.00)
Pot. hrs remote work × Time rel. treat. == 1 × Pandemic		0.03***	0.02***
		(0.00)	(0.00)
Pot. hrs remote work × Time rel. treat. == 2 × Pandemic		0.03***	0.02***
		(0.00)	(0.00)
Pot. commuting gains × Time rel. treat. == -1 × Pandemic			0.01***
			(0.00)
Pot. commuting gains × Time rel. treat. == 1 × Pandemic			0.00**
			(0.00)
Pot. commuting gains × Time rel. treat. == 2 × Pandemic			0.00***
			(0.00)
Time rel. treat. == -1	0.69***	0.69***	−0.2***
	(0.03)	(0.03)	(0.03)
Time rel. treat. == 1	0.65***	0.95***	0.49***
	(0.02)	(0.03)	(0.03)
Time rel. treat. == 2	1.38***	1.88***	1.28***
	(0.03)	(0.04)	(0.03)
Pandemic	1.44***	1.70***	0.51***
	(0.03)	(0.04)	(0.04)
Time rel. treat. == -1 × Pandemic	−0.49***	−0.55***	0.33***
	(0.03)	(0.05)	(0.04)
Time rel. treat. == 1 × Pandemic	−1.15***	−1.42***	−0.95***
	(0.03)	(0.04)	(0.04)
Time rel. treat. == 2 × Pandemic	−1.62***	−1.92***	−1.31***

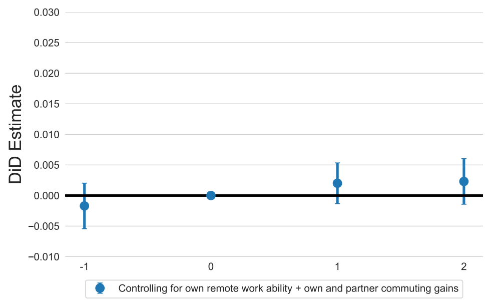
	(0.04)	(0.05)	(0.05)
Pot. hrs remote work		0.13***	0.12***
		(0.00)	(0.00)
Pot. hrs remote work × Pandemic		−0.03***	0.01***
		(0.00)	(0.00)
Pot. hrs remote work × Time rel. treat. == -1		−0.0	0.02***
		(0.00)	(0.00)
Pot. hrs remote work × Time rel. treat. == 1		−0.03***	−0.02***
		(0.00)	(0.00)
Pot. hrs remote work × Time rel. treat. == 2		−0.05***	−0.04***
		(0.00)	(0.00)
Partner: Pot. hrs remote work	−0.06***	−0.1***	−0.09***
	(0.00)	(0.00)	(0.00)
Partner: Pot. hrs remote work × Pandemic	−0.02***	−0.01***	−0.02***
	(0.00)	(0.00)	(0.00)
Partner: Pot. hrs remote work × Time rel. treat. == -1	−0.01**	−0.0*	−0.01***
	(0.00)	(0.00)	(0.00)
Partner: Pot. hrs remote work × Time rel. treat. == 1	−0.01***	−0.0	−0.01***
	(0.00)	(0.00)	(0.00)
Partner: Pot. hrs remote work × Time rel. treat. == 2	−0.02***	−0.0	−0.01***
	(0.00)	(0.00)	(0.00)
Partner: Pot. commuting gains			−0.02***
			(0.00)
Partner: Pot. commuting gains × Pandemic			−0.01***
			(0.00)
Partner: Pot. commuting gains × Time rel. treat. == -1			0.00
			(0.00)
Partner: Pot. commuting gains × Time rel. treat. == 1			0.01***
			(0.00)
Partner: Pot. commuting gains × Time rel. treat. == 2			0.01***
			(0.00)
Pot. commuting gains			−0.02***
			(0.00)
Pot. commuting gains × Pandemic			−0.0***
			(0.00)
Pot. commuting gains × Time rel. treat. == -1			−0.01***
			(0.00)
Pot. commuting gains × Time rel. treat. == 1			0.00***
			(0.00)
Pot. commuting gains × Time rel. treat. == 2			0.01***
			(0.00)
N children == 2	0.26***	0.20***	0.15***

	(0.02)	(0.02)	(0.02)
N children == 3	0.12***	0.04*	0.03
	(0.02)	(0.02)	(0.02)
N children == 4	−0.17***	−0.14**	−0.03
	(0.06)	(0.06)	(0.06)
N children >4	−0.14	0.06	0.14
	(0.16)	(0.15)	(0.15)
Age	−0.05***	−0.05***	−0.04***
	(0.00)	(0.00)	(0.00)
Partner: Age	−0.01***	−0.04***	−0.05***
	(0.00)	(0.00)	(0.00)
Youngest child age	0-15	0-15	0-15
Observations	2,593,590	2,593,590	2,471,927
R^2	0.01	0.021	0.022

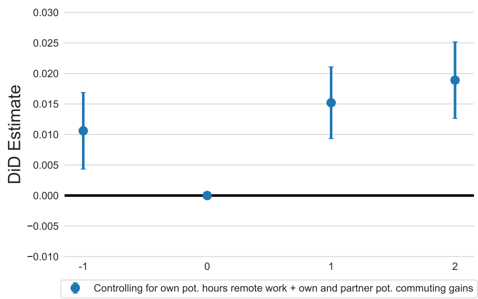
Notes: The table displays the coefficients of the event-study DiD pooled for fathers with a youngest child below 6. The first column includes only the partner remote work capability. The second column includes both the own and the partner remote work capability. The third column includes both the own and the partner remote work capability and the own and partner potential commuting gains. Remote work capability is calculated by multiplying the imputed remote work capability with actual working hours for the periods -1 and 0 and then taking the mean of it. Potential commuting gains are calculated by multiplying the imputed remote work share with the commuting distance for the periods -1 and 0, assuming that all individuals commute on five working days per week, and then taking the mean of it. The exact procedure of the variable generation and pooling is described in Section 4.4.2. All specifications include controls for own and partner age and fixed effects for the age of the youngest child and the number of children. Standard errors are obtained by clustering on the individual level.



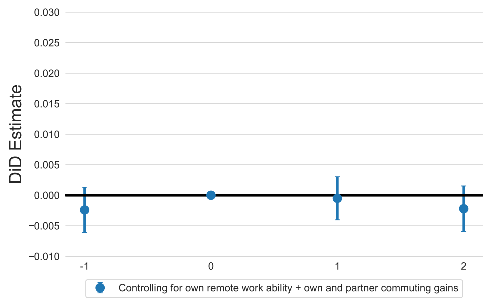
(a) Effect of fathers' potential remote work hours on mothers' working hours, youngest child < 6



(b) Effect of fathers' potential commuting gains on mothers' working hours, youngest child < 6



(c) Effect of mothers' potential remote work hours on fathers' working hours, youngest child < 6



(d) Effect of mothers' potential commuting gains on fathers' working hours, youngest child < 6

Figure 4.B.1. Direct and indirect effect of potential hours of remote work of the partner on own working hours, child below 6

Notes: The figure separates the event-study DiD estimate for the effect of potential hours of remote work of the partner on own working hours in a direct effect through potential hours of remote work and an indirect effect through potential commuting gains. Figures 4.B.1a and 4.B.1c show the event-study DiD estimates for the direct effect of potential hours of remote work of the partner on own working hours relative to the year of the Covid/Placebo shock. Figures 4.B.1b and 4.B.1d show the event-study DiD estimates for the effect through potential commuting gains of the partner on own working hours relative to the year of the Covid/Placebo shock. Results are reported separately for mothers and fathers with a youngest child below 6. The specification includes own and partner potential hours of remote work and own and partner potential commuting gains into the regression. Potential hours of remote work are calculated by multiplying the imputed remote work capability with actual working hours for the periods -1 and 0 and then taking the mean of it. Potential commuting gains are calculated by multiplying the imputed remote work capability with the commuting distance (in km), assuming that all individuals commute the same number of working days per week, for the periods -1 and 0 and then averaging it. All specifications include controls for own and partner age and fixed effects for the age of the youngest child and the number of children. Standard errors are obtained by clustering on the individual level. Complete regression results for mothers can be found in table 4.B.8. Complete regression results for fathers can be found in table 4.B.10.