Essays on the Macroeconomics of Labor Markets

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
MAREK IGNASZAK
aus Koszalin, Polen

Bonn 2021

Dekan:	Prof. Dr. Jürgen von Hagen
Erstreferent:	Prof. Dr. Petr Sedláček
Zweitreferent:	Prof. Dr. Keith Küster

Tag der mündlichen Prüfung: 17.08.2020

Contents

List of Figures

List of Tables

Introduction

Constant flow of resources across firms is vital to the economy's ability to adapt to changing circumstances, recover from adverse shocks, and generate economic growth. Therefore, the extent to which modern economic systems can deliver a steady improvement in the standards of living depends crucially on the efficacy of the markets for production inputs. In this context, the market for labor stands out as particularly important. Not only is this the largest market of all,[1] but also the one that is directly affecting virtually every member of society.

In this thesis, I aim to deepen our understanding of the functioning of the labor market. I study various frictions emerging in the market for labor and analyze how these imperfections shape the macro-economy. In Chapter 1, I put particular emphasis is on how changes in the composition of labor supply shape the characteristics of firms in the economy. Chapter 2, which a joint work with Philip Jung and Keith Kuester, examines how to use labor-market policy instruments and international transfers to stabilize business cycle fluctuations within a federation of countries. Finally, in Chapter 3, jointly with Petr Sedláček, we investigate how aggregate demand can shape long-run economic growth.

Chapter 1 is motivated by the observation that population aging and the increase in the relative supply of college-educated workers have transformed the labor force in developed economies. How do these secular trends affect the characteristics of firms in the economy? To answer this question, I develop a general equilibrium model in which both workers and firms are heterogeneous. In the model, firms of different sizes rely on different types of workers due to capital-skill complementarity in production. I estimate the model using administrative linked employer-employee data from Germany. The model predicts that the changes in the labor force composition entail the reallocation of production towards firms with a larger capital stock, which tend to be older and less dynamic. The quantitative results indicate that the demographic trends can account for most of the recently documented shift in the size distribution of firms, the falling number of new firms, and the increasing market concentration. The patterns of business dynamism across German industries provide reduced-form empirical support for the model's predictions.

In Chapter 2, we consider a union of atomistic member states, each faced with idiosyncratic business-cycle shocks. Private international risk sharing is limited. We analyze welfare gains from a federal unemployment reinsurance (RI) scheme when the member states have authority over domestic labor-market policies (layoff restrictions, hiring subsidies, and local unemployment benefits). We calibrate the economy to a stylized European Monetary Union. Focusing on the long run only, optimal federal RI is generous, provided it indexes payouts to past unemployment rates. Accounting for the transition phase, federal RI is more limited.

[1] As measured, for instance, by the share of the national income allocated to labor.

Once allowing for domestic countercyclical labor-market policies, gains from federal RI are minute.

Aggregate economic growth has traditionally been thought of as a supply-side phenomenon. In Chapter 3, we present a framework in which aggregate productivity growth is to a large extend demand-driven. In our model, heterogeneous firms can invest into research and development (R&D) but also expand their demand by lowering prices and attracting more customers. A powerful feedback loop emerges: higher firm-level productivity allows lower prices which attract more customers and, in turn, raise the incentives to innovate. Our quantitative analysis suggests that more than half of U.S. aggregate growth is in fact demand-driven.

CHAPTER 1

Composition of Labor Supply and Business Dynamism

1.1. Introduction

In recent decades, the structure of the labor force in developed economies has been fundamentally reshaped by two secular trends: population aging and the increase in the relative supply of college-educated workers. A large body of work documents the far-reaching impact of these long-run tendencies on various aspects of the economy, including income inequality, technological progress, monetary policy transmission, among many others. However, very little is known about the consequences of these demographic trends for the production side of the economy. How do these changes in the characteristics of workers affect the characteristics of firms in the economy?

In this chapter, I argue that the population aging and the increasing college attainment can lead to the concentration of production in larger, older, and less dynamic businesses. I develop and empirically validate a theory in which the composition of the labor force interacts with the life-cycle dynamics of firms. In the model, I incorporate both worker and firm heterogeneity, allowing firms of different sizes to employ different types of workers.[1] Through this channel, the changes in the composition of labor supply have heterogeneous effects on individual firms. The theory rests on complementarities in production between the physical capital of a firm and the human capital of its employees that are well-documented at the aggregate level (Krusell et al. 2000; Jaimovich et al. 2013). I propose a method to estimate these complementarities at the plant level using linked employer-employee data.

The theory helps to understand causes underlying recently documented secular decline in various measures of business dynamism in many developed economies. For example, the number of business startups has dropped, job creation and destruction rates have decreased, while economic activity has become more concentrated in large firms.[2] The model suggests that all these macroeconomic tendencies can be accounted by changes in the demographic structure of the workforce.

[1]As is standard in the literature on business dynamism, I use the term *firm size* to refer to the number of currently employed workers. I use the terms firm, plant, and production unit interchangeably. The empirical evidence presented in the paper is based on establishment-level data. In the developed model, I consider single establishment firms.

[2]Davis et al. (2006) document a secular decline in measures of job creation and destruction in the U.S. economy. Haltiwanger et al. (2011) provide evidence of a secular decline in the rate of firm creation. Decker et al. (2016) document a decline in the number of high-growth firms. See Decker et al. (2014) and Akcigit and Ates (2019) for an overview. Calvino et al. (2015) document trends in firm creation across OECD countries. Bajgar et al. (2019) show that industries in the U.S. and in Europe are becoming more concentrated.

How can the secular trends in the composition of the labor force lead to a decline in business dynamism? As the population gets older and college education becomes more prevalent, there are more experienced and educated individuals in the labor market. Their labor becomes relatively less expensive, leading firms to increase the share of experienced and educated workers in the workforce.[3] The data reveal that experience and education are *complementary* to capital. At the plant level, the change in the composition of the workforce makes capital more productive, hence firms decide to accumulate more capital and increase employment. However, capital accumulation takes time. Incumbent firms tend to be larger and have much higher capital stock than entrants[4]. As a result, incumbent firms tend to benefit much more from the increasing supply of the experienced and educated labor. As existing firms become larger and accumulate more capital, there is less space left for the startup businesses and the entry rate drops. Young firms tend to employ few workers at the beginning, but then grow quickly, creating most of the new jobs in the economy. Therefore, the falling number of new firms increases the average firm size, reduces job creation and leads to higher employment concentration in large firms. The new equilibrium features smaller number of larger, older, and less dynamic businesses.

The intuition discussed above is based on a general equilibrium model in which *both* firms and workers are heterogeneous. The model of heterogenous production units facing capital adjustment costs builds on Hopenhayn and Rogerson (1993a), Khan and Thomas (2008), and Clementi and Palazzo (2016a). However, in contrast to these papers, my model also includes worker heterogeneity. I assume that households supply three types of labor: raw labor, experienced labor, and educated labor.[5] Firms endogenously enter the economy and gradually accumulate physical capital over their life cycle. In every period, they face persistent idiosyncratic productivity shocks and can decide to shut down when they are no longer profitable. Production requires capital and the three types of labor, while the production technology allows for capital-experience and capital-education complementarities.

I parametrize the model using linked employer-employee panel data from Germany. The dataset is based on administrative records of employees and covers all establishments existing in Germany between 1976 and 2017.[6] I follow the literature and use a worker's age as a proxy for experience (Katz and Murphy 1992; Jaimovich et al. 2013). I classify workers who are 45 or older as experienced. Following Krusell et al. (2000), I assume that workers with a college or advanced degree supply educated labor.

[3]The results are driven by a fall in wages relative to the marginal productivity of labor. The mechanism remains the same even if, due to technological change, the wage level increases, as long as its growth is slower than the growth of the marginal productivity.

[4]See for instance Haltiwanger et al. (2013a) for an overview of the life-cycle patterns of firm growth. Panel (A) in Figure 1.1 summarizes these patterns among German establishments.

[5]In what follows, I use the terms "skills" and "skilled labor" to denote skills acquired by doing, or experienced labor and well as skills acquired thorough formal education. I associate a skilled worker with a either an experienced worker or an college-educated worker. I also use the terms "plant", "establishment" and "production unit" interchangeably. In the context of the model developed in Section 1.3, the terms "production units" and "firms" are equivalent and I use them interchangeably.

[6]I use the Establishment History Panel created by the Institute for Employment Research. The dataset is described in detail in Section 1.4.

I estimate the parameters governing firm entry, exit, and life-cycle dynamics using the simulated method of moments. The production complementarities are estimated in the following way. Firstly, I estimate a non-parametric relationship between firm size and workforce composition, controlling for firm characteristics (industry, age, the cohort of birth). Secondly, I choose the parameters of the production complementarities so that this relationship is the same in the model as in the data. The model replicates the German economy in the period 1976 - 1985.

The model allows me to analyze how the balanced growth path equilibrium is affected by exogenous changes in the supply of raw, experienced, and skilled labor. I modify the model parameters to reproduce the trends in the German labor market between the 1980s and 2010s. In the main experiment, I simultaneously alter the following aspects of labor supply: (i) the growth rate of the labor force, (ii) the relative supply of experienced workers, and (iii) the relative supply of college-educated workers.

The results of the main experiment show that the changes in the structure of the labor force can fully explain the increase in the average firm size and account for two-thirds of the drop in the startup rate. Moreover, they are responsible for 85% of the increase in market concentration, measured as the share of plants larger than 100 employees. Almost the entire effect is driven by the increase in the relative supply of experienced and educated workers. To understand how the labor force composition shapes business dynamism, I change one aspect at a time and examine the adjustments in the economy.

First, I consider a decline in the growth rate of the labor force, keeping the demographic composition intact.[7] The direct effect is that labor becomes scarcer and wages increase. Higher labor costs discourage potential entrants. A drop in the number of entrants leads to a lower job creation rate and a higher average firm size. The result echoes Hopenhayn et al. (2018) and Karahan et al. (2018), who use a model with homogenous workers.

The current paper, however, highlights additional general equilibrium effects. The slowdown in the labor force growth rate alone, without the concurrent changes in the age structure, would *not* lead to an increase in the average size of production units. Due to the production complementarities, the rise in the share of old firms in the economy induces a higher demand for experience and education. Since the structure of the labor supply is fixed, the higher demand is accommodated by an increase in wages of the two types of skilled labor. In response, firms switch to unskilled labor, accumulate less capital, and reduce their size. Therefore, the average size of production units decreases in the aftermath of the slowdown in the labor force growth rate. In the presence of the production complementarities, the skill composition of the labor force puts a constraint on the size distribution of firms in the economy.

Another important implication of my model is that, even for a fixed total size of the labor force, the changes in the relative supply of experienced and educated workers reshape the production side of the economy. According to the model, the increase in the relative supply of experienced or educated workers entails an increase in the average firm size and a drop in the startup rate. A similar increase in the relative supply of unskilled labor would have *opposite* consequences.

[7] Decoupling the age structure of the population from the labor force growth rate may seem counterintuitive. However, this thought experiment is useful to illustrate the mechanism.

The crux of the matter lies in the production complementarities that I estimate using the micro-level data. The estimated relationship reveals that large firms rely heavily on experienced and educated workers. When the supply of the two types of workers increases, the large, capital-rich firms benefit the most from the ensuing changes in relative wages. On the other hand, young firms do not have much capital; hence they rely on unskilled labor that becomes relatively more expensive. They anticipate that in the future they will accumulate capital, employ more educated and experienced workers, and start benefiting from the demographic change. However, only half of all firms survive the first five years.[8] Because of that, potential entrants heavily discount the future benefits against today's high prices of unskilled labor. As a result, the number of young firms in the economy falls.

I provide empirical support for the model using linked employer-employee data from Germany. Firstly, I demonstrate that the trends in the average firm size, entry, and concentration of employment are consistent with the model predictions. Secondly, I show that conditional upon age, the production units have become larger. Finally, I aggregate the establishment-level data into 3-digit industries and analyze at the industry level the relationship between the supply of experienced and educated workers and business dynamism. The data reveal that the industries that use experience and skills more intensively tend to be more concentrated, have a lower share of young firms, and are characterized by a higher average firm size.

In a related work, Engbom (2018) investigates how the population aging affects business dynamism. He shows that older workers are matched to better jobs, hence less likely to switch employers or become entrepreneurs. Due to lower worker mobility, *all firms* face a higher effective cost of job creation. As a result, business dynamism declines. The focus of the current paper is on the mechanism behind the rise in the average firm size and the increasing employment concentration in large firms, rather than worker reallocation per se. As mentioned above, the model predicts that the increase in the share of experienced and educated workers effectively increases the labor costs for startups, while reducing them for older production units.

In addition to a novel mechanism, the contribution of the paper is to account for the increase in relative supply of college-educated labor in addition to population aging.[9] I argue that, due to the imperfect substitutability of experienced and educated labor in production, accounting for the trends along both dimensions of human capital is essential for understanding the broad range of changes observed in developed economies. The model demonstrates that the reallocation of production towards large, productive, low-labor-share firms, recently documented in the U.S. and other economies by Kehrig and Vincent (2018) and Autor et al. (2019), was facilitated by concurrent trends in the composition of labor supply.

My work is broadly related to the rich literature on determinants of the college-wage premium and the returns to experience (Tinbergen 1956; Katz and Murphy 1992; Card and Lemieux 2001; Autor et al. 2003; Jeong et al. 2015). A common assumption in this strand of

[8]More precisely, 46% of establishments close within the first five years after birth. This statistic is based on the German plant-level data for the period 1976-1985. See section 1.2 for details on the data set. Similar regularity holds in the U.S., as documented in the Business Dynamics Statistics Database.

[9]Jiang and Sohail (2019) and Salgado (2019) use a model of occupational choice to argue that the falling firm creation rate in the U.S. can be attributed to the rising returns to skills in the labor market. The latter rises entrepreneurs' outside options and discourages firm entry.

the literature is that the production side of the economy can be characterized by a representative firm. Consequently, the returns to human capital are determined by the aggregate supply of different groups of workers, the aggregate stock of physical capital, and the productivity of various types of labor. I show that the changes in the demographics of *firms* play an important role in explaining trends in wage distribution and income inequality.

Outline. Section 1.2 describes the data and documents trends in the composition of the labor and business dynamism in Germany between 1975 and 2017. Section 1.3 describes the model. Section 1.4 deals with the model parameterization, while Section 1.5 discusses how the parameterized model is used to quantify the macroeconomic impact of changes in the composition of the labor force. Section 1.6 contains further empirical support for model predictions. Section 1.7 concludes.

1.2. Worker and Plant Demographics in Germany

In this section, I present the data set and give an overview of the relevant aspects of the German labor market. I begin with describing the main source of data and discussing the most important measures and definitions. Secondly, I report the changes in the composition of labor supply and trends in business dynamism in Germany between 1976 and 2017. Finally, I present the result of the regression analysis carried out to capture the relationship between worker and plant demographics.

1.2.1. Establishment History Panel (BHP). The main source of data is the Establishment History Panel (BHP). The panel was created by the Institute for Employment Research (IAB) and is based on administrative records on health, pension and unemployment insurance of private sector employees.[10] The individual-level data is then aggregated into establishments based on unique establishment identification numbers.[11] Accordingly, the data contains the establishment-level information about the demographic structure of employees, wages, occupation, and education. The panel is a 50% random sample of all German establishments with at least one employee subject to social security as of 30 June of a given year. The sample consists of between 640,000 and 1.5 million establishments per year and covers the period between 1975 and 2017. In the analysis, I restrict attention to the establishments in West Germany with at least one employee (full-time or part-time).[12]

1.2.2. Measurement and Definitions. I follow the standard approach in the literature concerning the estimation of capital-skill complementarity (see , for instance, Griliches 1969; Krusell et al. 2000) and define *skilled workers* as employees having a college degree.[13] As is common in the literature, I use age as a proxy for experience (see, for instance, Katz and Murphy 1992; Jaimovich et al. 2013). The *experienced workers* are defined as employees of

[10] Civil servants, self-employed, and students are not recorded in the data set.

[11] The establishment is defined as "a regionally and economically delimited unit in which employees work. An establishment may consist of one or more branch offices or workplaces belonging to one company." (Schmucker et al. 2018, p. 17)

[12] For more details on the data set and its construction see Schmucker et al. (2018).

[13] Some degrees that typically would be earned at universities in other countries are obtained through vocational training in Germany. In order to be consistent with the literature on capital-skill complementarity, I include the vocational training as a part of college education.

age 45 or above.[14] This is a parsimonious way to capture learning-by-doing over the course of worker's life. Most of the life-cycle increase in earnings takes place before age 45, suggesting that most of experience is accumulated before that age. Bayer and Kuhn (2018) document Using the German data that virtually all of the life-cycle wage growth attributed to worker's characteristics occurs before age 45.[15]

I define *plant size* as the total number of employees (both part-time and full-time). In order to correct for the mean-reversion bias in the estimates, I use the definition of size proposed by Davis et al. (1996). That is, I calculate plant size in period t as a simple average of the employee count in periods t and $t-1$. Formally, size of plant i in year t is defined as

$$n_{i,t} = 0.5\left(H_{i,t} + H_{i,t-1}\right), \tag{1.1}$$

where $H_{i,t}$ is the total number of employees in plant i in year t.

I define *education share* in a given plant as the share of workers who hold a college degree. Formally, education share in plant i in year t is defined as

$$S_{i,t} = \frac{H^s_{i,t} + H^s_{i,t-1}}{H_{i,t} + H_{i,t-1}}, \tag{1.2}$$

where $H^s_{i,t}$ marks the number of college-educated employees in plant i in year t. The *experience share* $E_{i,t}$ is defined analogously, as the share of employees who are 45 years old or older.

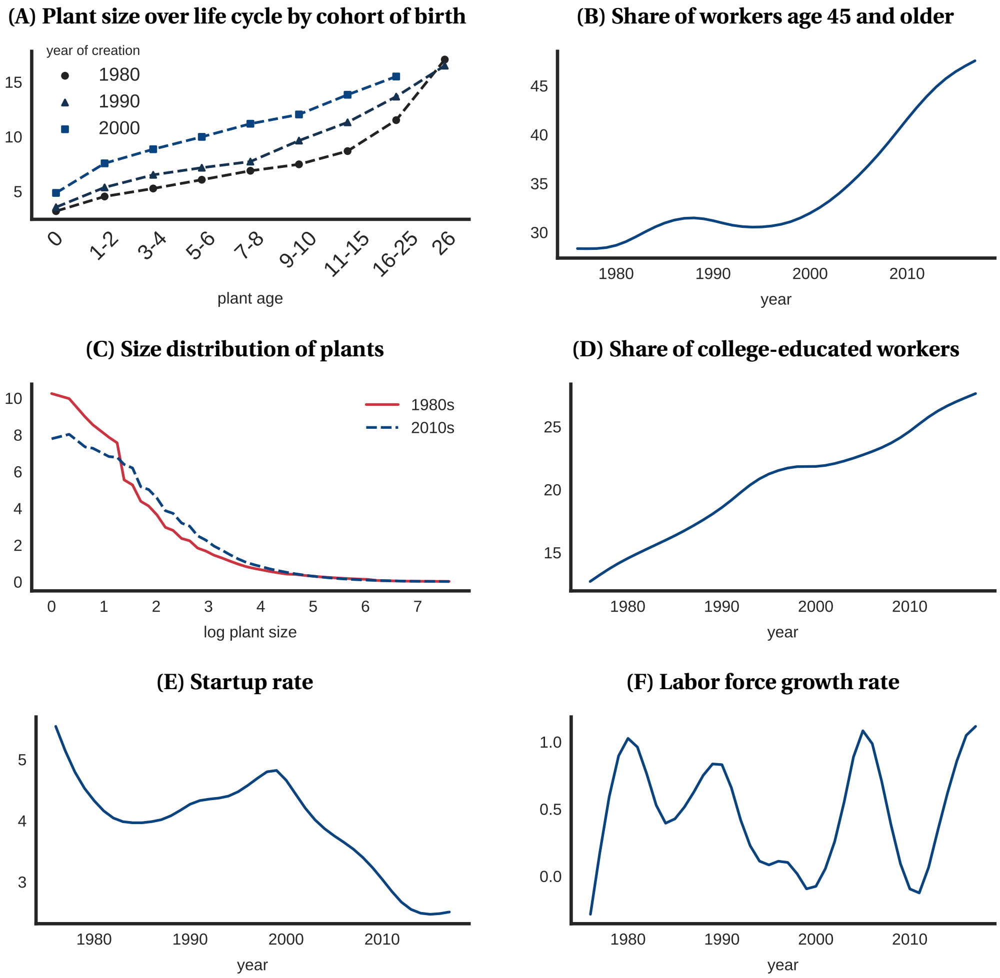
1.2.3. Changes in Labor Supply and Labor Demand in Germany. Figure 1.1 summarizes the trends in the composition of labor supply (Panels B, D, F), as well as the size distribution of plants in Germany between 1976 and 2017 (Panels A, C, E). In Panel (A) I present the relationship between plant age (the horizontal axis) and plant size. Each line corresponds to the set of plants set up in the year 1976, 1986, 1996, or 2016. In all age categories, plants established more recently are on average larger than plants established in the 1970s and the 1980s. Panel (C) reveals that, when comparing the size distribution of plants in the 1980s and 2010s, there was a significant shift towards larger units. Panel (E) illustrates that the startup rate (the share of plants of age 0 in the entire population of plants) has declined from above 5% in the 1980s to less than 3% in the 2010s. Similar secular tendencies have been documented in the U.S. and led scholars to worry about the performance of the American economy (see Decker et al. 2014; Akcigit and Ates 2019 for overview).

The right-hand side of Figure 1.1 presents the trends in the supply side of the labor market. The demographic structure of the German labor force shifted towards older and more educated individuals. As depicted in Panels (B) and (D), the share of workers of age 45 or

[14] For the purpose of estimating the capital-experience complementarity, Jaimovich et al. (2013) define workers of age 30 or older as experienced. My results hold qualitatively for experience cutoff values of 30 years and 40 years old. I also performed robustness checks using potential experience defined as a difference between the current age of a worker and the approximate age of graduation. The qualitative results hold using this alternative measure of experience.

[15] This is also consistent with the literature on life-cycle earings profiles based on the U.S data. For instance, Guvenen et al. (2017) document median wage by age for cohorts born between 1957 and 1983. Averaging their data on mean log income across all cohorts reveals that 98% of all lifetime increase occurs up to age 45.

FIGURE 1.1. Firm and worker demographics in Germany between 1976 and 2017



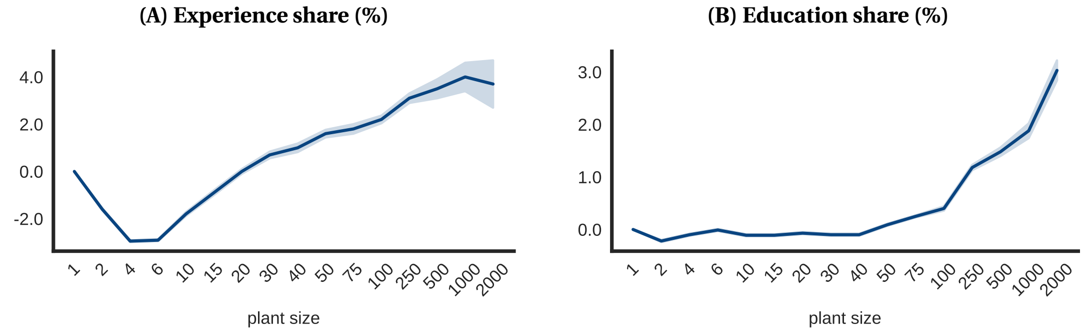
Notes: Calculations are based on the Establishment History Panel created by the IAB institute. Plant size is defined as the total number of employees. The share of college-educated workers is calculated as the total number of employees with a university degree or an advanced vocational training divided by the total number of employees. The startup rate is defined as the share of plants of age 0 in the total number of plants. All time series are smoothed with the HP filter with a smoothing parameter of 6.25.

older has increased from 30% in 1980 to more than 45% in 2017, and the share of college-educated workers has increased from 15% to 26% during this period. At the same time, there is no visible trend in the growth rate of the labor force (Panel F).[16]

1.2.4. Relationship Between Worker and Plant Demographics. In this section, I document the relationship between the demographic structure of employees and plant size. To this end, I estimate the establishment-level regressions in which the share of experienced and

[16]The aging of the German labor force is predominantly a result of the increasing participation rate and longer working life. Moreover, there is significant migration affecting the labor force growth rate.

FIGURE 1.2. The relationship between the plant size and employee demographics in the data.



Notes: The blue solid lines present the estimated coefficients β_g from regressions (1.3) and (1.4) using the establishment panel data. In the regressions the share of the experienced workers (left panel) or the share of the college-educated workers (right panel) is projected on the set of dummies indicating plant size. The coefficients are normalized such that $\beta_1 = 0$. The shaded area represents 95% confidence intervals.

college-educated employees is projected on a set of dummies indicating establishment's size. In order to control for the establishment characteristics, I include the following fixed effects: cohort of birth, industry, age, and year. I estimate the following regressions

$$E_{i,t} = \sum_g \beta_g^e \mathbf{1}\left(\text{size}_{i,t} \in g\right) + \left\{\text{year}_t,\ \text{industry}_i,\ \text{age}_i\right\} \text{ FE}, \tag{1.3}$$

$$S_{i,t} = \sum_g \beta_g^s \mathbf{1}\left(\text{size}_{i,t} \in g\right) + \left\{\text{year}_t,\ \text{industry}_i,\ \text{age}_i\right\} \text{ FE}, \tag{1.4}$$

where the depend variables $E_{i,t}$ and $S_{i,t}$ represent the shares of experienced and college-educated workers in plant i in year t. The indicator variable $\mathbf{1}\left(\text{size}_{i,t} \in g\right)$ is equal to one if the size of establishment i in year t falls into bin g.

Figure 1.2 presents the estimated coefficients β_g of the dummy variables indicating size bins. The relationship between plant size and employee experience is non-monotonic: it is decreasing on the interval 1 to 4 and increasing for larger establishments. As presented in Panel (B), the share of educated workers is a convex function of plant size, increasing sharply at the upper tail of the size distribution.[17]

I incorporate the observed regularities in a reduced-form in my model. The documented patterns may stem from the capital-experience complementarity (see Jaimovich et al. 2013) or the capital-skill complementarity (as in Krusell et al. 2000). More recently, Blanas et al. (2019) analyse 30 developed countries and show that software and robots raised the demand for high-skilled and older workers, suggesting that college education and experience are complementary to this type of capital. Furthermore, it may be that in larger and more sophisticated organizations, more difficult problems arise in the production process. Consequently, these organization require more experienced and educated employees whose task is to solve these

[17]These results are in line with the empirical literature studying relationship between workers' human capital and firm's characteristics. For example, Haltiwanger et al. (2007) find positive association between firm productivity and worker skill profile.

problems and manage the organization (see Garicano and Rossi-Hansberg 2006; Caicedo et al. 2019).

One may worry that the observed relationship between employee age (education) and employer size captures some unobserved worker characteristics that are unrelated to experience (schooling). In Appendix 1.A.3, I study a subset of establishments for which a more detailed information on employee characteristics is available. I show that the relationship between plant size and employee experience (education) holds even after controlling for additional worker characteristics, including occupation and year of birth.

In conclusion, the changes in the composition of the labor force in Germany have been accompanied by declining business dynamism. Interestingly, there is a strong relationship between the characteristics of workers and the characteristics of the production units in the economy: larger units tend to employ much more experienced and much more educated workers. This suggests that the demographic trends may be one of the factors underlying observed changes in the production side of the economy. To explore this hypothesis, in the following sections I develop a general equilibrium model of firm dynamics; I estimate the model using the establishment-level data and use the model to quantify the impact of the demographic trends on business dynamism.

1.3. Model

This section describes the model of interactions between heterogeneous plants and heterogeneous workers. I specify the household side of the model to allow for a simple representation of the following secular trends in the composition of the labor force: a slowdown in the labor force growth rate, population aging, and an increasing supply of college-educated workers.

The production side of the economy builds on Clementi and Palazzo (2016a). The model features production units indexed by productivity and the stock of capital. Production units endogenously enter the economy, gradually accumulate capital and can decide to shut down. They face persistent idiosyncratic productivity shocks. Production requires capital and different types of labor. The production function allows for the complementarities in production between the labor type and the plant type. All inputs for the production are traded in competitive markets and there is no aggregate uncertainty. I introduce the general equilibrium following Khan and Thomas (2008).

I begin with the description of households, then follow with the production side, aggregation, and equilibrium.

1.3.1. Households. Time is discrete. Next period's variables are denoted with primes. The economy is populated by a large family consisting of measure N of infinitely-lived, identical members. Household size grows over time at rate g_n so that $N' = (1 + g_n)N$. Household members derive utility from consumption and suffer disutility from supplying labor. Each household member is endowed with a stock of human capital. There are three aspects of human capital. Household members can supply raw labor l, experienced labor e, and educated labor s. The family head decides on the labor supply of each worker.

The instantaneous utility function of each household member is

$$U_l(c, n_l, n_e, n_s) = \log c - \frac{\psi_l}{1+\eta} n_l^{1+\eta} - \frac{\psi_e}{1+\eta} n_e^{1+\eta} - \frac{\psi_s}{1+\eta} n_s^{1+\eta},$$

where c denotes consumption and n_x marks the supplied hours of labor of type $x \in \{l, e, s\}$. $1/\eta$ is the Frish elasticity of labor supply. Parameters ψ_l, ψ_e, ψ_s govern the steady state supply of the three types of labor.

The family stores its wealth in one-period shares in plants. Measure $b(z, k)$ describes the number of shares in plants of type $(z, k) \in \mathcal{S}$ that the household owns, where z and k denote the plant-level productivity and capital stock. The production units are described in detail below. The household chooses the level of consumption per capita, the supply of the three types of labor, and the firm equity holdings, while taking as given the price $q_0(z, k)$ of the current shares (which includes dividends), the price $q_1(z, k)$ of the new shares, the wages w_l, w_e, w_s and the price of the final good p.

The household solves the following maximization problem

$$V^H(b) = \max_{c, n_l, n_s, n_s, b'} N \times U(c, n_l, n_s, n_s) + \beta V^H(b'), \tag{1.5}$$

subject to the budget constraint

$$pNc + \int_{\mathcal{S}} q_1(z, k) b'(\mathrm{d}[z \times k]) = N(n_l w_l + n_e w_e + n_s w_s) + \int_{\mathcal{S}} q_0(z, k) b(\mathrm{d}[z \times k]). \tag{1.6}$$

The optimal choice of labor supply equalizes the utility of an additional wage income with the disutility of an additional hour of work. The first-order conditions describing the optimal labor supply by the household members are given by

$$w_l = pc\psi_l n_l^{\eta}, \quad w_e = pc\psi_e n_e^{\eta}, \quad w_s = pc\psi_s n_s^{\eta}. \tag{1.7}$$

The total supply of raw, experienced, and skilled labor services is $N \times n_l$, $N \times n_e$, $N \times n_s$, respectively.

Let λ denote the Lagrange multiplier on the budget constraint. The first-order condition for consumption is

$$\frac{1}{c} = \lambda p, \tag{1.8}$$

meaning that the marginal gain from an additional unit of income is equal to the marginal utility from consumption. The first-order condition for the equity holdings satisfies

$$\lambda q_1(z, k) = \beta \lambda' q_0'(z, k), \tag{1.9}$$

for all shares (z, k) and equates the marginal cost of foregoing consumption with the future returns on investment in equity. Conditions (1.8) and (1.9) give rise to the Euler equation

$$\frac{c'}{c} = \beta \frac{p}{p'} \frac{q_0'(z, k)}{q_1(z, k)}, \tag{1.10}$$

for all (z, k). The Euler equation states that at the optimum, the household is indifferent between allocating resources to consumption in the current period or to consumption in the next period (through savings in equity).

1.3.2. Plants. This section describes the production side of the economy. At the beginning of each period, there is an endogenous mass of incumbent plants. Incumbents can decide whether to continue operating or to exit the market. Continuing plants choose investment subject to capital adjustment costs. Additionally, in each period there is an endogenous mass of entrants.

The plants are characterized by the idiosyncratic productivity $z \in [\underline{z}, \bar{z}]$ and by the beginning-of-period capital stock $k \in [\underline{k}, \bar{k}]$. Plant-specific productivity evolves according to the following AR(1) process

$$\log(z') = \bar{\mu}_z + \rho_z \log(z) + \sigma_z \varepsilon, \qquad \varepsilon \sim \mathcal{N}(0,1), \tag{1.11}$$

where $\bar{\mu}_z$ denotes the mean level of productivity, ρ_z is the persistence of the process, and σ_z is the standard deviation of the productivity shocks. In what follows, $F_z(z'|z)$ denotes the conditional distribution of next period's productivity z', conditional on current period's productivity z.

Production requires capital and the three types of labor. Let $L(l,e,s;z,k)$ denote a composite of labor services supplied by all workers employed in a plant of type (z,k) in period t. As described in detail below, the functional form of the labor composite depends on the plant characteristics. This assumption captures the production complementarities between employees' human capital and plant's type. Each plant has access to the following production function

$$y(z,k,l,e,s) = zk^{\alpha} L_t(l,e,s;z,k)^{\nu},$$

where $\alpha, \nu \in (0,1)$ govern the elasticities of output with respect to capital and labor, respectively, and L is the labor composite. The latter is given by

$$L(l,s,e;z,k) = \left[l^{\frac{\theta-1}{\theta}} + \bar{A}_e \times A_e(z,k) e^{\frac{\theta-1}{\theta}} + \bar{A}_s \times A_s(z,k) s^{\frac{\theta-1}{\theta}} \right]^{\frac{\theta}{\theta-1}}, \tag{1.12}$$

where $\bar{A}_e \times A_e(z,k)$ and $\bar{A}_s \times A_s(z,k)$ capture the marginal productivity of experienced and skilled labor employed at a firm of type (z,k). The productivity schedules consist of two ingredients. Firstly, $\bar{A}_e$ and $\bar{A}_s$ capture the time-varying factor-specific aggregate efficacy of the two types of labor. The second ingredient captures the capital-skill complementarities, $A_e(z,k)$ and $A_s(z,k)$, and reflect the comparative advantage of experienced and educated workers, respectively, when employed at a firm of type (z,k). Productivity of raw labor is normalized to unity in all periods. Parameter θ captures the elasticity of substitution between different types of labor. I restrict attention to the case $\theta \geq 1$ in which the labor inputs are imperfect substitutes.[18]

I capture the capital-skill complementarities by allowing the productivity schedules $A_e(z,k)$ and $A_s(z,k)$ depend on the firm type (z,k). The assumption states that the output generated by one unit of skilled or experienced labor depends on plant's capital stock and productivity. For example, the productivity of a computer scientist (skilled labor) depends on the quantity

[18]This is a standard assumption in the labor economics literature. The estimates of the elasticity of substitution between college- and high-school-educated labor vary from 1.4 in Katz and Murphy (1992), to 1.5 in Johnson (1997), to 2.5 in Card and Lemieux (2001). The latter work provides estimates of the elasticity of substitution between experienced and inexperienced workers in the range of 4 to 6.

and quality of the IT equipment at her disposal, whereas the value added of a manager (experienced labor) depends on the scale of the organization and the complexity of problems she is assigned to solve.

This functional form of the production function allows me to capture in a tractable way the key intuition behind the capital-skill complementarities: the productivity of experienced and educated labor depends on the quantity and quality of the capital. The above specification of the production function is very flexible and puts little a priori restrictions on the shape of the production complementarities. One important restriction is that the shape of complementarities, the functions $A_e(z,k)$ and $A_s(z,k)$, is time-invariant. This will be addressed in a grater detail in the quantitative experiments, Section 1.5.

The approach is also agnostic to the microeconomic mechanism underlying the complementarities. Explaining *why* skills acquired by learning-by-doing and formal education are complementary to capital is beyond the scope of this paper. I take the existence of these complementarities as given and I estimate them using the matched employer-employee data from Germany. In the estimation procedure, I choose the values of the productivity schedules $A_e(z,k)$ and $A_s(z,k)$, for all z, k, such that the model replicates the relationship between plant size and the demographic structure of its employees in the data. Thanks to the tractable specification, the production function can be estimated using only the information about the the skill composition of employees across establishments.

An alternative way of modeling the capital-skill complementarity is a nested CES structure in the production function. This allows the elasticity of substitution between capital and skilled labor to be different from the elasticity between capital and unskilled labor (see Griliches 1969; Krusell et al. 2000). In the current paper, the skills of workers are two dimensional – experience and education – making the nested CES structure less tractable. Moreover, the standard nested CES structure cannot easily replicate the salient features of the data that large firms tend to employ much more experienced and much more educated workers (see Figure 1.3). The reason is that, in the heterogeneous firm dynamics model developed in the current paper, there are two reasons for a firm to be large: a high capital stock or a high productivity. However, in a standard nested CES structure, for a given level of capital, the more productive the firm, the *lower* is the share of skilled workers it employs.[19] Finally, the specification of the production function assumed in the current paper makes it possible to estimate the production function using only information on the number and the demographic composition of employees. I do not need to measure capital at the firm level which has proven to be a notoriously difficult task. See Section 1.4 for more details on the estimation procedure.

Static problem of incumbent plants. At the beginning of each period, all incumbent plants produce the final good. To this end, they solve the following static maximization problem

$$\pi = \max_{L} \left\{ pzk^{\alpha} L^{\nu} - W(z,k)L \right\}, \tag{1.13}$$

[19] Similar point is made by Jaimovich et al. (2013) in the context of the aggregate production function and heterogeneity in the cyclical volatility of employment. To understand this intuition, consider a limiting case of Leontiev production function. In this case, the efficiency units of skilled labor used in the production are always exactly equal to the efficiency units of capital. Consequently, for a given capital stock, all increase in size stemming from productivity improvements will be accommodated by unskilled labor.

where $\pi(z,k)$ is the current profit, and $W(z,k)$ is the minimal cost of employing one unit of the composite labor L. The optimal choice of the composite labor satisfies

$$L^* = \left(\frac{pzk^{\alpha}\nu}{W(z,k)}\right)^{\frac{1}{1-\nu}}.$$

The above condition states that the plant hires additional workers up to the point in which the marginal gain of an additional unit of labor equals its marginal cost. The gain is proportional to plant's effective productivity zk^{α}. The marginal cost $W(z,k)$ depends on the firm type (z,k), since the production complementarities imply that establishment's characteristics (z,k) determine the skill composition of employees.

Given the optimal choice of the labor composite L^*, the plant decides how much services of the three labor types to hire to minimize the total labor cost. The cost-minimizing allocation satisfies[20]

$$e = \left(\bar{A}_e A_e(z,k)\right)^{\theta}\omega_e^{-\theta} l, \qquad s = \left(\bar{A}_s A_s(z,k)\right)^{\theta}\omega_s^{-\theta} l, \tag{1.14}$$

where

$$l = L^* \times \left[1 + \left(\bar{A}_e A_e(z,k)\right)^{\theta}\omega_e^{1-\theta} + \left(\bar{A}_s A_s(z,k)\right)^{\theta}\omega_s^{1-\theta}\right]^{-\frac{\theta}{\theta-1}} \tag{1.15}$$

and $\omega_e = \frac{w_e}{w_l}$, $\omega_s = \frac{w_s}{w_l}$ denote the experience and skill wage premium, respectively. The above conditions define the most cost-effective way of splitting the total labor input L^* into raw, experienced, and skilled labor. Note that the allocation of the labor demand between the three types of labor depends on the plant type (z,k). For a plant of type (z,k), the minimal cost of hiring one unit of the composite labor is

$$W(z,k) = w_l\left[1 + \left(\bar{A}_e A_e(z,k)\right)^{\theta}\omega_e^{1-\theta} + \left(\bar{A}_s A_s(z,k)\right)^{\theta}\omega_s^{1-\theta}\right]^{-\frac{1}{\theta-1}}. \tag{1.16}$$

The unit cost of composite labor is a weighted average of the wages of the three labor types. The weights depend on productivity $\left(\bar{A}_e A_e(z,k)\right)$ and $\left(\bar{A}_s A_s(z,k)\right)$ which, in turn, depend on plant's type. The type of labor that is the most productive receives the highest weight in the total wage cost. The more a plant relies on one type of labor, the more sensitive it is to changes in the corresponding wage.

Continuation and investment decisions. In each period, after producing the final good, incumbent plants incur a stochastic, i.i.d. overhead cost $c_f \sim G_f$ expressed in terms of output. After observing the realization of the shock, incumbents decide whether to shut down or to pay the cost and continue operating. Upon exit, the plant sells the remaining stock of capital $(1-\delta)k$ net of the destruction costs $g(-(1-\delta)k,k)$. The exit value is given by

$$V^x(k) = p\left[(1-\delta)k - g(-(1-\delta)k,k)\right].$$

For large realizations of the cost c_f, the continuation value of the plant may fall below the value of selling its capital stock. In this case, the plant will decide to exit.

[20]The derivations can be found in Appendix 1.B.1.

Plant that decided to continue operating invests i units of capital. The capital stock evolves according to

$$k' = (1-\delta)k + i. \tag{1.17}$$

The plant with capital k undertaking investment i pays the adjustment costs of $g(i,k)$ units of output.

At the beginning of period, the value of the incumbent plant $V(z,k)$ equals the sum of the current profit $\pi(z,k)$ and plant's continuation value that depends on the decision whether to continue operating or to exit

$$V(z,k) = \pi(z,k) + \int_{\mathbb{R}} \max\left\{V^x(k), \tilde{V}(z,k) - pc_f\right\} G_f(\mathrm{d}c_f). \tag{1.18}$$

$\tilde{V}(z,k)$ denotes the value of the plant that decided to continue operating. The integral stems from the stochastic nature of the operating costs c_f. The plant exits whenever the continuation value net of the operating costs $\tilde{V}(z,k) - c_f$ falls below the value of exit $V^x(k)$. The value of the continuing plant is given by

$$\tilde{V}(z,k) = \chi V^x(z,k) + (1-\chi)\max_i \Big[-pg(i,k) + \frac{1}{1+r}\int_Z V(z',k',\mu)F_z(\mathrm{d}z'|z)\Big], \tag{1.19}$$

where χ denotes an exogenous destruction probability (time-invariant and common across plant). Plant's discount factor $\frac{1}{1+r}$ is determined in the equilibrium.

Let $\bar{c}_f(z,k)$ be the threshold value of the cost at which the plant decides to exit. The threshold is given implicitly by

$$\bar{c}_f(z,k) = \frac{\tilde{V}(z,k) - V^x(k)}{p}. \tag{1.20}$$

The exit probability X equals the probability that the cost realization exceeds the above threshold

$$X(z,k) \equiv 1 - G_f\left(\bar{c}_f(z,k)\right). \tag{1.21}$$

In expectations, the cost paid by the plant satisfies

$$\tilde{c}_e(z,k) = \int_0^{\bar{c}_f(z,k)} c_f\, G_f(\mathrm{d}c_f). \tag{1.22}$$

Entry. Let me now describe how firms are created in the economy. A salient feature of the firm-level data in all developed economies is that the start-up business are on average much smaller than incumbents. However, conditional on survival, young firms grow rapidly creating most of the new jobs in the economy (Haltiwanger et al. 2013a; Decker et al. 2014). I specify the entry problem in the economy to capture these features of the data.

One explanation for the small size and the subsequent rapid growth of young firms is a financial friction (Albuquerque and Hopenhayn 2004; Clementi and Hopenhayn 2006; Buera 2009; Buera and Shin 2011). Due to market imperfections, prospective entrepreneurs face a constraint on the amount of capital they can borrow. As a result, young firms tend to operate at a scale that is below the optimal one. However, as firms become larger and more established, they slowly overcome market imperfections and gain easier access to capital.

With this intuition in mind, I specify the entry problem as follows. I assume that in each period, there is a mass M^e of potential entrants. Each prospective entrepreneur decides whether to pay a fixed cost of $c_e \geq 0$ units of the final good and enter the market. After paying the fixed cost and choosing the value of capital, the entrant receives a draw of initial productivity $z_e \in Z_e$ from the cdf F_e. Next, each potential entrant chooses the level of initial capital, subject to adjustment costs $g_e(k)$, an increasing and convex function. From then on, the entrant behaves like one of the incumbent production units described earlier.

Therefore, entrants differ with respect to their initial productivity. Although more productive entrants tend to choose a higher capital stock, the adjustment costs g_e induce all startups to choose capital that is lower than the unconstrained optimum. Intuitively, one can think of this environment as a reduced-form penalty function approach to approximating an imperfect capital market in which contracts are not fully enforceable and prospective entrepreneurs face collateral constraints (Marcet and Marimon 1992; Cagetti and De Nardi 2006).

Let V_e denote the value of entry defined as

$$V_e(z_e) = \frac{1}{1+r} \max_k \left[-pk - pg_e(k) + \int_Z V(z,k) F_z(\mathrm{d}z|z_e) \right]. \tag{1.23}$$

The mass of entrants is determined endogenously by the free entry condition. The potential entrant with initial productivity z_e decides to enter if and only if $V_e(z_e) \geq pc_e$. In what follows, $k^*(z_e)$ denotes the optimal capital choice of entrant receiving and initial productivity draw z_e.

1.3.3. Aggregation. The aggregate state of the economy consists of the plant measure μ describing the distribution of plants over the idiosyncratic state: the current productivity $z \in [\underline{z}, \bar{z}]$ and the beginning-of-period capital stock $k \in [\underline{k}, \bar{k}]$. The measure μ is defined on the Borel algebra $\mathscr{S}$ for the product space $[\underline{z}, \bar{z}] \times [\underline{k}, \bar{k}]$.

The measure μ includes surviving incumbents as well as startups, and evolves according to the following law of motion: for any measurable set $A \subset \mathscr{S}$ such that $z' \in A$

$$\begin{aligned} \mu'(A) = (1-\chi) & \underbrace{\int_{(z,k):\ (k^*(z,k)) \in A} (1 - X(z,k))\, F_z(z'|z)\, \mu(\mathrm{d}[z \times k])}_{\text{incumbents that choose } k^* \in A \text{ and transition to } z' \in A} \\ + M & \underbrace{\int_{z':\ z' \in A} \int_{z_e:\ k^*(z_e) \in A, V_e(z_e) \geq pc_e} G_e(\mathrm{d}z_e) F_z(z'|z_e)}_{\text{entrants that draw } z' \in A \text{, choose } k^* \in A}. \end{aligned} \tag{1.24}$$

The first line captures incumbents in the current state (z,k) that decided to continue, chose capital k^* and transitioned from z to z', for all pairs $(z', k^*) \in A$. The second line adds the mass of entrants that drew initial productivity z_e, choose capital k^* and draw next period productivity z' such that $(z', k^*) \in A$.

The aggregate variables are defined as follows. The real aggregate output is given by the production net of the operating and adjustment costs

$$\begin{aligned} Y = &\int_{\mathscr{S}} z k^{\alpha} L^*(z,k)^{\nu} \mu(\mathrm{d}[z\times k]) \\ &- (1-\chi)\int_{\mathscr{S}} (1-X(z,k)) \left[\tilde{c}_f(z,k) + g(i(z,k),k)\right] \mu(\mathrm{d}[z\times k]) \\ &- \chi \int_{\mathscr{S}} (1-X(z,k)) \left[g(-(1-\delta)k,k)\right] \mu(\mathrm{d}[z\times k]) \\ &- \int_{\mathscr{S}} X(z,k) \left[g(-(1-\delta)k,k)\right] \mu(\mathrm{d}[z\times k]) \\ &- M c_e - M\int_{\underline{z}}^{\bar{z}} g_e(k(z_e))\, G_e(\mathrm{d}z_e). \end{aligned}$$

The first line captures the output of a firm of type (z,k) and then integrates over all possible types. The integration is with respect to equilibrium measure of firms μ that dictates the "number" of firms of each type. The following lines subtract capital adjustment costs for continuers and exitors, and subtracts the entry costs and adjustment costs paid by startups.

The aggregate net investment equals the sum of investments of incumbents and entrants, net of capital sold by exitors

$$\begin{aligned} I = &(1-\chi)\int_{\mathscr{S}} (1-X(z,k)) \left(k^*(z,k) - (1-\delta)k\right) \mu(\mathrm{d}[z\times k]) + M\int_{\underline{z}}^{\bar{z}} k(z_e) G_e(\mathrm{d}z_e) \\ &- \int_{\mathscr{S}} \left[\chi(1-X(z,k))(1-\delta)k + X(z,k)(1-\delta)k\right] \mu(\mathrm{d}[z\times k]). \end{aligned} \tag{1.25}$$

The aggregate resource constraint in the economy is

$$Nc = Y - I.$$

1.3.4. Recursive Equilibrium. A recursive competitive equilibrium is a set of functions V, $\tilde{V}$, $\pi(z,k)$, k^*, L^*, X, l, e, s, c, b, n_l, n_e, n_s, firm measure μ and prices p, w_l, w_e, w_s such that, given prices,

(1) V, $\tilde{V}$ and $\pi(z,k)$ solve the plant's optimization problems (1.13), (1.18), (1.19), and X, k^*, L^*, l, e, s are the associated policy functions.
(2) V^H solves the household's optimization problem (1.5) and c, b, n_l, n_e, n_s are the associated policy functions.
(3) Labor markets clear

$$\begin{aligned} N(1-\upsilon_e-\upsilon_s)\, n_l &= \int_{\mathscr{S}} l(z,k)\mu(\mathrm{d}[z\times k]), \\ N\upsilon_e n_e &= \int_{\mathscr{S}} e(z,k)\mu(\mathrm{d}[z\times k]), \\ N\upsilon_s n_s &= \int_{\mathscr{S}} s(z,k)\mu(\mathrm{d}[z\times k]). \end{aligned}$$

(4) Equity market clears

$$b'(z',k') = \mu'(z',k'), \qquad \text{for all } (z',k')\in\mathscr{S}.$$

(5) Final good market clears by Walras law.

Balanced growth. In what follows I restrict attention to balanced growth equilibria. An equilibrium is said to be balanced growth path if the prices p, w_l, w_e, w_s are time-invariant and the measure μ satisfies

$$\frac{\mu'(A)}{N'} = \frac{\mu(A)}{N}, \qquad \text{for all measurable } A \in \mathscr{S},$$

meaning that the plant measure normalized by population is stationary. Let $\hat{\mu}$ denote the normalized firm measure. Given this normalization, all variables in the economy are stationary.

Following Hopenhayn (1992b), I rewrite the law of motion for the normalized firm measure

$$\hat{\mu}' = \frac{(1-\chi)}{1+g_n} P\hat{\mu} + \hat{M}'v, \tag{1.26}$$

where P is a bounded linear operator such that, conditional on policy functions $X(z,k)$ and $k^*(z,k)$, for every measurable set $A \in S$

$$P(z,k,A;X,k^*) = \begin{cases} \int_{z' \in A} F(\mathrm{d}z'|z) & \text{if } X(z,k) = 0 \text{ and } k^*(z,k) \in A \\ 0 & \text{otherwise} \end{cases} \tag{1.27}$$

and the measure of entrants satisfies

$$v(A) = \int_{z \in A} \int_{k \in A} G_e(\mathrm{d}z) G_k(\mathrm{d}k). \tag{1.28}$$

Intuitively, one can think about plants in the economy as aggregate resources. The larger the "stock" of plants (i.e. the number of plants), the more output can be produced. The plant stock is accumulated according to the law of motion (1.26). The population growth rate g_n can be interpreted as a "depreciation" of the plant stock.

From the stationarity of $\hat{\mu}$, it follows that

$$\hat{\mu} = \hat{M}\left(I - \frac{1-\chi}{1+g_n}P\right)^{-1} v = \hat{M}\sum_{t=0}^{\infty}\left(\frac{1-\chi}{1+g_n}P\right)^t v, \tag{1.29}$$

where P^t is the t-fold composition of P with itself and P^0 is the identity operator.

The aggregate output per capita is defined as

$$\begin{aligned} \hat{Y} = & \int_{\mathscr{S}} z k^{\alpha} L^*(z,k)^{\nu} \hat{\mu}(\mathrm{d}[z \times k]) \\ & - (1-\chi)\int_{\mathscr{S}} (1 - X(z,k)) \left[\tilde{c}_f(z,k) + g(i(z,k),k)\right] \hat{\mu}(\mathrm{d}[z \times k]) \\ & - \chi \int_{\mathscr{S}} (1 - X(z,k)) \left[g(-(1-\delta)k,k)\right] \hat{\mu}(\mathrm{d}[z \times k]) \\ & - \int_{\mathscr{S}} X(z,k) \left[g(-(1-\delta)k,k)\right] \hat{\mu}(\mathrm{d}[z \times k]) - \hat{M}^e c_e. \end{aligned} \tag{1.30}$$

The remaining aggregates are defined analogously. In per capital terms, the clearing of the markets for the three types of labor requires

$$n_l = \int_{\mathscr{S}} l(z,k)\hat{\mu}(\mathrm{d}[z\times k]),$$

$$n_e = \int_{\mathscr{S}} e(z,k)\hat{\mu}(\mathrm{d}[z\times k]),$$

$$n_s = \int_{\mathscr{S}} s(z,k)\hat{\mu}(\mathrm{d}[z\times k]). \tag{1.31}$$

1.4. Estimation and Model Fit

In this section, I describe how I bring the model to the data. The goal of the paper is to investigate the macroeconomic impact of the changes in the demographic structure of the labor force in Germany between the 1980s and 2010s. To this end, I parameterize the model so as to replicate the German economy in the period 1976 - 1985. Firstly, I discuss the assumptions concerning the functional forms used in the quantitative model. Secondly, I describe the calibration strategy. Finally, I discuss the model fit.

1.4.1. Functional Forms. The distribution of the operating costs G_f is assumed to be log-normal with the mean $\bar{\mu}_f$ and the standard deviation σ_f. The capital adjustment cost has a fixed and a convex part. Following Cooper and Haltiwanger (2006a), I assume the functional form

$$g(i,k) = \begin{cases} \zeta_0 k + \zeta_1 \frac{i^2}{k} & \text{if } \left|\frac{i}{k}\right| > 0.01, \\ \zeta_1 \frac{i^2}{k} & \text{otherwise.} \end{cases} \tag{1.32}$$

The first component of the adjustment costs does not depend on the value of investment. It aims to capture any disruption to the production process or other costs resulting from undertaking new investment, irrespective of its scale. It is paid only if the investment rate is larger than 1%. The second component is quadratic in the value of investment capturing the intuition that larger investment projects tend to be more than proportionally more difficult to complete. For entrants, the adjustment costs reads

$$g_e(k) = \zeta_e \exp(k) \tag{1.33}$$

The fixed component of the investment costs for startups is subsumed in the entry cost c_e. The initial productivity distribution for entrants, G_e, is assumed to be Pareto with the shape parameter ξ_e

$$F_e(z_e) = 1 - \left(\frac{z}{z_e}\right)^{\xi_e}.$$

I discretize the productivity grid using the Rouwenhorst method with 19 grid points. I use 501 points for the grid of capital spaced logarithmically over $[0.001, 1]$. I assume that the set of feasible initial productivity draws is $Z_e = [\mu_z - 3\sigma_z, \mu_z + 3\sigma_z]$ and is discretized using 501 equidistant points.

1.4.2. Calibration Strategy. A subset of parameters is calibrated externally, based on the literature. The remaining parameters are then estimated jointly with the simulated method

TABLE 1.1. Externally Calibrated Parameters

Parameter		Value	Target
I. Households			
β	discount factor	0.96	real interest of 4%
η	inverse of Frish elast.	0.25	standard RBC value (King and Rebelo 1999)
v_e	share of experienced workers	0.62	value for Germany in the 1980s
v_e	share of skilled workers	0.15	value for Germany in the 1980s
ψ_l	disutility, raw labor	1.13	normalization, 1 efficiency unit per worker
ψ_e	disutility, experienced labor	1.06	normalization, 1 efficiency unit per worker
ψ_s	disutility, skilled labor	1.78	normalization, 1 efficiency unit per worker
II. Incumbent plants			
α	output elasticity of capital	0.21	model-free estimates in Bachmann and Bayer (2014)
ν	output elasticity of labor	0.56	model-free estimates in Bachmann and Bayer (2014)
δ	depreciation rate	0.09	model-free estimates in Bachmann and Bayer (2014)
χ	exogenous exit rate	0.01	exit rate of plants with 250+ employees
III. Entrants			
c_e	entry costs	101.54	normalization, unit price of final good

Notes: I define experience share as the share workers of age 45 or older. I define education share as the share of workers who hold a college degree. The calculations are based on the BHP panel. Self-employed, unemployed, and public sector employees are not included in the data set.

of moments (SMM) using the linked employer-employee data from Germany. The parameters of the production complementarities between worker type and firm type are estimated through indirect inference. The procedure ensures that the relationship between the demographic structure of employees and the plant size in the model replicates the patterns estimated using the BHP establishment panel.

Externally calibrated parameters. The externally calibrated parameters are summarized in Table 1.1. One period in the model corresponds to one year. I set the discount factor β to 0.96, implying an annual interest rate of 4%.

I choose the parameters governing shares of household types such that the relative supply of experience equals 0.30 and the relative supply of skills equals 0.15, matching the shares of experienced and college-educated workers in the German data. The parameters ψ_l, ψ_e, ψ_s governing the disutility of labor of different workers are set such that each worker supplies one efficiency unit of labor. Frish elasticity of labor supply is set to 4, implying $\eta = 0.25$. The elasticity of output with respect to capital is set to $\alpha = 0.2075$ and with respect to labor to $\nu = 0.5565$. These values are directly estimated using the German firm-level data by Bachmann and Bayer (2014). The depreciation of physical capital is set to $\delta = 0.09$, the value calculated directly using the German national accounting data by Bachmann and Bayer (2014). The cost of entry is normalized to ensure that the price of the final good equals one. The exogenous destruction probability $\chi = 0.01$ is set to the value of the exit rate among plants larger than 250 employees.

Production complementarities. The interactions between the demographic structure of the labor force and business dynamism depend crucially on the parameters of the production complementarities. I estimate the productivity schedules $A_e(z,k)$ and $A_s(z,k)$ to ensure that the model replicates the relationship between plant size and employee composition in the BHP panel in the period 1976 - 1985.

The demand for experienced and educated labor services is given by equations (1.14) can be written as

$$\log e(z,k) - \log l(z,k) = \theta \log \bar{A}_e - \theta \log w_e + \theta \log A_e(z,k) \tag{1.34}$$

and

$$\log s(z,k) - \log l(z,k) = \theta \log \bar{A}_s - \theta \log w_s + \theta \log A_s(z,k) \tag{1.35}$$

The firm type (z,k) is unobservable in the data. However, in the model there is a one-to-one mapping between the firm type and the firm's total number of employees, l. The latter is directly measurable in the data.

Consider an establishments i in year t, operating in sector j. The establishments was created in year c. Let G denote a family of size bins, where a size bin $g \in G$ is an interval $[g_1, g_2) \subset \mathbb{R}^2$ such that $0 \le g_1 \le g_2 < \infty$. I estimate the following regression

$$\log e_{i,t} - \log l_{i,t} = \alpha + \sum_{g \in G} \beta_g^e \mathbf{1}\left(l_{i,t} \in g\right) + \delta_t + \delta_j + \delta_c + \varepsilon_{i,t}, \tag{1.36}$$

where δ_t, δ_j, and δ_c mark time, industry, and cohort of birth fixed effects. $e_{i,t}$ is the number of employees of age 45 years old and above, while $l_{i,t}$ is the number of employees overall. Indicator variables $\mathbf{1}\left(l_{i,t} \in g\right)$ are equal to unity if the total size of the establishment falls in bin g and zero otherwise. I estimate an equivalent regression for college-educated workers $s_{i,t}$.

Under what assumptions does the regression (1.36) identify the comparative advantage of experienced workers, $A_e(z,k)$? Intuitively, one can think about the identification in terms of the indirect inference. Equation (1.34) implies that in the model, in a given period t, the only reason for different firms to employ a different age composition of workers is the capital-experience complementarity, $A_e(z,k)$. Therefore, an estimate of $\beta^e_{g(z,k)}$ in equation (1.36) using the model-simulated data is exactly $\theta \log A_e(g(z,k)$, where $g(z,k)$ marks the size bin in which the establishments of type (z,k) falls. The indirect inference estimate of capital-experience complementarity requires choosing parameters $A_e(z,k)$ for all (z,k) such that $A_e(g(z,k)) = \exp\left(\hat{\beta}^e_{g(z,k)} \theta^{-1}\right)$. An analogous reasoning can be used to derive an estimate the capital-education complementarity $A_s(z,k)$.

This intuition illustrates that the following conditions need to hold to identify the productivity schedules:

i) elasticity of substitution between different types of labor, θ, is time-invariant.
ii) $\mathbb{E}\left(\varepsilon_{i,t}\right) = 0$ and $\mathbb{E}\left(\varepsilon_{i,t} \times \mathbf{1}\left(l_{i,t} \in g\right)\right) = 0$ for all i, all t, and all g. In other words, conditional on year, industry, and cohort of birth fixed effects, the only reason why larger firms

TABLE 1.2. Estimated Parameters and Model Fit

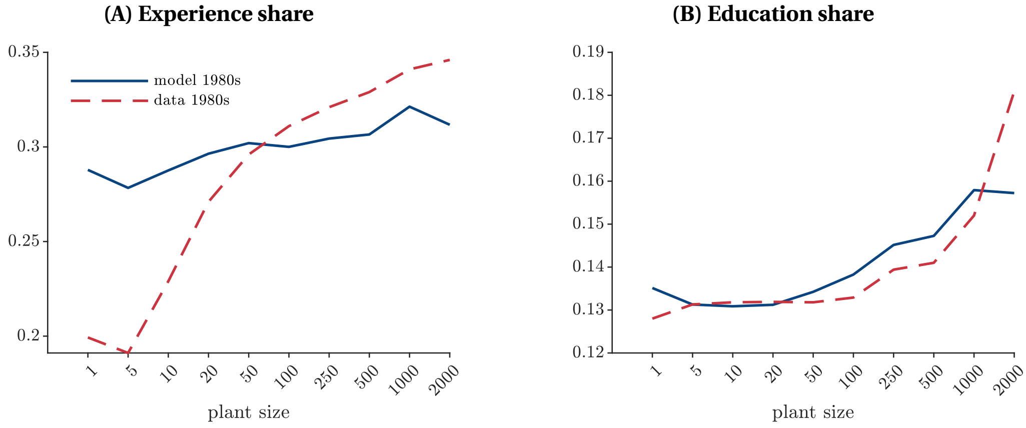
Parameter		Value	Target	Data	Model
I. Incumbent plants					
$\bar{\mu}_f$	mean of operation cost	17.86	exit rate of age-0 plant (%)	17.83	18.39
σ_f	st. dev. of operation cost	58.53	mean exit rate (%)	6.43	6.40
$\bar{\mu}_z$	mean productivity	2.35	mean size	14.66	14.69
σ_z	productivity st. dev.	0.33	share of plants of size 1000+ (%)	0.12	0.0070
ρ_z	productivity persistence	0.78	kurtosis of inv. rates	20.036	19.33
ζ_0	non-convex adj. cost	828.17	share of inv. rates above 20 (%)	13.80	14.61
ζ_1	convex adj. cost	920.35	skewness of inv. rates	2.192	3.04
θ	elast. of subst. between labor	1.60	share of firms 1-4 (%)	63.60	51.98
II. Entrants					
ξ_q	shape of init. productivity distr.	0.85	relative size of entrants (%)	19.44	21.13
ζ_e	entrant capital adj. cost	0.23	share of startups of size 100+ (%)	0.14	0.2773

Notes: The table presents the results of the SMM estimation. The three columns on the left show the parameters and their estimated values. The three columns on the right present the targeted moments and their values in the data and in the model. Moments related to the investment rates, skewness, kurtosis, and fraction of investment rates exceeding 20% are directly estimated using German data by Bachmann and Bayer (2013). The remaining targets are calculated using the BHP establishment panel.

hire more experienced and more educated workers is the time-invariant comparative advantage of those workers when working for more productive and more capital-intensive firms.

Simulated method of moments. Given the productivity schedules $A_e(z,k)$ and $A_s(z,k)$, the remaining 10 parameters are estimated jointly by minimizing the equally weighted sum of the squared percentage deviations between 10 moments in the model and their empirical counterparts. Although the parameters are estimated jointly, some moments are particularly informative about specific parameters. The mean plant size helps identify the mean idiosyncratic productivity $\bar{\mu}_z$. The establishment size distribution provides information about the persistence of the idiosyncratic productivity shock ρ_z, its standard deviation σ_z, as well as the elasticity of substitution between the three types of labor θ. The economy-wide exit rate and the exit rate of plants of age 1 provide information about the mean and the standard deviation of the operating costs $\bar{\mu}_f$ and σ_f. The size distribution of entrants informs the entrant capital cost ζ_e and the initial productivity distribution ξ_q. In order to estimate the parameters of the capital adjustment cost, I use moments of investment rates based on the German USTAN dataset, a firm-level balance sheet data base compiled by the German central bank. The moments are directly calculated and reported in Bachmann and Bayer (2013). The skewness and kurtosis of the investment rates, together with the fraction of the investment rates exceeding 20%, are used to pin down the parameters of the capital adjustment cost ζ_0 and ζ_1. Table 1.2 presents the estimated values of the parameters together with the values of the targeted moments in the data and in the model.

FIGURE 1.3. The relationship between the plant size and employee demographics in the model and in the data.



Notes: The blue solid lines present the share of the share of the experienced workers (left panel) or the share of the college-educated workers (right panel) in the model. The red dashed lines present the corresponding shares the BHP establishment panel in years 1976 - 1985.

1.4.3. Estimated Parameters and Model Fit. The right-hand side of Table 1.2 summarizes the model fit. Since there is a highly non-linear relationship between parameters and model moments, the match is not exact. Nevertheless, Table 1.2 reveals that the model matches well the size distribution of plants and their life-cycle dynamics. The estimated value of the persistence of the idiosyncratic shock is $\rho_z = 0.78$, in line with the literature.[21] The estimate of the standard deviation of the idiosyncratic shock $\sigma_z = 0.33$ is somewhat larger than typically estimated in the literature.[22] Without the ex-ante heterogeneity, a large dispersion of idiosyncratic shock is required to replicate the size distribution of plants in the data. The elasticity of substitution between the three types of labor is estimated at $\theta = 1.6$ within the range of the estimates in the literature.[23] The distribution of investment rates exhibits positive skewness and large kurtosis, while the ratio of skewness to kurtosis is similar to the one in the data. Investment rates are lumpy, as captured by frequent spikes (the investment rates exceeding 20%). The wages of experienced and college-educated workers implied by market-clearing in the baseline calibration are consistent with estimates for Germany in the 1980s (see Fuchs-Schündeln et al. 2010). The experience wage premium equals 20% and the college wage premium 42%.

[21]Using German USTAN data Bachmann and Bayer (2014) estimate the persistence of the idiosyncratic productivity component at 0.9675. The estimates for the U.S. vary from 0.43 in Castro et al. (2015) to 0.8 in Foster et al. (2008).

[22]Using German data Bachmann and Bayer (2013) estimate the average standard deviation of the idiosyncratic risk at 0.095. The estimates for the U.S. industries reported in Castro et al. (2015) vary between 0.067 to 0.352

[23]The elasticity of substitution between college- and high-school-educated workers ranges from 1.4 in Katz and Murphy (1992) for the U.S. to 2.5 in Card and Lemieux (2001) who study the U.S., the U.K. and Canada. Card and Lemieux (2001) estimate the elasticity of substitution between workers of different age groups at 4 to 6.

Figure 1.3 presents the shares of experienced (left panel) and educated (right panel) workers as a function of firm size. The blue solid lines present the shares in the model and the the red dashed line the shares in the BHP establishment panel. As a result of the production complementarities, the employment of experienced and educated workers is concentrated in the largest plants. The relationship between worker demographics and firm size is the key mechanism through which the labor force demographics interacts with business dynamism. The model matches the data very closely.

Figure 1.4 presents the life-cycle dynamics of the production units in the model. The blue solid line corresponds to the simulated data and the red dashed line to the BHP panel. All moments from the BHP panel are calculated using the sub-sample period from 1976 to 1985. Since the records begin in 1976, the oldest plant of known age is 10 years old in 1985. Panels (A) and (D) present how the exit rate varies with plant's age and size. Consistent with the data, the exit rate in the model tends to go down with plant's age and size. The shape and the magnitude of the exit rates resemble the rates in the data. It is evident from Panel (B) that the size distribution of plants in the model matches well the size distribution of plants in the BHP panel. Consistent with the data, the distribution is highly skewed, with more than 50% of establishments having less than 4 employees. Furthermore, there is a considerable mass in the right tail: more than 40% of all workers are employed in plants larger than 100, even though these plants constitute only 2.5% of plants.

The establishment growth rates decline steadily as plants get older and larger (Panels E and F).[24] Young firms in the model exhibit more rapid growth as compared to young firms in the data. The underlying reason is that in the model there is no persistent heterogeneity that may play an important role in reality (see Pugsley et al. 2018). In the model, young firms are endowed with some productivity and initial capital level. Therefore, all differences between startups are transitory and the time of birth the majority of startups are below their optimal size. As they converge towards the optimal size, young plants exhibit rapid growth. However, in reality, some small, young plants may already be at their optimal size.

To the extent that the model over-predicts the growth rate of small and young plants, it may overstate the effect of demographic trends on the aggregate job creation rate. Nevertheless, the model reproduces qualitatively the life-cycle dynamics of establishments in Germany in the 1980s (Figure 1.4) Most importantly, it replicates the relationship between plant demographics and employee demographics, as shown in Figure 1.3.

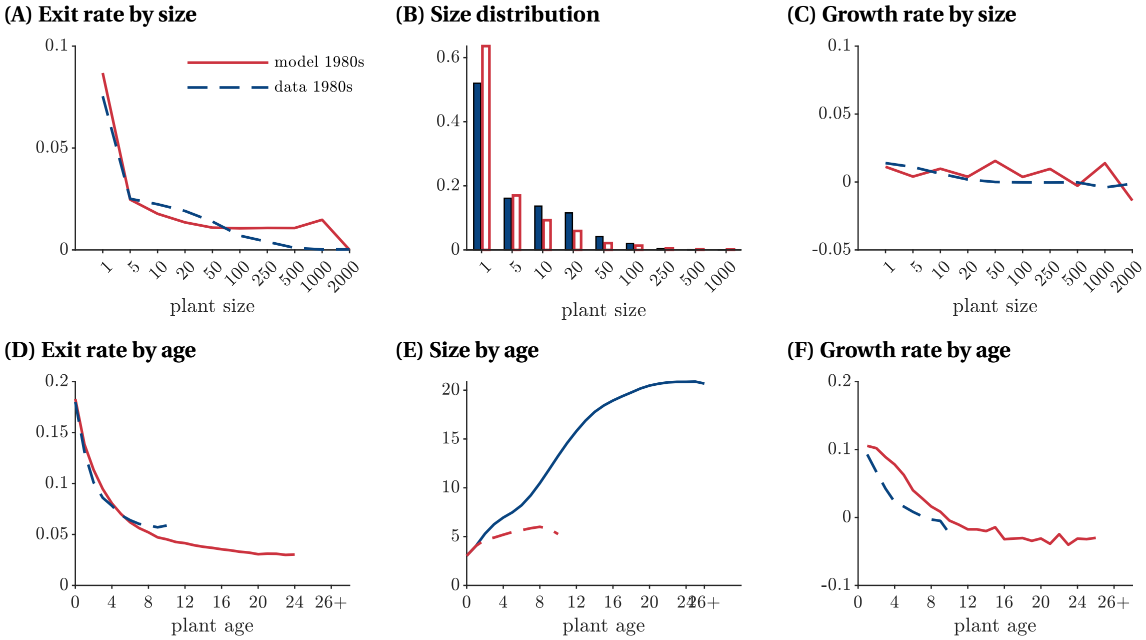
1.5. Impact of the Demographic Trends on Business Dynamism

In this section, I describe how I use the parameterized model to quantify the impact of the changes in the labor force composition on the size distribution of plants, firm entry, and job creation. As described earlier, the model replicates the West German economy in the period 1976 - 1985 (Section 1.4). By changing the parameters of household composition, I recreate the trends in the labor force observed between the 1980s and 2010s and study the new balanced growth equilibrium of the model.

I begin with an experiment in which I simultaneously change the structure of the labor force along three dimensions: the growth rate of the labor force, the relative supply of experienced labor, and the relative supply of college-educated workers. Next, I simulate the model

[24] This is also the case for U.S. firms. See Haltiwanger et al. (2013a) for an overview.

FIGURE 1.4. Business dynamism in the model and in the data.



Notes: The blue solid lines present the statistics implied by the model. The red dashed lines present the corresponding statstics in the BHP establishment panel estimated using the data from 1976 to 1985. The oldest plant of known age is 10 years old in 1985, hence the red dashed lines in Panels (D), (E), (F) do not extend beyond the age 10.

changing only one dimension at a time. This allows to look into the mechanism driving the results and to assess the relative importance of the individual trends on business dynamism.

Solving for the new equilibrium. In the model studied in the current paper, an increase in the relative supply of one labor type leads to a decrease in its price. However, despite the increase in the relative supply of experienced and college-educated workers, the data reveals that the price of experience has increased while the college wage premium has been stable (see Fuchs-Schündeln et al. 2010). Arguably, the prices of the two types of labor have been affected by the technological advancements biased towards experienced and educated labor.[25]

One may worry that the concurrent skill-biased changes in the technology invalidate the results presented in the current section. It is important to stress that the mechanism transmitting the changes in labor supply to the adjustments in the structure of labor demand does not hinge on the decline in the wage *level*. Rather, the results are driven by a fall in wages *relative* to the marginal productivity of labor. To demonstrate that the model accommodates technological changes, I design a model-consistent way to account for trends in productivity.

[25] Acemoglu (2002) and Acemoglu and Autor (2011) present an overview of evidence for the skill-biased technological change, while Caselli (2015) argues that the recent technological progress has also been biased towards experience.

Recall from equation (1.34) that the share of experienced employees in a firm (z,k) follows

$$\log e(z,k) - \log l(z,k) = \theta\left(\log \bar{A}_e - \log w_e + \log A_e(z,k)\right)$$

and analogously in the case of demand for educated labor services. To find the new stationary equilibrium that corresponds to the new labor supply composition, I proceed as follows. I fix the returns to experience and the college wage premium at the values estimated in Fuchs-Schündeln et al. (2010) for the year 2009. I then proceed to solve for the price of raw labor w_l and technology parameters $\bar{A}_e$ and $\bar{A}_e$ that guarantee the clearing of the markets for the three types of labor in the new stationary equilibrium. The production complementarities, $A_e(z,k)$ and $A_s(z,k)$, are held fixed at the estimated values.

1.5.1. Impact of All Three Trends. To quantify the effects of the demographic trends on business dynamism, I perform the following experiment. I replicate the developments in the German labor market between the 1980s and the 2010s. I change the value of the parameter describing the population growth rate from 0.64% to 0.43%. I increase the share of college-educated workers by 73%, from 15% to 26%, and the share of experienced workers by 46%, from 30% to 44%. Afterwards, I analyze the effects of these changes on the balanced growth path equilibrium. I solve for the new price of raw labor w_l, price of experience w_e, and price of skills w_s that clear the markets for the three types of labor.

The first set of results is presented in Figure 1.5. The black solid lines present the baseline simulation (the model's parameters reflect the German labor force in the 1980s), while the red dashed lines present the results under the new demographic structure. All plants benefit from the increase in the relative supply of experience and education, since these two types of labor are relatively more productive. Panel (A) in Figure 1.5 presents the mean firm value function for each admissible level of capital $\mathbb{E}_z V(z,k)$. The value function shifts upwards.

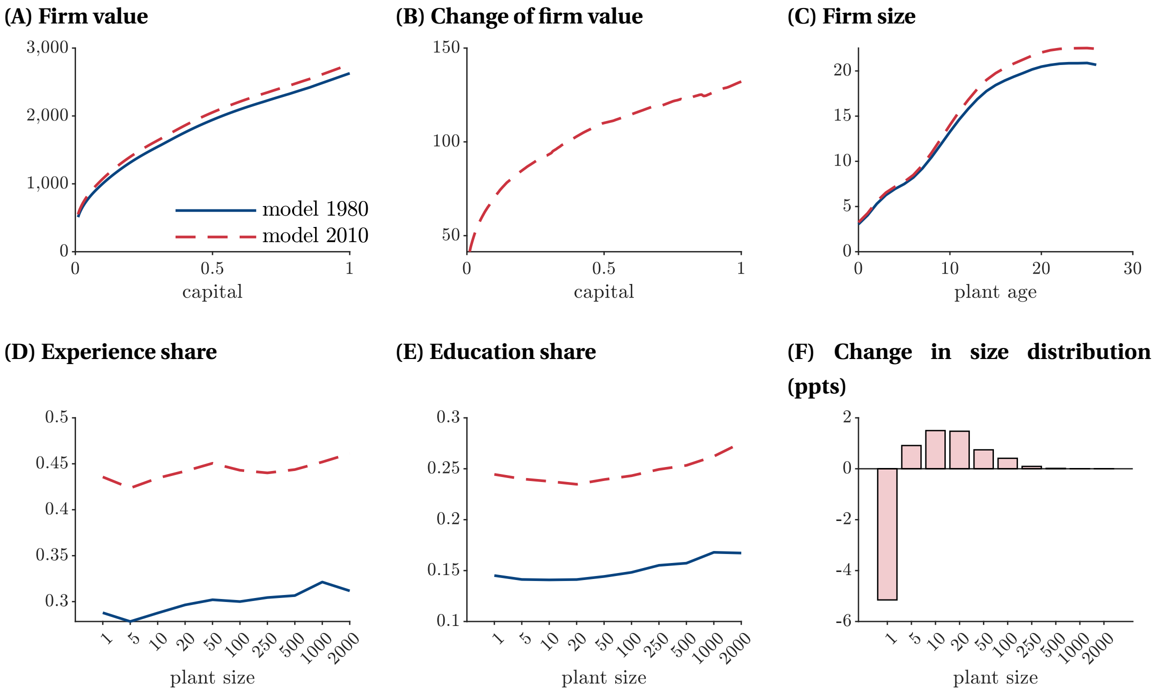
Importantly, these gains are more pronounced among the largest plants. To illustrate this point, in Panel (B) I plot the difference between $\mathbb{E}_z V(z,k)$ in the new equilibrium and in the baseline. The difference is increasing in capital. The reason is that the increase in the supply of experienced and skilled labor entails a fall in the corresponding wages. However, firms do not equally benefit from the falling wages. Recall that the unit cost of the labor composite is the weighted average of the wages of the three types of labor, with weights given by the productivity of each type of labor in a plant (z,k)

$$W(z,k) = w_l\left[1 + A_e(z,k)^\theta \omega_e^{1-\theta} + A_s(z,k)^\theta \omega_s^{1-\theta}\right]^{-\frac{1}{\theta-1}}. \tag{1.37}$$

Note that experienced and skilled labor is more productive in large plants, as $A_e(z,k)$ and $A_s(z,k)$ are increasing in z and k. Therefore, the fall in wages of experienced and educated workers brings the largest reduction in the labor costs in highly productive and capital-rich plants.

As a results of these changes in the firm value, the the size profile of plants over the life-cycle shifts upwards - plants of any age are now larger – as revealed in Panel (C) in Figure 1.5. The effect is more pronounced for the oldest enterprises since they rely heavily on experience and skills. This result is consistent with the empirical trends documented in Figure 1.1 in Panel (B).

FIGURE 1.5. Macroeconomic impact of the three secular demographics trends.

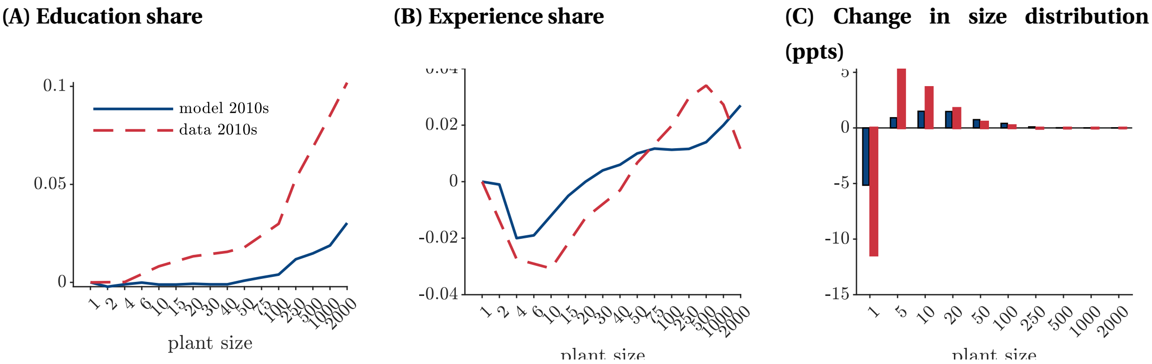


Notes: The blue solid lines present statistics from the model simualtion in the initial balanced growth equilibrium. The red dashed lines correspond to the outcomes under the new labor force composition. A general equilibrium response.

In Panel (F) I present the change (in percentage points) in the size distribution of plants. We can see that the fraction of production units employing less than 5 workers drops significantly, by 5 percentage points, while the mass shifts to the right as the number of plants in the remaining size bins increases. There are two forces driving the changes in the size distribution of plants. Firstly, all plants employ relatively more experienced and skilled workers (Panels D and E). These workers offer a type of labor complementary to capital, so the marginal product of capital rises. In consequence, all plants accumulate more capital and grow in size. Secondly, the in the new equilibrium there is a lower startup rate. Since older plants tend to be larger, this change in the number of entrants entails an increase in the average size of production units.

Figure 1.6 presents the out-of-sample predictions of the model. We see that the share of skilled workers (Panel A) and the share of experienced workers (Panel B) shift upwards in all size bins. Qualitatively, it is consistent with the model's predictions. Quantitatively, however, the model underestimates the shift of the demand towards skilled labor. Moreover, the model predicts a much higher share of experienced workers in small firms and a larger increase in the share of skilled workers in large firms. Furthermore, the model predicts the large drop in the percentage of small firms and the increase in the share of firms in the remaining size bins, as in the data (Panel C). This suggests that the shape complementarities between skilled

FIGURE 1.6. Out-of-sample Predictions of the Model.



Notes: The blue solid lines present statistics from the model under the labor force composition as in the 2010s. The red dashed lines present corresponding outcomes in the data.

labor and physical capital have changed over the last 4 decades. The impact of these changes on firm dynamics constitutes a promising avenue for future research.

Quantitatively, the three demographic trends can fully account for the increase in plant size and the decline in business creation we observe in the data. The results are summarized in Table 1.3. In the model, the average plant size increases from 14.69 to 17.13 employees, a little more than in the data. The three demographic trends account for 65% of the decline in the startup rate observed in the data. Since in the balanced growth equilibrium the exit rate equals the startup rate net of the growth rate of the labor force, the model predicts a large drop in the exit rate. This contrasts with the data, as the observed mean exit rate declines only slightly (from 5.8% to 5.6%).

Recall that large and old firms tend to grow at a slower pace and create less jobs on average (see Panels C and F in Figure 1.4). A drop in the startup rate results in an increase in the average firm age which leads to fall in the job creation. The results indicate that the demographic trends can account for half of the decline in the average growth rate and 10% of the decline in the aggregate job creation rate. On the other hand, the model predicts an increase in the job destruction rate, but this is not what we see in the data. Finally, the model accounts for 85% of the increase in concentration, measured by a share of plants larger than 100 employees. Since experienced and educated workers are more productive, the model also predicts a significant increase in the real GDP per capita.

1.5.2. Understanding the Mechanism. This section describes the effects of each of the three demographic trends in isolation. This allows me to shed light on the mechanism underlying the interactions between heterogeneous plants and heterogeneous workers.

Slowdown of the population growth rate. I begin with an experiment in which the relative supply of experience and education is held fixed, while the growth rate of the labor force

TABLE 1.3. Effects of the Demographic Trends on Business Dynamism

	1976-85		2008-17		
	data	model	data	model	$\frac{\Delta\text{ model}}{\Delta\text{ data}}$
I. Business Dynamism					
plant size	14.66	14.69	16.88	17.13	109.26%
mean growth	1.04	9.23	0.80	8.10	52.75%
job creation	14.12	17.82	13.03	17.68	10.12%
job destruction	13.60	9.72	12.80	9.79	-10.95%
startup rate	5.24	6.99	3.21	5.24	64.52%
II. Concentration					
share 100+	2.07	2.48	2.49	2.91	85.05%
III. Aggregate Outcomes					
GDP per capita	1.00	1.00	1.71	1.54	75.65%

Notes: Selected measures of business dynamism. The first two columns correspond to the initial balanced growth equilibrium, the following two columns report moments under the new structure of the labor force. The "data" columns present the moments in the BHP panel. The "model" columns present the moments implied by the model simulation. The last column tells us the what fraction of the change in the data is explained by the model. I normalize the real GDP per capita in the period 1976-1985 to unity both in the model and in the data (the actual values are taken from the World Bank database).

g_n declines from 0.64% to 0.43%.[26] All other parameters of the model are kept at their estimated values.

Firstly, I study the partial equilibrium response in which I fix the wages at the values corresponding to the levels observed in Germany in the 1980s. I will relax this assumption later on, in the general equilibrium experiment. I solve for the new stationary firm measure per capita $\hat{\mu}$ implied by the lower population growth rate and the policy functions corresponding to the prices from the baseline calibration.

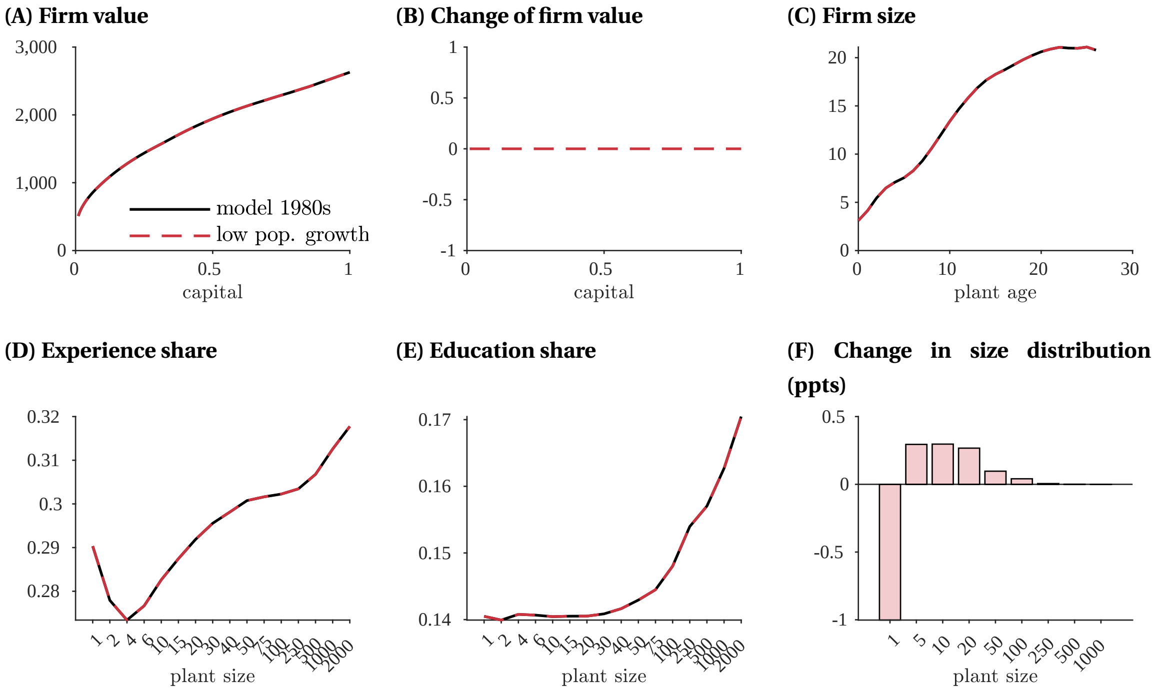
In response to the slowdown in the population growth rate, the startup rate decreases by 0.43% percentage points, from 6.99% to 6.56%. Along the balanced growth path the startup rate equals the aggregate exit rate plus the growth rate of the labor force, hence the following two forces are driving the decline in the startup rate in the model. Firstly, for a fixed aggregate exit rate, the startup rate declines by the amount equal to the difference in the population growth rates between the 1980s and 2010s, 0.64% − 0.43% = 0.21%. Furthermore, the firm distribution shifts towards larger establishments which are less likely to exit. The aggregate exit rate declines, leading to a proportional drop in the startup rate.

[26]In the U.S. the decline in the growth rate of the labor force over this time period is significantly more severe. As reported in Hopenhayn et al. (2018), in the U.S the labor force growth dropped from 1.64% in 1980s to 0.42% in 2010s.

Since entrants tend to be smaller than the incumbent plants, the drop in the startup rate causes an increase in the mean plant size from 14.69 to 15.60 employees. As older production units tend to be closer to their optimal size, they do not create as many jobs as young firms, hence the average job creation rate declines.

Figure 1.7 presents the results of this partial equilibrium experiment. As depicted in Panel (F), the size distribution of production units shifts towards larger entities. Since the wages are fixed, the firm value (depicted in Panel A) and the relative employment of experience and education (Panels D and E) remain the same as in the baseline.

FIGURE 1.7. Slowdown in the population growth rate in partial equilibrium.



Notes: The black solid lines present the statistics from the model simualtion in the 1980s. The red dashed line corresponds to the model outcomes under the lower population growth rate.

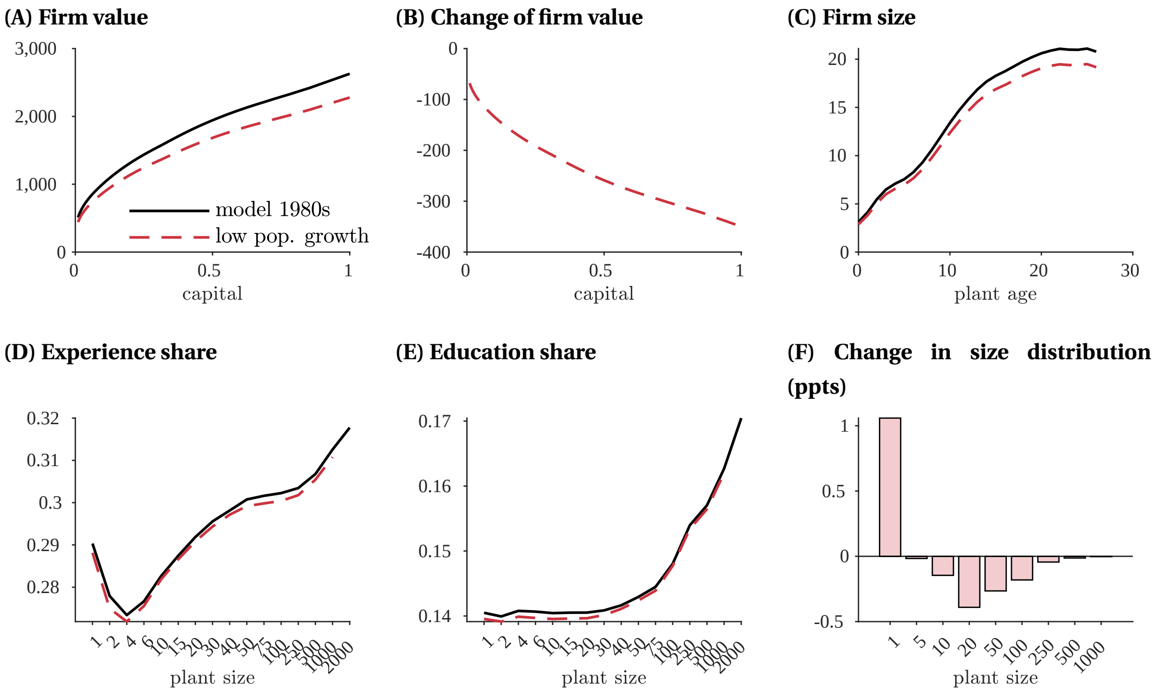
The partial equilibrium effects in the current model are similar to the results in Hopenhayn et al. (2018) and Karahan et al. (2018). Using a model with homogenous labor force, the authors show that the economy adjusts to the lower growth rate of the population entirely through changes in the entry rate, while prices remain intact. This corresponds to the partial equilibrium studied above.

However, the production complementarities in the current model generate novel general equilibrium effects. The availability of skilled and experienced labor puts a constraint on the size distribution of plants. To see this, I consider the following general equilibrium experiment. Again, I start with changing the population growth rate from 0.64% to 0.43%. However this time, in contrast to the previous exercise, I allow wages to adjust to clear the markets for the three types of labor. At the same time I keep the relative supply of experience and education fixed at the level of the baseline calibration. This may seem counterintuitive since, in reality, the slowdown in the growth rate of the labor force tends to lead to an older workforce. Nevertheless, I consider this hypothetical scenario to better illustrate the mechanism.

The economy adjusts to the new demographic structure in the following way. As explained above, the slowdown in the population growth rate implies a shift in the size distribution of plants towards larger units. Due to the production complementarities, a greater number of large firms entails a higher aggregate demand for experience and skills. Since the relative supply of the two types of labor is fixed, the returns to experience and the college wage premium have to rise to clear the labor market. The experience premium increases from 20% to 26% and the college wage premium from 40% to 44%.[27]

Affected by the rising prices of experience and skills, capital-rich establishments shrink. Young plants tend to have less capital and rely on raw labor that becomes relatively less expensive. Consequently, the life-cycle size profile shifts downwards and the size distribution of establishments tilts towards smaller units. These results are presented in Figure 1.8.

FIGURE 1.8. Slowdown in the population growth rate in the general equilibrium.



Notes: The black solid lines present statistics from the model simualtion in the initial balanced growth equilibrium. The red dashed line corresponds to the model outcomes under lower population growth rate.

[27] Fuchs-Schündeln et al. (2010) report a stable university wage premium in Germany between 1985 and 2004, and an increase in the price of experience from 20% to 40% in that period.

The model predicts a reduction in the average size of plants from 14.69 to 13.80 employees. In the general equilibrium, the startup rate declines slightly less than in partial equilibrium, from 6.99% to 6.65%. The direct effect of the lower growth rate of the labor force is partially offset by the indirect effect that stems from the reduction in plant size. Smaller units record lower profits and are more likely to leave the market. The aggregate exit rate increases and pushes upwards the rate of firm creation.

The above experiment illustrates that unless the slowdown in the growth rate of the labor force is accompanied by a sufficient increase in the stock of human capital, the life-cycle size profile of plants shifts downwards in response to the slowdown in the population growth. The patterns observed in the data seem to be consistent with the above reasoning. The average plant size conditional on age has declined in the U.S. (see Hopenhayn et al. 2018), in contrast to Germany. At the same time, the slowdown in the growth rate of the labor force has been much more severe in the U.S., making the induced shift towards older, larger firms more pronounced. However, the trends in the share of experienced and skilled workers are quantitatively similar. Interpreted through the lens of the current model, the decline in the average firm size conditional on age observed in the U.S., might be a result of an insufficient supply of skilled and experienced workers in the U.S. labor market.

Increase in the share of experienced and college-educated workers. Let us now consider the macroeconomic impact of an exogenous rise in the relative supply of experienced and college-educated workers. I increase the share of experienced workers from 30% to 44% and the share of skilled workers from 15% to 26%. I solve for the new wages of raw labor as well as the new skill-biased technology parameters to clear the labor market.

In response to changes in supply, the relative wages of experienced and skilled workers fall. The demand curve for the two types of labor shifts upwards. Due to the production complementarities, capital becomes more productive and firms accumulate more of it. The equilibrium distribution of firms shifts towards larger units.The quantitative results, summarized in Table 1.4, are very similar to the experiment in which all three demographic trends are considered jointly. We can see that almost the entire change in the size distribution of plants in Germany is driven by the trends in the supply of experience and skills. The slowdown in the growth rate of the population plays only a minor role.

The importance of plant heterogeneity. In this section, I decompose the changes in the aggregate variables into direct and indirect effects. The two effects are defined in the following way. As I explained in the previous sections, in response to the fall in wages, plants of all types (z, k) change their policy functions. I will call this the *direct effect.* In other words, the direct effect describes the changes in aggregate outcomes if there was no shift in the size distribution of firms. However, the important implication of my model is that the changes in wages lead to a shift in the equilibrium plant distribution and the following changes in aggregates, all of which I call the *indirect effect.* The indirect effect captures the changes in the aggregate outcomes stemming from differences in the number of firms of type (z, k), keeping the behavior of firms constant.

Table 1.5 presents the relative importance of the indirect effect for selected aggregate statistics. Each row corresponds to one of the general-equilibrium experiments discussed in the previous sections. In the first row I consider all three trends (see Section 1.5.1). The following rows correspond to the experiment in which I only adjust the population growth rate

TABLE 1.4. Impact of the changes in the demographic structure of the labor force

	1976-85		2008-17			
	data	model	low g_n	education & aging	all	data
I. Business Dynamism						
plant size	14.66	14.69	13.80	16.92	17.13	16.88
mean growth	1.04	9.23	9.28	8.03	8.10	0.80
job creation	14.12	17.82	17.83	17.67	17.68	13.03
job destruction	13.60	9.72	9.72	9.79	9.79	12.80
startup rate	5.24	6.99	6.65	5.54	5.24	3.21
II. Concentration						
share 100+	2.07	2.48	2.17	2.95	2.91	2.49
III. Aggregate Outcomes						
GDP per capita	1.00	1.00	1.03	1.40	1.54	1.71

Notes: Table presents selected measures of business dynamism for all quantitative experiements. The columns labeled "data" correspond to moments calculated using the BHP establishment panel. All results include general equilibrium effects.

(the second row), and supply of experience and education (the third row). Comparing the effects in the three experiments, the indirect effect accounts for around 60% of changes in the mean size of establishments. It is also responsible for 42% to 50% of the adjustments in the aggregate exit rate. The magnitude of the indirect effect is notable, and it demonstrates the importance of modelling explicitly the heterogeneity in the production side of the economy.

Moreover, the indirect effect accounts for more than 100% of the change in GDP per capita, meaning that the direct effect is negative. The negative sign of the direct effect hinges on the assumption of decreasing returns to scale, which implies that a large number of small firms generates output that is *higher* than the output generated by a small number of large firms (keeping the number of workers fixed). In this case, the entire increase in the GDP per capita is due to the indirect effect: the shift in the plant distribution towards entities that use workers' human capital more effectively. The opposite is true in the experiment altering the population growth rate (the second row in Table 1.5): firms get smaller and, all other things equal, more productive.

It is worth stressing that my results differ from the predictions of the homogenous-labor model studied in Hopenhayn et al. (2018) and Karahan et al. (2018). In their model, the demographic change is accommodated fully by the entry margin, while the equilibrium prices and the firm's policy functions remain intact. Using the terminology introduced in this section, the indirect channel is the only channel of adjustment. However, I argue that in the presence

TABLE 1.5. Percent of the aggregate change due to the indirect effect.

experiment	mean size	exit rate	GDP p.c.	demand for experience	demand for skills
all trends	62.80%	49.77%	115.24%	0.09%	-0.03%
lower pop. growth	-23.07%	116.78%	-2499.91%	0.66%	2.52%
education & aging	56.82%	41.88%	120.86%	0.09%	-0.01%

Notes: Decomposition of changes in selected measures of business dynamism into direct and indirect effect. I define the direct effect as the changes implied by the new policy functions, keeping constant the number of plants of each type. The indirect effect captures the changes stemming from the shift in the equilibrium distribution of plants $\hat{\mu}(z,k)$.

of the production complementarities, changes in the supply of human capital entail a revision of the firm's strategy. The model implies that plants born in later cohorts will be larger at any given age. This is indeed the case in Germany as documented in Panel (A) in Figure 1.1.

Interestingly, virtually none of the increase in the aggregate demand for experienced and educated labor results from the indirect effect. To understand why, recall the results presented in Panels (E) and (F) in Figure 1.5. A larger supply of experience and skills is accommodated by the parallel shift of the labor demand curve for all plants. The changes in the equilibrium distribution of plants become inconsequential.[28]

In conclusion, the estimated production complementarities imply that the changes in the demographic structure of the labor force have a large macroeconomic impact. The consequences of the rising supply of skilled and experienced labor include a higher average size of production units, a lower startup and exit rates, and a slower pace of worker reallocation.

1.6. Reduced-form Evidence for the Model's Predictions

The model developed in the current paper predicts that in the economy characterized by an older and more educated labor force, the number of business startups is lower, while economic activity is more concentrated in large firms. In this section, I provide a reduced-form evidence that the relationship predicted by the model holds across industries in Germany.

To this end, I aggregate the establishment-level data in the BHP panel into 301 industries using the 3-digit classification code. As for the measures of business dynamism, I consider five different characteristics of the industry: the average plant size, share of workers employed in plants larger than 100 employees, share of plants of age 11 or older (which I call mature), share of plants of age 0 (which I call startups), and share of workers employed in the 5 largest plants. I calculate these measures of each industry in each year and test the relationship between the business dynamism and the relative intensity of using experienced and

[28]In some cases the contribution of the indirect effect is negative. This is driven by an increase in the mass of plants in the firm size distribution where the demand for skilled labor becomes a downward-slopping function of size. The latter comes from the substitution between different types of labor.

educated labor. I estimate a set of regressions in the following form

$$\log\left(y_{i,t}\right) = \alpha_1 + \alpha_2 \log\left(E_{i,t}\right) + \alpha_3 \log\left(S_{i,t}\right) + \zeta_i + \zeta_t + \varepsilon_{i,t}, \tag{1.38}$$

where $y_{i,t}$ is a measure of business dynamism in industry i in year t, whereas $E_{i,t}$ marks the percent of experienced workers into the total number of employees in a given industry. The independent variable $S_{i,t}$ is defined as the percent of college-educated workers into the total number of employees in a given industry. I include time and industry fixed effects, ζ_t and ζ_i.

Firstly, I estimate (1.38) using OLS. However, the correlations between the industry characteristics and demographics of its employees may reflect common underlying factors that lead both to lower business dynamism and more intensive use of human capital. To alleviate these concerns, I use the aggregate shares experienced and skilled labor as an instrument for the industry-level shares. I build on Nekarda and Ramey (2011) who identify industry-level effects of the aggregate changes in government spending. The identification is based on the fact that different industries are differently exposed to changes in the aggregate supply of experienced and skilled labor.[29] Therefore, I also estimate regressions (1.38) using two-stage least squares. In the first stage, the share of experienced (educated) workers in a given industry is projected on the share of experienced (educated) workers in the aggregate economy, allowing for the elasticity to vary across industries. That is, I estimate

$$\begin{aligned} \hat{E}_{i,t} &= \beta_0 + \beta_i \bar{E}_t + \zeta_i + \zeta_t + \varepsilon_{i,t}, \\ \hat{S}_{i,t} &= \gamma_0 + \gamma_i \bar{S}_t + \xi_i + \xi_t + \nu_{i,t}, \end{aligned} \tag{1.39}$$

where $\bar{E}_t$ and $\bar{S}_t$ mark the shares of experienced and educated worker in the whole economy, ζ_i, ξ_i denote industry fixed effects and ζ_t, ξ_t time fixed effects. In the second stage, I estimate regressions (1.38) in which I instrument experience and education shares with the values predicted from (1.39).

The five measures of business dynamism define the five different specifications presented in columns (1)-(5) in Table 1.6 The table presents the 2SLS estimates, while the OLS results are similar and presented in Appendix 1.A.1.

German industries that were more exposed to the increase in the aggregate share of experienced and educated workers tend to exhibit larger average plant size, are characterized by a lower rate of entry and are more concentrated. These empirical results are consistent with the predictions of the model and quantitatively significant. For instance, a one-percent increase in the share of skilled workers translates on average into a one-half percent rise in the average plant size.

In Appendix 1.A.2, I show that across OECD countries there are similar correlations between the share of experienced and educated workers and the size distribution of manufacturing firms.

1.7. Conclusion

This paper develops and empirically validates a novel theory in which the composition of the labor force interacts with the life-cycle dynamics of firms. The interaction rests on

[29] The intuition behind the identification is similar to the Bartik shift-share instrument named after Bartik (1991).

TABLE 1.6. Employee demographics and business dynamism across German industries.

	(1) size	(2) 100+ share	(3) mature share	(4) startup share	(5) top 5 share
experience share	0.196^{***}	0.523^{***}	0.0566	-0.147^{*}	-0.0780^{*}
	(0.0912)	(0.0593)	(0.0368)	(0.0792)	(0.0433)
education share	0.575^{***}	0.802^{***}	0.0342	0.0977	0.720^{***}
	(0.106)	(0.107)	(0.0495)	(0.108)	(0.0799)
Observations	12774	11181	12749	10674	12774
Adjusted R^2	0.941	0.921	0.673	0.610	0.947

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table presents the results of regressions (1.38) for 301 industries in Germany and all years 1976 - 2017. Five columns correspond to five different different measures of business dynamism: mean size, the share of employees in plants larger than 100 employees, share of plants older than 11 years, the share of plants of age 0, and the employment share of the top 5 largest plants. The independent variables include experience share (first row) and education share (second row).

complementarities in production between firms' capital and workers' experience and education. I estimate the complementarities using linked employer-employee data from Germany. The results demonstrate that changes in the structure of the German labor force between the 1980s and the 2010s can account for the observed reallocation of production towards larger, older, and less dynamic businesses. Consistent with the model mechanism, German industries more exposed to the secular trends in labor force composition, tend to exhibit larger plant size, higher concentration, and lower dynamism.

My results prove that the demographic structure of the labor force determines the types of firms operating in the market. In conclusion, to fully assess the macro-economic impact of demographic trends, it is crucial to account for the interactions between worker and firm heterogeneity.

Appendices to Chapter 1

1.A. Further Empirical Results

In this section I provide further empirical support for the main prediction of the model. Secondly, I document the cross-country correlations between the labor force composition and business dynamism in a sample of OECD countries. Finally, I show that the discussed relationships between firm's size and workers' human capital hold, even conditional on additional characteristics of workers, such age, occupation and cohort of birth.

1.A.1. OLS Results of Cross-industry Regressions. In this section, I present results of regressions (1.38) estimated by OLS. The results are qualitatively similar to 2SLS. Notable exception is positive relationship between the startup rate and share of skilled workers. This suggest that dynamic sector with large number of startups attracts skilled workers.

TABLE 1.7. Employee demographics and business dynamism across German industries.

	(1) size	(2) 100+ share	(3) mature share	(4) startup share	(5) top 5 share
experience share	0.228***	0.313***	0.116***	-0.283***	-0.0152***
education share	-0.143	0.458***	0.0869	0.343***	0.672***
Observations	12774	11181	12749	10674	12774
Adjusted R^2	0.940	0.918	0.677	0.610	0.948

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

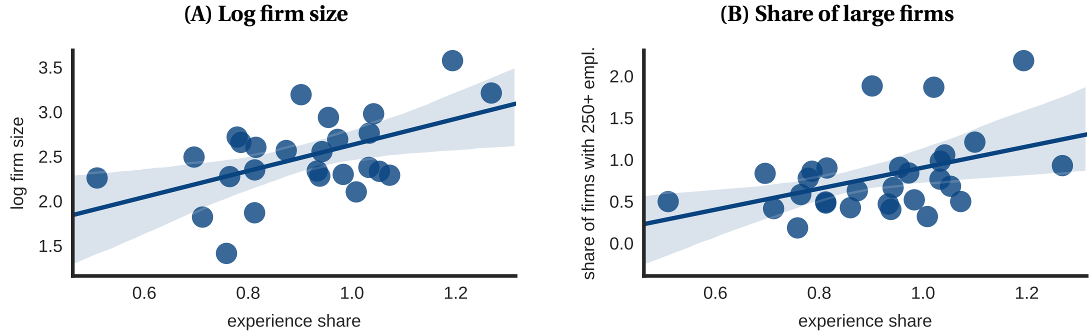
Notes: The table presents the OLS estimates of regressions (1.38). Columns (1) - (5) correspond to five endogenous variables which refer to different measures of business dynamism: mean plant size, the employment share of plants larger than 100 employees, the share of plants older than 11 years, the share of plants of age 0, and the employment share of the top 5 largest plants. The independent variables include experience share (first row) and education share (second row). The shares are instrumented with the variable corresponding to the exposure of the industry to the changes in the aggregate changes in labor supply.

1.A.2. Worker and Firm Demographics: Cross-country Correlations. In this section I present the cross-country correlations between the size distribution of firms in a given country and demographics of the labor force. I use OECD Structural and Demographic Business Statistics database that covers 27 countries.[30] Each dot in Figure 1.9 represents one OECD

[30]These countries are AUT, BEL, CZE, DEU, DNK, ESP, EST, FIN, FRA, GBR, GRC, HUN, IRL, ISL, ISR, ITA, JPN, LTU, LVA, NLD, NOR, POL, PRT, SVK, SVN, SWE, TUR. The dataset covers manufacturing firms only.

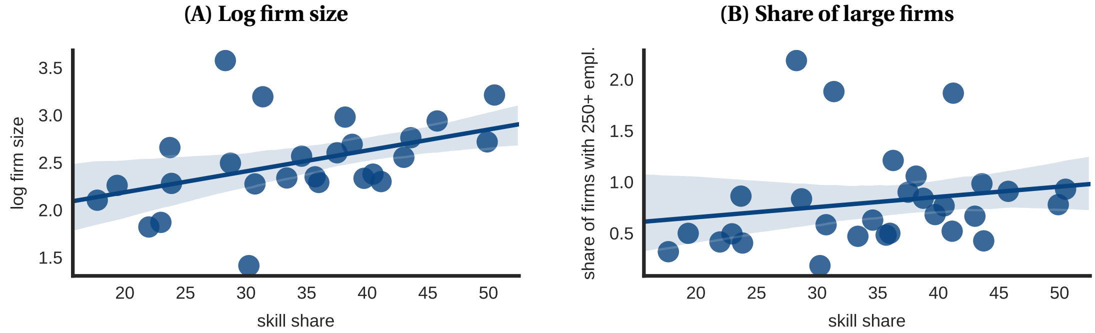
country in year 2016. Panel (A) presents the log of the mean firm size calculated by dividing the total number of employees by the total number of firms. Panel (B) presents the fraction of firms exceeding 250 employees relative to the total number of firms. The blue solid line depicts the regression line and the shaded gray area marks the 95% confidence interval. There is a positive correlation between the share of the experienced workers in the labor force and the average firm size and the share of large firms in the economy. Figure 1.10 shows that there is a similar, albeit weaker, relationship between the share of the college-educated workers and the firm size distribution (unconditionally).

FIGURE 1.9. Experienced workers and business dynamism.



Notes: Data from OECD Structural and Demographic Business Statistics. Each dot represents a country in 2016. The blue solid line denotes the regression line and the shaded gray area marks the 95% confidence interval. The dataset covers manufacturing firms only.

FIGURE 1.10. College-educated workers and business dynamism.



Notes: Data by OECD Structural and Demographic Business Statistics. Each dot represents a country in 2016. The blue solid line denotes the regression line and the shaded gray area marks the 95% confidence interval. The dataset covers manufacturing firms only.

Table 1.8 presents the results of the linear regressions in which log of the average firm size (left column) and share of firms larger than 250 employees (right column) is projected on the size the share of experienced and the share of the college-educated workers.

TABLE 1.8. Labor force demographics and business dynamism in OECD countries

	(1) log firm size	(2) share of large firms
experience share	0.0329***	0.000574***
	(0.00812)	(0.000144)
skill share	0.0128***	-0.000107
	(0.00368)	(0.0000653)
Observations	252	252
Adjusted R^2	0.174	0.053

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The table presents cross-country correlations between the structure of the labor force and firm size and a measure of concentration. The table reports ceofficients in linear regressions in which average firm size and fraction of firms larger than 250 employees is projected on the fraction of workers 45 years or older (experience share) and fraction of workers with college degree or above (skill share). The estimates use pooled data for all countries and all years available in the OECD database.

The results suggest that the relationship between the supply of skills and business dynamism predicted by the model holds across OECD countries. Although the presented evidence is not sufficient for a causal interpretation, it provides a further support for model's key predictions.

1.A.3. Worker and Plant Demographics. In this section I use a more detailed data source to alleviate the concern that the relationship between plant size and employee demographics is spurious. Specifically, I show that the relationship between age, education, and employer size holds also conditional on additional workers' characteristics.

Linked-Employer-Employee-Data of the IAB (LIAB).. In order to investigate in more detail the relationship between firm and worker demographics, I use the LIAB data set which combines establishment information from the Establishment History Panel (BHP) with a detailed employment biographies of individuals from an additional survey. The dataset contains detailed biographical information of all employees for a 2% sample of selected establishments surveyed in the BHP panel.[31] The information about the workers includes age, education, and occupation. Crucially, the data contains a direct measure of experience: a number of days of recorded employment. The data spans between 1994 and 2014.[32]

I use the LIAB data to study the relationship between employee demographics and the employer size in greater detail. I estimate a set of regressions in which the log size of the employer is projected on employee education, age and experience. In the regressions I add the following fixed effects: worker's occupation and birth cohort as well as plant's age, cohort, and industry. Table 1.9 presents the results. The left column shows the coefficients of a

[31] Establishments included in LIAB tend to be much older and larger than the population. The mean size of the establishment in LIAB is 99.05 employees (15.59 in BHP). There are no plants younger than 28 years old.

[32] For more details on the data set and its construction see Klosterhuber et al. (2016).

dummy variable indicating whether the employee is older than 45 years, and a dummy indicating whether she has a college degree. Older individuals tend to work in 7.73% larger production units as compared to the younger workers (the mean and the standard deviation of the plant size is 99.05 and 553 employees, respectively). Having a college degree is associated with working in a 1.73% larger plant. The overall effect is small, but highly significant. Note that the magnitude of coefficient is influenced by inclusion of the control variables, including fixed effects for worker's occupation and birth cohort, as well as plant's industry, cohort, and age.

The right column in Table 1.9 presents the results of a regression in which the experienced measured as the number of years in employment (rather than proxied by age). Again, individuals with more experience tend to work in larger plants.

TABLE 1.9. The employment of skilled labor and business dynamism across German industries.

	(1) log plant size	(2) log plant size
age ≥ 45	0.0773***	
	(0.000825)	
experience		0.0147***
		(0.000003)
experience^2		0.0004***
		(0.000001)
university degree	0.0173***	0.0648***
	(0.0006)	(0.0006)
worker and plant FE	yes	yes
Observations	47928986	47928986
Adjusted R^2	0.316	0.321

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Table presents the results of linear regressions in which employer size is projected on employee characteristics. Left column corresponds to a regression in which is proxied by age, while the right column to a regression in which experience is measured as a number of years a given worker spent in employment. The first column corresponding to a coefficient of a discrete variable attaining one if individual is weakly older than 45 years. The second (third) row corresponds to a continous variable measuring number of years of recorded employment spell (squared). The fourth row corresponding to a discrete variable attaining one if the individual has a university degree. Regressions control for the following fixed effects: worker's occupation and birth cohort as well as plant's age, cohort, and industry.

1.B. Model Details

1.B.1. Labor demand. Consider a problem of obtaining L units of composite labor at the minimal cost possible. Formally, consider a solution to the following problem

$$\min_{e,s} - w_l l - w_e e - w_s s,$$

s. t.

$$L = \left[A_l l^{\frac{\theta-1}{\theta}} + \bar{A}_e A_e(z,k) e^{\frac{\theta-1}{\theta}} + \bar{A}_s A_s(z,k) s^{\frac{\theta-1}{\theta}} \right]^{\frac{\theta}{\theta-1}}$$

Input allocation satisfies

$$\frac{x}{y} = \left(\frac{\bar{A}_x A_x(z,k)}{A_y(z,k)} \right)^{\theta} \left(\frac{w_x}{w_y} \right)^{-\theta}$$

and hence

$$s = a_s(z,k) l, \qquad e = a_e(z,k) l, \tag{1.40}$$

where $a_x(z,k) = \left(\frac{\bar{A}_x A_x(z,k)}{A_l(z,k)} \right)^{\theta} \left(\frac{w_x}{w_l} \right)^{-\theta}$ It follows that

$$l = L \left[A_l + \bar{A}_e A_e(z,k) a_e^{\frac{\theta-1}{\theta}} + \bar{A}_s A_s(z,k) a_s^{\frac{\theta-1}{\theta}} \right]^{-\frac{\theta}{\theta-1}}$$

Normalize $A_l \equiv 1$. Then $a_x(z,k) = \left(\bar{A}_x A_x(z,k) \right)^{\theta} \left(\frac{w_x}{w_l} \right)^{-\theta}$, $\bar{A}_x A_x(z,k) a_x(z,k)^{\frac{\theta-1}{\theta}} = \bar{A}_x A_x(z,k)^{\theta} \left(\frac{w_x}{w_l} \right)^{1-\theta}$ and

$$l = L \left[1 + \bar{A}_e A_e(z,k)^{\theta} \left(\frac{w_e}{w_l} \right)^{1-\theta} + \bar{A}_s A_s(z,k)^{\theta} \left(\frac{w_s}{w_l} \right)^{1-\theta} \right]^{-\frac{\theta}{\theta-1}}$$

The cost of hiring one unit of composite labor is

$$\begin{aligned} W(z,k) &= \left[1 + \bar{A}_e A_e(z,k)^{\theta} \left(\frac{w_e}{w_l} \right)^{1-\theta} + \bar{A}_s A_s(z,k)^{\theta} \left(\frac{w_s}{w_l} \right)^{1-\theta} \right]^{-\frac{\theta}{\theta-1}} \left(w_l + a_e(z,k) w_e + a_s(z,k) w_s \right) \\ &= w_l \left[1 + \bar{A}_e A_e(z,k)^{\theta} \omega_e^{1-\theta} + \bar{A}_s A_s(z,k)^{\theta} \omega_s^{1-\theta} \right]^{-\frac{1}{\theta-1}} \end{aligned}$$

1.B.2. Solving for productivity schedules. Since each worker supplies one efficiency unit of labor, the total size is

$$n = l \left(1 + a_e(z,k) + a_s(z,k) \right),$$

where $a_x(z,k) = \left(\frac{\bar{A}_x \bar{A}_x A_x(z,k)}{A_l(z,k)} \right)^{\theta} \left(\frac{w_x}{w_l} \right)^{-\theta}$. Moreover, (1.40) implies that

$$\frac{s}{n} = \left[1 + \left(\frac{\bar{A}_e A_e(z,k)}{\bar{A}_s A_s(z,k)} \frac{w_s}{w_e} \right)^{\theta} + \bar{A}_s A_s(z,k)^{-\theta} \omega_s^{\theta} \right]^{-1} \tag{1.41}$$

and

$$\frac{e}{n} = \left[1 + \left(\frac{\bar{A}_e A_e(z,k)}{\bar{A}_s A_s(z,k)} \frac{w_s}{w_e} \right)^{-\theta} + \bar{A}_e A_e(z,k)^{-\theta} \omega_e^{\theta} \right]^{-1} \tag{1.42}$$

Let $\hat{e}(n(z,k))$, $\hat{s}(n(z,k))$ mark the factor intensity in the data, where I explicitly indicated that the experience and skill shares in the data are function of plant size. We can re-write (1.41) and (1.42) as follows

$$\hat{s}(n(z,k))^{-1} - 1 = XY^{-1}\left(\frac{\omega_s}{\omega_e}\right)^{\theta} + Y^{-1}\omega_s^{\theta}$$
$$\hat{e}(n(z,k))^{-1} - 1 = X^{-1}Y\left(\frac{\omega_s}{\omega_e}\right)^{-\theta} + X^{-1}\omega_e^{\theta}$$

where $X = \hat{A}_e(z,k)^{\theta}$ and $Y = \hat{A}_s(z,k)^{\theta}$. The above system of equations yields the estimates of productivity schedules $\hat{A}_e$, $\hat{A}_s$. The solution is

$$X = \hat{A}_e(z,k)^{\theta} = \frac{\hat{s}(n(z,k))^{-1}\omega_e^{\theta}}{\hat{e}(n(z,k))^{-1}\hat{s}(n(z,k))^{-1} - \hat{e}(n(z,k))^{-1} - \hat{s}(n(z,k))^{-1}}$$
$$Y = \hat{A}_s(z,k)^{\theta} = \frac{\hat{e}(n(z,k))^{-1}\omega_s^{\theta}}{\hat{e}(n(z,k))^{-1}\hat{s}(n(z,k))^{-1} - \hat{e}(n(z,k))^{-1} - \hat{s}(z,k)^{-1}}$$

CHAPTER 2

Federal Unemployment Reinsurance and Domestic Labor-Market Policies

```
Economic and monetary union ...  means that the principal decisions of economic
policy will be taken at the Community level and therefore that the necessary powers
will be transferred from the national plane to the Community.
[Conclusions, 1970 Werner report on Economic and Monetary Integration]
```

2.1. Introduction

Having entered a monetary union, member states no longer have access to monetary stabilization policy. An established view is that, therefore, monetary union requires some form of federal fiscal capacity. This would provide countercyclical transfers so as to help stabilize member states' economies in the face of idiosyncratic shocks or to compensate for the welfare loss of fluctuations; a classic reference being Kenen 1969, more recently Farhi and Werning 2017. At the same time, if member states have authority over domestic policies, their behavioral responses may limit the very stabilization that can be provided; Persson and Tabellini 1996.

This paper analyzes the gains from a federal scheme that provides transfers to each member state based on local unemployment – "federal unemployment reinsurance (RI)," for short. How should such a scheme be designed when it is layered on top of local labor market policies (including local unemployment insurance)? How severe are the limitations induced by incentives to free ride? How much of a welfare gain can such a scheme still provide? We provide theory and a quantitative analysis for a stylized European monetary union.[1]

We find that concerns about persistent free-riding by member states can effectively be addressed by indexing RI to a member state's past unemployment. Focusing on the long run only, optimal federal RI is generous, does not have thresholds, and helps stabilize employment substantially. All of this amid sizable welfare gains. Short-run considerations, instead, are a potential limit to federal RI, even if transfers are indexed to past unemployment. We find that accounting for the transition path after the introduction of federal RI may limit the optimal generosity of transfers, as may an interaction of federal RI with a member state's stabilization policies.

[1]The European monetary union has little private and fiscal risk sharing, Furceri and Zdzienicka 2015. We consider federal RI schemes that, once in place, work without discretion, Juncker 2015.

The paper models a union of *ex-ante* identical and atomistic member states. Each member state is subject to idiosyncratic and possibly marked business-cycle risk.[2] There is no private risk sharing between member-state households, nor can member-state governments borrow from each other. The lack of liquidity and international insurance provides the rationale for a federal fiscal capacity. The capacity is introduced under full commitment. Labor markets in each member state are characterized by search and matching frictions. Federal transfers can help address excessive employment fluctuations by providing the fiscal space to keep taxes on firms low in recessions. The union as a whole is modeled as a small open economy that has access to international borrowing and lending.

Member states set local labor-market policies; namely, layoff taxes (think more widely, layoff restrictions) and hiring subsidies (think more widely, active labor-market policies), and local unemployment benefits. Taxes on production balance the member state's budget. We show analytically that the introduction of federal RI affects the local labor-market policy mix in a direction that implies less employment in the long run. At the same time, a lack of international borrowing and lending means that – absent a federal RI scheme – member states will tend to self-insure through *too much* employment, an implication reminiscent of Aiyagari 1994. So some fall in long-run employment induced by federal RI need not be detrimental. A quantitative assessment is needed.

Therefore, next we calibrate the model to a stylized European Monetary Union (with fluctuations of the same size as witnessed in the crisis years). In the baseline, member states have authority over labor-market policies, but do not have a system of automatic countercyclical policies in place. Payouts from the federal RI scheme are financed through a flat contribution. Payouts can condition on current and historical unemployment rates according to a flexible parametrized functional form. This, for example, allows for federal payouts to increase linearly in unemployment (such as implicit in member states' unemployment insurance schemes today) or for payouts to a member state only when a certain threshold of unemployment is exceeded (the U.S. federal unemployment insurance system).[3] In this class, we look for "optimal" schemes that maximize *ex-ante* welfare of the constituents of the union under the condition that the federal RI budget be balanced in net present value.

If we abstract from member state's incentive effects, the optimal federal RI scheme is linear in unemployment and federal transfers almost entirely make up for the output lost in a recession. The welfare gains would amount to about 0.4 percent of life-time consumption. Next to this, the scheme would notably smooth employment fluctuations. The reason is simple: federal RI provides insurance to member states, and it allows for fiscal space in recessions (liquidity). This means that taxes on firms (the balancing item in the member state fiscal balance) can be kept low in recessions, which helps smooth employment.

Against this background, we go through a sequence of different setups that illustrate how member states' incentives to adjust local policies affect the optimal shape and scope of federal RI. Toward this end, we first allow member states to adjust local labor-market policies to

[2] We look at productivity shocks that are uncorrelated across member states. The idea is that union-wide monetary policy would address the common component. With this, we abstract from modeling nominal rigidities. At the level of the member state, the modeling follows Jung and Kuester 2015.

[3] These are the main characteristics of various schemes proposed in practice, such as Beblavý et al. 2015, or Bénassy-Quéré et al. 2018. We do not consider federal RI as a mere credit facility. And we do not consider ways of fostering private risk sharing.

a new level once the federal RI scheme is introduced. That level, then, remains in place permanently. We look at two different designs of federal RI. We discuss long-run considerations first, and then turn to the transition phase. Our findings are as follows.

First, member states' authority over local labor-market policies is crucial for the design of federal RI. If payouts are linked to current unemployment only, the optimal RI resembles a trigger system where payouts are made only in severe recessions. What is more, even if member states cannot change the generosity of their local unemployment insurance benefits (say, because of union-wide unemployment-insurance standards), as long as they control other labor-market policies (hiring subsidies and layoff taxes), quantitatively the welfare gains from federal RI are virtually zero.

Second, a small element of accountability brings back the power of federal RI in the long run. Namely, if federal payouts are indexed to member state's historical average unemployment, we find that optimal federal RI would be linear in deviations of unemployment from past unemployment; and there is no need for thresholds. Transfers would be large enough to – on average – replace income losses in a recession. The gains from introducing federal RI amount to about 0.2 percent of life-time consumption. Gains are smaller than absent free-riding by member states since indexation not only means accountability but also that federal payments are not always well-timed.

Next, and third, we account for the transition phase. The transition phase bears notably on the shape of federal RI. The optimal scheme continues to feature sizable transfers from member states in booms. Remarkably, however, transfers to member states in recession hardly rise with unemployment. Key to this are the incentives on the transition path where indexation to past unemployment does not eliminate the incentives to free ride. So even when indexing payouts to past unemployment, the optimal federal RI scheme achieves at best a welfare gain of 0.06 percent of life-time consumption, or about a third of the gains that would prevail under a long-run-only view.

Last, we evaluate how federal RI interacts with member states' incentives to engage in countercyclical labor-market stabilization policy. Having member states choose their labor-market policies in a (self-interested) Ramsey-optimal fashion,[4] the welfare gains of such stabilization policy are of the same or larger magnitude as the gains from from federal RI allowing from member-state's behavioral responses.[5] Introducing a federal RI scheme over and above the member states' use of such counter-cyclical stabilization policy, the additional gains are minute. *Vice versa*, a federal RI scheme designed under the assumption that member states do not engage in stabilization policy reduces welfare if they do. There are two reasons. Countercyclical labor-market policies, in our setting, notably reduce the size and cost of fluctuations. Once countercyclical labor-market policies are in place, then, there is less to gain from federal RI to start with. Instead, federal RI crowds out member states' stabilization and facilitates free-riding on the transition path.

In sum, we analyze simple optimized federal RI schemes. The schemes condition payouts on a few indicators only, and do not directly condition on policy choices past and present or on shocks. Our aim is to spell out important trade-offs in the design of implementable federal

[4]That is, member states choose optimal state-contingent Pigouvian taxes and labor-market instruments

[5]The model economy features wage rigidity, following Shimer 2004 This renders employment excessively responsive to cyclical shocks, giving rise to a welfare-improving role of domestic labor-market stabilization policies; Jung and Kuester 2015.

RI schemes. For this, we go through different perspectives on what the horizon is over which policymakers optimize and what the policy options are for member states. We have little insight into which of these scenarios is the most applicable in practice. Regardless of the reader's priors, we hope to provide valuable information, though.

Next, we put the paper into the context of the literature. Section 2.2 spells out the model and the member states' and federal government's problems. Section 2.3 provides analytical insights. Section 2.4 provides the calibration and discusses our numerical implementation. Section 2.5 derives the quantitative results discussed above. A final section concludes.

Related literature

To the best of our knowledge, ours is the first paper that studies the design of implementable federal unemployment reinsurance schemes in a quantitative business-cycle model, allowing for optimal labor-market policies by member states. The following aims to make clear what we do (and what we do not do).

We introduce a federal unemployment-based transfer scheme in a union of member states that have authority over local labor-market policies. The need for local policies arises from local frictions: idiosyncratic unemployment risk combined with moral hazard by households gives rise to imperfect unemployment insurance and wage rigidities amplify cyclical responses. Considerations for labor-market policies for single countries are discussed, for example, in Landais et al. 2018, Mitman and Rabinovich 2015, and Jung and Kuester 2015. Birinci and See 2019 focus on economies with self-insurance by households, from which we abstract. Jung and Kuester 2015 have assessed the optimal labor-market policy mix over the business cycle (and in steady state) for a single country. The current paper adds to this a consideration of federal RI amid behavioral responses by member states. By way of reference, empirical considerations for cyclical labor market policies are discussed in Chodorow-Reich et al. 2018 and Hagedorn et al. 2013 (for benefits) and Cahuc et al. 2018 (for hiring subsidies). The latter find that, if well-targeted, such subsidies are an effective countercyclical stabilization policy amid rigid wages.

The current paper's insights link to an extensive literature on fiscal federalism, see Alesina and Wacziarg 1999 or Oates 1999 for references. A central reference for us are Persson and Tabellini 1996 who ask if fiscal risk sharing can induce local governments to under–invest in programs that alleviate local risk. They have a static setting with two countries and two states per country. One way to think about our paper is as an extension of the literature to a dynamic environment. Dynamics allow to calibrate the potential gains from risk-sharing, and to separate considerations for the short and the long run.

The current paper abstracts from nominal rigidities. In Farhi and Werning (2017), instead, for given policies at the member-state level, these rigidities give rise to aggregate-demand externalities and scope for federal fiscal transfers. The current paper, instead, focuses on how the member states' behavioral response shapes the scope for federal transfers. For this, we takes labor-market rigidities as the central element that amplifies the welfare costs of business cycles. Needless to say that also with nominal rigidities, one could ask which fiscal mix helps overcome these, for example, Farhi et al. 2013.

In terms of solution techniques, we employ fourth-perturbation so as to account for the possible non-linearity of the federal RI scheme. We extend the moment formulae in Andreasen et al. 2018 to that order and then follow Levintal 2017. While we have not done so, the solution techniques should allow extensions to more complex environments. Nominal rigidities are one case in point. Another would be accommodating precautionary motives of households as in Ravn and Sterk, forthcoming.

We look at deliberately simple federal transfer mechanisms triggered by unemployment. Our motivation is that more information (such as policy choices) may either not be observable, or not contractible. The setting allows us to handle cyclical economies with endogenous unemployment fluctuations. Another advantage is that we can allow for a wide-range of policy choices for the member states. Necessarily, there are downsides as well. Namely, in contrast to the endogenous incomplete-markets literature, such as Atkeson 1991 and Phelan and Townsend 1991, we take a strong form of international market incompleteness as given. And we assume commitment to the federal RI scheme. Interestingly, even under commitment, once accounting for the transition phase, the optimal federal RI scheme inherits a property from the endogenous market-incompleteness literature (for example, Krueger and Uhlig, 2006): pre-financing the scheme during the transition phase.

On the empirical side, Furceri and Zdzienicka 2015 find that the European Monetary Union to date has limited mechanisms for risk sharing. We analyze if and when federal unemployment reinsurance is a potential means of overcoming this. Interestingly, the amounts of risk-sharing that emerge from our theory-based exercise are roughly in line with estimates of fiscal risk sharing for the US. Feyrer and Sacerdote 2013, and earlier Asdrubali et al. 1996, estimate that between 13 and 25 cents of every dollar of a state-level income shock is offset by federal fiscal policy.

Our work considers transitory *ex-post* heterogeneity, but abstracts from *ex-ante* heterogeneity. So do Cooper and Kempf 2004. They show in a two-period setting that a scheme that involves federal unemployment-based transfers can help overturn the Mundellian (1961) argument that monetary unions are a straight-jacket. They take unemployment in each member state as exogenous to member-state's policy choices. Ábrahám et al. 2017, instead, put great effort into measuring labor-market flow rates across the different euro-area members. They, then, build a heterogeneous-agent macro model to study the optimal level of federal unemployment benefits, taking the heterogeneity across member states as given. The authors find that a harmonized unemployment insurance scheme is Pareto-improving relative to the *status quo*. In our setting, member states choose an optimal local unemployment insurance system already in autarky. We, then, ask what the benefits of a federal RI scheme are on top of this. Moyen et al. 2016 consider a heterogeneous monetary union of two countries. A federal planner can mandate the cyclical level of benefits in each member state and the member state cannot adjust other labor-market policies. They find considerable scope for such an unemployment insurance scheme in smoothing the business cycle. The same is true for Enders and Vespermann 2019 who analyze the stabilizing effect of federal unemployment insurance in a New Keynesian two-country DSGE model, and for Dolls et al. (2018) who do micro simulations for the previous euro-area recession. Also in a two-country DSGE model, Evers 2015 assesses the risk-sharing and welfare properties of several exogenously-given fiscal risk-sharing schemes. House et al. 2018 study the role of labor migration as a shock absorber in a monetary union, for constant labor-market instruments. Relative to all these paper, we focus on the optimal scope of federal unemployment-based transfers when member state retain and use authority over local policies.

2.2. The model

There is a federal union that consists of a unit mass of atomistic, *ex-ante* identical member states that are subject to member-state specific shocks. Member states are marked by subscript i. Member states control their own UI benefit system, layoff taxes, hiring subsidies, and production taxes leveled on domestic firms, adhering to a balanced-budget requirement. There is no international borrowing and lending, nor is there self-insurance by households. A federal UI scheme makes unemployment-dependent transfers to the member-state government.[6] Time is discrete and runs from $t = 0$ to infinity.

2.2.1. The member-state economy. There are three types of agents in each member state: a unit mass of infinitely-lived workers, an infinite mass of potential one-worker firms that produce a homogenous final good, and the member-state government.

2.2.1.1. *Workers.* A worker's lifetime utility is given by

$$E_0 \left\{ \sum_{j=0}^{\infty} \beta^t [\mathrm{u}(c_t) + \overline{h} \cdot I(\text{not working}_t) - \iota \cdot I(\text{search}_t)] \right\}. \tag{2.1}$$

E_0 denotes the expectation operator. $\beta \in (0,1)$ is the time-discount factor. The worker draws utility from consumption, c_t. Felicity function $\mathrm{u}(c) : \mathscr{R} \to \mathscr{R}$ is twice continuously differentiable, strictly increasing and concave. I is the indicator function. If not employed, the worker enjoys an additive utility of leisure $\overline{h}$. Search for a job is a 0-1 decision. Workers are differentiated by a utility cost of search, ι, incurred only if the worker searchers for a new job. Both across workers and time $\iota \sim F_\iota(0, \sigma_\iota^2)$ is iid, where $F_\iota(\cdot,\cdot)$ marks the logistic distribution with mean 0 and variance $\sigma_\iota^2 = \pi \frac{\psi_\iota^2}{3}$, with $\psi_\iota > 0$ and π being the mathematical constant.

Workers own all firms in their member state in equal proportion. Ownership of firms is not traded. Workers cannot self-insure against income fluctuations through saving or borrowing. Letting Π_t^i mark the dividends that the member state's firms pay and w_t^i the wage that an employed worker earns, consumption of the worker is given by

$$\begin{aligned} c_{u,t}^i &:= b_t^i + \Pi_t^i \quad \text{if unemployed at the beginning of } t, \\ c_{e,t}^i &:= w_t^i + \Pi_t^i \quad \text{if employed at the beginning of } t. \end{aligned} \tag{2.2}$$

If the worker enters the period unemployed, the worker receives an amount b_t^i of unemployment benefits. The assumption is that the member-state's government cannot observe the search effort of workers. The government, by assumption, conditions payment only on the worker's current employment status. A worker who enters the period employed receives wage income or, if separated, a severance payment equal to the period's wage.

Value of an employed worker

Let ξ_t^i be the separation rate of existing matches. Before separations occur, the value of an employed worker is

$$V_{e,t}^i = \mathrm{u}(c_{e,t}^i) + [1 - \xi_t^i] \beta E_t V_{e,t+1}^i + \xi_t^i [V_{u,t}^i - \mathrm{u}(c_{u,t}^i)]. \tag{2.3}$$

[6]At the level of the member state, the model and exposition build extensively on Jung and Kuester 2015.

The worker consumes $c^i_{e,t}$. With probability $1-\xi^i_t$ the match does not separate and continues into $t+1$. With probability ξ^i_t, instead, the match separates. The worker can immediately start searching for new employment. The only difference of its value to the value of a worker, who was unemployed to start with, therefore, is that the separated worker receives the severance payment while the unemployed worker receives unemployment benefits (and, thus, lower consumption). $V^i_{u,t}$ is the value of a worker who starts the period unemployed.

Value of an unemployed worker and search
An unemployed worker needs to actively search in order to find a job. All workers whose disutility of search falls below a cutoff value $\iota^{s,i}_t$ do so.[7] At the cutoff, the utility costs of search just balances with the expected gain from search:

$$\iota^{s,i}_t = f^i_t \beta E_t \left[\Delta^i_{t+1}\right]. \tag{2.4}$$

Here $\Delta^i_t = V^i_{e,t} - V^i_{u,t}$ marks the gain from employment and f^i_t marks the job-finding rate. For tractability, the paper assumes that $F_\iota(0,\sigma^2_\iota)$ is the logistic distribution with mean 0 and variance $\sigma^2_\iota := \pi\frac{\psi^2_s}{3}$, where a lower-case π refers to the mathematical constant and parameter $\psi_s > 0$. Using the properties of the logistic distribution, the share of unemployed workers who search is given by

$$s^i_t = Prob(\iota \le \iota^{s,i}_t) = 1/[1+\exp\{-\iota^{s,i}_t/\psi_s\}]. \tag{2.5}$$

With this, the value of an unemployed worker at the beginning of the period, before the search preference shock has realized, is given by

$$\begin{aligned} V^i_{u,t} \quad = \quad & \mathrm{u}(c^i_{u,t}) + \overline{h} \\ & + \int_{-\infty}^{\iota^{s,i}_t} dF_\iota(\iota) + s^i_t \left[f^i_t \beta E_t V^i_{e,t+1} + [1-f^i_t]\beta E_t V^i_{u,t+1}\right] \\ & + (1-s^i_t)\beta E_t V^i_{u,t+1}. \end{aligned} \tag{2.6}$$

In the current period, the worker consumes $c^i_{u,t}$ and enjoys utility of leisure $\overline{h}$ (first row). If the worker decides to search (second row), the utility cost is ι. The term with the integral is the expected utility cost of search. With probability f^i_t the searching worker will find a job. In that case, the worker's value at the beginning of the next period will be $V^i_{e,t+1}$. With probability $(1-f^i_t)$ the worker remains unemployed in the next period. If the worker does not search (third row), the worker remains unemployed.

2.2.1.2. *Firms.* Profits in the firm sector accrue to the workers, all of whom hold an equal amount of shares in the domestic firms. The decisions made by firms are dynamic and involve discounting future profits. We assume that firms discount the future using discount factor $Q^i_{t,t+s}$, where $Q^i_{t,t+s} := \beta\frac{\lambda^i_{t+s}}{\lambda^i_t}$, with λ^i_t being the weighted marginal utility of the workers (the firms' owners):

$$\lambda^i_t := \left[\frac{e^i_t}{\mathrm{u}'(c^i_{e,t})} + \frac{u^i_t}{\mathrm{u}'(c^i_{u,t})}\right]^{-1}. \tag{2.7}$$

[7]The s in $\iota^{s,i}_t$ stands for the *s*earch cutoff.

This reflects that mass e_t^i of workers are employed at the beginning of the period and mass $u_t^i := 1 - e_t^i$ are unemployed.

Firms need a worker to produce output. A firm that enters the period matched to a worker can either produce or separate from the worker. Production entails a firm-specific resource cost, ϵ_j. This fixed cost is independently and identically distributed across firms and time with distribution function $F_\epsilon(\mu_\epsilon, \sigma_\epsilon^2)$. $F_\epsilon(\cdot,\cdot)$ is the logistic distribution with mean μ_ϵ and variance $\sigma_\epsilon^2 = \pi \frac{\psi_\epsilon^2}{3}$, with $\psi_\epsilon > 0$. The firm separates from the worker (first line) whenever the idiosyncratic cost shock, ϵ_j, is larger than a state-dependent threshold $\epsilon_t^{\xi,i}$. Using the properties of the logistic distribution, conditional on the threshold, the separation rate can be expressed as

$$\xi_t^i = Prob(\epsilon_j \geq \epsilon_t^{\xi,i}) = 1/[1 + \exp\{(\epsilon_t^{\xi,i} - \mu_\epsilon)/\psi_\epsilon\}]. \tag{2.8}$$

Ex ante, namely, before the idiosyncratic cost shock ϵ_j is realized, the value of a firm that has a worker is given by

$$\begin{aligned} J_t^i \quad = \quad & -\xi_t^i \left[\tau_{\xi,t}^i + w_t^i\right] \\ & - \int_{-\infty}^{\epsilon_t^{\xi,i}} \epsilon_j \, dF_\epsilon(\epsilon_j) + (1 - \xi_t^i) \left[\exp\{a_t^i\} - w_t^i - \tau_{J,t}^i + E_t Q_{t,t+1}^i J_{t+1}^i\right]. \end{aligned} \tag{2.9}$$

Upon separation, the firm is mandated to pay layoff tax $\tau_{\xi,t}^i$ to the government and a severance payment of a period's wage w_t^i to the worker. Instead, if ϵ_j does not exceed the threshold, the firm will not separate, the firm will pay the resource cost, and the match will produce (second line). a_t^i is a member-state specific labor-productivity shock. This is the source of *ex-post* heterogeneity of member states. The firm produces $\exp\{a_t^i\}$ units of the good and pays the wage w_t^i to the worker. In addition, the firm pays a production tax $\tau_{J,t}^i$. A match that produces this period continues into the next.

The labor-productivity shock, a_t^i, evolves according to

$$a_t^i = \rho_a \, a_{t-1}^i + \varepsilon_{a,t}^i, \quad \rho_a \in [0,1), \; \varepsilon_{a,t}^i \sim N(0, \sigma_a^2).$$

An employment-services firm that does not have a worker can post a vacancy. If the firm finds a worker, the worker can start producing from the next period onward. Accounting for subsidies by the member state, the cost to the firm of posting a vacancy is $\kappa_v(1 - \tau_{v,t}^i)$. $\kappa_v > 0$ marks a resource cost, and $\tau_{v,t}^i$ the government's subsidy for hiring. In equilibrium, employment-services firms post vacancies until the after-tax cost of posting a vacancy equals the prospective gains from hiring:

$$\kappa_v(1 - \tau_{v,t}^i) = q_t^i E_t \left[Q_{t,t+1}^i J_{t+1}^i\right], \tag{2.10}$$

where q_t^i is the probability of filling a vacancy.

Let v_t^i be the mass of vacancies posted. Matches m_t^i evolve according a constant-returns matching function:

$$m_t^i = \chi \cdot \left[v_t^i\right]^\gamma \cdot \left[[\xi_t^i e_t^i + 1 - e_t^i] s_t^i\right]^{1-\gamma}, \; \gamma \in (0,1). \tag{2.11}$$

Here, $\chi > 0$ is match efficiency. The mass of workers who potentially search is $\xi_t^i e_t^i + 1 - e_t^i$, with $\xi_t^i e_t^i$ being workers separated at the beginning of the period. s_t^i is the share of those who do actually search. With this, employment evolves according to

$$e^i_{t+1} = [1 - \xi^i_t] \cdot e^i_t + m^i_t. \tag{2.12}$$

Total production of output is given by

$$y^i_t = e^i_t (1 - \xi^i_t) \exp\{a^i_t\}, \tag{2.13}$$

where $e^i_t(1 - \xi^i_t)$ is the mass of existing matches that are not separated in t.

For subsequent use, define labor-market tightness as $\theta^i_t := v^i_t / ([\xi^i_t e^i_t + 1 - e^i_t] s^i_t)$, the job-finding rate as $f^i_t := m^i_t / ([\xi^i_t e^i_t + 1 - e^i_t] s^i_t) = \chi [\theta^i]^\gamma$, and the job-filling rate as $q^i_t := m^i_t / v^i_t = \chi [\theta^i]^{\gamma - 1} = f^i_t / \theta^i_t$. Also, define unemployment as $u^i_t := 1 - e^i_t$.

Dividends

Dividends in each member state arise from the profits generated by the member state's firms, namely,

$$\begin{aligned} \Pi^i_t \quad = \quad & -e^i_t \left[\int_{-\infty}^{\epsilon^{\xi,i}_t} \epsilon dF_\epsilon(\epsilon)\right] + e^i_t (1 - \xi^i_t) \left[\exp\{a^i_t\} - w^i_t - \tau^i_{J,t}\right] - e^i_t \xi^i_t \left[w^i_t + \tau^i_{\xi,t}\right] \\ & - [\kappa_v - \tau^i_{v,t}] v^i_t. \end{aligned} \tag{2.14}$$

2.2.1.3. *Bargaining between firm and worker.* At the beginning of the period, matched workers and firms observe the aggregate shock, a^i_t. Conditional on this, and *prior* to observing match-specific cost shock ϵ_j, firms and workers bargain over the wage and the severance payment as well as over a state-contingent plan for separation. Anticipating that the firm will insure the risk-averse worker against the idiosyncratic risk associated with ϵ_j so that the wage, w_t, is independent of the realization of ϵ_j and that the severance payment equals the wage, firm and worker solve

$$(w^i_t, \epsilon^\xi_t) = \arg\max_{w^i_t, \epsilon^{\xi,i}_t} (\Delta^i_t)^{1-\eta_t} (J^i_t)^{\eta_t}, \tag{2.15}$$

where η_t measures the bargaining power of the firm. We shall assume that η_t is linked to productivity as $\eta^i_t = \eta \cdot \exp\{\gamma_w \cdot a^i_t\}$, with $\gamma \geq 0$. If $\gamma_w > 0$, the bargaining power of firms is low in recessions and high in booms.

The first-order condition for the wage is as follows

$$(1 - \eta^i_t) J^i_t = \eta^i_t \frac{\Delta^i_t}{\mathrm{u}'(c^i_{e,t})}. \tag{2.16}$$

It states that after adjusting for the bargaining weights, the value of the firm equals the surplus of the worker from working expressed in units of consumption when employed.

The first-order condition for the separation cutoff yields

$$\epsilon^{\xi,i}_t \quad = \quad \left[\exp\{a^i_t\} - \tau^i_{J,t} + \tau^i_{\xi,t} + E_t Q^i_{t,t+1} J^i_{t+1}\right] + \frac{\beta E_t \Delta^i_{u,t+1} + \psi_s \log(1 - s^i_t) - \overline{h}}{\mathrm{u}'(c^i_{e,t})}. \tag{2.17}$$

2.2.1.4. *Federal RI scheme and market clearing.* As for the federal unemployment reinsurance scheme, let $\mathbf{B}_F \left(u^i_t; u^{i,avg}_t\right)$ mark transfers of final goods from the federal level to the governments of member states. These transfers condition on the member state's current unemployment, u^i_t. They may also condition on a moving index of past unemployment,

$u_t^{i,avg} := \delta u_{t-1}^{i,avg} + (1-\delta) u_{t-1}^i$, with $\delta \in (0,1)$. All member states are subject to the same structure of the federal RI scheme. Let τ_F mark a flat, time-independent contribution toward the federal RI scheme, paid by each member state.

Anticipating that member-state governments do not have access to international borrowing or lending (see Section 2.2.2.2), goods market clearing in each member state requires that in each of them

$$y_t^i + \mathbf{B}_F\left(u_t^i; u_t^{i,avg}\right) - \tau_F = e_t^i c_{e,t}^i + u_t^i c_{u,t}^i + e_t^i \int_{-\infty}^{\epsilon_t^{\xi,i}} \epsilon \, dF_\epsilon(\epsilon) + \kappa_v v_t^i. \tag{2.18}$$

The left-hand side has goods produced in the member state plus the net transfers received under the federal RI scheme. In equilibrium, goods are used for either consumption (the first two terms on the right-hand side), for production costs, or for vacancy-posting costs. Once markets clear in all the member states, they also clear for the union as a whole.

2.2.2. Government sector. There are two levels of government, the federal level and the local (member-state) level. At the beginning of period $t = 0$, before idiosyncratic shocks to member states have materialized, the federal government can set up a federal unemployment reinsurance ("RI") scheme, knowing the initial distribution of member states in the state space, and anticipating the response by member states and households. Period $t = 0$ is the first period in which the scheme will make payouts and collect contributions. The federal government is a first mover. Member-state governments, when choosing labor-market policies, take the federal RI scheme as given. The federal RI scheme is implemented in a permanent manner and under full commitment. This choice of timing seems a reasonable first pass for many countries; even more so for the European Monetary Union/European Union, we believe, where changes to binding agreements often require unanimity. We describe each government level in turn.

2.2.2.1. *The federal government's problem.* Let μ_t mark the distribution of member states across the possible states of the economy, in period t. Let $\tilde{\mu}_t$ be the induced distribution over the payout-relevant characteristics $(u_t^i, u_t^{i,avg})$. Note that, once conditioning on the initial distribution of member states at the beginning of time, by the law of large numbers, both μ_t and $\tilde{\mu}_t$, $t = 0, 1, \ldots$, are measurable at the beginning of period $t = 0$.[8]

The federal government has access to international borrowing and lending at a fixed gross interest rate $1 + r = 1/\beta$. The federal RI scheme has to be self-financing in the sense that payouts or any debt be financed completely by the federal RI taxes. Assuming that there is no initial debt, this is the case if

$$\sum_{t=0}^{\infty} (1+r)^{-t} \int \left(\mathbf{B}_F(u_t; u_t^{avg}) - \tau_F\right) d\tilde{\mu}_t = 0. \tag{2.19}$$

Here, the integral is over the distribution of $(u_t^i, u_t^{i,avg})$ in all member states in the respective period t.

Weighting all households in a member state equally, using (2.1) and the logistic distribution, after shocks have realized in period t a member state's utilitarian welfare function can

[8]The position of each member state i in the distribution is random. Since all risk is idiosyncratic, though, and member states are given equal weight in the federal planner's welfare function, it does not matter which member state is in what position of the distribution.

be written as (see Jung and Kuester 2015 for a derivation)

$$W_t^i := E_t \sum_{k=t}^{\infty} \beta^k \left[e_k^i \mathrm{u}(c_{e,k}^i) + u_k^i \mathrm{u}(c_{u,k}^i) + \ (e_k^i \xi_k^i + u_k^i)(\Psi_s(s_k^i) + \overline{h}) \right]. \tag{2.20}$$

The first term is the consumption-related utility of employed workers. The second term is the consumption-related utility of unemployed workers. The third term refers to the value of leisure and the utility costs of search.[9]

The federal government's problem is to

(2.21)

$$\max_{\mathbf{B}_F(\cdot;\cdot),\tau_F} \int W_0 \, d\mu_0$$

s.t. member states' policy response (see Section 2.2.2.2)
induced law of motion of member states' economies (earlier sections)
financing constraint (2.19),

where the maximization over $\mathbf{B}_F(\cdot;\cdot)$ indicates that the federal government chooses the shape of the payout function of the federal RI scheme. Throughout, we assume that $\mathbf{B}_F(\cdot;\cdot)$ will be continuously differentiable. In choosing the payout function, the federal planner anticipates the response to the scheme of both the member member states' governments and of the constituents of each member state. Last, the federal RI scheme is restricted to break even.

2.2.2.2. *The member-state government's problem.* The member-state government does not have access to international financial markets, nor does it issue debt to local residents. The member-state government faces the budget constraint

$$e_t^i(1-\xi_t^i)\tau_{J,t}^i + e_t^i \xi_t^i \tau_{\xi,t}^i + \mathbf{B}_F\left(u_t^i; u_t^{i,avg}\right) = u_t^i b_t^i + \kappa_v \tau_{v,t}^i v_t^i + \tau_F. \tag{2.22}$$

The left-hand side shows the revenue from the production and layoff taxes, and the transfers received under the federal RI scheme. The right-hand side has unemployment benefits and vacancy subsidies paid by the member state, as well as the federal RI contribution.

We model the member-state government as a utilitarian Ramsey planner, that acts in the interest of its own constituency. It chooses unemployment benefits, Pigouvian layoff taxes, and hiring subsidies (or a subset of the three thereof), next to taxes on production. The exposition below assumes that the member state chooses all labor-market policy instruments. The member-state's problem is analogous when the member state only has access to some of the instruments, a case that we explore in the numerical analysis in Section 2.5.

The member state sets policies in period 0, before shocks have realized, but after the federal level has announced the shape of its federal RI scheme. We look at two different scenarios. Either, the member state chooses fixed levels of policies once and for all, or the

[9] Here $\Psi_s(s_k^i) := -\psi_s \left[(1-s_k^i)\log(1-s_k^i) + s_k^i \log(s_k^i) \right]$. $\Psi_\xi(\xi_k^i)$, which is used further below, is defined in an analogous manner.

member state chooses state-contingent policies. In the former case, the member-state government's problem is to choose state-and-time-independent labor-market policies, with production taxes balancing the budget

(2.23)
$$\max_{\{\tau_v^i, \tau_\xi^i, b^i, \tau_{J,t}^i\}} \int W_0 \, d\mu_0$$

s.t. a given federal RI scheme $\mathbf{B}_F, \tau_F$
the induced law of motion of the member state's economy (earlier sections)
the member-state government's budget constraint (2.22),

where W_0 is given by (2.20). Being atomistic, the member-state planner takes the federal RI scheme as given. Also, there is no strategic interaction with other member states' governments. Each member state's government does anticipate, however, how its choice of labor-market instruments affects the local economy. We model a one-time choice of labor-market instruments, with commitment to these values afterward.

An alternative scenario for the member-state government, that we also look at, is that the member state chooses state-contingent labor-market policies.

(2.24)
$$\max_{\{\tau_{v,t}^i, \tau_{\xi,t}^i, b_t^i, \tau_{J,t}^i\}} \int W_0 \, d\mu_0$$

s.t. a given federal RI scheme $\mathbf{B}_F, \tau_F$
the induced law of motion of the member state's economy (earlier sections)
the member-state government's budget constraint (2.22).

The main difference between the problems in (2.23) and (2.24) is that in the latter all instruments are chosen in a state-contingent way. That is the value of instruments changes with the state of the economy. What remains fixed over time and states of nature, however, is the shape of the federal UI scheme, which the federal government sets in at the beginning of period $t = 0$. In evaluating welfare in this scenario, we restrict ourselves to optimal Ramsey policies from a time-less perspective.

2.2.3. GDP. We view the resources spent retaining the match as intermediate goods, such that the definition of GDP is

$$gdp_t^i = y_t^i - e_t^i \int_{-\infty}^{\epsilon_t^{\xi,i}} \epsilon \, dF_\epsilon(\epsilon). \tag{2.25}$$

Market clearing expressed in units of GDP then is

$$gdp_t^i + \mathbf{B}_F\left(u_t^i; z_t^i\right) - \tau_F = e_t^i c_{e,t}^i + u_t^i c_{u,t}^i + \kappa_v v_t^i. \tag{2.26}$$

2.3. Analytical insights

A federal unemployment reinsurance scheme may affect the member states' labor-market policy mix, in the long run, during the transition phase after its introduction, and over the business cycle. This section seeks to derive intuition for the numerical results that follow later. For simplicity, in this section, we focus on the case that the member-state Ramsey planner can choose the entire labor-market policy mix (local UI benefits, layoff taxes, and hiring subsidies) in a state-contingent way. We focus on the case in which payouts only condition on current unemployment, $\mathbf{B}_F(u_t^i; u_t^{i,avg}) = \mathbf{B}_F(u_t^i)$. We start with results for the steady state and, thereafter, discuss considerations for the business cycle.

2.3.1. Intuition for the steady state. The steady state is symmetric, so steady-state values neither carry a super-script for the country nor a time index. $\mathbf{B}'_F(u)$ below marks the marginal increase in transfers payed by the federal level, evaluated at steady-state unemployment.

PROPOSITION 1. *Consider the economy described in Section 2.2. Let* $\Omega := \frac{\eta}{\gamma}\frac{1-\gamma}{1-\eta}$ *be the Hosios measure of search externalities and* $\zeta = \frac{\psi_s}{f(1-s)} \frac{1-e}{[\xi e+(1-e)]} \frac{\mathrm{u}'(c_u)-\mathrm{u}'(c_e)}{\mathrm{u}'(c_u)\mathrm{u}'(c_e)}$ *be a measure of tension between moral hazard and insurance of the* unemployed *in each member state. Suppose there is a given federal RI scheme that is characterized by* $\mathbf{B}_F(u_t^i)$ *and* τ_F. *Let* $\mathbf{B}'_F(u_t^i)$ *mark the first derivative of the federal transfer function. Focus on the steady state. The following labor-market policies and taxes implement the allocation that the member-state planner chooses*

$$\tau_v = [1-\Omega] + \frac{\eta}{1-\eta}\frac{\zeta}{\kappa_v \frac{\theta}{f}}, \tag{2.27}$$

$$\tau_\xi = \tau_J + \tau_v \kappa_v \frac{\theta}{f} + \zeta(1-sf) - \mathbf{B}'_F(u), \tag{2.28}$$

$$\begin{aligned} b = \ & \frac{(1-\beta)}{\beta}\tau_v\kappa_v\frac{\theta}{f}e + \zeta e\frac{[1-\beta(1-sf)(1-\xi)]}{\beta} \\ & + [\mathbf{B}_F(u) - \tau_F] \ + e\mathbf{B}'_F(u), \end{aligned} \tag{2.29}$$

$$\tau_J = \frac{1-e}{e}[b - \zeta sf] \ - \ \frac{\mathbf{B}_F(u) - \tau_F}{e}, \tag{2.30}$$

PROOF. The proof is an extension of the results in Jung and Kuester 2015 to the case of federal unemployment reinsurance, and presented in the online appendix to the current paper. □

Two elements of the federal RI scheme figure prominently. First, the steady-state net transfers per period, $\mathbf{B}_F(u) - \tau_F$. Second, the generosity of the federal RI scheme at the margin, $\mathbf{B}'_F(u)$. All else equal, the higher the net transfers, the more fiscal space there is for the member state and the higher will be the unemployment benefits b and the lower taxes on firms τ_J. In turn, the more generous federal RI is at the margin, the more generous a replacement rate will the member state set, see equation (2.29), and the less stringent will be the layoff restrictions, (2.36). The proposition, therefore, highlights that the federal RI scheme affects not only member-state's choice of unemployment benefits (b), but the wider labor-market policy mix that the member state considers optimal. Indeed, the numerical results later in the paper highlight precisely this. There, one finding is that a federal RI scheme will improve welfare the most if it restricts (or harmonizes) a wider range of policies (and not only unemployment benefits). Alternatively, the federal RI scheme will need an element of accountability. When we consider federal RI schemes that index payouts to the average unemployment level later, we do so precisely to take care of the long-run incentives to free-ride that arise at the margin.

Other effects can be seen most clearly under somewhat less general conditions:

COROLLARY 2. *Consider the same conditions as in Proposition 1. Define the average duration of an unemployment spell as* $D \equiv \frac{1}{sf}$, *and the average time of receipt of unemployment benefits as* $D_2 \equiv D - 1$.[10] *Define the elasticity of the latter with respect to unemployment benefits as* $\epsilon_{D_2,b} = \frac{D}{D_2}\frac{f(1-s)}{\psi_s}$. *Assume* $\beta \to 1$, *that utility is logarithmic,* $\mathrm{u}(c) = \log(c)$, *and the Hosios condition* $(\Omega = 1)$ *holds. Then the following characterize the policies that implement the allocation that the member-state planner chooses*

$$\frac{b}{w} = \frac{1}{1+D\epsilon_{D_2,b}} + \frac{D\epsilon_{D_2,b}}{1+D\epsilon_{D_2,b}}\left[\frac{\mathbf{B}_F(u)-\tau_F}{w} + e\frac{\mathbf{B}'_F(u)}{w}\right], \tag{2.31}$$

$$\frac{\tau_\xi}{w} = \frac{\frac{D}{1-\eta}-1}{1+D\epsilon_{D_2,b}} - \frac{(D\epsilon_{D_2,b}+\frac{D}{1-\eta})}{1+D\epsilon_{D_2,b}}\left[\frac{\mathbf{B}_F(u)-\tau_F}{w} + e\frac{\mathbf{B}'_F(u)}{w}\right], \tag{2.32}$$

$$\kappa_v\frac{\theta}{f}\frac{\tau_v}{w} = \frac{\eta}{1-\eta}\frac{D}{1+D\epsilon_{D_2,b}} - \frac{\eta}{1-\eta}\frac{D}{1+D\epsilon_{D_2,b}}\left[\frac{\mathbf{B}_F(u)-\tau_F}{w} + e\frac{\mathbf{B}'_F(u)}{w}\right], \tag{2.33}$$

$$\frac{\tau_J}{w} = -\left[\frac{\mathbf{B}_F(u)-\tau_F}{w} - u\frac{\mathbf{B}'_F(u)}{w}\right]. \tag{2.34}$$

$$\underbrace{\qquad\qquad\qquad}_{\textit{autarky}}\quad\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}_{\textit{federal RI}}$$

PROOF. The proof is a special case of Proposition 1. It is presented in the online appendix to the paper. □

In each of the rows, the terms labeled "autarky" are those that would appear absent a federal RI scheme, while "federal RI" marks the terms introduced by the scheme. We go through each of the instruments in turn. In the absence of a federal RI scheme, the steady-state replacement rate $\frac{b}{w}$ is given by a version of the closed-economy Baily-Chetty formula (Baily 1978 and Chetty 2006). This aligns the private gains to a worker of an increase in benefits with the social costs. See equation (2.31). Layoff taxes make domestic firms internalize the *domestic* fiscal costs of layoffs, namely, the payment of RI benefits over a typical unemployment spell and the additional cost of hiring subsidies, see equation (2.32). The latter, in turn make firms and workers internalize the search externality, equation (2.33). Absent a federal RI scheme, production taxes being zero, equation (2.34). For a detailed discussion of the case without federal RI, see Jung and Kuester 2015.

The novel results here, instead, concern how the federal RI scheme influences the policy instruments, the second terms. The terms in square brackets show how the federal RI scheme

[10] These differ because the first period of joblessness is covered by severance payments.

affects the choice of labor-market instruments in the long run. Once again, there are two effects, one pertaining to the average cost of financing, $\mathbf{B}_F(u) - \tau_F$. And one pertaining to the incentive effects for member states at the margin, $\mathbf{B}'_F(u)$. Let us turn to the former first. The federal RI schemes have to be self-financing in net present value terms, compare (2.19). This means that net transfers per period in the long run, $\mathbf{B}_F(u) - \tau_F$ need not necessarily be equal to zero. To the extent that net transfers at the steady state are positive, member states will cut production taxes. This will make hiring more attractive for firms, necessitating fewer hiring subsidies, lower lay-off taxes, and allowing for somewhat more generous unemployment benefits; equations (2.34) to (2.31), in that order.

As discussed earlier, though, the federal RI scheme not only has direct pecuniary effects. Rather, it also works to affect member states' optimal labor-market policies at the margin. These are the terms $\mathbf{B}'_F(u)$ in the above. For the sake of exposition, and somewhat realistically, suppose $\mathbf{B}'_F(u) > 0$, so higher unemployment in the member state means higher federal RI transfers. All else equal, the steady-state replacement rate rises linearly in the marginal payout. This is the more so, the more elastic search is with respect to benefits (the larger $\epsilon_{D_2,b}$) and the larger the pool of employed workers is that – through higher benefits – could be moved into a federally-cushioned unemployment spell, see equation (2.31). If the federal scheme pays at the margin, not only will this affect the member states' choice of unemployment benefits, but also will the member state tend to opt for lower layoff taxes, (2.32), and expend less on hiring subsidies (2.33). The production tax, τ_J, will be somewhat higher, reflecting that after adjusting the other instruments, the state still has funding needs. Taken together, the corollary suggests that federal RI benefits may induce member states to implement less employment-friendly policies. Indeed, numerically, we find in Section 2.5 that absent indexation to past unemployment, and when accounting for incentives to free-ride by the member state, the optimal federal RI scheme does not feature *any* payouts at or near the steady state. The first-order effect laid out here explains why.

It is intriguing to observe that federal RI matters even if it does not induce the member state to adjust the generosity of its local unemployment benefits. To see this in Proposition 2, suppose, for example, that unemployed workers' search intensity for a new job does not depend on their outside option. In other words, suppose that the micro-elasticity of search $\epsilon_{D_2,b} = 0$. In this case, there would be no moral hazard on the part of the unemployed in the member state. Member states would grant full consumption insurance, setting $b/w = 1$. This is so irrespective of the design of the federal RI system. Still, federal RI benefits will affect the rest of the labor-market policy mix. The reason is that the federal RI scheme changes the marginal costs for the member state of having a worker unemployed. The member-state government, therefore, reduces the hiring subsidy. In addition, it reduces the layoff tax. With the unemployed perfectly insured throughout, this comes at a fiscal cost even after the federal transfers received. The total effect of the measures, thus, has budgetary costs, and the member state raises the tax on firms. This explains why our numerical results in Section 2.5 suggest that constraining unemployment benefits alone may not be sufficient to render generous federal RI optimal.

Last, the corollary is suggestive of ways to contain the impact of federal RI on steady-state unemployment. Namely, what distorts instruments at the margin is $\mathbf{B}'_F(u) > 0$. Suppose, instead, that transfers were indexed to long-run unemployment, $\mathbf{B}_F(u_t, u_t^{avg})$. In the long-run $u^{avg} = u$. So $\frac{d}{du}\mathbf{B}_F(u, u^{avg}) = \frac{\partial}{\partial u}\mathbf{B}_F(u, u^{avg}) + \frac{\partial}{\partial u^{avg}}\mathbf{B}_F(u, u^{avg})$, where the scheme could be designed such that the two terms on the right-hand side cancel.

2.3.2. Business cycle and transition. Another contribution of the current paper is that we spell out the effect that federal RI could have on the member state's own business-cycle stabilization policies. Still for a given federal RI scheme, the following proposition summarizes the tax and benefit rules that emerge as optimal for the member state, over the business cycle and during the transition. We continue to focus on a federal RI scheme without indexation to long-run unemployment. All the variables mentioned below refer to the member-state level. Nevertheless, we drop superscript ^{i} for the sake of readability. For the same reason, the exposition here focuses on log utility $\mathrm{u}(c_t) = \log(c_t)$.

PROPOSITION 3. *Consider the member-state economy described in Section 2.2. Suppose there is a given federal RI scheme characterized by* $\mathbf{B}_F(u_t)$ *and* τ_F. *Let* $\mathbf{B}_F'(u_t)$ *be the first derivative of federal RI transfers in unemployment. In addition, assume that the bargaining power* η_t *is measurable* $t-1$. *Define the effective replacement rate as* $\widetilde{b}_t := c_{u,t}/c_{e,t}$. *Define* $\Omega_t := \frac{\eta_t}{\gamma}\frac{1-\gamma}{1-\eta_t}$. *Furthermore, suppose that the government implements the following policies for all periods* $t \geq 0$*:*

$$
\begin{aligned}
\widetilde{b}_{t+1}E_t\left[\beta\frac{\lambda_{t+1}}{\lambda_t}\frac{c_{e,t+1}}{e_{t+1}}\right] &= \tau_{v,t}\kappa_v\frac{\theta_t}{f_t}+\zeta_t-E_t\left[\beta\frac{\lambda_{t+1}}{\lambda_t}\zeta_{t+1}(1-s_{t+1}f_{t+1})(1-\xi_{t+1})\right]\\
&\quad -E_t\left[\beta\frac{\lambda_{t+1}}{\lambda_t}\tau_{v,t+1}\kappa_v\frac{\theta_{t+1}}{f_{t+1}}\frac{e_{t+2}}{e_{t+1}}\right]+E_t\left[\beta\frac{\lambda_{t+1}}{\lambda_t}\frac{\Pi_{t+1}}{e_{t+1}}\right]\\
&\quad +E_t\left[\beta\frac{\lambda_{t+1}}{\lambda_t}\left[\frac{\mathbf{B}_F(u_{t+1})-\tau_F}{e_{t+1}}+\mathbf{B}_F'(u_{t+1})\right]\right],
\end{aligned} \tag{2.35}
$$

$$
\tau_{\xi,t} = \tau_{J,t}-\mathbf{B}_F'(u_t)+\tau_{v,t}\kappa_v\frac{\theta_t}{f_t}+\zeta_t(1-s_tf_t), \tag{2.36}
$$

$$
\tau_{v,t} = \left[1-\frac{\Omega_{t+1}}{1+\varsigma_t}\right]+\frac{\eta_{t+1}}{1-\eta_{t+1}}\frac{\zeta_t}{(1+\varsigma_t)\kappa_v\frac{\theta_t}{f_t}}, \tag{2.37}
$$

$$
\begin{aligned}
\tau_{J,t} &= \frac{1-e_t}{e_t}\left[\widetilde{b}_tc_{e,t}-\Pi_t-\mathbf{B}_F(u_t)+\tau_F\right]+\kappa_v\tau_{v,t}\theta_t\left[\frac{1-e_t}{e_t}s_t-\xi_t\frac{1-s_tf_t}{f_t}\right]\\
&\quad -\zeta_t\xi_t(1-s_tf_t)
\end{aligned} \tag{2.38}
$$

where the two wedges are given by

$$
\varsigma_t = \frac{e_t(1-\xi_t)}{[\xi_te_t+(1-e_t)]f_ts_t}\left[1-\frac{\widetilde{b}_{t+1}(1-e_{t+1})+e_{t+1}}{\widetilde{b}_t(1-e_t)+e_t}\right]. \tag{2.39}
$$

$$
\begin{aligned}
\varsigma_t := \quad & \frac{\psi_s}{f_t(1-s_t)} \frac{1}{\lambda_t} \frac{1-e_t}{\xi_t e_t + (1-e_t)} \frac{1-\widetilde{b}_t}{\widetilde{b}_t(1-e_t)+e_t} \\
& + \frac{\psi_s}{f_t(1-s_t)} \frac{1}{s_t f_t[\xi_t e_t + (1-e_t)]} \frac{1}{\lambda_t} e_{t+1} \left[\frac{1}{\widetilde{b}_{t+1}(1-e_{t+1})+e_{t+1}} - \frac{1}{\widetilde{b}_t(1-e_t)+e_t} \right].
\end{aligned}
\tag{2.40}
$$

Note that both wedges are measurable in t. Then the following is true:

(1) These tax rules are consistent with the member-state government's budget constraint.

(2) The equilibrium allocations in the decentralized economy of the member state with federal transfers satisfy the first-order conditions in the member-states' Ramsey planner problem (2.24)*, and* vice versa.

PROOF. The proof is an extension of the results in Jung and Kuester 2015 and presented in the online appendix to the current paper. □

ζ_t is a measure of the tension between moral hazard and insurance of the *unemployed* in each member state. ς_t measures the wedge between the member-state planner's marginal utility of wealth and the employed workers' marginal utility. This term is zero in the steady state. Unemployment is a state variable that with the business cycle, and gradually over the transition. To the extent that the federal RI scheme both loads on unemployment and affects unemployment, Proposition 3, thus, highlights that the federal RI scheme affects the entire policy mix over the business cycle and the transition. Our numerical analysis seeks to evaluate to what extent each of these dimensions matters quantitatively over and above the effects that federal RI has in the long run.

As a step toward quantitative assessment, the next section presents the calibration. Section 2.5 presents the quantitative results.

2.4. Calibration and computation

This section calibrates the model to a stylized European Monetary Union ("EMU" henceforth). We wish to be clear: As of the time of writing, there is notable heterogeneity across EMU. Rather than taking a stand on the extent to which this heterogeneity is exogenous or due to policy choices, we deliberately abstract from *ex-ante* heterogeneity altogether. Rather, we wish to ask under which circumstances a federal RI scheme would be expected to provide notable welfare gains in a union of *ex-ante* identical generic member states, and what shape such a federal RI scheme should have in the presence of moral hazard by member states.

2.4.1. Calibration. We assume log utility $\mathrm{u}(c) = \log c$. One period in the model is a month. Our aim is to parametrize the model so as to replicate first and second moments of the data for a generic member state of EMU (with the 19 member states as of 2019Q4). The sample period is 1995Q1 to 2019Q4. This includes both tranquil periods and the deep recessions during and after the financial and debt crises.

As regards authority over policies, we target the *status quo*. There is no federal RI. Next to this, our understanding is that, today, in most member states the labor-market policy mix is kept constant over the business cycle. In the baseline, therefore, member states unilaterally choose a labor-market policy mix (that is, the replacement rate, the layoff tax, the hiring subsidy) that is fixed over the business cycle.

Under those assumptions, we calibrate model parameters so as to match average fluctuations and long-run moments for an "average" euro-area country. Toward this end, we obtain country-level data for 14 euro-area member states from Eurostat.[11] Of course, there would be endless different ways of assigning the movements in variables to the trend or cycle, or of assigning fluctuations to the individual country or to the aggregate level. We make two judicious choices here. First, in extracting the cyclical component for each time series we apply a linear trend (or, if possible, simply demean). We do so such that the drop in GDP in several member states after 2008 and the commensurate rise in unemployment is left in the cyclical component of the time series. Second, we assign all fluctuations in the data to country-specific shocks, rather than to the union level. We calculate moments for selected time series and calibrate the member state's economy in our model to a population-weighted average of these moments. We treat the parameters that emerge from the calibration as structural in the policy experiments that we conduct later.

Second moments of the data. The business-cycle properties of the data are reported in Table 2.1. All data are reported at quarterly frequency. Their model counterparts are quar-

TABLE 2.1. Business-cycle properties of the data

		y	c	$lprod$	e	$urate$	w
Stand. dev.		3.87	3.45	1.96	3.04	26.16	1.90
Autocorr.		0.96	0.96	0.93	0.99	0.99	0.92
Correlations	y	1.00	0.84	0.57	0.81	-0.61	0.29
	c	-	1.00	0.50	0.66	-0.50	0.27
	$lprod$	-	-	1.00	0.15	-0.14	0.36
	e	-	-	-	1.00	-0.74	0.28
	$urate$	-	-	-	-	1.00	-0.10

Notes: Summary statistics of the data (quarterly). Series are labeled like their counterparts in the model or as described in the text. All data are quarterly aggregates, in logs (including the unemployment rate), and multiplied by 100. We report the cyclical component after applying a linear trend. The exception is the log unemployment rate, which we demean only. Entries can be interpreted as percent deviation from the steady state. The first block reports the standard deviations and autocorrelations. The second block reports cross-correlations of time series within the typical country. The sample is 1995Q1 to 2019Q4. All entries are population-weighted averages of member state-level moments.

terly averages of the monthly observations in the model. The data are seasonally adjusted. The table reports percentage deviations from a linear trend of 100 times the log series (the log unemployment rate is only demeaned).

[11]The EA14: Belgium, Germany, Ireland, Greece, Spain, France, Italy, Cyprus, Luxembourg, Malta, Netherlands, Austria, Portugal, and Finland.

The first block of the table reports the standard deviation and first-order autocorrelation, the second block the cross-correlation of the main aggregates at the country level. All series are from Eurostat. Output y is real gross domestic product (chain-linked volumes). Consumption c is the consumption by households and non-profits divided by the GDP deflator. Labor productivity, $lprod := \frac{y}{e(1-\xi)}$, is measured as our series of GDP divided by employment (heads). The unemployment rate, $urate$ is taken directly from Eurostat. The model counterpart of the unemployment rate is $urate := (e\xi + u)s/[(e\xi + u)s + e(1-\xi)]$ (the mass of non-employed workers who search divided by the labor force). The counterpart of the model's wage, w, is taken to be the ratio of wages and salaries from the national accounts per employee and deflated by the GDP price index.

Targets and Parameters. Three of the model's parameters are directly linked to the business cycle: the standard deviation of the productivity shock, σ_a, the dispersion of the continuation costs, ψ_ϵ, and the wage rigidity parameter γ_w. We choose these so as to bring the model as close as possible to matching three business-cycle targets: the standard deviation of measured labor productivity, the standard deviation of the unemployment rate, and the relative standard deviation of the job-finding and separation rate. In our calibration the separation rate is 60 percent as volatile as the job-finding rate, in line with the findings for European OECD countries in Elsby et al. 2013. The other parameters are chosen directly based on outside evidence or using targets for the steady state of the model. The calibrated parameter values are summarized in Table 2.2. The monthly discount factor β equals .996, a customary value. In order to match an average unemployment rate ($urate$) of 9.5 percent, we adjust parameter $\overline{h}$ such that the value of leisure is $\Psi_s(s) + \overline{h} = 0.52$ or 91 percent of the wage. We set $\psi_s = 0.04$ with a view to matching the micro-elasticity of unemployment with respect to benefits. The value chosen here implies an elasticity of the average duration of unemployment with respect to UI benefits of 0.8, in line with micro estimates such as Meyer 1990. The vacancy posting cost of $\kappa_v = 0.86$ replicates the EMU-average monthly job finding rate of 7.5 percent derived from Elsby et al. 2013. For reference, this gives an average cost per hire net of the hiring subsidy, $\frac{v\kappa_v(1-\tau_v)}{m}$ of a little less than one monthly wage, in line with estimates of recruiting costs (Silva and Toledo 2009). We set the elasticity of the matching function with respect to vacancies to $\gamma = .3$, within the range of estimates deemed reasonable by Petrongolo and Pissarides 2001. The matching-efficiency parameter is set to $\chi = .12$ so as to match a quarterly job-filling rate of 71 percent. We take the latter target from den Haan et al. 2000.

The bargaining power of firms in steady state is set to $\eta = 0.3$, with an eye on the Hosios 1990 condition. Parameter γ_w governs the rigidity of wages with respect to fluctuations in productivity. We set this to $\gamma_w = 13.33$ to match the variability of unemployment. This implies that for a 1 percent negative productivity shock the bargaining power of firms falls by 13.33 percent, from a steady-state value of .3 to .26.

The average idiosyncratic cost of retaining a match is set to $\mu_\epsilon = .28$. This parameter governs the average costs of continuing a match. We set the parameter such that in steady state GDP equals output. Next, we set the dispersion parameter for the idiosyncratic cost shock to $\psi_\epsilon = 1.74$, with a view toward matching the relative volatility of job-finding and job-separation rates in Elsby et al. (2013). Last, we set the serial correlation of the productivity shock to $\rho_a = 0.98$. This translates into a quarterly persistence of the productivity of 0.94, within the range of values entertained in the literature. The standard deviation of the shock

TABLE 2.2. Parameters for the baseline

	description	value	target
Preferences			
β	time–discount factor	0.996	putative real rate of 4% p.a.
$\Psi_s(s) + \bar{h}$	value of leisure.	0.52	st.-st. u rate of 9.5 %
ψ_s	dispers. search cost	0.04	micro-elasticity, Meyer 1990.
Vacancies and matching			
κ_v	vac. posting cost	0.86	EMU-avg. monthly job finding rate.
γ	match elasti. wrt v	0.30	Petrongolo and Pissarides 2001.
χ	match-efficiency	0.12	qtrly job fill rate 71%, den Haan et al. 2000.
Wages			
η	firms' st.-st. barg. p.	0.30	Hosios condition.
γ_w	cyclic. barg. power	13.33	unemployment volatility.
Production and layoffs			
μ_ϵ	mean idios. cost	0.28	share of depreciation in GDP of 20%.
ψ_ϵ	dispers. cost shock	1.74	rel. vola. job-f., sep. rate, Elsby et al. (2013).
ρ_a	AR(1) prod. shock	0.98	qtrly persistence of prod. shock of 0.96.
$\sigma_a \cdot 100$	std. dev.	0.51	standard deviation of measured $lprod$.
Labor market policy			
b	unemploym. benefits.	0.38	optimal policy in steady state, autarky.
τ_v	hiring subsidy.	0.80	optimal policy in steady state, autarky.
τ_ξ	layoff tax.	6.39	optimal policy in steady state, autarky.

Notes: The table reports the calibrated parameter values in the baseline economy.

is set to $\sigma_a = 0.0051$, with an eye on the business-cycle properties (standard deviations) of the model, as discussed above.

Implied business-cycle statistics of the model. Table 2.3 reports business cycle statistics for the calibrated model. The calibrated model matches the data reasonably well, compare to Table 2.1. GDP is about twice as volatile as productivity. The log unemployment rate is about six times as volatile as GDP.

Implied steady state. Table 2.4 reports selected steady-state values for the baseline. The optimal replacement rate in steady state (b/w) is 52 percent, a reasonable value for the euro area, compare Christoffel et al. 2009. The optimal vacancy subsidy is $\tau_v = 0.80$. This amounts to a subsidy per actual hire of roughly two and a half monthly wages. The optimal layoff tax equals approximately 9 monthly wages, reasonable given the long average duration of unemployment spells in EMU and the corresponding fiscal costs.

Impulse responses. Figure 2.2 on page 68 shows impulse responses to a negative one standard-deviation productivity shock for the baseline calibration (as blue dashed lines labeled "autarky"). To repeat, the baseline features no federal RI system, hence there are no

TABLE 2.3. Business-cycle properties of the model

		gdp	c	$lprod$	e	$urate$	w	f	ξ
Standard dev.		4.70	3.89	1.96	2.90	26.16	2.06	18.77	12.52
Autocorr.		0.98	0.99	0.93	0.99	0.99	0.99	0.96	0.98
Correlations	y	1.00	0.99	0.92	0.97	-0.96	0.99	0.98	-1.00
	c	-	1.00	0.86	1.00	-0.99	1.00	0.94	-0.98
	$lprod$	-	-	1.00	0.80	-0.77	0.87	0.98	-0.93
	e	-	-	-	1.00	-1.00	0.99	0.90	-0.96
	$urate$	-	-	-	-	1.00	-0.99	-0.88	0.95
	w	-	-	-	-	-	1.00	0.95	-0.99
	f	-	-	-	-	-	-	1.00	-0.99

Notes: Second moments in the model. All data are quarterly aggregates, in logs and multiplied by 100 in order to express them in percent deviation from the steady state. Note: the series for the unemployment rate is in logs as well. The first row reports the standard deviation, the next row the autocorrelation. Then follow contemporaneous correlations. Based on a first-order approximation of the model.

TABLE 2.4. Steady-state values

Labor market policy			Output and consumption		
b	local UI benefits	0.38	gdp, y	GDP and output	0.72
τ_v	vacancy posting subsidy	0.80	c_e	consumption employed	0.73
τ_ξ	layoff tax	6.39	c_u	consumption unemployed	0.38
Labor market			Other variables		
ξ	separation rate	0.008	Π	dividends	0.002
f	job finding rate	0.078	Δ	gain from employment	1.65
s	search intensity	0.97	J	value of employ.-serv. firm	0.52
e	employment	0.91	τ_J	tax on firms	0.002
u	unemployment	0.0907			
$urate$	unemployment rate	0.095			

Notes: Selected steady-state values for the baseline economy.

federal transfers. Benefits b, hiring subsidies τ_v, and layoff taxes τ_ξ are kept constant. By construction, productivity a_t falls by about half a percent on impact and then gradually recovers (not shown). This directly translates into lower GDP. The fall in GDP is amplified and propagated further by the labor-market response. Wages are rigid and fall only about half as much as the productivity shock. Employment, therefore, falls (the unemployment rate rises by 0.3 percentage points). The monthly job finding rate falls by about 0.3 percentage points, the separation rate rises by 0.015 percentage points, and the search intensity of the unemployed falls. With the labor-market policies fixed, higher separations and less recruiting means that the government raises layoff taxes and saves on hiring subsidies. In spite of the recession, the production tax, therefore, initially falls by 0.3 percent. This response is short-lived, however.

2.4.2. Approximating and evaluating the federal RI scheme. We will look at two cases. Either payments are not indexed to past unemployment, so that $\mathbf{B}_F(u_t^i; u_t^{i,avg}) = \mathbf{B}_F(u_t^i)$, or payments are made as a function of the *difference* between actual unemployment and a long-term average of unemployment, so that $\mathbf{B}_F(u_t^i; u_t^{i,avg}) = \mathbf{B}_F(u_t^i - u_t^{i,avg})$. In each case, we approximate function $\mathbf{B}_F(\cdot)$ by a fourth-order Chebychev polynomial. Let ϕ mark the parameters of the polynomial. Unless noted otherwise, the outer nodes are fixed at plus/minus two standard deviations of an unemployment increase (measured using the member-states economy in autarky). At the four fixed Chebychev nodes, we parametrize the amount of transfers by values $\phi = [\phi_1, \phi_2, \phi_3, \phi_4]' \in \mathbb{R}$.

In case there is indexation, a choice has to be made as to how to measure average historical unemployment. We choose δ so as to allow for persistent transfers in persistent recessions on the one hand. On the other hand, the measure has to eventually respond to changes in local unemployment rates eventually. For the rest of the paper, we take $u_t^{i,avg}$ to be the 10-year geometric average of unemployment ($\delta = 1 - 1/120$).

Once we evaluate the member state's optimal stabilization policy over the business cycle, the model has too many states to solve it by global methods. Instead, throughout we rely on perturbation methods. The polynomials for $\mathbf{B}_F(\cdot)$ are flexible enough to allow for notable non-linearity (such as a scheme that are flat in the left tail and have threshold-like behavior in the right tail). Such asymmetries in $\mathbf{B}_F$ are important. We therefore need to make sure that the asymmetries feed through to how we evaluate welfare. We solve the model through a fourth-order perturbation with pruning, for which we rely on the routines by Levintal 2017. Our goal is to find an equilibrium in which the federal authority chooses the shape of the RI scheme so as to maximize expected welfare of member-state households, balancing the federal budget. In choosing, the federal authority anticipates the member states' optimal self-interested policy choices. Toward evaluating welfare, we extend the moment formulae in Andreasen et al. 2018 to fourth order. The formulae are reported in Appendix 2.B.

Last, we provide details on how we numerically address the maximization problems of the member states and the federal government. Our aim is to obviate the need for Monte-Carlo evaluations (such as drawing from the ergodic distribution) when evaluating welfare. Rather, we wish to have numerical evaluations of both the federal government's and the member state's objective function that become available in closed form as a by-product of solving the model by perturbation. We have to separate two different types of exercises: exercises that focus on the long run only, and exercises that include the transition path after the introduction of the federal RI scheme.

When we focus on the long-run incentives only, we evaluate the federal government's and the member state's objective function using the unconditional (that is, long-run) mean under the respective mix of policies. This is the case, for example, in Sections 2.5.1 and 2.5.2.

However, an important contribution of the paper is, precisely, to highlight the role of member-states' incentives during the transition phase. In other exercises, we do therefore deliberately include the transition phase after federal RI is introduced. Here, too, for each policy we wish to have closed-form solution for the objective function values. When solving the federal program (2.21) and the member states' (2.23) or (2.24), we condition the welfare evaluations on all member states entering period $t = 0$ from the same state. That is, there is no *ex-ante* heterogeneity. This is the case in Sections 2.5.3 and 2.5.4. Unless noted otherwise, we assume that the state that all member states start from, both in evaluating member-state policies and in searching for the optimal federal RI scheme, is the non-stochastic steady state implied by our calibration (Table 2.4). Appendix A provides more details.

2.5. Optimal federal unemployment reinsurance

The current section analyzes the shape and scope of an *ex-ante* optimal federal RI scheme for the calibrated model of the euro area. As a point of departure, we focus on the long run. We first illustrate the potential scope for federal unemployment reinsurance and then illustrate how member states' incentives shape the scope. Important take-aways will be that the entire labor-market policy mix matters for the optimal shape of the federal RI scheme, and that in the long run accountability can check member state's incentives to free ride.

Thereafter, we turn toward a dynamic perspective. That is, a perspective that explicitly accounts for the incentives on the transition path. These matter two-fold. First, on the transition path member states can have an incentive to free-ride as well, even if elements of long-term accountability are present (trading transitory pecuniary gains against permanently higher unemployment). Second, in autarky member states are liquidity-constrained. That is, the federal RI scheme is one way for member states to accumulate a buffer for the future. This may lead to member states saving or borrowing on the transition path. The dynamic perspective provides important insights into the shape of optimal federal RI.

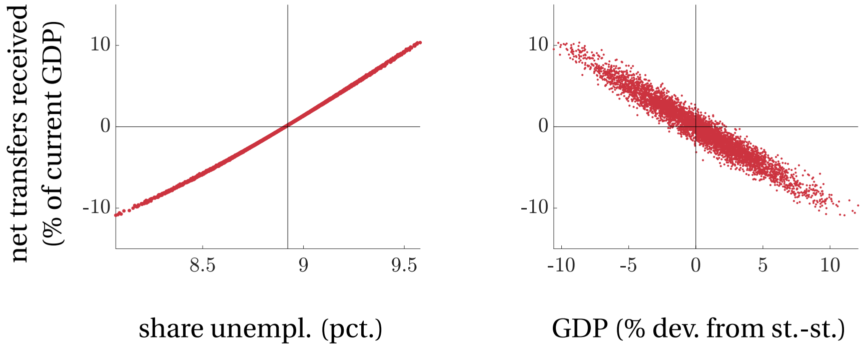
Table 2.5 in Section 2.5.5 collects the welfare gains for all scenarios, Table 2.7 the effect of federal RI on employment.

2.5.1. No response by member states: generous federal RI transfers. To set the stage, we want to know how generous federal unemployment reinsurance should be, abstracting from transition dynamics and from a response by the member state. That is, we hold member states' labor-market policies fixed. In order to compute the optimal federal RI scheme, in this section we take a long-run perspective. The optimal federal RI is computed as if the economy where immediately to jump to the new stochastic steady state. We choose the federal RI that maximizes the unconditional mean of welfare.

Figure 2.1 plots the net transfers that result under the optimal federal RI scheme, against unemployment (left panel) and GDP (right panel); results do not depend on whether there is indexation to past unemployment rates or not, so we only show one case here. The baseline economy is characterized by cyclical risk at the member-state level amid financial autarky. That is, the member state is both imperfectly insured and lacks the liquidity to smooth domestic consumption in the face of shocks. Absent a member state's response, these two elements shape the optimal long-run federal RI scheme. The optimal scheme is roughly linear in unemployment (left panel). And it is also generous: the implied transfers make up for virtually all of the output lost in recessions (right panel).

What is more, the optimal federal RI scheme not only stabilizes consumption. Rather, it also stabilizes the business cycle itself. Whereas in autarky a ten percent drop in GDP would be associated roughly with a six percentage-point rise in unemployment, with federal RI unemployment rises by only 0.7 percentage point (see the bounds in the left panel of Figure 2.1). Figure 2.2 illustrates the mechanism at work. The figures shows impulse responses to a one percent recessionary productivity shock if the optimal federal RI scheme is in place (red solid lines) and compares these to the impulse responses in autarky (blue dashed line). Under the

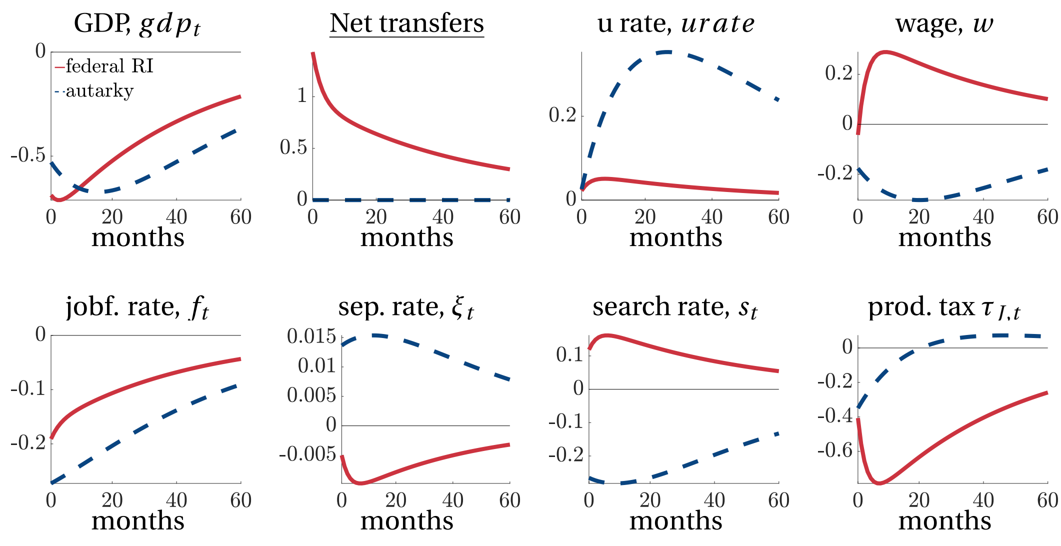
FIGURE 2.1. Net transfers received – fixed local policies



Notes: Net transfers received from the rest of the union $\mathbf{B}_F - \tau_F$ (y-axis) against share of unemployed workers (x-axis, left panel) or percent deviation of GDP from steady state (x-axis, right panel). Based on simulations under the optimal federal RI scheme for the case of fixed policies.

optimal European RI scheme, the layoff rate no longer rises in a recession and unemployment fluctuates by an order of magnitude less than in autarky.

FIGURE 2.2. Impulse responses, optimal federal RI – fixed local policies



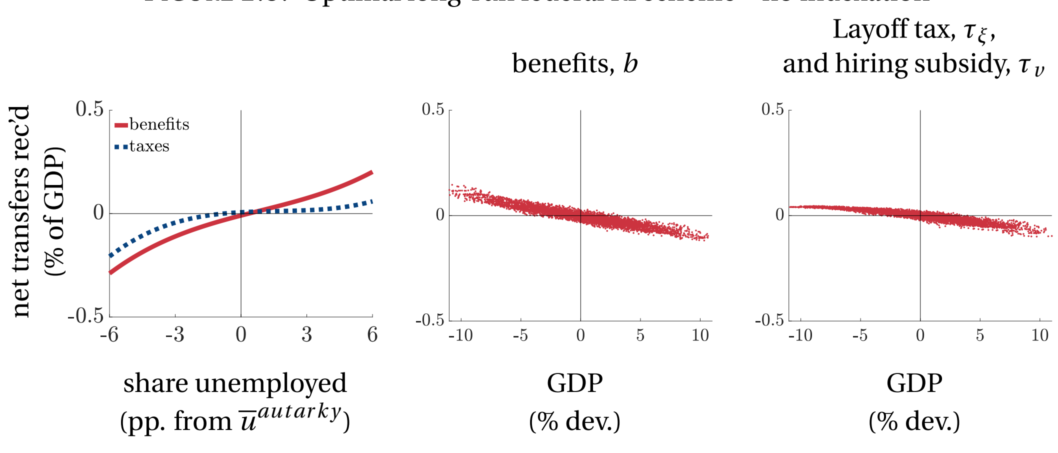
Notes: Impulse responses to a one standard deviation productivity shock. Shown is the case of autarky (blue dashed line) and the case with optimal federal RI (red solid line). Impulse responses are derived under the assumption that the local policy instruments do not react at all to the introduction of a federal RI scheme. All variables are expressed in terms of percent deviation from the steady state (a "1" meaning the variable is one percent above the steady-state level), except for the net transfers received and production tax, which are expressed in percent of GDP.

The key to this is the member state's fiscal policy. By assumption, the labor-market instruments (benefits, layoff taxes, and hiring subsidies) are fixed in this section. The generous federal transfers mean that the member state receives a sizable, persistent fiscal injection in a recession. In order to balance the government budget, this injection is distributed through cuts in taxes on production, τ_J. The persistent cuts in taxes raise the surplus of firms, stimulate hiring, and reduce layoffs; all of which stabilizes employment, and output. Indeed, average employment is slightly higher with federal RI and the standard of employment only a quarter the size under autarky, see Table (see entry "Section 5.1" in Table 2.7). The corresponding welfare gains from introducing federal RI for this scenario are large as well. The welfare costs of business cycles under autarky amount to 0.39 percent of life-time consumption. The federal RI scheme would bring welfare gains of exactly these magnitudes (see entry "Section 5.1" in Table 2.5).

2.5.2. One-time response by member states, the long-run view. To the extent that with the introduction of the federal RI scheme, member states do not centralize authority over labor-market policies at the federal level, member states might have an incentive to adjust policies after the federal RI scheme is introduced. Anticipation of this response would, then, change the shape of the optimal federal RI scheme. This is the focus of the current section. First, we focus on the no-indexation case. Then, we discuss the gains from indexation for the scope of federal RI policy. For now, we continue to focus on the optimal policies that emerge from a long-run perspective.

2.5.2.1. *No indexation: few transfers.* Suppose payouts under the federal scheme are not indexed to past unemployment rates, so payouts are a function only of current unemployment. For this case, Figure 2.3 presents the shape of the optimal federal RI scheme and the implied federal transfers. There are two important results. First, federal transfers are almost

FIGURE 2.3. Optimal long-run federal RI scheme – no indexation



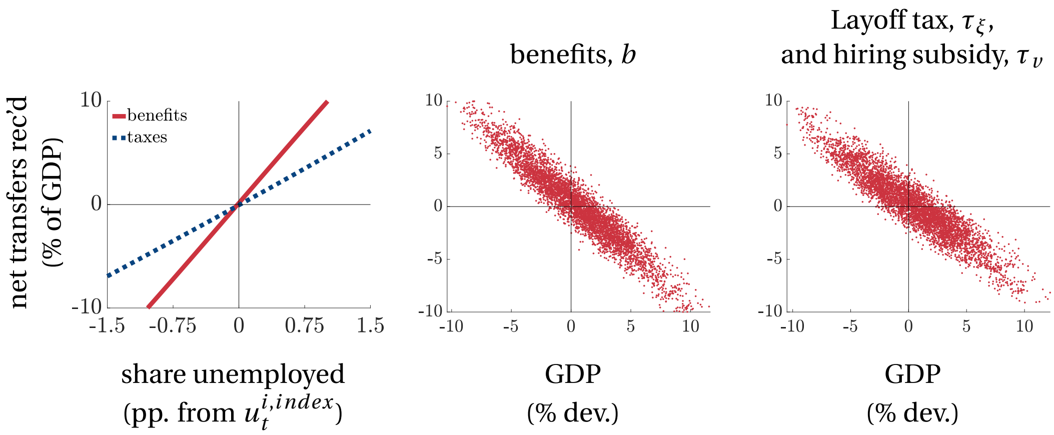
Notes: Optimal federal RI without indexation, evaluated using long-run welfare. Left: net transfers received from the rest of the union $\mathbf{B}_F(u_t^i) - \tau_F$ (y-axis), against share of unemployed workers (x-axis, left panel). The member state either can adjust local UI benefits once and for all (solid red line), or layoff taxes and hiring subsidies (brown dotted line). Center panel: net transfers as percent of GDP (y-axis), against GDP if the member state can adjust benefits; based on 10,000 simulations. Right panel: same as center panel but the member states can adjust only layoff taxes and hiring subsidies.

two orders of magnitude less generous here than absent a response by the member states (compare previous section). In line with this, also the welfare gains from the optimized federal RI scheme are minuscule (see the columns pertaining to entries "Section 5.2.1, no indexation" in Table 2.5). Figure 2.3 presents two different cases. In one ("benefits, b"), the member state can adjust benefits but not hiring subsidies and layoff taxes. In the other layoff tax and hiring subsidies can be adjusted, but benefits remain fixed at the level that the member state would choose in autarky. A red solid line in the left panel marks the case when the member state of all the labor-market instruments can adjust benefits only. The blue dots in the left panel show the case when the member state can adjust the other labor-market policies only, but not benefits. The other two panels simulate the corresponding transfers in GDP space. Next to the small size of payouts, the second key take-away is that the size of federal transfers tends to be small regardless of which instrument is allowed to adjust. This highlights that the entire mix of labor-market policies matters. Indeed, for the other labor-market instruments, the incentives to free ride are stronger than for benefits. Only if the unemployment rate rises by three percentage points above the steady-state level will the federal RI scheme pay any transfers; what emerges is a trigger system similar to the US federal unemployment insurance system. In itself, more unemployment is socially costly. In order to prevent the member state from choosing too high unemployment permanently, the optimal federal scheme pays transfers only when unemployment is sufficiently high; recall the discussion of the marginal payout's ($\mathbf{B}'_F$) effect on policy choices in Proposition 2. Even in the tails payouts remain small.

2.5.2.2. *Indexation: long-run gains.* The previous analysis illustrates that member states' behavioral responses could potentially severely limit the scope for federal unemployment reinsurance. This section introduces an element of accountability into the design of the federal RI scheme, namely, the indexation of payouts to past unemployment discussed in Section 2.4.2. For now, we keep the focus on the long run.

Figure 2.4 shows that indexation of payouts to past unemployment rates successfully controls the member states' incentives to free ride in the long run. Indeed, the payouts (center

FIGURE 2.4. Optimal long-run federal RI scheme – indexation to past unemployment



Notes: Same as Figure 2.3, but with indexation of payouts to a member state's past unemployment.

and right panel) are about as generous as they were absent any long-term incentives to free ride, compare to the center and right panel of Figure 2.1.[12] There is notable redistribution between economies in boom and recession. The welfare gains of introducing federal unemployment RI accordingly are large. Still, at about 0.2 percent of life-time consumption they are only half as large as absent any response by member states (Table 2.5). The reason is that unemployment is a slow-moving state variable. Indexation therefore makes federal payments less well-timed.

It is important to note that these welfare gains arise amid a *fall* in average employment (Table 2.7, entries for Section 5.2.2) relative to autarky. The reason is that federal RI provides insurance to member states. Absent federal RI, member states choose labor-market policies to partially self-insure against economic fluctuations. In the model at hand, self-insurance means fostering employment relationships. This effect can be seen in the first column of Table 2.7. This column reports average employment for all the respective scenarios, in autarky. The entry for Section 5.1 (an average employment rate of 0.901) is based on labor-market policies that are optimal for the steady-state economy (as in Section 2.4.1). That is, optimal if there is no cyclical risk. When member states, instead, choose policies in light of business-cycle risk, they opt for policies that induce higher average employment. If member states can optimize benefits only, the average employment rate rises by 2 percentage points to 0.925, and further still when member states can use layoff taxes and hiring subsidies (to 0.938). Federal unemployment reinsurance allows member states to reduce these costly efforts to self-insure. Employment fluctuations fall quite notably, too.

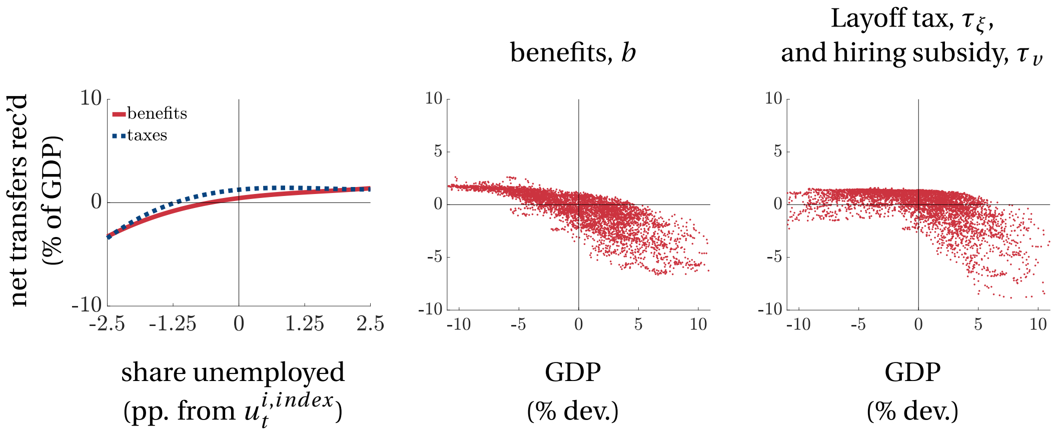
So far, all the calculations have assumed, that the economy would immediately jump from one steady state to the new stochastic steady state. That is, we have focused on the long-run welfare gains. After the introduction of a federal RI scheme, however, the member-state economies will only gradually move to the new non-stochastic steady state. This has important implications for the scope of federal RI, to which we turn next.

2.5.3. Accounting for the transition: limited generosity. This section explicitly allows for the transition period, both in evaluating welfare and in searching for the optimal federal RI scheme. We use the calibrated non-stochastic steady state as the initial state for the welfare evaluation. We highlight two important considerations that arise once a federal RI scheme is introduced. First, federal RI – to the extent it is pre-funded – gives member states a means of accumulating buffer stock savings. Second, and opposing this, the transition phase itself renders the indexation of benefits to past unemployment a weaker deterrent to free-riding. The reason is that a long-run average of unemployment by definition is slow to catch up if the member state engages in policies that persistently increase unemployment (reduce employment). The resulting incentives one misses looking at the long run only. It turns out that the latter is of first-order importance.

Figure 2.5 shows the optimal federal RI schemes that emerge when accounting for the transition period. All of the schemes shown here have indexation; so compare these to Figure 2.4. In spite of indexation, the schemes no longer are linear, or nearly as generous as depicted in Section 2.4. Accounting for the transition phase, the schemes have three features.

[12]Note that the left panel in Figure 2.4 looks different from the left panel in 2.1. The reason is that the x-axis units differ. In the left panel of Figure 2.4, we plot net transfers as a function of the deviation of unemployment from past average unemployment. In the corresponding panel of Figure 2.1 we plot transfers against current unemployment.

FIGURE 2.5. Accounting for the transition, with indexation

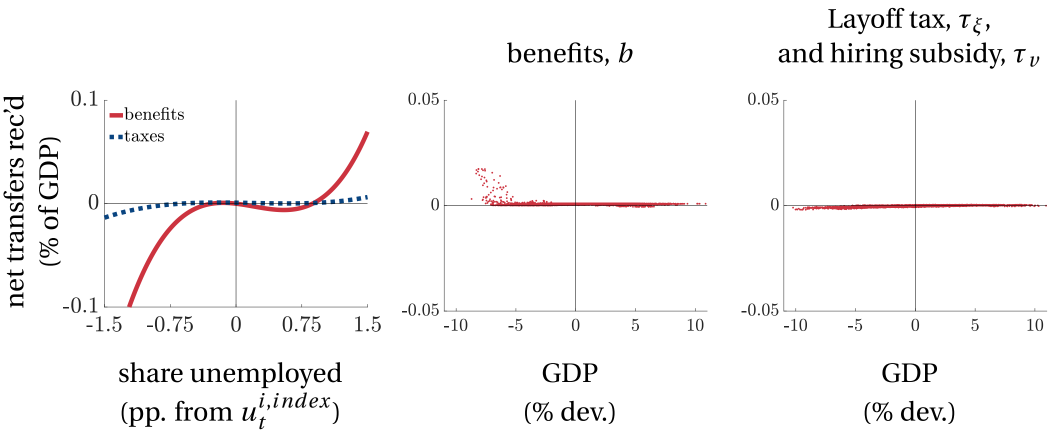


Notes: Same as Figure 2.4, with the difference that – in assessing welfare and in optimizing policies – both the member state and the federal government account for the transition phase. Member states continue to set benefits or layoff taxes and hiring subsidies once and for all in the initial period.

First, redistribution over the business cycle. Member states experiencing a boom pay into the schemes, member states experiencing a recession receive transfers. This redistribution is sizable. In a ten-percent of GDP boom, a member state would transfer about 5 percent of GDP to the rest of the union. Note that this is about half the transfer that prevailed abstracting from the transition path. Second, the schemes have positive net transfers at the steady state. These transfers amount to 0.32 and 1.17 percent to GDP, respectively, when member states can adjust benefits or when they can adjust layoff taxes and hiring subsidies. Third, and still more striking, the schemes see payouts rise very little with unemployment for recession countries (whenever current unemployment exceeds past unemployment), inspite of indexation of payouts to past unemployment. For this, the transition path is essential. By setting more generous unemployment benefits or a less employment-friendly mix of layoff taxes and hiring subsidies, the member state can raise unemployment rates permanently. Indexation means that this does not generate future transfers. In the shorter run, however, higher unemployment does generate transfers from the federal level. Since the federal unemployment RI scheme presented here implements transfers, rather than loans, these initial transfer the individual member state would not expect to repay ever. In other words, there are incentives to free-ride on the federal RI scheme shortly after implementation of the scheme. As a result, once accounting for the transition phase, the federal RI scheme takes on a characteristic from the scheme without indexation, shown in Figure 2.3: the flat inaction region is large. The optimal federal RI scheme resembles a transfer scheme. The scheme makes essentially flat payments to member states in recession, with little regard to the depths of the recession. Accounting for the transition, at 0.03 to 0.06 percent of life-time consumption, the welfare gains from introducing the federal RI scheme are still notable, but at most one third of those derived disregarding the transition; see entry "Section 5.3" in Table 2.5). With the federal RI scheme, the standard deviation of employment falls by about half; average employment is about 3 percent lower than absent federal RI; see Table 2.7.

2.5.4. Counter-cyclical labor-market policy and federal RI. So far, we have assumed that member states could not (did not) engage in countercyclical labor-market policies of their own. In this section, instead, we allow the member states to adjust their labor-market policy-mix not only in the long run but also over the business cycle.[13] Again, we study two scenarios: (i) member states adjust unemployment benefits and production taxes, while hiring subsidy and layoff taxes are fixed at the steady state level; (ii) member state adjust all taxes while benefits are fixed at the steady state level. As in Section 2.5.3, we allow for the transition period, both in evaluating welfare and in searching for the optimal federal RI scheme. We use the calibrated steady state as the initial point for the welfare evaluation. Other aspects of the exercise remain as in previous sections.[14]

FIGURE 2.6. Member states adjust policy over the business cycle, with indexation.



Notes: Member states adjust unilaterally policy-mix both in the long run and over the business cycle so as to maximize welfare of households within their jurisdiction. The benefits are indexed to long-run unemployment with δ corresponding to ten years.

Figure 2.6 reports the payouts under the federal RI schemes that arise as optimal. Note that the scale on the y-axis is about two orders of magnitude smaller than in the previous graphs. The main take-away, therefore is that – once member states optimize their cyclical

[13] For the Ramsey policies (the choice of state-contingent Pigouvian benefits, or layoff taxes and hiring subsidies), we have to make a decision as to the initial values of the multipliers on forward-looking constraints (past promises). The results shown here proceed under a time-less perspective, that is, as if past promises bind the Ramsey policy maker. Results were very similar, when we computed policies as if there were no past promises to start with.

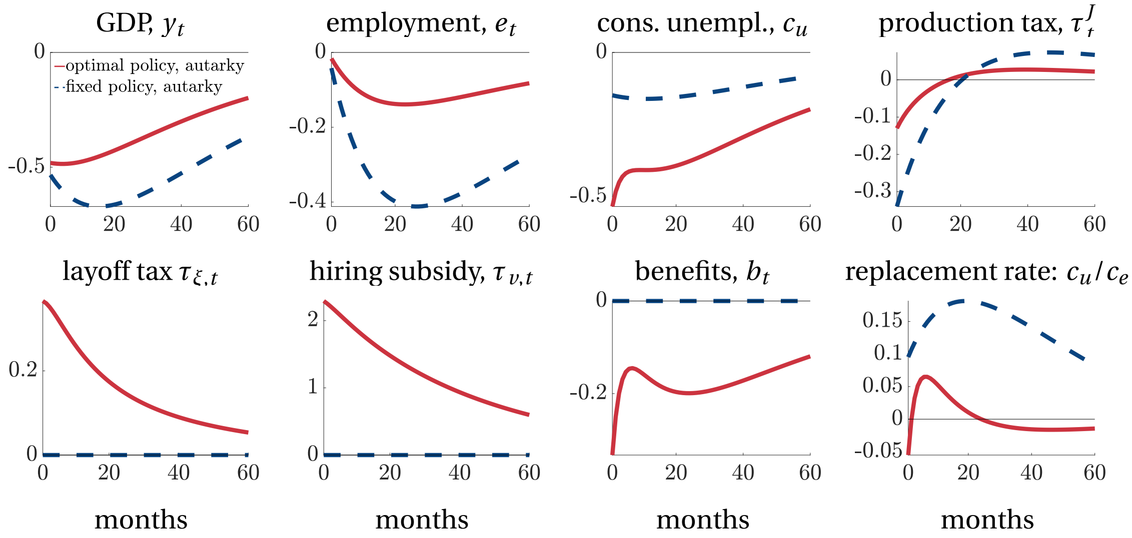
[14] The only change with respect to previous sections is the interval of unemployment levels over which we specify the nodes of the Chebyshev polynomial defining the federal reinsurance scheme; compare Section 2.4.2. Since the volatility of unemployment falls notably when countries adjust policies optimally over the business cycle, the domain over which we define the federal RI scheme needs to be adapted accordingly. Here, we choose the four Chebyshev nodes to cover the interval between -0.0175 and 0.0175, which is defined in terms of the deviation of the share of unemployed from its long-run level. This corresponds to ± 2 standard deviations of unemployment prevailing in the autarky when member states adjust all instruments optimally over the business cycle. We have experimented with different domains. Qualitatively the results are robust to the choice of domain: There is very little scope for federal reinsurance whenever member states adjust the policy mix over the business cycle.

stabilization policies – the optimal federal RI scheme features virtually no payouts to speak of. Accordingly, the welfare gains from federal RI are negligible; see entries "Section 5.4" in Table 2.5. This is so inspite of the fact that there is full long-run indexation of payouts to past unemployment. The reason for this is two-fold. First, federal RI was beneficial in the settings discussed above in part because it helped smooth member states' business cycles. In theory, an alternative to such stabilization would be counter-cyclical labor-market policies implemented unilaterally in member states. Indeed, in the current model environment, optimal labor-market stabilization policies at the member-state level already go a long way in stabilizing economic activity and in reducing the cost of fluctuations; see Jung and Kuester 2015. We show the welfare gains in Table 2.6 in Section 2.5.5. Optimal countercyclical labor-market policy removes about half to 3/4 of the costs of business cycles. Such policy can remove the amplification of shocks, but cannot directly address the country-specific risk, however. Section, while federal unemployment RI could be a useful complement, federal RI does crowds out member state's incentives to conduct stabilization policy on their own. We discuss these considerations in the next two sections.

2.5.4.1. *The gains from labor-market stabilization policies.* For the current model, this section discusses the gains (over the constant-policy baseline, and in the current model environment) that member states could reap from introducing countercyclical labor-market stabilization policy. We allow member state governments to adjust all their policy instruments every period, in a state-contingent way, so as to maximize welfare in its constituency. To prepare the ground, we abstract from any federal RI scheme. The analysis here, thus, follows closely Jung and Kuester's (2015) analysis for the closed economy; we apply their insights to the current calibration for the euro area. Figure 2.7 shows impulse responses to a negative one standard-deviation productivity shock. The response under optimal countercyclical stabilization policy is shown as a red solid line (assuming all labor-market policies are optimized over the business cycle). The blue dashed line refer to the response in the calibrated baseline, where member states help their policies fixed (the same scenario shown as blue dashed lines in Figure 2.2). Optimal self-interested countercyclical labor-market stabilization policy would notably stabilize employment and, thus, the business cycle. By raising layoff taxes and hiring subsidies, the member state can stabilize employment. Unemployment benefits would be adjusted so as to keep the replacement rate roughly constant. In particular, the fall in employment in the recession would be only 30 percent of the fall witnessed in the constant-policy baseline. This renders the business cycle less costly to start with. In the scenario shown here, the member state's optimal policy removes 85 percent of the cost of business cycles.

2.5.4.2. *Does federal RI crowd out countercyclical stabilization policy?* So as to further understand why, quantitatively, there remains little scope for federal unemployment reinsurance if member states adjust labor-market policies optimally over the business cycle, this section looks at how federal RI affects member state's policy responses to productivity shocks with and without federal RI. We impose one specific federal RI scheme, namely, the scheme that emerged as optimal in Section 2.5.3, where member states could choose layoff taxes and hiring subsidies once; fully accounting for the transition (the blue dashed line in the left panel of Figure 2.5). Figure 2.8 presents impulse responses to a negative productivity shock with and without that federal RI scheme; in each case with the member state implementing optimal countercyclical labor-market stabilization policy for the respective case.. The solid red

FIGURE 2.7. Member state's optimal stabilization policy – no federal RI



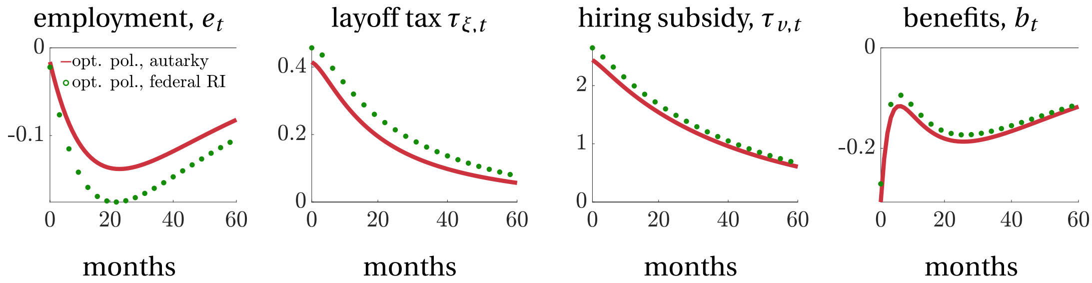
Notes: Impulse responses to a one standard deviation productivity shock. There is no federal insurance. Shown are two cases a) the member state optimally adjusts its labor-market policies over the business cycle (red solid line) and b) the member keeps labor-market policy instruments fixed (blue dashed line). The latter case corresponds to the blue-dashed case shown in Figure 2.2. All variables are in percent deviations from the steady state (a "1" meaning the variable is one percent above the steady-state level). The exception are those variables that are expressed in rates to start with. The responses of the latter are shown in percentage points. The production tax is expressed as a percent of GDP.

line is the case without federal RI (identical to the solid red line in Figure 2.7). The green circles mark the responses that the member state optimally implements when said federal RI scheme is in place. The figure suggests that federal RI induces member states to engage less actively in countercyclical labor-market stabilization. As a result, employment becomes more volatile (left-most panel). In addition, not only does federal RI mean more volatile employment, the federal RI scheme in the current example also means that average employment is considerably lower than without. Indeed, this is the case even if the member state adjusts only benefits or only the other labor-market instruments optimally (entries "Section 5.4.2" in Table 2.7 show this). Consequently, such a federal RI scheme would be detrimental to welfare, see the last row of Table 2.5.

2.5.5. Welfare. Table 2.5 summarizes the welfare effects of federal unemployment RI. The welfare costs are reported in terms of the percentage increase in equivalent steady-state life-time consumption (right column; an entry of 0.39 meaning the welfare gain amounts to 0.39 percent of life-time consumption). For reference, we also compare the welfare gain to the costs of business cycles in the baseline economy of Section 5.1 (first column; in percent; a 100 would mean that the gains from introducing federal RI in consumption-equivalent terms would be as large as the gains from abolishing business cycles).

The table shows that potential welfare gains heavily depend on what member states can do in terms of cyclical labor-market stabilization policy, and in terms of responding to the

FIGURE 2.8. Member state's stabilization policy – fixed federal RI



Notes: Impulse responses to a one standard deviation productivity shock. Red solid line: same as in Figure 2.7. Green circles: the federal RI scheme shown as a blue dashed line in the left panel of Figure 2.5) is in place, and the member state chooses all labor-market instruments optimally over the business cycle, taking the federal RI scheme into account.

federal RI scheme. At the lower end, the welfare gains from the optimal federal RI scheme are virtually nil. At the high end, they account for 0.39 percent of steady-state consumption. The earlier sections have discussed the reasons. Table 2.6 reports the welfare gains that arise from countercyclical labor-market policy at the member-state level alone. In terms of magnitude, the gains to be had could be about as large as the maximum gains from introducing federal RI.

2.5.6. Effect of federal unemployment reinsurance on employment. Table 2.7 summarizes the effect of federal RI on average employment (left column) and the fluctuation of employment (right column). Here, too, the effect of federal RI on employment heavily depends on what policy options member states have, and what policy whether one accounts for the transition or not. Average employment may rise by 1 percent or fall by up to 3 percent for the optimized federal RI schemes. Similarly, federal RI may notably reduce the standard deviation of employment (by up to 87 percent), or barely have an effect where federal UI is stingy. Once more, the earlier sections have discussed the reasons.

TABLE 2.5. Welfare gains from optimal federal unemployment reinsurance

		welfare gain relative to cost of B.C. (%)	welfare gain (cons. equivalent, %)
No transition	Section 2.5.1		
	no response by members	100	0.39
	Section 2.5.2.1, one-time response, no indexation		
	only b adjusts	1	0.004
	only τ_ξ and τ_v adjust	0.32	0.0013
	Section 2.5.2.2, one-time response, indexation		
	only b adjusts	48	0.19
	only τ_ξ and τ_v adjust	56	0.22
With transition	Section 2.5.3, one-time adjustment, indexation		
	only b adjusts	8	0.03
	only τ_ξ and τ_v adjust	15	0.06
	Section 2.5.4.1, optimal business-cycle policy		
	only b adjusts	$\approx 10^{-4}$	0.00017
	only τ_ξ and τ_v adjust	$\approx 10^{-4}$	0.00011
	Section 2.5.4.2, federal RI scheme from Section 2.5.3 amid optimal business-cycle policy		
	only b adjusts	-17	-0.07
	only τ_ξ and τ_v adjust	-59	-0.23

Notes: Welfare gains from the introduction of federal RI. The first column shows how big the gains are relative to the welfare cost of business cycles in autarky and in the calibrated baseline economy. An entry of "100" would mean that the welfare gains are of the same magnitude. The second column shows the consumption-equivalent welfare gains from introducing the optimal federal RI scheme (in percent of steady-state consumption; a "1" meaning one percent of steady-state consumption). Welfare gains are computed relative to the same scenario without federal RI. That is, relative to a scenario where member states have the the same policy options with and without federal RI. Entries have been rounded.

TABLE 2.6. Welfare gains from countercyclical labor-market policy alone

		welfare gain from introducing stabilization policy, relative to cost of B.C. without (%)	welfare gain (cons. equivalent, %)
With transition	Section 2.5.4.1, optimal business-cycle policy		
	only b	20	0.20
	only τ_ξ and τ_v	51	0.32
	all	85	0.33

Notes: Welfare gains from introducing countercylical labor-market policy. The table shows how big these gains are relative to the welfare cost of business cycles in the baseline economy. An entry of "100" would mean that they are as big. Welfare gains are computed relative to the same scenario without federal RI but the same policy options at the level of the member state.

TABLE 2.7. Effect of federal RI on employment

		Autarky employment	Effect of federal RI on employment	
		uncond. mean	uncond. mean	standard dev.
No transition	Section 2.5.1			
	no response by members	0.901	+1%	–76%
	Section 2.5.2.1, one-time response, no indexation			
	only b adjusts	0.925	–0.2%	+5%
	only τ_ξ and τ_υ adjust	0.938	–0.2%	+5%
	Section 2.5.2.2, one-time response, indexation			
	only b adjusts	0.925	–0.1%	–55%
	only τ_ξ and τ_υ adjust	0.938	–0.7%	–46%
With transition	Section 2.5.3, one-time adjustment, indexation			
	only b adjusts	0.917	–3%	+53%
	only τ_ξ and τ_υ adjust	0.923	–3%	+14%
	Section 2.5.4.1, optimal business-cycle policy			
	only b adjusts	0.909	+0.1%	–25%$^{(*)}$
	only τ_ξ and τ_υ adjust	0.909	–0.1%	–7%$^{(*)}$
	Section 2.5.4.2, federal RI scheme from Section 2.5.3 amid optimal business-cycle policy			
	only b adjusts	0.909	–1.6%	+28%
	only τ_ξ and τ_υ adjust	0.909	–4.5%	+33%

Notes: For each case discussed in the earlier sections, the table reports unconditional moments for employment both with the federal RI scheme in place and in autarky. Left: long-run mean under autarky (if member states can choose the same set of policies in autarky as after introduction of federal RI). Center: change in unconditional mean after introducing federal RI and allowing member states to react (+1% meaning that average employment with the federal RI scheme is one percent higher than in autarky). Right: change in the standard deviation of log employment relative to the calibrated baseline (-76% means that after introduction of federal RI the standard deviation of log employment is 76% lower than in the baseline; where the standard deviation is 2.9 – see Table 2.3. $^{(*)}$: throughout, the Table reports the effect of federal RI on the standard deviation of employment. With optimal business-cycle policy, already in autarky, the standard deviation of employment is considerably lower than with fixed instruments; namely 0.12% when only benefits adjust and 1.10% when taxes adjust.

TABLE 2.8. Instrument values for each scenario, and implied steady-state and fluctuation

		instruments			steady state		
		b	τ_ξ	τ_v	$\frac{\bar{c}_u}{\bar{c}_e}$	$\bar{e}$	std(e)
No transition	Section 2.5.1						
	instr. for non-stoch st. st.	0.379	6.39	0.796	0.518	0.91	2.90
	Section 2.5.2.1, one-time response, no indexation						
	b adjusts						
	autarky	0.367	6.39	0.796	0.503	0.927	2.03
	federal RI	0.369	6.39	0.796	0.506	0.924	2.13
	τ_ξ *and* τ_v *adjust*						
	autarky	0.379	7.22	0.813	0.520	0.941	1.81
	federal RI	0.379	7.12	0.811	0.520	0.938	1.81
	Section 2.5.2.2, one-time response, indexation						
	b adjusts						
	autarky	0.367	6.39	0.796	0.503	0.927	2.03
	federal RI	0.373	6.39	0.796	0.511	0.919	0.92
	τ_ξ *and* τ_v *adjust*						
	autarky	0.379	7.22	0.813	0.520	0.941	1.81
	federal RI	0.379	6.86	0.797	0.520	0.929	0.98
With trans.	Section 2.5.3, one-time adjustment, indexation						
	b adjusts						
	autarky	0.371	6.39	0.796	0.508	0.928	2.25
	federal RI	0.386	6.39	0.796	0.537	0.876	3.44
	τ_ξ *and* τ_v *adjust*						
	autarky	0.379	6.81	0.809	0.520	0.941	2.25
	federal RI	0.379	6.29	0.791	0.520	0.899	2.57

Notes: For each of the scenarios in which instruments are either fixed or chosen once and for all in the initial period, the table reports the value of the labor-market instruments (first three columns). The table also reports information for the implied steady state, namely, the effective replacement rate of consumption c_u/c_e and employment (next two columns). Last the table reports the value of the standard deviation of log employment (in percent, last column).

2.6. Conclusions

What is the scope of a federal unemployment reinsurance scheme in a union of member states that retain authority over local labor-market policies? The paper has provided a quantitative exploration for a stylized European Monetary Union. Our answer is: it heavily depends on what labor-market policies member states can implement on their own, and how they can react to the introduction of the federal scheme. We have shown scenarios in which the welfare gains were large; and scenarios in which their were miniscule.

To us, there are three important findings. First, for assessing federal unemployment reinsurance, it matters who has authority over the entire mix of labor-market policies, and that goes beyond the unemployment insurance system. Second, in the long run indexing payouts to past unemployment in the member state addresses free-riding successfully. Third, the short run is essential for assessing federal unemployment reinsurance scheme. This is so because, on the one hand, shortly after the introduction of the federal reinsurance scheme indexation is necessarily imperfect. And, on the other and to the extent such responses are feasible, because federal unemployment reinsurance may interfere with member states' own stabilization policies.

The reader may have several objections to the exercise, all of which are well taken. We focus on a union of atomistically-small member states, abstracting from strategic interaction. We do not model externalities across member states, or aggregate shocks at the level of the union. We also abstract from heterogeneity *ex ante* across member states, or more pronounced heterogeneity within. And we abstract from labor migration when federal social insurance would bring portability of benefits which might be conducive to efficient reallocation of labor. In all these dimensions, future work seems valuable to us.

In closing, an important note. We have looked at simple federal unemployment reinsurance schemes where payouts to member states depended on unemployment rates (or the rise in unemployment relative to a long-run average). We did so for two reasons: because this is the flavor of current policy proposals, and because we looked for federal schemes that might be implementable in the current European institutional setting. More complicated schemes (that condition payouts on member states' policy choices, or past sequences of shocks, say) could improve upon the schemes that we have analyzed, but might require a further transfer of political authority to the community level (such as giving the European Parliament or the Commission discretionary authority over payouts). That said, the analysis here could easily be extended to such scenarios.

Appendices to Chapter 2

2.A. Finding optimal policies and welfare gains

This section describes the algorithm that we use to derive the optimal federal RI scheme.

As discussed in Section 2.4.2, we wish to have closed-form expressions for the objective function values (for given member-state policies and a given federal RI scheme). To obtain optimal federal RI in Sections 2.5.1 and 2.5.2, we find federal and member-state policies that maximize the unconditional mean of the objective function. When accounting for the transition, in Sections 2.5.3 and 2.5.4, we find policies that maximize the conditional expectation of the objective. In these cases, we condition on the initial state being the non-stochastic steady state implied by our calibrated model (Table 2.4).

The next section describes how we find the optimal federal RI scheme. We describe this for the case when the member state can choose policies once and we account for the transitions. The other cases are handled analogously.

2.A.1. Finding the optimal federal RI scheme. The algorithm proceeds as follows.

(1) Fix Chebyshev nodes for the federal RI scheme as described in Section 2.4.2. Keep nodes fixed.
(2) The goal is to find values $\phi = [\phi_1, \phi_2, \phi_3, \phi_4]' \in \mathbb{R}$ and τ_F that solve the federal government's problem (2.21) anticipating the member-states' policy choices. The values of ϕ induce payout function $\mathbf{B}_F(\cdot;\cdot)$.
(3) Find ϕ by numerical optimization. For this, for each try ϕ, evaluate the federal government's objective function either using unconditional expectations (Sections 2.5.1 to 2.5.2) or conditional expectations for a given initial state (Sections 2.5.3 and 2.5.4).
(4) In order to evaluate the federal planner's objective function for given ϕ, make sure that the scheme is feasible in light of the member states' optimal response. In particular, a federal RI policy has to be self-financing in light of member-states' responses, recall (2.19). For fixed ϕ we iterateas follows.
   (a) mark the iteration by $^{(n)}$. Set $n = 0$. Start from an initial value of $\tau_F^{(-1)}$.
   (b) set $\tau_F = \tau_F^{(n-1)}$. For given federal RI policy ϕ and $\tau_F = \tau_F^{(n-1)}$, let the member state solve (2.23).
   (c) label the maximizing member-state policies $\{\tau_v^{i,(n)}, \tau_\xi^{i,(n)}, b^{i,(n)}\}$. These induce a law of motion $\mu_0^{(n)}$ and a value for the objective function of $\int W_0^{(n)} d\mu_0^{(n)}$.
   (d) for given member-state policies $\{\tau_v^{i,(n)}, \tau_\xi^{i,(n)}, b^{i,(n)}\}$, and given federal policy ϕ, find a value $\tau_F^{(n)}$ that solves the federal RI scheme's financing constraint (2.19) for these given policies and given the induced dynamics for the member-state economies.

(e) if $\tau_F^{(n)}$ is not sufficiently close to $\tau_F^{(n-1)}$, set $n = n+1$ and go to step 4b. Else, set $\tau_F = \tau_F^{(n)}$ and go to step 5.

(5) The federal policy implied by ϕ and τ_F is feasible. Set $\int W_0 d\mu_0 = \int W_0^{(n)} d\mu_0^{(n)}$.

(6) Continue numerical optimization started in step 3 until the maximum for the federal government's objective is found.

2.A.2. Welfare gains from federal UI. The consumption-equivalent welfare gains reported in Table 2.5 are computed as follows. For each policy setting, compute welfare under optimal federal RI as detailed in Section 2.A.1. This gives $\int W_0 d\mu_0$. Follow the program as in Section 2.A.1, but setting $\phi = 0$ ($\mathbf{B}_F(\cdot;\cdot) = 0$) and $\tau_F = 0$. That is, solve for optimal member-state policy in autarky. Denote the induced welfare by $\int W_0^{\text{aut}} d\mu_0^{\text{aut}}$. The consumption equivalent welfare gains are computed as the value of the direct transfer received every period which would make households indifferent between living in autarky and living under the federal RI. The welfare gain is expressed in terms of percent of the steady state consumption level.

2.B. Calculating fourth-order-accurate unconditional first moments

In this section we consider a pruned perturbation solution to a dynamic stochastic general equilibrium (DSGE) model. We derive closed-form solutions for fourth-order-accurate unconditional first moments of the model's endogenous variables. The exposition here heavily builds on Andreasen et al. (2018) who already provide the formulae up to third-order of accuracy. Our sole contribution is to provide formulae for fourth-order moments.

2.B.1. Preliminaries. We consider the following class of DSGE models. Let $y_t \in \mathbb{R}^{n_y}$ be a vector of control variables, $x_t \in \mathbb{R}^{n_x+1}$ a vector of state variables which includes a perturbation parameter $\sigma \geq 0$. Consider a perturbation solution to a DSGE model around the steady state $x_{SS} = 0$. The exact solution to the model is given by

$$\begin{aligned} y_t &= g(x_t), \\ x_{t+1} &= h(x_t) + \sigma\eta\epsilon_{t+1}, \end{aligned} \tag{2.41}$$

where ϵ_{t+1} follows n_ϵ dimensional multivariate normal distribution and is independently and identically distributed in each period. Solving a DSGE model amounts to finding unknown functions g and h.

For most DSGE models, the full solution to system (2.41) cannot be found explicitly. The perturbation solution approximates the true solution using a Taylor series expansion around the steady state, $x_t = x_{t+1} = 0$. Up to fourth order, we have

$$x_{t+1} = h_x x_t + \frac{1}{2} h_{xx} x_t^{\otimes 2} + \frac{1}{6} h_{xxx} x_t^{\otimes 3} + \frac{1}{24} h_{xxxx} x_t^{\otimes 4} + \sigma\eta\varepsilon_{t+1}, \tag{2.42}$$

where $h_x, h_{xx}, \ldots$ denote first, second, etc, order derivatives of function h with respect to vector x. Superscript $^{\otimes n}$ represents the n-th Kronecker power, i.e. $x^{\otimes n} = \overbrace{x \otimes x \otimes \ldots}^{n \text{ times}}$.

However, the system (2.42) may display explosive dynamics and may not have any finite unconditional moments (Andreasen et al. 2018) . The solution to this problem suggested by Kim et al. (2008) is to prune the state space of the approximated solution so as to remove explosive paths. As shown by Andreasen et al. (2018) the pruned 4th-order approximation to the perturbation solution reads

$$x_{t+1} = x_{t+1}^f + x_{t+1}^s + x_{t+1}^{rd} + x_{t+1}^{4th},$$

where

$$\begin{aligned} x_{t+1}^f &= h_x x_t^f + \sigma\eta\varepsilon_{t+1}, \\ x_{t+1}^s &= h_x x_t^s + \frac{1}{2} h_{xx} \left(x_t^f\right)^{\otimes 2}, \\ x_{t+1}^{rd} &= h_x x_t^{rd} + \frac{1}{2} h_{xx} \left(2\left((x_t^f \otimes x_t^s)\right)\right) + \frac{1}{6} h_{xxx} \left(x_t^f\right)^{\otimes 3}, \end{aligned}$$

and

$$x_{t+1}^{4th} = h_x x_t^{4th} + \frac{1}{2} h_{xx} \left(2\left(x_t^f \otimes x_t^{rd}\right) + \left(x_t^s\right)^{\otimes 2}\right) + \frac{1}{6} h_{xxx} \left(3\left(x_t^f\right)^{\otimes 2} \otimes x_t^s\right) + \frac{1}{24} h_{xxxx} \left(x^f\right)^{\otimes 4}.$$

Note that if the shock is drawn from the standard normal distribution, as is the case in the model developed in the current paper, then

$$\mathbb{E}(\varepsilon_t)^{\otimes 2} = vec(I_{n_e}),$$
$$\mathbb{E}(\varepsilon_t)^{\otimes 3} = 0, \text{and}$$
$$\mathbb{E}(\varepsilon_t)^{\otimes 5} = 0.$$

Let $M^4 \equiv \mathbb{E}(\varepsilon_t)^{\otimes 4}$ be the kurtosis of the standard multivariate normal distribution.

In the course of the proofs we will use extensively the following (well-known) properties of the Kronecker product (for example, Magnus and Neudecker 1999). These are:

$$\mathbf{A} \otimes (\mathbf{B} + \mathbf{C}) = \mathbf{A} \otimes \mathbf{B} + \mathbf{A} \otimes \mathbf{C},$$
$$(\mathbf{A} + \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes \mathbf{C} + \mathbf{B} \otimes \mathbf{C},$$
$$(k\mathbf{A}) \otimes \mathbf{B} = \mathbf{A} \otimes (k\mathbf{B}) = k(\mathbf{A} \otimes \mathbf{B}),$$
$$(\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C}),$$
$$(\mathbf{AC}) \otimes (\mathbf{BD}) = (\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}),$$
$$vec(ABC) = (C' \otimes A)vec(B).$$

We say that matrix $K_{m,n}$ of size $mn \times mn$ is a *commutation matrix* if it has the following property: Let A be $(m \times n)$ matrix and B is $(p \times q)$ matrix. Then

$$K_{m,p}(A \otimes B)K_{q,n} = B \otimes A.$$

That is, the commutation matrix reverses the order of Kronecker product. The commutation matrix $K_{m,n}$ can be defined explicitly as

$$K_{n,m} = \sum_{i=1}^{m}\sum_{j=1}^{n}\left((e_i^m(e_j^n)') \otimes (e_j^n(e_i^m)')\right),$$

where e_i^m is the ith unit column vector of order m. For any commutation matrix $K_{p,q} = K_{q,p}^{-1}$.

2.B.2. Analytical expressions for the first moments. We are ready to derive formulas for unconditional first moments of the endogenous variables. Our goal is to characterize the following expression

$$\mathbb{E}_0 x_t = \mathbb{E}_0 x_t^f + \mathbb{E}_0 x_t^s + \mathbb{E}_0 x_t^{rd} + \mathbb{E}_0 x_t^{4th},$$

Andreasen et al. (2018) showed that $\mathbb{E}_0 x_t^f = \mathbb{E}_0 x_t^{rd} = 0$. The first equality is the certainty equivalence of the linear approximation. The second equality, $\mathbb{E}_0 x_t^{rd} = 0$, results from the symmetry of the normal distribution (i.e. skewness is zero).

We write perturbation parameter σ as a separate variable, not included in state x_t. Note that $h_\sigma = h_{x\sigma} = h_{xx\sigma} = h_{xxx\sigma} = h_{xxxx\sigma} = 0$ (see, for example, Theorem 7 in Jin and Judd (2002)).

For completeness, we derive expressions for the unconditional first moments for the solution approximations of all orders from one up to four. Derivations for orders of approximation up to three are based on Andreasen et al. (2018). Formulas for the fourth order are our contribution.

2.B.2.1. *First and second-order of approximation.* We start with the formulae accurate up to the second order. We have

$$x^f_{t+1} = h_x x^f_t + \sigma\eta\varepsilon_{t+1}$$
$$x^s_{t+1} = h_x x^s_t + \frac{1}{2}h_{xx}\left(x^f_t\right)^{\otimes 2} + \frac{1}{2}h_{\sigma\sigma}\sigma^2$$

$$\mathbb{E}x^f_t = 0$$
$$\mathbb{E}\left(x^f_t\right)^{\otimes 2} = (I_{n_x^2} - h_x \otimes h_x)^{-1}(\sigma^2\eta\otimes\eta)vec(I))$$
$$\mathbb{E}x^s_t = (I_{n_x} - h_x)^{-1}\left[\frac{1}{2}h_{xx}(I_{n_x^2} - h_x\otimes h_x)^{-1}(\sigma^2\eta\otimes\eta)vec(I)) + \frac{1}{2}h_{\sigma\sigma}\sigma^2\right]$$

PROOF.

$$x^f_{t+1} = h_x x^f_t + \sigma\eta\epsilon_{t+1}$$
$$\mathbb{E}x^f_{t+1} = \mathbb{E}h_x x^f_t$$

(stationarity)

$$\mathbb{E}x^f_t(I - h_x) = 0 \Rightarrow \mathbb{E}x^f_t = 0.$$

$$\left(x^f_t\right)^{\otimes 2} = \left(h_x x^f_t + \sigma\eta\epsilon_{t+1}\right)\otimes\left(h_x x^f_t + \sigma\eta\epsilon_{t+1}\right)$$
$$= h_x x^f_t \otimes h_x x^f_t + (h_x x^f_t)\otimes(\sigma\eta\epsilon_{t+1}) + (\sigma\eta\epsilon_{t+1})\otimes(h_x x^f_t) + (\sigma\eta\epsilon_{t+1})\otimes(\sigma\eta\epsilon_{t+1})$$
$$\mathbb{E}\left(x^f_t\right)^{\otimes 2} = h_x\otimes h_x\mathbb{E}\left(x^f_t\right)^{\otimes 2} + \sigma^2\eta\otimes\eta\mathbb{E}(\epsilon_{t+1})^{\otimes 2}$$
$$\mathbb{E}\left(x^f_t\right)^{\otimes 2} = (I - h_x\otimes h_x)^{-1}(\sigma^2\eta\otimes\eta)vec(I))$$

$$x^s_{t+1} = h_x x^s_t + \frac{1}{2}h_{xx}\left(x^f_t\right)^{\otimes 2} + \frac{1}{2}h_{\sigma\sigma}\sigma^2$$
$$\mathbb{E}x^s_t = h_x\mathbb{E}x^s_t + \frac{1}{2}h_{xx}\mathbb{E}\left(x^f_t\right)^{\otimes 2} + \frac{1}{2}h_{\sigma\sigma}\sigma^2$$
$$\mathbb{E}x^s_t = (I - h_x)^{-1}\left[\frac{1}{2}h_{xx}(I - h_x\otimes h_x)^{-1}(\sigma^2\eta\otimes\eta)vec(I)) + \frac{1}{2}h_{\sigma\sigma}\sigma^2\right].$$

□

2.B.2.2. *Third order.* Next, we tackle the third-order approximation. We replicate the results in Andreasen et al. (2018) in the following

$$x^{rd}_{t+1} = h_x x^{rd}_t + \frac{1}{2}h_{xx}\left(2\left((x^f_t\otimes x^s_t)\right)\right) + \frac{1}{6}h_{xxx}\left(x^f_t\right)^{\otimes 3} + 3\cdot\frac{1}{6}h_{x\sigma\sigma}x^f_t\sigma^2 + \frac{1}{6}h_{\sigma\sigma\sigma}\sigma^3$$

$$
\begin{aligned}
\mathbb{E}\, x_{t+1}^{f} \otimes x_{t+1}^{s} &= \left(I - h_x^{\otimes 2}\right)^{-1} \left(h_x \otimes \frac{1}{2} h_{xx}\right) \mathbb{E}\left(x_t^{f}\right)^{\otimes 3} \\
\mathbb{E}\left(x_{t+1}^{f}\right)^{\otimes 3} &= 0 \\
\mathbb{E}\, x_{t+1}^{rd} &= 0.
\end{aligned}
$$

PROOF.

$$x^f_{t+1} \otimes x^s_{t+1} = \left(h_x x^f_t + \sigma\eta\epsilon_{t+1}\right) \otimes \left(h_x x^s_t + \frac{1}{2} h_{xx} \left(x^f_t\right)^{\otimes 2} + \frac{1}{2} h_{\sigma\sigma}\sigma^2\right)$$

$$= (h_x \otimes h_x)\left(x^f_t \otimes x^s_t\right) + \left(h_x \otimes \frac{1}{2} h_{xx}\right)\left(x^f_t\right)^{\otimes 3} + (h_x x_f) \otimes \frac{1}{2} h_{\sigma\sigma}\sigma^2$$

$$+ \left(\sigma\eta \otimes h_x\right)\left(\epsilon_{t+1} \otimes x^s_t\right) + \left(\sigma\eta \otimes .5 h_{xx}\right)\left(\epsilon_{t+1} \otimes \left(x^f_t\right)^{\otimes 2}\right) + (\eta\epsilon_{t+1}) \otimes \frac{1}{2} h_{\sigma\sigma}\sigma^3.$$

$$\mathbb{E}(x^f_t \otimes x^s_t) = \left(I - h_x^{\otimes 2}\right)^{-1}\left(h_x \otimes \frac{1}{2} h_{xx}\right)\mathbb{E}\left(x^f_t\right)^{\otimes 3}.$$

$$\left(x^f_{t+1}\right)^{\otimes 3} = \left(x^f_{t+1}\right)^{\otimes 2} \otimes \left(h_x x^f_t + \sigma\eta\epsilon_{t+1}\right)$$

$$= \left(h_x x^f_t + \sigma\eta\epsilon_{t+1}\right) \otimes \Big[(h_x \otimes h_x)\left(\left(x^f_t\right)^{\otimes 2}\right) + \left(\sigma\eta \otimes \sigma\eta\right)\left(\epsilon^{\otimes 2}_{t+1}\right)$$

$$+ \left(\sigma\eta \otimes h_x\right)\left(\epsilon_{t+1} \otimes x^f_t\right) + \left(h_x \otimes \sigma\eta\right)\left(x^f_t \otimes \epsilon_{t+1}\right)\Big]$$

$$= (h_x \otimes h_x \otimes h_x)\left(x^f_t\right)^{\otimes 3} + (\sigma\eta\epsilon_{t+1}) \otimes (h_x \otimes h_x)\left(\left(x^f_t\right)^{\otimes 2}\right)$$

$$+ (h_x x^f_t) \otimes \left(\sigma\eta \otimes \sigma\eta\right)\left(\epsilon^{\otimes 2}_{t+1}\right) + (\sigma\eta\epsilon_{t+1}) \otimes \left(\sigma\eta \otimes \sigma\eta\right)\left(\epsilon^{\otimes 2}_{t+1}\right)$$

$$+ (h_x x^f_t) \otimes \left(\sigma\eta \otimes h_x\right)\left(\epsilon_{t+1} \otimes x^f_t\right) + (\sigma\eta\epsilon_{t+1}) \otimes \left(\sigma\eta \otimes h_x\right)\left(\epsilon_{t+1} \otimes x^f_t\right)$$

$$+ (h_x x^f_t) \otimes \left(h_x \otimes \sigma\eta\right)\left(x^f_t \otimes \epsilon_{t+1}\right) + (\sigma\eta\epsilon_{t+1}) \otimes \left(h_x \otimes \sigma\eta\right)\left(x^f_t \otimes \epsilon_{t+1}\right)$$

$$= (h_x \otimes h_x \otimes h_x)\left(x^f_t\right)^{\otimes 3} + \text{ terms zero in expectation.}$$

The last equality follows from the fact that x^f_t and ϵ_{t+1} are independent, therefore $\mathbb{E}\left(x^f_t \otimes \epsilon_{t+1}\right) = \mathbb{E}x^f_t \otimes \mathbb{E}\epsilon_{t+1} = 0$, $\mathbb{E}(h_x x^f_t) \otimes \left(\sigma\eta \otimes \sigma\eta\right)\left(\epsilon^{\otimes 2}_{t+1}\right) = 0$ since $\mathbb{E}(h_x x^f_t) = 0$.

Therefore,

$$\mathbb{E}\left(\left(x^f_{t+1}\right)^{\otimes 3}\right) = \mathbb{E}\,(h_x \otimes h_x \otimes h_x)\left(x^f_t\right)^{\otimes 3}$$

and by stationarity $\mathbb{E}\,(h_x \otimes h_x \otimes h_x)\left(x^f_t\right)^{\otimes 3} = 0$,

$$x^{rd}_{t+1} = h_x x^{rd}_t + \frac{1}{2} h_{xx}\left(2\left((x^f_t \otimes x^s_t)\right)\right) + \frac{1}{6} h_{xxx}\left(x^f_t\right)^{\otimes 3} + 3 \cdot \frac{1}{6} h_{x\sigma\sigma} x^f_t \sigma^2 + \frac{1}{6} h_{\sigma\sigma\sigma}\sigma^3$$

$$\mathbb{E}\, x^{rd}_{t+1} = h_x \mathbb{E}\, x^{rd}_t + \frac{1}{6} h_{\sigma\sigma\sigma}\sigma^3$$

$$\mathbb{E}\, x^{rd}_t (I - h_x) = 0,$$

the last line following since $h_{\sigma\sigma\sigma} = 0$ for symmetric distribution (see Andreasen, 2012) □

2.B.2.3. *Fourth order.* Finally, we derive the solutions accurate up to fourth order.

We have that the fourth-order accurate law of motion of the states is given by

$$x_{t+1}^{4th} = h_x x_t^{4th} + \frac{1}{2} h_{xx} \left(2\left(x_t^f \otimes x_t^{rd}\right) + (x_t^s)^{\otimes 2}\right) + \frac{1}{6} h_{xxx} \left(3\left(x_t^f\right)^{\otimes 2} \otimes x_t^s\right) + \frac{1}{24} h_{xxxx} \left(x^f\right)^{\otimes 4}$$
$$+ \frac{3}{6} h_{\sigma\sigma x} \sigma^2 x_t^s + 6 \cdot \frac{1}{24} h_{\sigma\sigma xx} \sigma^2 \left(x_t^f\right)^{\otimes 2} + 4 \cdot \frac{1}{24} h_{\sigma\sigma\sigma x} \sigma^3 x_t^f + \frac{1}{24} h_{\sigma\sigma\sigma\sigma} \sigma^4.$$

The fourth raw moment of the state variables is given by.

$$\mathbb{E} x_t^{4th} = \left(I_{n_x} - h_x\right)^{-1} \Big[\frac{1}{2} h_{xx} \mathbb{E}\left(2\left(x_t^f \otimes x_t^{rd}\right) + (x_t^s)^{\otimes 2}\right)$$
$$+ \frac{1}{6} h_{xxx} \mathbb{E}\left(3\left(x_t^f\right)^{\otimes 2} \otimes x_t^s\right) + \frac{1}{24} h_{xxxx} \mathbb{E}\left(x^f\right)^{\otimes 4}$$
$$+ \frac{3}{6} h_{\sigma\sigma x} \sigma^2 \mathbb{E} x_t^s + 6 \cdot \frac{1}{24} h_{\sigma\sigma xx} \sigma^2 \mathbb{E}\left(x_t^f\right)^{\otimes 2} + \frac{1}{24} h_{\sigma\sigma\sigma\sigma} \sigma^4\Big].$$

Where the respective terms for each basis vector of the 4th-order pruned state space are listend in the following.

$$\mathbb{E}\left(x_{t+1}^f\right)^{\otimes 4} = \sigma^2 \left(I_{n_x^4} - h_x^{\otimes 4}\right)^{-1} \Big[\sigma^2 \eta^{\otimes 4} M^4 + \Big((h_x^{\otimes 2} \otimes \eta^{\otimes 2}) K_{n_e^2, n_x^2}$$
$$+ (h_x \otimes \eta \otimes h_x \otimes \eta)\left(I_{n_x} \otimes K_{n_x, n_e} \otimes I_{n_e}\right) K_{n_e^2, n_x^2}$$
$$+ (h_x \otimes \eta \otimes \eta \otimes h_x)\left(I_{n_x} \otimes K_{n_x, n_e^2}\right) K_{n_e^2, n_x^2}$$
$$+ (\eta \otimes h_x \otimes h_x \otimes \eta)\left(I_{n_e} \otimes K_{n_e, n_x^2}\right)$$
$$+ (\eta \otimes h_x \otimes \eta \otimes h_x)\left(I_{n_e} \otimes K_{n_e, n_x} \otimes I_{n_x}\right)$$
$$+ \left(\eta^{\otimes 2} \otimes h_x^{\otimes 2}\right)\Big)\left(vec(I_{n_e}) \otimes \mathbb{E}(x_t^f)^{\otimes 2}\right)\Big].$$

This we can calculate given $\mathbb{E}\left(x_t^f\right)^{\otimes 2} = \sigma^2 (I - h_x^{\otimes 2})^{-1} (\eta^{\otimes 2}) vec(I_{n_e}))$ from further above.

Regarding the remaining vectors that span the 4th-order pruned state space, we have

$$\mathbb{E}\left[(x_{t+1}^s)^{\otimes 2}\right] = \left(I_{n_x^2} - h_x^{\otimes 2}\right)^{-1} \Big(+.5\left(K_{n_x, n_x} + I_{n_x^2}\right)(h_{xx} \otimes h_x) \mathbb{E}\left(\left(x_t^f\right)^{\otimes 2} \otimes x_t^s\right) + \frac{1}{4} h_{xx}^{\otimes 2} \mathbb{E}(x_t^f)^{\otimes 4}$$
$$+ \left(K_{n_x, n_x} + I_{n_x^2}\right)\left[\frac{1}{2}[(h_x \otimes h_{\sigma\sigma})(\sigma^2 x_t^s)] + \frac{1}{4}(h_{xx} \otimes h_{\sigma\sigma})(\sigma^2 (x_t^f)^{\otimes 2})\right] + \frac{1}{4} h_{\sigma\sigma} \otimes h_{\sigma\sigma} \sigma^4\Big),$$

$$\mathbb{E}\left[\left(x_t^f\right)^{\otimes 2}\otimes x_t^s\right] = \left(I_{n_x^3} - h_x^{\otimes 3}\right)^{-1}\Big((\sigma^2\eta^{\otimes 2}\otimes h_x)(\mathbb{E}\left[\epsilon_{t+1}^{\otimes 2}\right]\otimes\mathbb{E}\,x_t^s) + \frac{1}{2}(h_x^{\otimes 2}\otimes h_{xx})\mathbb{E}(x_t^f)^{\otimes 4}$$
$$+\frac{1}{2}(\sigma^2\eta^{\otimes 2}\otimes h_{xx})(vec(I_{n_e})\otimes\mathbb{E}(x_t^f)^{\otimes 2}) + \frac{1}{2}(h_x^{\otimes 2}\otimes h_{\sigma\sigma})\sigma^2\,\mathbb{E}\left[(x_t^f)^{\otimes 2}\right]$$
$$+\frac{1}{2}(\eta^{\otimes 2}\otimes h_{\sigma\sigma})\sigma^4\,\mathbb{E}\epsilon_{t+1}^{\otimes 2}\Big),$$

$$\mathbb{E}\left[x_{t+1}^f\otimes x_{t+1}^{rd}\right] = \left(I_{n_x^2} - h_x^{\otimes 2}\right)^{-1}\Big((h_x\otimes h_{xx})\mathbb{E}((x_t^f)^{\otimes 2}\otimes x_t^s) + \frac{1}{6}(h_x\otimes h_{xxx})\mathbb{E}\left(x_t^f\right)^{\otimes 4}$$
$$+\frac{3}{6}(h_x\otimes h_{\sigma\sigma})\sigma^2\,\mathbb{E}\left[(x_t^f)^{\otimes 2}\right]\Big).$$

PROOF. We derive each of the terms.

$$\begin{aligned}\left(x_{t+1}^f\right)^{\otimes 4} &= \left(h_x x_t^f + \sigma\eta\epsilon_{t+1}\right)\otimes\left(x_{t+1}^f\right)^{\otimes 3}\\ &= (h_x x_t^f)\otimes\left(x_{t+1}^f\right)^{\otimes 3} + (\sigma\eta\epsilon_{t+1})\otimes\left(x_{t+1}^f\right)^{\otimes 3}.\end{aligned}\tag{2.43}$$

We tackle separately each of the summands in (2.43).

$$\begin{aligned}(h_x x_t^f)\otimes\left(x_{t+1}^f\right)^{\otimes 3} &= (h_x x_t^f)\otimes\left(\left[h_x x_t^f + \sigma\eta\epsilon_{t+1}\right]\otimes\left[h_x x_t^f + \sigma\eta\epsilon_{t+1}\right]\otimes\left[h_x x_t^f + \sigma\eta\epsilon_{t+1}\right]\right)\\ &= (h_x x_t^f)\otimes\Big(\left[h_x x_t^f + \sigma\eta\epsilon_{t+1}\right]\\ &\quad\otimes\left[h_x x_t^f\otimes h_x x_t^f + h_x x_t^f\otimes\sigma\eta\epsilon_{t+1} + \sigma\eta\epsilon_{t+1}\otimes h_x x_t^f + \sigma\eta\epsilon_{t+1}\otimes\sigma\eta\epsilon_{t+1}\right]\Big)\\ &= (h_x x_t^f)\otimes\Big(\left[h_x x_t^f + \sigma\eta\epsilon_{t+1}\right]\\ &\quad\otimes\left[h_x^{\otimes 2}(x_t^f)^{\otimes 2} + \sigma(h_x\otimes\eta)(x_t^f\otimes\epsilon_{t+1}) + \sigma(\eta\otimes h_x)(\epsilon_{t+1}\otimes x_t^f) + \sigma^2\eta^{\otimes 2}\epsilon_{t+1}^{\otimes 2}\right]\Big).\end{aligned}$$

Further,

$$\begin{aligned}(h_x x_t^f)\otimes\left(x_{t+1}^f\right)^{\otimes 3} &= (h_x x_t^f)\otimes\Big(h_x^{\otimes 3}(x_t^f)^{\otimes 3} + \sigma(h_x^{\otimes 2}\otimes\eta)((x_t^f)^2\otimes\epsilon_{t+1})\\ &\quad+\sigma(h_x\otimes\eta\otimes h_x)(x_t^f\otimes\epsilon_{t+1}\otimes x_t^f) + \sigma^2(h_x\otimes\eta^{\otimes 2})(x_t^f\otimes\epsilon_{t+1}^{\otimes 2})\\ &\quad+\sigma(\eta\otimes h_x^{\otimes 2})(\epsilon_{t+1}\otimes(x_t^f)^{\otimes 2}) + \sigma^2(\eta\otimes h_x\otimes\eta)(\epsilon_{t+1}\otimes x_t^f\otimes\epsilon_{t+1})\\ &\quad+\sigma^2(\eta^{\otimes 2}\otimes h_x)(\epsilon_{t+1}^{\otimes 2}\otimes x_t^f) + \sigma^3\eta^{\otimes 3}\epsilon_{t+1}^{\otimes 3}\Big).\end{aligned}$$

So that

$$(h_x x_t^f) \otimes \left(x_{t+1}^f\right)^{\otimes 3} = h_x^{\otimes 4}(x_t^f)^{\otimes 4} + \sigma^2 (h_x^{\otimes 2} \otimes \eta^{\otimes 2})((x_t^f)^{\otimes 2} \otimes \epsilon_{t+1}^{\otimes 2})$$
$$+ \sigma^2 \left(h_x \otimes \eta \otimes h_x \otimes \eta\right)\left(x_t^f \otimes \epsilon_{t+1} \otimes x_t^f \otimes \epsilon_{t+1}\right)$$
$$+ \sigma^2 \left(h_x \otimes \eta^{\otimes 2} \otimes h_x\right)\left(x_t^f \otimes \epsilon_{t+1}^{\otimes 2} \otimes x_t^f\right) + \text{ terms zero in expectation.} \tag{2.44}$$

The last equality follows since $\epsilon_{t+1} \perp x_t^f$ and $\mathbb{E} x_t^f = \mathbb{E}\left((x_t^f)^{\otimes 3}\right)\mathbb{E}\epsilon_{t+1} = 0$.

Using the commutation matrix K_{n_x,n_e} to change the order of $\otimes$. For instance

$$\begin{aligned} x_t^f \otimes \epsilon_{t+1} \otimes \epsilon_{t+1} \otimes x_t^f &= x_t^f \otimes I_{n_e}\epsilon_{t+1} \otimes K_{n_x,n_e}(x_t^f \otimes \epsilon_{t+1}) \\ &= x_t^f \otimes (I_{n_e} \otimes K_{n_x,n_e})(\epsilon_{t+1} \otimes x_t^f \otimes \epsilon_{t+1}) \\ &= x_t^f \otimes (I_{n_e} \otimes K_{n_x,n_e})(K_{n_x,n_e}(x_t^f \otimes \epsilon_{t+1}) \otimes \epsilon_{t+1}) \\ &= x_t^f \otimes (I_{n_e} \otimes K_{n_x,n_e})((K_{n_x,n_e} \otimes I_{n_e})(x_t^f \otimes \epsilon_{t+1} \otimes \epsilon_{t+1})) \\ &= \left[I_{n_x} \otimes (I_{n_e} \otimes K_{n_x,n_e})(K_{n_x,n_e} \otimes I_{n_e})\right]\left[(x_t^f)^{\otimes 2} \otimes \epsilon_{t+1}^{\otimes 2}\right]. \end{aligned}$$

This can be simplified further using $(I_r \otimes K_{m,s})(K_{m,s} \otimes I_s) = K_{m,rs}$. Thus

$$\mathbb{E}\left[x_t^f \otimes \epsilon_{t+1} \otimes \epsilon_{t+1} \otimes x_t^f\right] = \left(I_{n_x} \otimes K_{n_x,n_e^2}\right)\mathbb{E}\left((x_t^f)^{\otimes 2} \otimes \epsilon_{t+1}^{\otimes 2}\right).$$

Similarly,

$$\mathbb{E}\left(x_t^f \otimes \epsilon_{t+1} \otimes x_t^f \otimes \epsilon_{t+1}\right) = \left(I_{n_x} \otimes K_{n_x,n_e} \otimes I_{n_e}\right)\mathbb{E}\left((x_t^f)^{\otimes 2} \otimes \epsilon_{t+1}^{\otimes 2}\right).$$

Going back to (2.44) we have

$$\begin{aligned} \mathbb{E}\left[(h_x x_t^f) \otimes \left(x_{t+1}^f\right)^{\otimes 3}\right] = h_x^{\otimes 4}\,\mathbb{E}(x_t^f)^{\otimes 4} + \sigma^2 \Big[&(h_x^{\otimes 2} \otimes \eta^{\otimes 2}) \\ &+ (h_x \otimes \eta \otimes h_x \otimes \eta)\left(I_{n_x} \otimes K_{n_x,n_e} \otimes I_{n_e}\right) \\ &+ (h_x \otimes \eta \otimes \eta \otimes h_x)\left(I_{n_x} \otimes K_{n_x,n_e^2}\right)\Big]\left(\mathbb{E}\left[(x_t^f)^{\otimes 2}\right] \otimes vec(I_{n_e})\right), \end{aligned}$$

since $E\left((x_t^f)^{\otimes 2} \otimes \epsilon_{t+1}^{\otimes 2}\right) = \mathbb{E}\left[(x_t^f)^{\otimes 2}\right] \otimes vec(I_{n_e})$.

Regarding the other summand in (2.43),

$$
\begin{aligned}
(\sigma\eta\epsilon_{t+1})\otimes\left(x^f_{t+1}\right)^{\otimes 3} &= (\sigma\eta\epsilon_{t+1})\otimes\Big(h_x^{\otimes 3}(x^f_t)^{\otimes 3}+\sigma(h_x^{\otimes 2}\otimes\eta)((x^f_t)^2\otimes\epsilon_{t+1})\\
&\quad +\sigma(h_x\otimes\eta\otimes h_x)(x^f_t\otimes\epsilon_{t+1}\otimes x^f_t)+\sigma^2(h_x\otimes\eta^{\otimes 2})(x^f_t\otimes\epsilon^{\otimes 2}_{t+1})\\
&\quad +\sigma(\eta\otimes h_x^{\otimes 2})(\epsilon_{t+1}\otimes(x^f_t)^{\otimes 2})+\sigma\eta\epsilon_{t+1}\otimes\sigma(h_x\otimes\eta)(x^f_t\otimes\epsilon_{t+1})\\
&\quad +\sigma\eta\epsilon_{t+1}\otimes\sigma(\eta\otimes h_x)(\epsilon_{t+1}\otimes x^f_t)+\sigma^3\eta^{\otimes 3}\epsilon^{\otimes 3}_{t+1}\Big)\\
&= \sigma^2\left(\eta\otimes h_x\otimes h_x\otimes\eta\right)\left(\epsilon_{t+1}\otimes x^f_t\otimes x^f_t\otimes\epsilon_{t+1}\right)\\
&\quad +\sigma^2\left(\eta\otimes h_x\otimes\eta\otimes h_x\right)\left(\epsilon_{t+1}\otimes x^f_t\otimes\epsilon_{t+1}\otimes x^f_t\right)\\
&\quad +\sigma^2\left(\eta^{\otimes 2}\otimes h_x^{\otimes 2}\right)\left(\epsilon^{\otimes 2}_{t+1}\otimes(x^f_t)^{\otimes 2}\right)+\sigma^4\eta^{\otimes 4}\epsilon^{\otimes 4}_{t+1}\\
&\quad +\text{ terms that are zero in expectation.}
\end{aligned}
\tag{2.45}
$$

Using the commutation matrices

$$
\begin{aligned}
\mathbb{E}\left(\epsilon^{\otimes 2}_{t+1}\otimes(x^f_t)^{\otimes 2}\right) &= vec(I_{n_e})\otimes\mathbb{E}(x^f_t)^{\otimes 2}\\
\mathbb{E}\left(\epsilon_{t+1}\otimes x^f_t\otimes\epsilon_{t+1}\otimes x^f_t\right) &= \left(I_{n_e}\otimes K_{n_e,n_x}\otimes I_{n_x}\right)\left(vec(I_{n_e})\otimes\mathbb{E}(x^f_t)^{\otimes 2}\right)\\
\mathbb{E}\left(\epsilon_{t+1}\otimes x^f_t\otimes x^f_t\otimes\epsilon_{t+1}\right) &= (I_{n_e}\otimes K_{n_e,n_x^2})(vec(I_{n_e})\otimes\mathbb{E}(x^f_t)^{\otimes 2}).
\end{aligned}
$$

Plugging into (2.45) delivers

$$
\begin{aligned}
\mathbb{E}\left[(\sigma\eta\epsilon_{t+1})\otimes\left(x^f_{t+1}\right)^{\otimes 3}\right] &= \sigma^2\Big[\left(\eta\otimes h_x\otimes h_x\otimes\eta\right)\left(I_{n_e}\otimes K_{n_e,n_x^2}\right)\\
&\quad +\left(\eta\otimes h_x\otimes\eta\otimes h_x\right)\left(I_{n_e}\otimes K_{n_e,n_x}\otimes I_{n_x}\right)\\
&\quad +\left(\eta^{\otimes 2}\otimes h_x^{\otimes 2}\right)\Big]\left(vec(I_{n_e})\otimes\mathbb{E}(x^f_t)^{\otimes 2}\right)+\sigma^4\eta^{\otimes 4}M^4,
\end{aligned}
$$

where $M^4\equiv\mathbb{E}\left[\epsilon^{\otimes 4}_{t+1}\right]$.

Going back to the original formula (2.43),

$$
\begin{aligned}
\mathbb{E}\left(x^f_{t+1}\right)^{\otimes 4} &= \mathbb{E}\left[(h_x x^f_t)\otimes\left(x^f_{t+1}\right)^{\otimes 3}\right] + \mathbb{E}\left[(\sigma\eta\epsilon_{t+1})\otimes\left(x^f_{t+1}\right)^{\otimes 3}\right] \\
&= h_x^{\otimes 4}\,\mathbb{E}(x^f_t)^{\otimes 4} + \sigma^2\Big[(h_x^{\otimes 2}\otimes\eta^{\otimes 2}) \\
&\quad + (h_x\otimes\eta\otimes h_x\otimes\eta)\left(I_{n_x}\otimes K_{n_x,n_e}\otimes I_{n_e}\right) \\
&\quad + (h_x\otimes\eta\otimes\eta\otimes h_x)\left(I_{n_x}\otimes K_{n_x,n_e^2}\right)\Big]\left(\mathbb{E}\left[(x^f_t)^{\otimes 2}\right]\otimes vec(I_{n_e})\right) \\
&+\sigma^2\Big[\left(\eta\otimes h_x\otimes h_x\otimes\eta\right)\left(I_{n_e}\otimes K_{n_e,n_x^2}\right) \\
&\quad + \left(\eta\otimes h_x\otimes\eta\otimes h_x\right)\left(I_{n_e}\otimes K_{n_e,n_x}\otimes I_{n_x}\right) \\
&\quad + \left(\eta^{\otimes 2}\otimes h_x^{\otimes 2}\right)\Big]\left(vec(I_{n_e})\otimes\mathbb{E}(x^f_t)^{\otimes 2}\right) + \sigma^4\eta^{\otimes 4}M^4.
\end{aligned}
$$

Hence, using stationarity and the fact that $\mathbb{E}\left[(x^f_t)^{\otimes 2}\right]\otimes vec(I_{n_e}) = K_{n_e^2,n_x^2}\left(vec(I_{n_e})\otimes\mathbb{E}(x^f_t)^{\otimes 2}\right)$, we have that

$$
\begin{aligned}
\mathbb{E}\left(x^f_{t+1}\right)^{\otimes 4} = \sigma^2\left(I_{n_x^4}-h_x^{\otimes 4}\right)^{-1}\Big[\sigma^2\eta^{\otimes 4}M^4 + &\Big((h_x^{\otimes 2}\otimes\eta^{\otimes 2})K_{n_e^2,n_x^2} \\
&+ (h_x\otimes\eta\otimes h_x\otimes\eta)\left(I_{n_x}\otimes K_{n_x,n_e}\otimes I_{n_e}\right)K_{n_e^2,n_x^2} \\
&+ (h_x\otimes\eta\otimes\eta\otimes h_x)\left(I_{n_x}\otimes K_{n_x,n_e^2}\right)K_{n_e^2,n_x^2} \\
&+ \left(\eta\otimes h_x\otimes h_x\otimes\eta\right)\left(I_{n_e}\otimes K_{n_e,n_x^2}\right) \\
&+ \left(\eta\otimes h_x\otimes\eta\otimes h_x\right)\left(I_{n_e}\otimes K_{n_e,n_x}\otimes I_{n_x}\right) \\
&+ \left(\eta^{\otimes 2}\otimes h_x^{\otimes 2}\right)\Big)\left(vec(I_{n_e})\otimes\mathbb{E}(x^f_t)^{\otimes 2}\right)\Big].
\end{aligned}
$$

Continued on next page.

Next, we calculate $\mathbb{E}\left(\left(x^s_{t+1}\right)^{\otimes 2}\right)$.

$$
\begin{aligned}
\left(x^s_{t+1}\right)^{\otimes 2} &= \left(h_x x^s_t + \frac{1}{2}h_{xx}\left(x^f_t\right)^{\otimes 2} + \frac{1}{2}h_{\sigma\sigma}\sigma^2\right) \otimes \left(h_x x^s_t + \frac{1}{2}h_{xx}\left(x^f_t\right)^{\otimes 2} + \frac{1}{2}h_{\sigma\sigma}\sigma^2\right) \\
&= h_x^{\otimes 2}(x^s_t)^{\otimes 2} + (h_x x^s_t) \otimes (\frac{1}{2}h_{xx}\left(x^f_t\right)^{\otimes 2}) + (\frac{1}{2}h_{xx}\left(x^f_t\right)^{\otimes 2}) \otimes (h_x x^s_t) \\
&\quad + \left(\frac{1}{2}h_{xx}\left(x^f_t\right)^{\otimes 2}\right) \otimes \left(\frac{1}{2}h_{xx}\left(x^f_t\right)^{\otimes 2}\right) \\
&\quad + \frac{1}{2}(h_{\sigma\sigma} \otimes h_x)(\sigma^2 x^s_t) + \frac{1}{4}(h_{\sigma\sigma} \otimes h_{xx})(\sigma^2 (x^f_t)^{\otimes 2}) + \frac{1}{4}h_{\sigma\sigma} \otimes h_{\sigma\sigma}\sigma^4 \\
&\quad + \frac{1}{2}(h_x \otimes h_{\sigma\sigma})(\sigma^2 x^s_t) + \frac{1}{4}(h_{xx} \otimes h_{\sigma\sigma})(\sigma^2 (x^f_t)^{\otimes 2}) \\
&= h_x^{\otimes 2}(x^s_t)^{\otimes 2} + (\frac{1}{2}h_x \otimes h_{xx})\left(x^s_t \otimes \left(x^f_t\right)^{\otimes 2}\right) + \left(\frac{1}{2}h_{xx} \otimes h_x\right)\left(\left(x^f_t\right)^{\otimes 2} \otimes x^s_t\right) \\
&\quad + \frac{1}{4}h_{xx}^{\otimes 2}(x^f_t)^{\otimes 4} + \frac{1}{2}\left(h_x \otimes h_{\sigma\sigma} + h_{\sigma\sigma} \otimes h_x\right)(\sigma^2 x^s_t) \\
&\quad + \frac{1}{4}\left(h_{\sigma\sigma} \otimes h_{xx} + h_{xx} \otimes h_{\sigma\sigma}\right)(\sigma^2 (x^f_t)^{\otimes 2}) + \frac{1}{4}h_{\sigma\sigma} \otimes h_{\sigma\sigma}\sigma^4 \\
&= h_x^{\otimes 2}(x^s_t)^{\otimes 2} + (\frac{1}{2}h_x \otimes h_{xx})K_{n_x^2,n_x}\left(\left(x^f_t\right)^{\otimes 2} \otimes x^s_t\right) + \left(\frac{1}{2}h_{xx} \otimes h_x\right)\left(\left(x^f_t\right)^{\otimes 2} \otimes x^s_t\right) \\
&\quad + \frac{1}{4}h_{xx}^{\otimes 2}(x^f_t)^{\otimes 4} + \frac{1}{2}\left(K_{n_x,n_x} + I_{n_x^2}\right)[(h_x \otimes h_{\sigma\sigma})(\sigma^2 x^s_t)] \\
&\quad + \frac{1}{4}\left(K_{n_x,n_x} + I_{n_x^2}\right)([h_{xx} \otimes h_{\sigma\sigma})(\sigma^2 (x^f_t)^{\otimes 2})] + \frac{1}{4}h_{\sigma\sigma} \otimes h_{\sigma\sigma}\sigma^4 \\
&= h_x^{\otimes 2}(x^s_t)^{\otimes 2} + \frac{1}{2}(h_{xx} \otimes h_x)\left(K_{n_x^2,n_x} + I_{n_x^2}\right)\left(\left(x^f_t\right)^{\otimes 2} \otimes x^s_t\right) + \frac{1}{4}h_{xx}^{\otimes 2}(x^f_t)^{\otimes 4} \\
&\quad + \left(K_{n_x,n_x} + I_{n_x^2}\right)\left[\frac{1}{2}[(h_x \otimes h_{\sigma\sigma})(\sigma^2 x^s_t)] + \frac{1}{4}(h_{xx} \otimes h_{\sigma\sigma})(\sigma^2 (x^f_t)^{\otimes 2})\right] + \frac{1}{4}h_{\sigma\sigma} \otimes h_{\sigma\sigma}\sigma^4.
\end{aligned}
$$

where the last line follows since $(h_x \otimes h_{xx})K_{n_x^2,n_x} = K_{n_x,n_x}(h_{xx} \otimes h_x)$ (Note: $K^{-1}_{p,q} = K_{q,p}$, so $K_{n_x,n_x} = K^{-1}_{n_x,n_x}$ and $K_{1,n} = K_{n,1} = I_n$).

Hence

$$
\begin{aligned}
\mathbb{E}\left[\left(x^s_{t+1}\right)^{\otimes 2}\right] &= \left(I_{n_x^2} - h_x^{\otimes 2}\right)^{-1}\left(+ .5\left(K_{n_x,n_x} + I_{n_x^2}\right)(h_{xx} \otimes h_x)\mathbb{E}\left(\left(x^f_t\right)^{\otimes 2} \otimes x^s_t\right) + \frac{1}{4}h_{xx}^{\otimes 2}\mathbb{E}(x^f_t)^{\otimes 4}\right. \\
&\quad \left. + \left(K_{n_x,n_x} + I_{n_x^2}\right)\left[\frac{1}{2}[(h_x \otimes h_{\sigma\sigma})(\sigma^2 x^s_t)] + \frac{1}{4}(h_{xx} \otimes h_{\sigma\sigma})(\sigma^2 (x^f_t)^{\otimes 2})\right] + \frac{1}{4}h_{\sigma\sigma} \otimes h_{\sigma\sigma}\sigma^4\right).
\end{aligned}
$$

Continued on next page.

Next we calculate $\mathbb{E}\left(\left(x_{t+1}^f\right)^{\otimes 2}\otimes x_{t+1}^s\right)$.

$$\begin{aligned}
\left(x_{t+1}^f\right)^{\otimes 2}\otimes x_{t+1}^s &= \left(h_x x_t^f\otimes h_x x_t^f + (h_x x_t^f)\otimes(\sigma\eta\epsilon_{t+1}) + (\sigma\eta\epsilon_{t+1})\otimes(h_x x_t^f) + (\sigma\eta\epsilon_{t+1})\otimes(\sigma\eta\epsilon_{t+1})\right)\\
&\quad\otimes\left(h_x x_t^s + \frac{1}{2}h_{xx}\left(x_t^f\right)^{\otimes 2} + \frac{1}{2}h_{\sigma\sigma}\sigma^2\right)\\
&= h_x^{\otimes 3}((x_t^f)^{\otimes 2}\otimes x_t^s) + (\sigma^2\eta^{\otimes 2}\otimes h_x)(\epsilon_{t+1}^{\otimes 2}\otimes x_t^s) + \frac{1}{2}(h_x^{\otimes 2}\otimes h_{\sigma\sigma})\sigma^2(x_t^f)^{\otimes 2}\\
&\quad + \frac{1}{2}(\eta^{\otimes 2}\otimes h_{\sigma\sigma})\sigma^4\epsilon_{t+1}^{\otimes 2} + \frac{1}{2}(h_x^{\otimes 2}\otimes h_{xx})(x_t^f)^{\otimes 4} + \frac{1}{2}(\sigma^2\eta^{\otimes 2}\otimes h_{xx})(\epsilon_{t+1}^{\otimes 2}\otimes(x_t^f)^{\otimes 2})\\
&\quad + \text{ terms zero in expectation.}
\end{aligned}$$

Hence

$$\begin{aligned}
\mathbb{E}\left[\left(x_t^f\right)^{\otimes 2}\otimes x_t^s\right] &= \left(I_{n_x^3} - h_x^{\otimes 3}\right)^{-1}\Big((\sigma^2\eta^{\otimes 2}\otimes h_x)(\mathbb{E}\left[(\epsilon_{t+1})^{\otimes 2}\right]\otimes\mathbb{E}\,x_t^s) + \frac{1}{2}(h_x^{\otimes 2}\otimes h_{xx})\,\mathbb{E}(x_t^f)^{\otimes 4}\\
&\quad + \frac{1}{2}(\sigma^2\eta^{\otimes 2}\otimes h_{xx})(vec(I_{n_e})\otimes\mathbb{E}(x_t^f)^{\otimes 2}) + \frac{1}{2}(h_x^{\otimes 2}\otimes h_{\sigma\sigma})\sigma^2\,\mathbb{E}\left[(x_t^f)^{\otimes 2}\right]\\
&\quad + \frac{1}{2}(\eta^{\otimes 2}\otimes h_{\sigma\sigma})\sigma^4\,\mathbb{E}\epsilon_{t+1}^{\otimes 2}\Big),
\end{aligned}$$

where the equality follows since ϵ_{t+1} and x_t^s are orthogonal.

For the last missing terms.

$$\begin{aligned}
x_{t+1}^f\otimes x_{t+1}^{rd} &= \left(h_x x_t^f + \sigma\eta\varepsilon_{t+1}\right)\otimes\Big(h_x x_t^{rd} + \frac{1}{2}h_{xx}\left(2\left((x_t^f\otimes x_t^s)\right)\right) + \frac{1}{6}h_{xxx}\left(x_t^f\right)^{\otimes 3}\\
&\quad + \frac{3}{6}h_{\sigma\sigma x}\sigma^2 x_t^f + \frac{1}{6}h_{\sigma\sigma\sigma}\sigma^3\Big)\\
&= h_x^{\otimes 2}(x_{t+1}^f\otimes x_{t+1}^{rd}) + (h_x\otimes h_{xx})((x_t^f)^{\otimes 2}\otimes x_t^s) + \frac{1}{6}(h_x\otimes h_{xxx})\left(x_t^f\right)^{\otimes 4}\\
&\quad + \frac{3}{6}(h_x\otimes h_{\sigma\sigma})\sigma^2(x_t^f)^{\otimes 2} + \text{ terms zero in expectation}
\end{aligned}$$

so that

$$\begin{aligned}
\mathbb{E}\left[x_{t+1}^f\otimes x_{t+1}^{rd}\right] &= \left(I_{n_x^2} - h_x^{\otimes 2}\right)^{-1}\Big((h_x\otimes h_{xx})\,\mathbb{E}((x_t^f)^{\otimes 2}\otimes x_t^s) + \frac{1}{6}(h_x\otimes h_{xxx})\,\mathbb{E}\left(x_t^f\right)^{\otimes 4}\\
&\quad + \frac{3}{6}(h_x\otimes h_{\sigma\sigma})\sigma^2\,\mathbb{E}\left[(x_t^f)^{\otimes 2}\right]\Big)
\end{aligned}$$

This completes the proof. □

2.B.2.4. *Simplifying the 4th-order expressions.* For the model used in the current paper, the 4th-order Kronecker products can be simplified further. Namely,

$$\left(h_x^{\otimes 2}\otimes\eta^{\otimes 2}\right)K_{n_e^2,n_x^2}=K_{n_x^2,n_x^2}\left(\eta^{\otimes 2}\otimes h_x^{\otimes 2}\right)$$

For the current model we have $n_e=1$ so that $K_{n_e,n_x}=K_{n_x,n_e}=I_{n_x}$, $I_{n_e}=vec(I_{n_e})=1$.

$$\begin{aligned}
\mathbb{E}\left(x_{t+1}^f\right)^{\otimes 4} &= \sigma^2\left(I_{n_x^4}-h_x^{\otimes 4}\right)^{-1}\left[\sigma^2\eta^{\otimes 4}M^4+K_{n_x^2,n_x^2}\left(\eta^{\otimes 2}\otimes h_x^{\otimes 2}\right)\right.\\
&\quad +\left(h_x\otimes\eta\otimes h_x\otimes\eta\right)\left(I_{n_x}\otimes I_{n_x}\right)\\
&\quad +\left(h_x\otimes\eta\otimes\eta\otimes h_x\right)\left(I_{n_x}\otimes I_{n_x}\right)\\
&\quad +\left(\eta\otimes h_x\otimes h_x\otimes\eta\right)\\
&\quad +\left(\eta\otimes h_x\otimes\eta\otimes h_x\right)\left(I_{n_x}\otimes I_{n_x}\right)\\
&\quad \left.+\left(\eta^{\otimes 2}\otimes h_x^{\otimes 2}\right)\right)\left(\mathbb{E}(x_t^f)^{\otimes 2}\right)\Big]\\
&= \sigma^2\left(I_{n_x^4}-h_x^{\otimes 4}\right)^{-1}\left[\sigma^2\eta^{\otimes 4}M^4\right.\\
&\quad +\left(h_x\otimes\eta\otimes h_x\otimes\eta\right)\left(I_{n_x}\otimes I_{n_x}\right)\\
&\quad +\left(h_x\otimes\eta\otimes\eta\otimes h_x\right)\left(I_{n_x}\otimes I_{n_x}\right)\\
&\quad +\left(\eta\otimes h_x\otimes h_x\otimes\eta\right)\\
&\quad +\left(\eta\otimes h_x\otimes\eta\otimes h_x\right)\left(I_{n_x}\otimes I_{n_x}\right)\\
&\quad \left.+\left(K_{n_x^2,n_x^2}+I_{n_x^2}\right)\left(\eta^{\otimes 2}\otimes h_x^{\otimes 2}\right)\right)\left(\mathbb{E}(x_t^f)^{\otimes 2}\right)\Big].
\end{aligned}$$

Note that $I_{n_x}\otimes I_{n_x}=I_{n_x^2}$ so that

$$\begin{aligned}
\mathbb{E}\left(x_{t+1}^f\right)^{\otimes 4} &= \sigma^2\left(I_{n_x^4}-h_x^{\otimes 4}\right)^{-1}\left[\sigma^2\eta^{\otimes 4}M^4\right.\\
&\quad +\left(h_x\otimes\eta\otimes h_x\otimes\eta\right)\\
&\quad +\left(h_x\otimes\eta\otimes\eta\otimes h_x\right)\\
&\quad +\left(\eta\otimes h_x\otimes h_x\otimes\eta\right)\\
&\quad +\left(\eta\otimes h_x\otimes\eta\otimes h_x\right)\\
&\quad \left.+\left(K_{n_x^2,n_x^2}+I_{n_x^2}\right)\left(\eta^{\otimes 2}\otimes h_x^{\otimes 2}\right)\right)\left(\mathbb{E}(x_t^f)^{\otimes 2}\right)\Big].
\end{aligned}$$

Moreover

$$
\begin{aligned}
h_x \otimes \eta \otimes h_x \otimes \eta &= h_x \otimes \left[K_{n_x,n_x}\left(h_x \otimes \eta\right) K_{n_e,n_x}\right] \otimes \eta \\
&= \left(I_{n_x} \otimes K_{n_x,n_x} \otimes I_{n_x}\right)\left(h_x^{\otimes 2} \otimes \eta^{\otimes 2}\right) \\
&= \left(I_{n_x} \otimes K_{n_x,n_x} \otimes I_{n_x}\right)\left(K_{n_x^2,n_x^2}\left(\eta^{\otimes 2} \otimes h_x^{\otimes 2}\right)\right),
\end{aligned}
$$

$$
\begin{aligned}
h_x \otimes \eta \otimes \eta \otimes h_x &= h_x \otimes \eta \otimes \left[K_{n_x,n_x}\left(h_x \otimes \eta\right) K_{n_e,n_x}\right] \\
&= \left(I_{n_x^2} \otimes K_{n_x,n_x}\right)\left[h_x \otimes \eta \otimes h_x \otimes \eta\right] \\
&= \left(I_{n_x^2} \otimes K_{n_x,n_x}\right)\left[\left(I_{n_x} \otimes K_{n_x,n_x} \otimes I_{n_x}\right)\left(h_x^{\otimes 2} \otimes \eta^{\otimes 2}\right)\right] \\
&= \left(I_{n_x^2} \otimes K_{n_x,n_x}\right)\left[\left(I_{n_x} \otimes K_{n_x,n_x} \otimes I_{n_x}\right)\left(K_{n_x^2,n_x^2}\left(\eta^{\otimes 2} \otimes h_x^{\otimes 2}\right),\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
\eta \otimes h_x \otimes \eta \otimes h_x &= \eta \otimes \left[K_{n_x,n_x}\left(\eta \otimes h_x\right) K_{n_x,n_e}\right] \otimes h_x \\
&= (I_{n_x} \otimes K_{n_x,n_x} \otimes I_{n_x})(\eta^{\otimes 2} \otimes h_x^{\otimes 2}),
\end{aligned}
$$

$$
\begin{aligned}
\eta \otimes h_x \otimes h_x \otimes \eta &= \eta \otimes h_x \otimes [K_{n_x,n_x}(\eta \otimes h_x) K_{n_x,n_e}] \\
&= (I_{n_x^2} \otimes K_{n_x,n_x})(\eta \otimes h_x \otimes \eta \otimes h_x) \\
&= (I_{n_x^2} \otimes K_{n_x,n_x})\left[(I_{n_x} \otimes K_{n_x,n_x} \otimes I_{n_x})(\eta^{\otimes 2} \otimes h_x^{\otimes 2})\right].
\end{aligned}
$$

Hence

$$
\begin{aligned}
\mathbb{E}\left(x_{t+1}^f\right)^{\otimes 4} = \sigma^2 &\left(I_{n_x^4} - h_x^{\otimes 4}\right)^{-1}\Big[\sigma^2 \eta^{\otimes 4} M^4 \\
&+ \left(I_{n_x} \otimes K_{n_x,n_x} \otimes I_{n_x}\right)\left(K_{n_x^2,n_x^2}\left(\eta^{\otimes 2} \otimes h_x^{\otimes 2}\right)\right) \\
&+ \left(I_{n_x^2} \otimes K_{n_x,n_x}\right)\left[\left(I_{n_x} \otimes K_{n_x,n_x} \otimes I_{n_x}\right)\left(K_{n_x^2,n_x^2}\left(\eta^{\otimes 2} \otimes h_x^{\otimes 2}\right)\right)\right] \\
&+ (I_{n_x^2} \otimes K_{n_x,n_x})\left[(I_{n_x} \otimes K_{n_x,n_x} \otimes I_{n_x})(\eta^{\otimes 2} \otimes h_x^{\otimes 2})\right] \\
&+ (I_{n_x} \otimes K_{n_x,n_x} \otimes I_{n_x})(\eta^{\otimes 2} \otimes h_x^{\otimes 2}) \\
&+ \left(K_{n_x^2,n_x^2} + I_{n_x^2}\right)\left(\eta^{\otimes 2} \otimes h_x^{\otimes 2}\right)\Big)\left(\mathbb{E}(x_t^f)^{\otimes 2}\right)\Big],
\end{aligned}
$$

or

$$\begin{aligned}\mathbb{E}\left(x_{t+1}^{f}\right)^{\otimes 4} &= \sigma^2\left(I_{n_x^4}-h_x^{\otimes 4}\right)^{-1}\Big[\sigma^2\eta^{\otimes 4}M^4 \\ &+\Big\{\left(I_{n_x}\otimes K_{n_x,n_x}\otimes I_{n_x}\right)K_{n_x^2,n_x^2} \\ &+\left(I_{n_x^2}\otimes K_{n_x,n_x}\right)\left[\left(I_{n_x}\otimes K_{n_x,n_x}\otimes I_{n_x}\right)\left(K_{n_x^2,n_x^2}\right)\right] \\ &+\left(I_{n_x^2}\otimes K_{n_x,n_x}\right)\left[\left(I_{n_x}\otimes K_{n_x,n_x}\otimes I_{n_x}\right)\right] \\ &+\left(I_{n_x}\otimes K_{n_x,n_x}\otimes I_{n_x}\right) \\ &+\left(K_{n_x^2,n_x^2}+I_{n_x^4}\right)\Big\}\left(\eta^{\otimes 2}\otimes h_x^{\otimes 2}\right)\Big)\left(\mathbb{E}(x_t^f)^{\otimes 2}\right)\Big],\end{aligned}$$

$$\begin{aligned}\mathbb{E}\left(x_{t+1}^{f}\right)^{\otimes 4} &= \sigma^2\left(I_{n_x^4}-h_x^{\otimes 4}\right)^{-1}\Big[\sigma^2\eta^{\otimes 4}M^4 \\ &+\Big\{\left(I_{n_x^2}\otimes K_{n_x,n_x}\right)\left(I_{n_x}\otimes K_{n_x,n_x}\otimes I_{n_x}\right)+I_{n_x^4} \\ &+I_{n_x}\otimes K_{n_x,n_x}\otimes I_{n_x}\Big\}\left(K_{n_x^2,n_x^2}+I_{n_x^4}\right)\left(\eta^{\otimes 2}\otimes h_x^{\otimes 2}\right)\left(\mathbb{E}(x_t^f)^{\otimes 2}\right)\Big],\end{aligned}$$

hence

$$\begin{aligned}\mathbb{E}\left(x_{t+1}^{f}\right)^{\otimes 4} &= \sigma^2\left(I_{n_x^4}-h_x^{\otimes 4}\right)^{-1}\Big[\sigma^2\eta^{\otimes 4}M^4 \\ &+\Big\{\left(I_{n_x^2}\otimes K_{n_x,n_x}+I_{n_x^4}\right)\left(I_{n_x}\otimes K_{n_x,n_x}\otimes I_{n_x}\right)+I_{n_x^4}\Big\}\left(K_{n_x^2,n_x^2}+I_{n_x^4}\right)\left(\eta^{\otimes 2}\otimes h_x^{\otimes 2}\right)\left(\mathbb{E}(x_t^f)^{\otimes 2}\right)\Big].\end{aligned}$$

CHAPTER 3

Demand-Driven Growth

"Taking U.S. performance over the past 50 years as a benchmark, the potential for welfare gains from better long-run, supply-side policies exceeds by far the potential from further improvements in short-run demand management."
Robert E. Lucas, Jr. (2003)

3.1. Introduction

Aggregate economic growth has traditionally been thought of as a supply-side phenomenon.[1] In this paper, we argue that in fact aggregate economic growth is to a large extent demand-driven. Importantly, failing to account for this demand channel changes our understanding of the macroeconomic, supply-side, policies affect aggregate growth.

We build a tractable, general equilibrium, model with heterogeneous firms. This type of environment has been extensively used for questions related to industry dynamics (see e.g. the seminar work of Hopenhayn and Rogerson 1993b) and more recently also for understanding aggregate growth (see e.g. Acemoglu et al. 2018). However, while demand accumulation and disturbances have been shown to be key drivers of firm-level dynamics (see e.g. Foster et al. 2016), aggregate growth has remained in the domain of the supply side in the form of investments into improving the efficiency level of production (see e.g. Klette and Kortum 2004). Our framework brings together these two forces.

Specifically, in our model households consume a bundle of imperfectly substitutable differentiated goods varieties. The latter are produced by individual firms which differ not only in the efficiency with which they produce goods ("productivity"), but also in the amount of customers they can sell to ("customer base"). Importantly, firms can invest into improving both productivity and the customer base. The former happens through costly research and development (R&D). Investing resources into R&D gives the firm a chance to innovate. Successful innovations lead to productivity increases. It is assumed that firms can attract customers by charging lower relative prices. Given the imperfect substitutability between goods varieties, firms can afford to charge a markup over their marginal costs. Therefore, in order to attract additional customers, firms lower their markups.

Considering both endogenous margins at the same time is crucial because they form an endogenous feedback loop. Higher productivity allows firms to charge lower prices and attract more customers. In turn, attracting more customers allows firms to sell greater amounts of their output increasing the payoffs from innovation.

[1] Exceptions to this include recent contributions of "New-Keynesian endogenous growth models" of e.g. Benigno and Fornaro (2018).

We parametrize our model to match salient features of U.S. data pertaining to firm growth and R&D expenditures. Importantly, our model matches the life-cycle profile of average employment, the R&D intensity and aggregate productivity growth. Armed with this model, we conduct two distinct exercises to highlight the importance of our novel channel.

First, we decompose aggregate growth into it's contributions. As in existing endogenous growth models, also our framework features the following drivers of aggregate growth: (i) *selection*, whereby resources reallocate from less to more productive firms over the life-cycle, (ii) *creative destruction*, whereby relatively unproductive firms shut down and give way to entrants and (iii) *incumbents' R&D*, whereby incumbent firms invest into improving their own efficiency. All three factors push the technological frontier also in our framework. However, demand is a new driver of growth in our model because it increases the incentives to conduct R&D. A counterfactual exercise in which the possibility of accumulating demand is "switched off" results in aggregate growth dropping from an annual rate of 1.5 to 0.7 percent. In other words, 55% of aggregate growth is demand-driven.

Second, we investigate how the presence of endogenous demand accumulation changes our understanding of the effectiveness of supply-side policies in promoting growth. Towards this end we recalibrate our counterfactual economy with fixed demand. This counterfactual economy matches exactly the same moments as our baseline, including the life-cycle pattern of employment, aggregate R&D intensity and aggregate growth. Therefore, the two economies are observationally equivalent from the point of view of these parametrization targets.

Nevertheless, the two economies feature important differences. Most notably, the distribution of firm values is different. Specifically, endogenous demand accumulation provides an additional boost to the value of a business. The counterfactual economy with fixed demand is, therefore, characterized by a much more condensed firm value distribution. Importantly, the distribution of firm values is crucial for the sensitivity of the economy to policy instruments. The reason for this is that all forward-looking decisions, such as R&D investment but also firm exit, depend in one way or another on firm values. A different distribution of firm values implies a different sensitivity to otherwise identical policy interventions. This relates to the results found in Pugsley et al. (2019) and suggests an interesting avenue for future research.

In what follows, we provide an extensive literature review. We then describe the model and its parametrization after which we present our main results. The last section concludes.

3.2. Related literature

Our work unifies two major strands of the literature: the body of work on the causes and consequences of firm heterogeneity and the literature on the firm-level origins of aggregate economic growth.

Recent advances in Schumpeterian endogenous growth theory suggest that accounting for firm-level heterogeneity is necessary to fully understand the underlying sources of economic growth (Luttmer 2010; Aghion et al. 2014; Akcigit and Nicholas 2019). The evidence from European (Lentz and Mortensen 2008) and U.S. (Acemoglu et al. 2018; Akcigit and Kerr 2018) firm-level data reveals significant heterogeneity in R&D activity across firms. These studies highlight that the reallocation of production towards the most innovative firms, exit of the least productive units, and the creation of new businesses are key forces generating aggregate economic growth. In the current work, we contribute to the literature by deepening our understanding of how these forces interact with firm lifecycle dynamics.

Even within narrowly defined industries, there are large differences in terms of firm size, productivity, and growth between young firms and more mature businesses (Davis et al. 1996; Foster et al. 2001; Cabral and Mata 2003; Haltiwanger et al. 2013b; Foster et al. 2016). Researchers have studied various aspects of firm and industry dynamics to understand the underlying causes of these lifecycle patterns. A body of evidence points towards the selection on idiosyncratic productivity. Over time, as less productive firms decide to exit, the average productivity among surviving firms gradually increases giving rise to a positive relationship between firm age and average size (Hopenhayn 1992a; Lentz and Mortensen 2005; Luttmer 2007). This lifecycle dynamics induced by selection can be amplified by various frictions: labor adjustment costs (Hopenhayn and Rogerson 1993b); capital adjustment costs (Cooper and Haltiwanger 2006b; Clementi and Palazzo 2016b); information frictions and learning (Jovanovic 1982; Arkolakis et al. 2018); organizational capital (Atkeson and Kehoe 2005); or financial frictions (Albuquerque and Hopenhayn 2004; Clementi and Hopenhayn 2006). Since young and small firms are affected by these market imperfections to a much larger extent than large and mature enterprises, these frictions can account for the lifecycle dynamics of firms documented in micro-data.

In the current paper, we focus on a complementary source of market imperfections that can affect firm lifecycle dynamics. When new firms are created, few consumers are aware of their presence. Various informational or reputational frictions can hinder firms' effort to acquire more clients. Over time, these frictions gradually abate, young firms build a customer base, and catch up with the established businesses (Caminal and Vives 1999; Drozd and Nosal 2012; Gourio and Rudanko 2014). Foster et al. (2016) provide evidence that differences in idiosyncratic demand levels can account for a considerable portion of the relationship between plant age and average size. We argue that demand accumulation, unlike the other frictions to firm growth mentioned above, is a force that boosts economic growth.

We build a general-equilibrium model of firm dynamics with endogenous entry and exit in the spirit of Hopenhayn and Rogerson (1993b). We contribute to the literature by augmenting the workhorse model in two significant ways. Firstly, following Foster et al. (2016), we allow firms to passively accumulate customer base which becomes a state variable for firms. Each product sold rises firm's customer base and affects its future demand. Secondly, we assume that the firm's idiosyncratic productivity is endogenously determined by its research and development investment as in Acemoglu et al. (2018). We share the insight of Schumpeterian growth theory that selection on productivity and reallocation of production towards the most efficient innovators is a crucial driver of aggregate growth (Acemoglu et al. 2018). A successful innovation reduces the firm's marginal cost, allows it to charge lower prices, and to increase sales in the current period. Since demand accumulates, all future sales are affected positively as well. This, in turn, raises the returns to innovation creating a dynamic complementarity. This process helps to drive less productive firms out of the market, reallocates production towards innovators, and lifts aggregate growth. This suggests that forces slowing down the firm-level growth are not necessarily detrimental for aggregate growth.

Cavenaile et al. (2019) find that the advertising expenses and R&D expenses are substitutes and argue more effective marketing technology leads to lower growth. In Cavenaile et al. (2019) firms make a static decision about resources invested in advertising. In contrast, we allow for *dynamic* demand accumulation through sales, consistent with empirical evidence in Foster et al. (2016). This dynamic aspect makes demand accumulation complementary to R&D. Our model highlights that the flow of customers towards the least expensive goods facilitates the flow of resources towards the most innovative firms.

Our work is related to recent advances in endogenous growth new Keynesian literature. There, the innovation effort of individual firms constitutes an endogenous amplification mechanism for demand shocks (Benigno and Fornaro 2018; Anzoategui et al. 2019) or financial shocks (Bianchi et al. 2019; Queralto 2019) at business cycle frequency. In the current work, we share the insight that the weak aggregate demand can have a detrimental impact on firm-level R&D and, consequently, on aggregate economic growth. However, in contrast to this strand of the literature, we focus on understanding the origins of aggregate economic growth in the long run, using a model that is consistent with micro evidence on firm lifecycle dynamics and innovation.

3.3. Model

In this section, we describe the model which helps us understand how aggregate demand can shape aggregate economic growth. The model aims to capture the interactions between a firm's ability to accumulate customer base and its efforts to create innovative products.

In the model, households supply labor in a competitive labor market and consume a bundle of imperfectly substitutable product varieties. Each variety is supplied by a single monopolistic producer. The demand for these goods is endogenous and accumulates over time as firms gradually build their customer base. In addition, firms can spend resources on research and development (R&D). Each successful innovation reduces the firm's marginal production cost. This, therefore, impacts not only its current revenue (by reducing the costs necessary to produce a unit of output) but also the future revenues through the endogenous demand accumulation channel. Finally, incumbent firms must also pay stochastic operation costs which induce endogenous exit.

New firms endogenously enter the economy. In the beginning, very few households are aware of products offered by startup businesses. With each sale, however, the brand recognition rises and so does the demand the firm faces at any given price. Firms, in turn, use their pricing power in a strategic way to attract more customers. Young firms, with a relatively small customer base, set low markups to quickly gain demand. As firms get older, they gradually increase markups to harvest their growing customer base. The strategic pricing behavior has important implications for firm innovation over the lifecycle and, ultimately, for aggregate economic growth.

We start by describing the household side of the economy which we try to keep as simple as possible. We put more emphasis on the theory of the production side of the economy, described next. We follow with aggregation and the equilibrium definition.

3.3.1. Households. There is a unit mass of identical, infinitely lived households that derive utility from consumption and suffer disutility from hours worked. We assume the following utility function

$$\mathbb{U}(C_t, N_t) = \ln C_t - \nu N_t, \tag{3.1}$$

where the consumption bundle C_t is a Dixit-Stiglitz aggregator over a consumption $c_{t,j}$ of each product j out of set of all available products $\mathscr{J}_t$. Households have a preference towards each product, $d_{t,j}$, which is endogenously determined and can be shaped by firms through their production and pricing decisions. We describe in detail the behavior of firms below. Before that, let us state formally the definition of the consumption bundle,

$$C_t = \left[\int_{j \in \mathscr{J}_t} d_{j,t}^{\frac{1}{\eta}} c_{j,t}^{\frac{\eta-1}{\eta}} \, \mathrm{d}j \right]^{\frac{\eta}{\eta-1}}. \tag{3.2}$$

The parameter η captures the elasticity of the demand for each variety with respect its price $p_{j,t}$. The households maximize their lifetime utility, $\sum_{t=1}^{\infty} \beta^t \mathbb{U}(C_t, N_t)$, where β is the subjective discount factor.

The households rent their labor at the competitive wage W_t. They can save in firm equity A_t. Let $\tilde{p}_t^a$ be the price of assets in period t and $\tilde{\pi}_t^a$ the dividend. All prices are relative to the aggregate price index P_t which we take as the numeraire. Formally, the budget constraint is

$$\tilde{p}_t^a A_{t+1} + \int_{j \in \mathscr{J}_t} p_{j,t} c_{j,t} \, \mathrm{d}j = W_t N_t + \left(\tilde{p}_t^a + \tilde{\pi}_t^a \right) A_t. \tag{3.3}$$

The optimality conditions for the assets, labor, and consumption choices, respectively, can be written as

$$\frac{1}{C_t} = \beta \mathbb{E}_t (1 + r_{t+1}) \frac{1}{C_{t+1}}, \tag{3.4}$$

$$\nu = \frac{W_t}{C_t}, \tag{3.5}$$

$$c_{j,t} = d_{j,t} p_{j,t}^{-\eta} C_t. \tag{3.6}$$

where the return on equities is $1 + r_{t+1} = \frac{\tilde{p}_{t+1}^a + \tilde{\pi}_{t+1}^a}{\tilde{p}_t^a}$. Equation (3.4) is the standard intertemporal Euler equation. Equation (3.5) pins down the labor supply by equating marginal utility of consumption with marginal disutility from hours worked. Finally, equation (3.6) describes the demand for each product j which depends negatively on its price $p_{j,t}$ and positively on its appeal $d_{j,t}$.

3.3.2. Firms. The timing of events is the following: at the beginning of the period, firms decide whether to remain in operation or not. Conditional on survival, firms produce, invest in innovation, and pay their workers. At the end of the period, innovations are realized, demand accumulates and firms face exogenous destruction shock.

Incumbents. Each firm is characterized by idiosyncratic productivity $q_{j,t}$. Varieties are produced with labor $n_{c,j,t}$, according to a linear production function

$$c_{j,t} = q_{j,t} n_{c,j,t}. \tag{3.7}$$

There are no labor adjustment cost, hence the employment choice is static.

The specification of research and development process is similar to Acemoglu et al. (2018). Firms employ $n_{r,j,t}$ workers to conduct the research and development. In each period, these researchers generate new innovations with probability $\tilde{x}_{j,t}$. The R&D effort translates into innovation probability according to

$$\tilde{x}_{j,t} = \left(\gamma_j n_{r,j,t}\right)^{1/\psi}, \tag{3.8}$$

where $\psi > 1$ and $\gamma_j \geq 0$ govern the elasticity of innovation probability with respect to labor, and the innovation efficiency, respectively. Parameter γ_j is firm-specific, while ψ common to all firms. A successful innovation increases the firm's productivity by a factor λ, that is $q_{j,t+1} = (1+\lambda) q_{j,t}$.

In addition to active R&D effort, we assume that there is a passive diffusion of technology from market leaders to less innovative firms. Concretely, the probability of innovation for each firm is

$$x_{j,t} = \tilde{x}_{j,t} + x_0, \tag{3.9}$$

where $x_0 > 0$ is a parameter that captures the rate of technology diffusion.

The firm-specific demand follows

$$d_{j,t} = \alpha_j^0 \left(b_{j,t}\right)^{\alpha_j^1}, \tag{3.10}$$

where $b_{j,t}$is firm's customer base and $\alpha_j = \left(\alpha_j^0, \alpha_j^1\right)$ is a vector of firm-specific parameters governing the impact of the customer base on demand.[2] The customer base accumulates endogenously and evolves according to the following law of motion

$$b_{j,t+1} = (1-\delta)\left(b_{j,t} + \frac{c_{j,t}}{C_t}\right), \tag{3.11}$$

where δ is the rate at which the customer base depreciates over time. This means that a higher sale today leads to a higher demand in the future.

Firms maximize their value by choosing prices $p_{j,t}$ and the probability of innovation $x_{j,t}$. Formally, at the beginning of the period firms maximize

$$V_j(q,b) = \max\left\{\int_{\Re_+} \left[\tilde{V}_j(b,q) - W\phi \, \mathrm{d}F_\phi(\phi)\right], 0\right\}. \tag{3.12}$$

The outer "max" operator captures the exit decision of the firm. Upon exit, firms receive an outside option which value is normalized to zero. If a firm decides to continue operating, it faces a stochastic operation cost ϕ distributed according to the cdf F_ϕ. The cost is expressed in terms of labor. We mark firm value with a subscript j to indicate that the firm's value depends on its idiosyncratic characteristics α_j and γ_j. $\tilde{V}_j$ is the firm's continuation value that satisfies

$$\tilde{V}_j(b_{j,t}, q_{j,t}) = \max_{p_{j,t}, x_{j,t}, b_{j,t+1}} \left[p_{j,t} c_{j,t} - W(n_{c,j,t} + n_{r,j,t}) + \tilde{\beta}_t \mathbb{E} V_j(b_{j,t+1}, q_{j,t+1})\right] \tag{3.13}$$

s. t.

$$b_{j,t+1} = (1-\delta)\left(b_{j,t} + \frac{c_{j,t}}{C_t}\right), \tag{3.14}$$

$$b_{j,t+1} - (1-\delta) b_{j,t} \geq 0, \quad p_{j,t} \geq 0, \quad x_{j,t} \in [0,1]. \tag{3.15}$$

[2]A micro-foundation via information acquisition can be found in Sedláček and Sterk (2017).

The effective discount factor satisfies $\tilde{\beta}_t = \frac{(1-\rho)}{(1+r_t)(1+g_t)}$, where ρ is the exogenous exit rate and g_t is the aggregate growth rate of the economy.

Pricing and innovation. The firm exits whenever the expected operating cost are too large relative to the continuation value. Let $\tilde{\phi}$ mark the the cost at which a firm is indifferent between continuation and exit. We have

$$\tilde{\phi}_j(b_{j,t}, q_{j,t}) = \frac{\tilde{V}_j(b_{j,t}, q_{j,t})}{W_t}. \tag{3.16}$$

We denote the survival probability by

$$\Phi_j(b_{j,t}, q_{j,t}) \equiv F_\phi\left(\tilde{\phi}_j(b_{j,t}, q_{j,t})\right). \tag{3.17}$$

With this notation at hand, we can write the firm's value as

$$V(b_{j,t}, q_{j,t}) = \Phi_j(b_{j,t}, q_{j,t}) \max_{\substack{p_{j,t}, x_{j,t}, \\ b_{j,t+1}}} \begin{bmatrix} p_{j,t} c_{j,t} - W_t(n_{c,j,t} + n_{r,j,t} + \hat{\phi}_j) \\ +\tilde{\beta}_t\left[x_{j,t} V_j(b_{j,t+1}, q_{j,t}(1+\lambda)) + (1-x_{j,t}) V_j(b_{j,t+1}, q_{j,t})\right] \end{bmatrix} \tag{3.18}$$

where $\hat{\phi}_j(b_{j,t}, q_{j,t}) \equiv \mathbb{E}\left(\phi | \phi \le \tilde{\phi}_j(b_{j,t}, q_{j,t})\right)$ is the expected operating cost paid by a surviving firm. In equation (3.18) we used the fact that $q_{j,t+1} = (1+\lambda) q_{j,t}$ if the innovation is successful and $q_{j,t+1} = q_{j,t}$ otherwise.

Next, we will derive the optimal pricing strategy of the firm. Let $\mu_j\left(b_{j,t}, q_{j,t}\right) \equiv \frac{p_{j,t} q_{j,t}}{W_t}$ denote the markup a firm charges over its marginal cost $q_{j,t}/W_t$. In what follows, we leave implicit the dependance of all objects on $\left(b_{j,t}, q_{j,t}\right)$. First-order condition for the optimal price yields the following formula for markup

$$\mu_{j,t} = \frac{\eta}{\eta-1} - \frac{\eta}{\eta-1}\frac{q_{j,t}}{W_t}\frac{\tilde{\beta}_t(1-\delta)}{C_t}\left[x_{j,t}\partial_{b_{j,t+1}} V_j(b_{j,t+1}, q_{j,t+1}) + (1-x_{j,t})\partial_{b_{j,t+1}} V_j(b_{j,t+1}, q_{j,t})\right]. \tag{3.19}$$

There are two components of the markup formula: static and dynamic. The former is the well-known constant markup implied by the Dixit-Stiglitz preferences, $\frac{\eta}{\eta-1}$. In addition, firms take into account the impact of the current sales on the demand they will face in the future. These strategic pricing considerations are reflected in the term involving the derivatives of the firm's value with respect to the customer base.

Firms also choose how much resources to spend on research and development. The first-order condition for the innovation probability is

$$x_{t,j} = \left[\frac{\gamma_j}{\psi W_t}\tilde{\beta}_t\left(V_j(b_{j,t+1}, q_{j,t}(1+\lambda)) - V_j(b_{j,t+1}, q_{j,t})\right)\right]^{\frac{1}{\psi-1}} \tag{3.20}$$

Incumbents aim to achieve the higher innovation success rate, devote the more resources to R&D, the larger is gain from productivity improvement, $V_j(b_{j,t+1}, q_{j,t}(1+\lambda)) - V_j(b_{j,t+1}, q_{j,t})$. Crucially, the gains from innovation also depend on the customer base $b_{j,t+1}$. Finally, the resources spend on R&D depend positively on the idiosyncratic research efficacy γ_j.

Entrants. In each period there is a bounded mass of prospective entrepreneurs. Firm entry happens in two phases. First, potential entrepreneurs must pay an entry cost, ϕ_e, to get a chance to start up. This cost represents the necessary investment to start a businesses, such as creating a business plan, red tape etc. Conditional on paying this cost, entrepreneurs get an exogenous draw of firm characteristics α and γ determining their long-run capabilities of attracting demand and doing R&D. We denote by $P_{i,j}$ the probabilities of obtaining a particular combination of firm characteristics α_i and γ_j. Thereafter, they can invest into R&D as incumbents. Production happens only in the next period, following entry, where all businesses start with an initial customer base of b_e. Formally, the free entry condition is given by

$$\phi_e = \sum_{i,j} P_{i,j} V_e(\alpha_i, \gamma_j), \tag{3.21}$$

where the value of entrants, $V_e(\alpha,\gamma)$ is given by

$$V_e(\alpha,\gamma) = \max\left\{0, \max_{x\in[0,1]} \left[-W_t n_r + \tilde{\beta}_t \left(xV(b_e, \bar{q}_t(1+\lambda); \alpha,\gamma) + (1-x)V(b_e, \bar{q}_t; \alpha,\gamma)\right)\right]\right\}, \tag{3.22}$$

where we assume that entrants start up at the average productivity level q_t, unless they manage to innovate. Upon entry, startups become incumbent businesses.

3.3.3. Aggregation and Balanced Growth Path. We assume that the economy evolves along a balanced growth path (BGP) where all variables grow at the same, time-invariant rate g. Using the definition of consumption aggregate C_t in (3.2) and the production function (3.7) we have

$$C_t = \left[\int_{j\in\mathcal{J}_t} d_{j,t}^{\frac{1}{\eta}} c_{j,t}^{\frac{\eta-1}{\eta}} \,\mathrm{d}j\right]^{\frac{\eta}{\eta-1}} = \bar{q}_t \left[\int_{j\in\mathcal{J}_t} d_{j,t}^{\frac{1}{\eta}} \left(\frac{q_{j,t}}{\bar{q}_t} n_{c,j,t}\right)^{\frac{\eta-1}{\eta}} \,\mathrm{d}j\right]^{\frac{\eta}{\eta-1}}, \tag{3.23}$$

which shows that C_t grows at the same rate as the average productivity $\bar{q}_t$. The labor supply optimality condition (3.5) implies that the aggregate wage, and hence other aggregate variables, grows at the same rate as C_t.

Let $M_{\mathcal{J}}$ be total measure of firms in the economy. Let $\mathcal{I}$ and $\mathcal{E}$ denote the set of incumbents and entrants, respectively, with $M_{\mathcal{I}}$ and $M_{\mathcal{E}}$ being the associated measures. By construction $\mathcal{J} = \mathcal{I} \cup \mathcal{E}$. The growth rate of the average productivity can be expressed as a sum of contributions of incumbents and entrants

$$1+g_t = \frac{\bar{q}_{t+1}}{\bar{q}_t} = \frac{1}{M_{\mathcal{J},t+1}} \int_{j\in\mathcal{J}_{t+1}} \frac{q_{j,t+1}}{\bar{q}_t} \,\mathrm{d}j = \frac{1}{M_{\mathcal{J},t+1}} \left(\int_{j\in\mathcal{I}_{t+1}} \frac{q_{j,t+1}}{\bar{q}_t} \,\mathrm{d}j + \int_{j\in\mathcal{E}_{t+1}} \frac{q_{j,t+1}}{\bar{q}_t} \,\mathrm{d}j\right)$$

Recall that for incumbents $q_{j,t+1} = (1+\lambda)q_{j,t}$ if the innovation is successful and $q_{j,t+1} = q_{j,t}$ otherwise. For entrants, on the other hand, $q_{j,t+1} = (1+\lambda)\bar{q}_t$ if the innovation is successful and $q_{j,t+1} = \bar{q}_t$ otherwise. The growth rate satisfies

$$1+g_t = \frac{1-\rho}{M_{\mathcal{J},t+1}} \int_{j\in\mathcal{I}_{t+1}} \Phi_{j,t} \left(x_{j,t} \frac{q_{j,t}}{\bar{q}_t}\lambda + \frac{q_{j,t}}{\bar{q}_t}\right) \mathrm{d}j + \frac{1}{M_{\mathcal{J},t+1}} \int_{j\in\mathcal{E}_{t+1}} (x_{e,j,t}\lambda + 1) \,\mathrm{d}j \tag{3.24}$$

For the sake of notational clarity, let us introduce $\tilde{q}_{j,t} \equiv \frac{q_{j,t}}{\bar{q}_t}$. Note further that

$$\frac{1}{M_{\mathcal{J},t+1}} \int_{j\in\mathcal{E}_{t+1}} 1 \,\mathrm{d}j = \frac{M_{\mathcal{E},t+1}}{M_{\mathcal{J},t+1}}.$$

Now, we can express the aggregate growth rate as a sum of growth rates driven by surviving incumbents and a creative destruction term

$$1+g_t = \frac{1-\rho}{M_{\mathscr{I},t+1}} \int_{j\in\mathscr{I}_{t+1}} \Phi_{j,t}\left(x_{j,t}\tilde{q}_{j,t}\lambda + \tilde{q}_{j,t}\right) \mathrm{d}j \; + \; \frac{1}{M_{\mathscr{I},t+1}}\lambda \int_{j\in\mathscr{E}_{t+1}} x_{e,j,t}\,\mathrm{d}j + \frac{M_{\mathscr{E},t+1}}{M_{\mathscr{I},t+1}}. \tag{3.25}$$

To shed more light on the sources of the aggregate economic growth let us re-write equation (3.25). Firstly, let us impose balanced growth path and drop the time subscripts. Next, recall that behavior of each firm j is fully characterized by a tuple $(\alpha_j, \gamma_j, b_j, q_j)$ which we call firm state. For each state, let $\mathbb{F}(\alpha,\gamma,b,q)$ be the measure of firms that prevails in the economy along BGP. Let $\mathbb{F}_e(\alpha,\gamma,b,q)$ be the analogous measure among entrants. Note that since $\mathbb{F}$ and $\mathbb{F}_e$ are measures, $F \equiv \frac{\mathbb{F}}{M_{\mathscr{I}}}$ and $F_e \equiv \frac{\mathbb{F}_e}{M_{\mathscr{E}}}$ are distributions.

With this notation at hand, we can re-write equation (3.25) as

$$1+g = \frac{M_{\mathscr{E},t+1}}{M_{\mathscr{I},t+1}} \underbrace{\left(\lambda \int x_e \,\mathrm{d}F_e + 1\right)}_{\text{creative destruction}} + \frac{M_{\mathscr{I},t+1}}{M_{\mathscr{I},t+1}} \underbrace{\int \hat{\Phi}\left(\lambda x \tilde{q} + \tilde{q}\right) \mathrm{d}F}_{\text{contribution of incumbents}} \quad , \tag{3.26}$$

where we leave implicit the dependance of all objects on the firm state and where $\hat{\Phi} \equiv (1-\rho)\Phi$ is the effective survival rate.

There are two terms weighted by relative importance of entrants and incumbents. The first term, the creative destruction, captures the innovation effort by entrants who replace the least productive incumbents that decided to exit. In addition, the growth is positively affected by reallocation of production towards more innovative businesses.

3.3.4. Equilibrium. The equilibrium consist of consumer choices N_t, $c_{j,t}$, A_t, firm values V, $\tilde{V}$, V_e, policy $p_{j,t}$, $x_{j,t}$, $x_{e,j,t}$, $b_{j,t+1}$, prices $\tilde{p}^a_t$, W_t, P_t such that, taking prices as given,

- N_t, $c_{j,t}$, A_t solve household problem of maximizing $\sum_{t=1}^{\infty}\beta^t \mathbb{U}(C_t, N_t)$ subject to the budget constraint (3.3), where $\mathbb{U}$ is given by (3.1).
- V, $\tilde{V}$ $p_{j,t}$, $x_{j,t}$, $b_{j,t+1}$ solve incumbent problem (3.18).
- V_e, $x_{e,j,t}$ solve entrant problem (3.22).
- The labor market clears, $N_t = \int_{j\in\mathscr{I}_t}(n_{c,j,t} + n_{r,j,t})\,\mathrm{d}i$.
- The asset market clears, $A_t = \int_{j\in\mathscr{I}_t} V_j\,\mathrm{d}i$.
- The goods market clears at price $P_t \equiv 1$ by Walras law.

3.3.5. Parametrization. In this section we discuss how we bring the model to the data. Let us first describe the data being used and then the procedure with which we assign values to structural parameters of the model.

3.3.5.1. *Data.* In what follows, we use two primary data sources: Compustat and the Business Dynamics Statistics (BDS). The former has important information pertaining to research and development activities and to markups. However, it is not a representative sample of firms. The latter, on the other hand, contains basic information on effectively the universe of U.S. employers.

For information regarding research and development activities and markups we use Compustat. Specifically, we use all non-financial public companies in the U.S. economy over the period between 1977 and 2016.[3] In order to compute markups at the firm-level, we use the methodology of De Loecker et al. (2018).

The BDS dataset is used for information regarding firm dynamics. Specifically, we use it for life-cycle profiles of firm employment and exit rates. Importantly, because Compustat is a highly selected sample of firms, we also use the BDS to reweigh the observations in Compustat to reflect the population weights. The weights attached to each observation correspond to a number of similar firms in the whole economy. We use weights based on the firm size. That is, we approximate the weight of each firm i by the number of firms in the U.S. economy with the same number of employees.[4]

To approximate the size distribution of firms in the economy with the help of the BDS data, we fit a set of Pareto distributions to each size bin in the BDS dataset. We choose the shape of each fitted Pareto distribution to match mean firm size in the bin.[5]

Finally, let us note that growth rates of all firm-level variables are calculated according to the Davis et al. (1996) (DHS) convention. That is, in order to correct for a mean-reversion bias in our estimates, we calculate each outcome variable in period t as a simple average of the outcomes in periods t and $t-1$. For any variable X_t, its growth rate in period t reads

$$\Delta X_t = \frac{X_t - X_{t-1}}{0.5 * (X_t + X_{t-1})} \tag{3.27}$$

3.3.5.2. *Parameter values.* In this subsection we describe the parametrization of the model. We begin by a set of parameters related to the household and then we move on to the set of parameters which govern firm-specific outcomes. Here, for convenience of the exposition, we first describe parameters which are common to all businesses and then we move on to those related to different firm types. All externally calibrated parameters, their values, and targets are in Table 3.1. Estimated parameters, together with targeted and model implied moments are presented in Table 3.2.

Following the annual frequency of the BDS, we choose one period in the model to be one year. Therefore, we set the discount factor to $\beta = 0.97$. The elasticity of substitution between goods varieties is set to $\eta = 4$ as is standard in the literature. This value implies a 33% markup in the case of fixed demand. The disutility of labor, υ, is set such that wages are normalized to $W = 1$. Finally, the customer base is assumed to depreciate at a rate of $\delta = 0.2$. This is within the range used in Gourio and Rudanko (2014).

Structural parameters which are common to all firms include the elasticity of the innovation probability with respect to R&D expenditure (ψ), the innovation rate (λ), the exogenous rate of firm exit (ρ), entry cost (ϕ_e), the rate of technological diffusion (x_0), the mean and

[3]This range is dictated by the BDS. We drop all companies in utilities and finance. We keep only companies incorporated in the U.S. and traded in the U.S. stock exchange. We remove observations with negative values for employment, sales or R&D expenses. We remove observation if for a given firm-year pair acquisition constitute more that 10% of the total assets. We remove top 1% observation with the largest R&D expenses – some values are extremely large indicating reporting issues.

[4]Ideally, we would use a register of firms in order to find the number of firms of particular size in the population of firms. Since we do not have access to one, we use the BDS data set which contains numbers of firms across several size bins.

[5]In some cases the mean size within a BDS bin falls below the lower end of the size bin indicating reporting issues. In these cases, we use Pareto coefficient of 0.72, which a value estimated in Kondo et al. (2018) using the universe of firms in the U.S.

TABLE 3.1. Externally calibrated parameters

Parameter		Value	Target
I. households			
β	discount factor	0.97	real interest of 3%
ν	labor disutility	1	normalization $W_t = 1$
η	elasticity of demand	4	static markup of 33%
II. firms			
δ	customer base depreciation	0.2	Gourio and Rudanko (2014)
ψ	elasticity of innovation prob.	2	Akcigit and Kerr (2018)
ρ	exogenous exit rate	0	normalization
ϕ_e	entry cost	2.27	normalization $C_t = 1$
b_e	initial customer base for entrants	1	normalization

Notes: Starting from the leftmost column, the table presents the symbol of a parameter, its meaning, assigned value, and the calibration target.

dispersion of operational costs (μ_H and σ_H). While $\psi = 2$ is taken from the literature (see e.g. Akcigit and Kerr 2018), firms are assumed to exit only endogenously, i.e. $\rho = 0$ and entry cost is set such that aggregate output is normalized to 1. The remaining parameters are set such that the model matches the annual growth rate of labor productivity of 1.5%, and the average exit rate of all and young (less than six years old) businesses. The rate of technological diffusion is important for life-cycle growth and is pinned down, together with other parameters described below, they the life-cycle profile of average employment.

We assume that each firm differs in its (permanent) ability to (i) conduct R&D and (ii) accumulate demand. These two characteristics therefore jointly determine four firm types. Specifically, a firm could either be a "non-R&D" firm or an "R&D" firm. This is determined by the value of γ_i, the efficiency of R&D in equation (3.8), which takes on two values. In addition, firms can either be "small" or "large", which is governed by $\alpha_{1,i}$, the elasticity of demand with respect to the customer base. Once again, $\alpha_{1,i}$ takes on two values. In addition to the elasticity of demand, $\alpha_{1,i}$, the level of demand is influenced by the scale parameter α_0. Moreover, related to firm types is initial entrant distribution, $P_{(\alpha,\gamma)}$.

We parameterize these four parameters (γ, $\alpha_{1,S}$, $\alpha_{1,L}$ and α_0) to match the life-cycle profile of average employment and the aggregate R&D intensity.[6] While α_0 is closely related to average employment, the convex shape of the life-cycle employment profile then pins down the values of $\alpha_{1,S}$ and $\alpha_{1,L}$. Type-specific entry costs are then informative about the firm shares in each type. We parametrize these in order to match firm shares of small, large firms (not) conducting R&D. We choose to define small firms as those with fewer than 250 workers. This results in roughly equal employment shares of small and large firms in the BDS.

[6]The R&D efficiency of non-R&D firms is a normalization s.t. R&D expenditures are 0. R&D intensity is defined as R&D expenditures relative to sales.

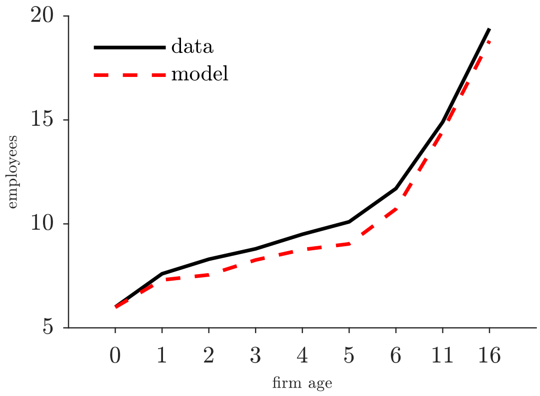
TABLE 3.2. Internally estimated parameters

Parameter		Value	Target	Data	Model
I. common to all firms					
λ	innovation rate	6.30%	labor prod. annual growth	1.5%	1.5%
μ_H	mean operational costs	1.90	mean exit rate	8.6%	9.1%
σ_H	std. operational costs	15.32	exit rate of firms < 6 years	15.2%	11.0%
x_0	technology diffusion	9.55%	firm lifecycle dynamics	see Fig. 3.2	
II. idiosyncratic					
γ	R&D efficacy, small R&D	0.10	mean R&D/sales	10.4%	9.7%
$\alpha_{1,S}$	customer base elas., small	0.15	firm lifecycle dynamics	see Fig. 3.2	
$\alpha_{1,L}$	customer base elas., large	0.54	firm lifecycle dynamics	see Fig. 3.2	
α_0	demand scaling	14.13	firm lifecycle dynamics	see Fig. 3.2	
P_{α_S}	entry share, small non-R&D	0.445	respective firm share	36.6%	36.0%
P_{α_L}	entry share, large non-R&D	0.0005	respective firm share	0.1%	0.1%
$P_{\alpha_S,\gamma}$	entry small, small R&D	0.55	respective firm share	62.5%	63.6%

Notes: Starting from the leftmost column, the table presents the symbol of a parameter, its meaning, estimated value, description of the targeted moment, followed by value of the targeted moment in the data, and the corresponding value in the model.

R&D conducting firms are those that continuously report having positive R&D expenditures in Compustat.

FIGURE 3.1. Firm lifecycle profile of average employment: data and model



Notes: Average firm size by age, where 0 refers to startups, 6 refers to 6-10 year old firms, 11 refers to 11-15 year old firms and 16 refers to 16-20 year old firms. Data is taken from the BDS.

3.3.5.3. *Restricted model.* To highlight the importance of demand-driven growth, we will make use of a counterfactual analysis. Specifically, we will compare our baseline model to one in which firms cannot accumulate demand. On the contrary, all businesses face fixed levels of demand, b_i. Apart from this feature, the rest of this restricted model version is exactly the same as the baseline.

Our primary comparison will be a restricted version of the model in which firms face the average levels of *type-specific* demand from the baseline model. Note, however, that considering instead the stationary distribution of demand levels from the baseline changes very little. Importantly, with fixed demand there is no incentive to vary markups and therefore they are constant at $\eta/(\eta-1)$.

Finally, there are two versions of the restricted model which we will consider. First, we will consider the restricted model *without* recalibrating it. This means that we literally take the baseline model and replace firm-specific demand (and investments into it) with fixed demand values. We will use this version to show the importance of the demand channel for growth.

However, one may think that it is possible to recalibrate the restricted version of the model to still match all the observable targets. This is indeed the case and we consider this version of the restricted model as well. However, as we will argue, this *recalibrated* version of the restricted model exhibits markedly different properties from our baseline.

3.3.5.4. *Untargeted moments.* Before moving on to the main results, this subsection shows that the model does well on a range of untargeted moments relevant for our mechanism.

Demand. First, let us focus on characteristics pertaining to demand accumulation and markup dynamics. While Compustat does not allow one to measure consistently firm age, there is evidence that markups increase with age (Peters 2019; Hosono et al. 2020). Figure 3.2 shows that markups are indeed increasing in firm age in the model. The intuition for this is straightforward. Firms enter, on average, small and need to accumulate demand in order to grow. They achieve this by charging low markups initially and increasing them as they age and broaden their customer base.[7]

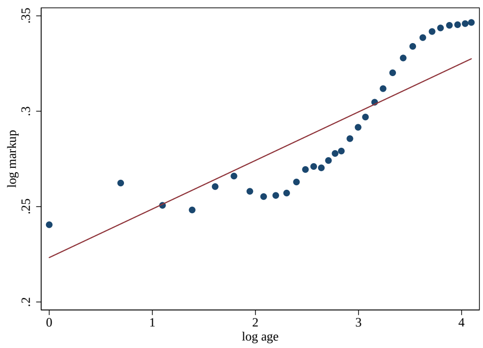
Demand and R&D.. We now move on to features of the model related to the interaction between demand and research and development. Towards this end, we estimate the following regressions, both in the data and, in the model for each firm f in year t

$$\begin{aligned}\log\left(\mu_{f,t}\right) = \alpha_0 \log(\text{emp}_{f,t}) + \alpha_1 \log(\text{emp}_{f,t}) \times \mathbf{1}_{\text{R\&D}} \\ + \alpha_2 \log(\text{age}_{f,t}) + \alpha_3 \log(\text{age}_{f,t}) \times \mathbf{1}_{\text{R\&D}} \\ + \boldsymbol{\beta}\Gamma_{f,t} + \delta_j + \delta_t + \varepsilon_{f,t},\end{aligned} \tag{3.28}$$

where the dependent variable μ is the firm-level markup, "emp" stands for firm size (total employment) and "age" for the number of years in the sample. $\times \mathbf{1}_{\text{R\&D}}$ indicates an interaction with dummy variables equal to one for R&D firms and zero otherwise. $\Gamma_{f,t}$ stands for controls, including measures of size (book value of total assets and sales) and leverage (short-term and long-term debt divided by total equity). δ_j and δ_t capture industry and year fixed effects,

[7] This positive relationship between markups and age is observed for both non-R&D and R&D conducting firms.

FIGURE 3.2. Relationship markup and age in the model



Notes: The figure visualizes the relationship between age and markups in the stationary equilibrium using binned scatterplots (Stepner 2013; Chetty et al. 2014). That is, we group firm age into 40 equal-sized bins, compute the mean of age and markups within each bin, then create a scatterplot of these data points. In addition, we plot a linear fit line using OLS.

respectively. The errors are clustered at the level of the individual firm industry. We estimate equivalent regression on the model-simulated data.[8]

The Table 3.3 illustrates that the model captures the relationship between markups, firm size and age relatively well, including the respective differences between R&D-conducting and non-R&D firms. In particular, both in the data and in the model markups are decreasing with size, but increasing with age and R&D conducting firms can afford to charge higher markups albeit these decline with their age.

Importance of non-R&D firms. Finally, to highlight the role of non-R&D conducting firms, we describe the dynamics of employment and productivity growth at the firm level.[9] Table 3.4 shows the correlations between size and productivity growth in the data, baseline model and in the recalibrated restricted model. As can be seen from the table, the correlation between labor productivity growth and employment growth at the firm level is far from perfect (0.32). In the restricted model, despite the fact that it is recalibrated to match the same moments as the baseline, including the life-cycle pattern of size, the correlation between labor productivity and size growth is essentially equal to 1. The intuition is simple. Firms grow only if they manage to innovate. On the contrary, in the baseline model, firms can grow also by attracting new customers. The baseline model does very well in matching the data in this sense, despite it not being part of the parametrization.

3.4. Results

This section presents our main results. We begin with quantifying the importance of our demand-channel for aggregate growth. Towards this end, we turn to comparisons of our

[8]In the model we abstract from year or industry effects as well as liquidity. In the model we can measure firm age directly.

[9]In Compustat we measure labor productivity as sales to employment ratio.

TABLE 3.3. Markups, size, age and R&D

	$\log(\mu)$	
	Compustat	Model
log(emp)	-0.0120	-0.339
	(0.0119)	
R&D×log(emp)	0.0384***	0.0168
	(0.00825)	
log(age)	0.0529***	0.0200
	(0.0133)	
R&D×log(age)	-0.0907***	-0.0111
	(0.0173)	
Observations	29706	751446
Adjusted R^2	0.197	0.794

Notes: Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. We do not report standard errors implied by the model simulated data since there is no notion of measurement uncertainty in the model. Markups follow the baseline specification in De Loecker et al. (2018). Age in Compustat is measured as the number of years in the sample.

TABLE 3.4. Correlation between productivity and size growth rates

	data	baseline	restricted
$corr(\Delta q, \Delta n)$	0.32	0.40	0.99

Notes: Table presents correlations between growth rates of firm size and labor productivity in the model and in Compustat. In Compustat, labor productivity q is measured as the ratio of sales to the total number of employees.

baseline model with that of the restricted version which assumes fixed demand levels at the firm level. For the purpose of this comparison, we do not recalibrate the restricted model. Instead, we keep all the structural parameters at the values they hold in the baseline model and simply replace firm-level demand with the type-specific averages observed in the stationary distribution of the baseline model.[10]

After quantifying the extent to which growth is demand-driven, and understanding the channels through which this happens, we turn to highlighting the importance of our channel for other questions. In particular, we document that our model framework behaves differently in important dimensions compared to endogenous growth model which do not consider endogenous demand accumulation.

[10]This means that we retain the structure of four firm types. As mentioned earlier, using instead the whole stationary distribution of firm-level demand, i.e. extending the number of firm types, changes very little in terms of results. More details can be found in the Appendix 3.A.

3.4.1. Demand-driven growth. Table 3.5 depicts our main result. Demand drives about 54% of aggregate growth. The remaining 46% can be further decomposed into contributions of (i) selection, (ii) creative destruction and (iii) incumbents' R&D efforts.

TABLE 3.5. Sources of aggregate growth

	percentage points	share of aggregate (%)
aggregate growth (%)	1.5	100
of which		
- *demand*	0.81	54
- *selection*	-0.002	-0.1
- *creative destruction*	0.06	3.9
- *incumbent innovation*	0.63	42.2

Notes: The table presents a decomposition of the aggregate growth rate into various channels. The demand channels corresponds to a difference between growth rate in the baseline economy and a counterfactual the economy in which firms' customer base is fixed. The selection channel corresponds to a share of aggregate growth that can be attributed to a difference in distribution of firms types between stationary equilibrium and at the time of entry. Creative destruction is the productivity gain stemming from the innovations of entrants that replace exiting incumbents.

The importance of firm selection (i.e. between types of firms) was highlighted by Lentz and Mortensen (2005). The same mechanism is present also in our model. Because firms differ in their (long-run) ability to accumulate demand and do R&D, they also differ in their survival rates. Therefore, as firms age, resources get reallocated from less productive, smaller businesses, towards more productive and larger firms. However, as will become clear below R&D conducting firms have higher exit rates in our counterfactual economy with fixed demand. Therefore, the selection channel actually *decreases* growth slightly.

Note, however, that this does not imply that selection is inconsequential in our model. On the contrary and as will be discussed in more detail below, firm selection is important for aggregate growth. Specifically, the share of R&D conducting firms rises from about 56% at startup to 64% in the stationary distribution. This precisely highlights the fact that resources get reallocated towards more productive, R&D-conducting, firms. Our results suggest, however, that this process is entirely demand-driven in our baseline economy. Without the option to accumulate demand, firm selection effectively stops.

Another source of aggregate growth in our baseline model is creative destruction - the process by which relatively unproductive incumbents give way to new businesses. In our baseline model, only about 4% of overall growth can be attributed to this channel. This again contrasts existing studies which show a considerably larger contribution.[11] There are two important difference of our framework compared to existing studies. First, not all businesses conduct research and development in our model. Specifically, about 40 percent of all firms do not do any R&D. Second, as with firm selection, creative destruction as well is partly driven by demand which is not captured in our decomposition exercise.

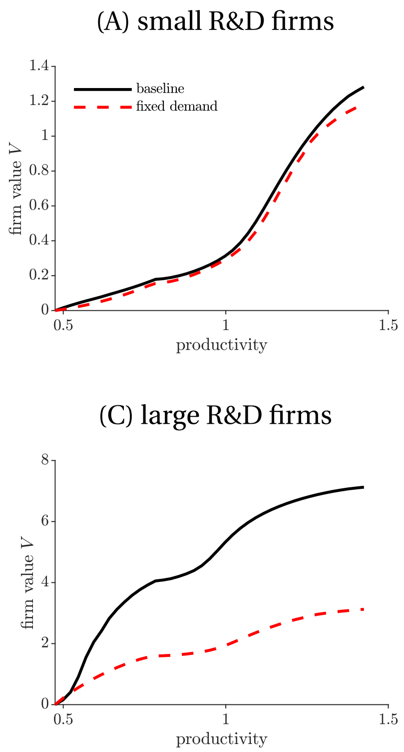
[11] For instance, Lentz and Mortensen (2005) suggest that about a quarter of aggregate growth can be attributed to creative destruction.

Finally, the second largest contributor to aggregate growth are the efforts of incumbent businesses in doing R&D. These alone account for about 42% of aggregate growth.

3.4.1.1. *Underlying channels.* Let us now turn to further analyzing how the option to accumulate demand affects aggregate growth. Table 3.6 shows several margins through which endogenous demand accumulation impacts aggregate growth. The table shows values for these margins in the baseline and the restricted models as well as broken down by firms (not) conducting R&D. There are three important channels at play.

First, the top panel of Table 3.6 shows that forcing demand to be fixed raises overall firm exit. This is because firm values drop, as shown in Figure 3.3. Importantly, the loss of the potential to accumulate demand is, unsurprisingly, particularly painful for large businesses which require sizeable customer bases to operate efficiently.

FIGURE 3.3. Firm values in the baseline model and the restricted model with fixed demand.



Notes: Each panel plots the natural logarithm of the firm value as a function of productivity $q_{j,t}$. Solid black line corresponds to the baseline model and dashed red line to the fixed demand counterfactual. The four panels correspond to four types of firms. The values in the baseline model are weighted averages over all admissible values of customer base where weights are given by distribution of customer base in the stationary distribution. In the fixed demand scenario, there is only one admissible value of customer base for each firm type. Value functions in all panels are re-scaled such that the value at the lowest productivity level is zero and hence the values in the figure should be interpreted as a percentage point increase relative to the value at the lowest productivity point.

The increase in overall firm exit means that as relatively more businesses give way to new entrants, the creative destruction channel of growth strengthens. This, all else equal, would have the potential to *increase* aggregate growth in the restricted model.

Second, there is a strong composition effect. Notice from the lower two panels of Table 3.6 that the increase in firm exit is almost entirely driven by R&D conducting firms. This fact has important implications for firm selection. Specifically, the fact that R&D conducting firms shut down relatively more often in the restricted version of the model means that their firm share is lower compared to the baseline (it drops from about 64% to 53%). Therefore, the selection channel of firm growth is weakened in the restricted model.

Third, the rate at which firms innovate drops dramatically as well. Specifically, because the prospects of higher demand in the future is absent in the restricted model, incentives to innovate also slump. The probability of successfully innovating goes down from 21% in the baseline to just 2% in the restricted model.[12] Hence, the contribution of incumbents innovation to growth is weaker.

TABLE 3.6. Firm-level outcomes in the baseline model and fixed demand counterfactual

	baseline	restricted
aggregate growth (%)	1.5	0.7
	All firms	
exit rate (%)	8.6	10.9
R&D prob., x (%)	13.8	1.3
	Non-R&D firms	
exit rate (%)	11.3	11.3
entrant firm share (%)	44.5	44.5
stationary firm share (%)	36.1	46.7
	R&D firms	
exit rate (%)	7.1	10.7
entrant firm share (%)	55.5	55.5
stationary firm share (%)	63.9	53.3
R&D prob. (%)	21.5	2.3

Notes: Comparison between the baseline model and the model in which the customer base is fixed. The fixed demand counterfactual is not re-calibrated.

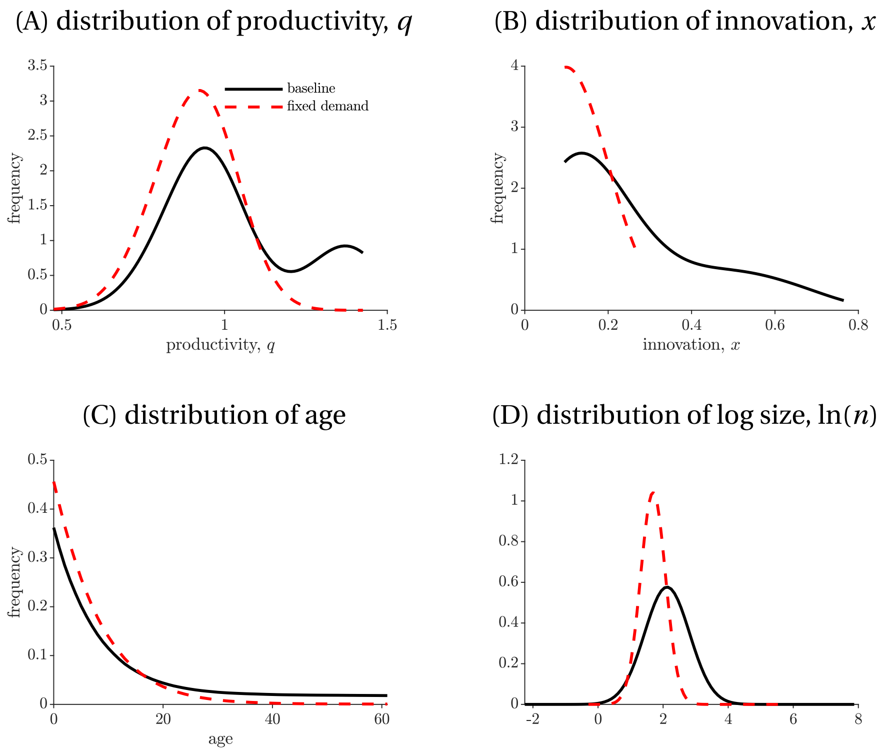
As a result of the above effects the equilibrium distribution of firms shifts towards less productive, less innovative, younger and smaller businesses. This is depicted in Figure 3.4 where panel (A), we present the stationary distribution of idiosyncratic productivity $q_{j,t}$ under the two scenarios.[13]. Panel (B) indicates that mass of firms shifts towards less innovative

[12]Recall that there is a rate of technological diffusion, x_0, with which all (including non-R&D) firms obtain productivity improvements. This is what drives much of the growth in the restricted version of the model.

[13]The figure presents (normal) kernel density estimator (with bandwidth 0.1) of the stationary distribution of productivity implied by simulated data.

units. Panel (C) illustrates that firms are on average younger, confirming the rising exit rates in Table 3.6. Finally, panel (D) shows that firm have become smaller on average. Decline in firm size is a result of higher exit rates, leading to younger firm population, and lower incentives to innovate.

FIGURE 3.4. In the fixed demand scenario, the equilibrium distribution of firms shifts towards less productive units.



Notes: The figure presents the stationary distribution in the baseline model (black solid) and the fixed demand counterfactual (red dashed). The figure presents the distribution fitted using kernel estimator with bandwidth 0.1 with the exception of size distribution which presents a fitted log-normal distribution.

3.4.2. Discussion. The previous subsection highlighted that endogenous demand accumulation is important for understanding the sources of aggregate growth. In particular, more than half of aggregate growth is demand-driven. However, the previous exercises were decompositions which "shut down" the demand channel without recalibrating any other model parameters. One may wonder whether our understanding of the macroeconomy changes fundamentally if we were to recalibrate the restricted version of the model to match the same moments as the baseline economy.

The answer to the above question is yes. The key difference between the baseline and the restricted model (re-calibrated to match the same moments as the baseline economy) lies in the distribution of firm values. In particular, the distribution of firm values is much more condensed in the restricted model, owing to the fact that firms cannot expand via demand accumulation.

This features is similar in spirit to Pugsley et al. (2019). The latter studies the difference in heterogeneous firm models which match the same observable moments, but are based on different underlying structural shocks. The authors find that the distribution of firm values is crucial for the responsiveness of the economy to policies and frictions.

Therefore, our results suggest that a promising avenue for future research may lie in the analysis of the impact of demand (accumulation) on the efficacy of growth policies and/or frictions disrupting it.

3.5. Conclusion

In this paper we argue that growth, a traditionally supply-side concept, is to a large extent demand-driven. Specifically, we build a tractable endogenous growth model with heterogeneous firms in which businesses can grow either because they become more productive or because they manage to attract more customers. While the former has been the subject of study of the growth literature, the latter has been the domain if research into firm and industry dynamics. Crucially, allowing for endogenous investment into demand *and* productivity creates a powerful feedback loop. Higher productivity allows firms to sell at lower prices. Lower prices attract more customers. And finally a larger customer base raises the returns from investment into productivity.

Our baseline model suggests not only that growth is to a large extent demand driven, but they also raise the possibility that neglecting this demand channel can change our udnerstanding of the efficacy or R&D subsidies or the impact of frictions on aggregate growth. The reason for this is that the baseline model and the restricted version with fixed demand are characterized by a vastly different distribution of firm values. This is despite being parametrized to match the same set of observable moments. This difference is important because firm values determine (forward-looking) choices of businesses and as such impact the responsiveness of the economy to policies and disruptions. This line of research we, however, leave for the future.

Appendices to Chapter 3

3.A. Details on the restricted demand model

3.A.1. Restricted model. To quantify the importance of demand accumulation for aggregate growth, we consider a restricted version of the model in which the customer base is time-invariant. We keep all the structural parameters at the values they hold in the baseline model and simply replace firm-level demand with the type-specific averages observed in the stationary distribution of the baseline model. These averages are $b = 20$ for small firms and $b = 1350$ for large firms.

Table 3.7 compares moments targeted in the estimation procedure in the data, baseline model, and restricted model. Since the customer base does not accumulate, firms tend to be smaller, less innovative, and more likely to exit.

TABLE 3.7. Externally calibrated parameters, the restricted model

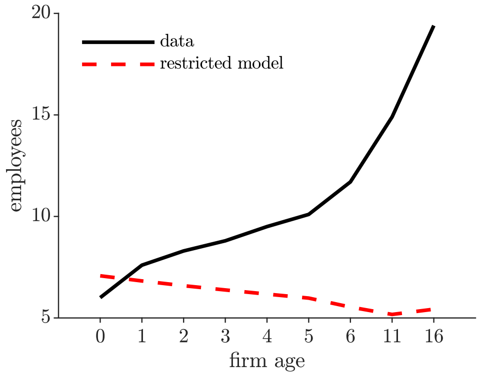
		Model	
Target	Data	Baseline	Restricted
labor prod. annual growth	1.5%	1.5%	0.7%
exit rate of firms < 6 years	15.2%	11.0%	10.7%
mean exit rate	8.6%	9.1%	11.4%
mean R&D/sales	10.4%	9.7%	3.8%
firm share, small non-R&D	36.6%	36.0%	45.8%
firm share, large non-R&D	0.1%	0.1%	0.2%
firm share, small R&D	62.5%	63.6%	53.2%

Notes: Starting from the leftmost column, the table presents the symbol of a parameter, its meaning, assigned value, and the calibration target.

Figure 3.5 presents the lifecycle profile of firm size in the restricted model and in the BDS data. Since the incentives to innovation are severely diminished, and the customer base does not accumulate, firms tend to get smaller as they age.

3.A.2. Re-calibrated restricted model. We also study a version of the baseline model that in addition to constraining the customer base to time-invariant values $b = 20$ for small firms and $b = 1350$ for large firms, we re-calibrate the remaining parameters to match the same moments in the data as in the case of baseline calibration. Table 3.8 presents calibrated parameters, Table 3.9 estimated parameters and targeted moments in the data followed by corresponding values in the model. Figure 3.6 presents age profile of the average firm size.

FIGURE 3.5. Firm lifecycle profile of average employment: data and restricted model



Notes: Average firm size by age, where 0 refers to startups, 6 refers to 6-10 year old firms, 11 refers to 11-15 year old firms and 16 refers to 16-20 year old firms. Data is taken from the BDS.

TABLE 3.8. Externally calibrated parameters, the re-calibrated restricted model

Parameter		Value	Target
I. households			
β	discount factor	0.97	real interest of 3%
ν	labor disutility	1	normalization $W_t = 1$
η	elasticity of demand	4	static markup of 33%
II. firms			
δ	customer base depreciation	0.2	Gourio and Rudanko (2014)
ψ	elasticity of innovation prob.	2	Akcigit and Kerr (2018)
ρ	exogenous exit rate	0	normalization
ϕ_e	entry cost	0.7	normalization $C_t = 1$
b_e	initial customer base for entrants	1	normalization

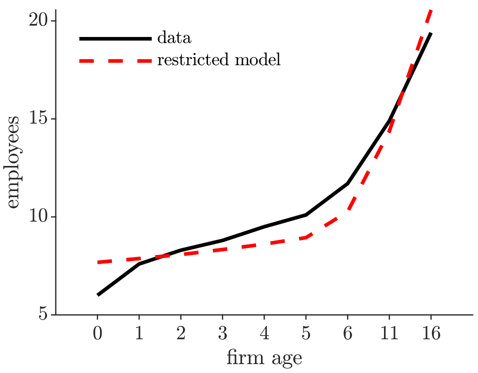
Notes: Starting from the leftmost column, the table presents the symbol of a parameter, its meaning, assigned value, and the calibration target.

TABLE 3.9. Internally estimated parameters, the re-calibrated restricted model

Parameter		Value	Target	Data	Model
I. common to all firms					
λ	innovation rate	2.90%	labor prod. annual growth	1.5%	1.5%
μ_H	mean operational costs	1.75	exit rate of firms < 6 years	15.2%	11.0%
σ_H	std. operational costs	16.07	mean exit rate	8.6%	9.1%
x_0	technology diffusion	50.00%	firm lifecycle dynamics	see Fig. 3.1	
II. idiosyncratic					
γ	R&D efficacy	0.3	mean R&D/sales	10.4%	9.7%
α_1	customer base elas., large	1.00	fixed demand		
α_0	demand scaling	1.00	fixed demand		
P_{α_S}	entry share, small non-R&D	0.33	respective firm share	36.6%	36.0%
P_{α_L}	entry share, large non-R&D	0.0003	respective firm share	0.1%	0.1%
$P_{\alpha_{S,\gamma}}$	entry share, small R&D	0.66	respective firm share	62.5%	63.6%

Notes: Starting from the leftmost column, the table presents the symbol of a parameter, its meaning, estimated value, description of the targeted moment, followed by value of the targeted moment in the data, and the corresponding value in the model.

FIGURE 3.6. Firm lifecycle profile of average employment: data and the re-calibrated restricted model



Notes: Average firm size by age, where 0 refers to startups, 6 refers to 6-10 year old firms, 11 refers to 11-15 year old firms and 16 refers to 16-20 year old firms. Data is taken from the BDS.

3.B. Numerical appendix

We can simplify and re-write the firm's problem as follows

$$V(b,q) = \max_{b',x} \Phi(b,q) \times \left\{ \frac{W}{q} C \left[\left(\frac{b'}{(1-\delta)} - b \right)^{1-\frac{1}{\eta}} \left(\frac{q}{W} \right) b^{\alpha/\eta} - \frac{b'}{(1-\delta)} + b \right] \right.$$

$$\left. - W \left(\frac{x^{\psi}}{\gamma} + \hat{\phi}(b,q) \right) + \tilde{\beta} \left(x V(b', q^{+}) + (1-x) V(b', q) \right) \right\}$$

s.t.

$$x \in [0, 1-x_0], \quad (1-\delta)b - b' \leq 0 \tag{3.29}$$

where primes ′ denote next period variables and we leave implicit the dependence of all variables on firm type j. There are two state variables of the firm problem: productivity q and customer base b. We parameterize the grid for productivity using 41 points. Values of productivity q are normalized such that $q = 1$ corresponds to average productivity in the economy. The grid for customer base uses 61 points. Both grids are equally-spaced. Moreover, both grids are type-specific. The upper and lower bounds of the grids corresponding to each type are choses as follows. The upper bound of productivity grid for non-R&D firm is $q = 1$, since these firms enter at the average productivity $q = 1$ and afterwards slide downwards as the economy grows. The lower bound is set such that it takes a non-R&D firm 61 years to reach the lower bound of the grid, $q = 0.475$. The q grid for R&D firms uses the same lower bound. The upper bound is set at three times the lower bound value, in line with evidence on productivity dispersion Bartelsman and Doms (2000). The lower bound on grid of customer base corresponds to $b_e = 1$. The upper bound is set such that the largest of small firms has 250 employees, and the largest of large firms 10000 employees.

We solve the firm problem using value function iteration. Towards this end, we begin with an initial guess for $V(b,q)$ for each state (b,q) and solve firm problem (3.29) given the current guess for V. The maximization problem yields policy $b'(b,q)$, $x(b,q)$. This allows us to derive the survival probabilities $\Phi(b,q)$ using equation (3.17) as well as pricing policy

$$\mu(b,q) = \left(\frac{b'}{(1-\delta)} - b\right)^{-1/\eta} \left(\frac{q}{W}\right) b^{\alpha/\eta}.$$

The optimal policy, in turn, allows as to calculate the implied new value function $V(b,q)$ for each state (b,q). We iterate this procedure until convergence.

To approximate the stationary firm distribution, we simulate 75 thousand firms. We follow each firm until it decides to exit, but no longer than 61 years. Given the approximated stationary firm distribution, we compute all aggregate variables.

Bibliography

Ábrahám, Árpád, João Brogueira de Sousa, Ramón Marimon, and Lukas Mayr. 2017. *On the Design of a European Unemployment Insurance Mechanism.* Draft. European University Institute, Florence.

Acemoglu, Daron. 2002. "Technical Change, Inequality, and the Labor Market". *Journal of Economic Literature* 40 (1): 7–72.

Acemoglu, Daron, Ufuk Akcigit, Harun Alp, Nicholas Bloom, and William Kerr. 2018. "Innovation, reallocation, and growth". *American Economic Review* 108 (11): 3450–3491.

Acemoglu, Daron, and David H. Autor. 2011. *Skills, tasks and technologies: Implications for employment and earnings,* 4:1043–1171. PART B.

Aghion, Philippe, Ufuk Akcigit, and Peter Howitt. 2014. "What do we learn from Schumpeterian growth theory?" In *Handbook of economic growth,* 2:515–563. Elsevier.

Aiyagari, S. Rao. 1994. "Uninsured Idiosyncratic Risk and Aggregate Saving". *The Quarterly Journal of Economics* 109 (3): 659–684.

Akcigit, Ufuk, and Sina T Ates. 2019. *Ten Facts on Declining Business Dynamism and Lessons from Endogenous Growth Theory.* Working Paper, Working Paper Series 25755. National Bureau of Economic Research.

Akcigit, Ufuk, and William R Kerr. 2018. "Growth through heterogeneous innovations". *Journal of Political Economy* 126 (4): 1374–1443.

Akcigit, Ufuk, and Tom Nicholas. 2019. "History, Microdata, and Endogenous Growth". *Annual Review of Economics* 11:615–633.

Albuquerque, Rui, and Hugo A Hopenhayn. 2004. "Optimal lending contracts and firm dynamics". *The Review of Economic Studies* 71 (2): 285–315.

Alesina, Alberto, and Romain Wacziarg. 1999. "Is Europe going too far?" *Carnegie-Rochester Conference Series on Public Policy* 51:1–42.

Andreasen, Martin M, Jesús Fernández-Villaverde, and Juan F Rubio-Ramírez. 2018. "The Pruned State-Space System for Non-Linear DSGE Models: Theory and Empirical Applications". *The Review of Economic Studies* 85 (1): 1–49.

Anzoategui, Diego, Diego Comin, Mark Gertler, and Joseba Martinez. 2019. "Endogenous technology adoption and R&D as sources of business cycle persistence". *American Economic Journal: Macroeconomics* 11 (3): 67–110.

Arkolakis, Costas, Theodore Papageorgiou, and Olga A Timoshenko. 2018. "Firm learning and growth". *Review of Economic Dynamics* 27:146–168.

Asdrubali, Pierfederico, Bent E. Sørensen, and Oved Yosha. 1996. "Channels of Interstate Risk Sharing: United States 1963–1990". *The Quarterly Journal of Economics* 111 (4): 1081–1110.

Atkeson, Andrew. 1991. "International Lending with Moral Hazard and Risk of Repudiation". *Econometrica* 59 (4): 1069–1089.

Atkeson, Andrew, and Patrick J Kehoe. 2005. "Modeling and measuring organization capital". *Journal of political Economy* 113 (5): 1026–1053.

Autor, David H., David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen. 2019. "The Fall of the Labor Share and the Rise of Superstar Firms".

Autor, David H., Frank Levy, and Richard J. Murnane. 2003. "The Skill Content of Recent Technological Change: An Empirical Exploration". *The Quarterly Journal of Economics* 118 (4): 1279–1333.

Bachmann, Rüdiger, and Christian Bayer. 2013. "'Wait-and-See' business cycles?" *Journal of Monetary Economics* 60 (6): 704–719.

— . 2014. "Investment dispersion and the business cycle". *American Economic Review* 104 (4): 1392–1416.

Baily, Martin N. 1978. "Some aspects of optimal unemployment insurance". *Journal of Public Economics* 10 (3): 379–402.

Bajgar, Matej, Giuseppe Berlingieri, Sara Calligaris, Chiara Criscuolo, and Jonathan Timmis. 2019. "Industry Concentration in Europe and North America". (Paris), OECD Productivity Working Papers, no. 18.

Bartik, Timothy. 1991. *Who Benefits from State and Local Economic Development Policies?* Tech. rep. W.E.Upjohn Institute.

Bayer, Christian, and Moritz Kuhn. 2018. "Which ladder to climb? Wages of workers by job, plant, and education".

Beblavý, Miroslav, Gabriele Marconi, and Ilaria Maselli. 2015. *A European Unemployment Benefits Scheme: The rationale and the challenges ahead.* CEPS Papers 10952. Centre for European Policy Studies.

Bénassy-Quéré, Agnès, et al. 2018. *Reconciling risk sharing with market discipline: A constructive approach to euro area reform.* CEPR POLICY INSIGHT NO. 91. Centre for Economic Policy Research.

Benigno, Gianluca, and Luca Fornaro. 2018. "Stagnation traps". *The Review of Economic Studies* 85 (3): 1425–1470.

Bianchi, Francesco, Howard Kung, and Gonzalo Morales. 2019. "Growth, slowdowns, and recoveries". *Journal of Monetary Economics* 101:47–63.

Birinci, Serdar, and Kurt See. 2019. *How Should Unemployment Insurance Vary over the Business Cycle?* Working Papers 2019-022. Federal Reserve Bank of St. Louis.

Blanas, Sotiris, Gino Gancia, and Sang Yoon Tim Lee. 2019. "Who Is Afraid of Machines". *Economic Policy*.

Buera, Francisco J. 2009. "A dynamic model of entrepreneurship with borrowing constraints: theory and evidence". *Annals of finance* 5 (3-4): 443–464.

Buera, Francisco J, and Yongseok Shin. 2011. "Self-insurance vs. self-financing: A welfare analysis of the persistence of shocks". *Journal of Economic Theory* 146 (3): 845–862.

Cabral, Luis, and Jose Mata. 2003. "On the evolution of the firm size distribution: Facts and theory". *American economic review* 93 (4): 1075–1090.

Cagetti, Marco, and Mariacristina De Nardi. 2006. "Entrepreneurship, Frictions, and Wealth". *Journal of Political Economy* 114 (5): 835–870.

Cahuc, Pierre, Stéphane Carcillo, and Thomas Le Barbanchon. 2018. "The Effectiveness of Hiring Credits". *The Review of Economic Studies* 86, no. 2 (): 593–626.

Caicedo, Santiago, Robert E Jr. Lucas, and Esteban Rossi-Hansberg. 2019. "Learning, Career Paths, and the Distribution of Wages". *American Economic Journal: Macroeconomics* 11, no. 1 (): 49–88.

Calvino, Flavio, Chiara Criscuolo, and Carlo Menon. 2015. "Cross-country evidence on start-up dynamics", OECD Technology and Industry Working Papers.

Caminal, Ramon, and Xavier Vives. 1999. "Price dynamics and consumer learning". *Journal of Economics & Management Strategy* 8 (1): 95–131.

Card, David, and Thomas Lemieux. 2001. "Can falling supply explain the rising return to college for younger men? A cohort-based analysis". *Quarterly Journal of Economics* 116 (2): 705–746.

Caselli, Francesco. 2015. "Experience-biased Technical Change".

Castro, Rui, Gian Luca Clementi, and Yoonsoo Lee. 2015. "Cross sectoral variation in the volatility of plant level idiosyncratic shocks". *The Journal of Industrial Economics* 63 (1): 1–29.

Cavenaile, Laurent, et al. 2019. "Advertising, innovation and economic growth".

Chetty, Raj. 2006. "A General Formula for the Optimal Level of Social Insurance". *Journal of Public Economics* 90, **numbers** 10-11 (): 1879–1901.

Chetty, Raj, John N Friedman, and Jonah E Rockoff. 2014. "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood". *American economic review* 104 (9): 2633–79.

Chodorow-Reich, Gabriel, John Coglianese, and Loukas Karabarbounis. 2018. "The Macro Effects of Unemployment Benefit Extensions: a Measurement Error Approach". *The Quarterly Journal of Economics* 134, no. 1 (): 227–279.

Christoffel, Kai, Keith Kuester, and Tobias Linzert. 2009. "The role of labor markets for euro area monetary policy". *European Economic Review* 53, no. 8 (): 908–936.

Clementi, Gian Luca, and Hugo A Hopenhayn. 2006. "A theory of financing constraints and firm dynamics". *The Quarterly Journal of Economics* 121 (1): 229–265.

Clementi, Gian Luca, and Berardino Palazzo. 2016a. "Entry, Exit, Firm Dynamics, and Aggregate Fluctuations". *American Economic Journal: Macroeconomics* 8 (3): 1–41.

— . 2016b. "Entry, exit, firm dynamics, and aggregate fluctuations". *American Economic Journal: Macroeconomics* 8 (3): 1–41.

Cooper, Russell W., and John Haltiwanger. 2006a. "On the Nature of Capital Adjustment Costs On the Nature Adjustment of Capital Costs". *Review of Economic Studies* 73 (3): 611–633.

Cooper, Russell W, and John C Haltiwanger. 2006b. "On the nature of capital adjustment costs". *The Review of Economic Studies* 73 (3): 611–633.

Cooper, Russell, and Hubert Kempf. 2004. "Overturning Mundell: Fiscal policy in a monetary union". *Review of Economic Studies* 71 (2): 371–396.

Davis, Steven J., John Haltiwanger, Ron S. Jarmin, and Javier Miranda. 2006. "Volatility and dispersion in business growth rates: publicly traded versus privately held firms". In *NBER Macroeconomics Annual 2006, Volume 21*, 21:107–180. NBER Chapters, May. National Bureau of Economic Research, Inc.

Davis, Steven J, John C Haltiwanger, Scott Schuh, et al. 1996. *Job creation and destruction*, vol. 1. The MIT Press.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger. 2018. "The Rise of Market Power and the Macroeconomic Implications".

Decker, Ryan A., John Haltiwanger, Ron S. Jarmin, and Javier Miranda. 2016. "Where has all the skewness gone? The decline in high-growth (young) firms in the U.S." *European Economic Review* 86:4–23.

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda. 2014. "The Role of Entrepreneurship in US Job Creation and Economic Dynamism". *Journal of Economic Perspectives* 28 (3): 3–24.

den Haan, Wouter J., Garey Ramey, and Joel Watson. 2000. "Job Destruction and Propagation of Shocks". *American Economic Review* 90, no. 3 (): 482–498.

Dolls, Mathias, Clemens Fuest, Dirk Neumann, and Andreas Peichl. 2018. "An unemployment insurance scheme for the euro area? A comparison of different alternatives using microdata". *International Tax and Public Finance* 25, no. 1 (): 273–309.

Drozd, Lukasz A, and Jaromir B Nosal. 2012. "Understanding international prices: Customers as capital". *American Economic Review* 102 (1): 364–95.

Elsby, Michael W L, Bart Hobijn, and Aysegül Sahin. 2013. "Unemployment Dynamics in the OECD". *The Review of Economics and Statistics* 95, no. 2 (): 530–548.

Enders, Zeno, and David Vespermann. 2019. *Effects of a European Unemployment Insurance.* MIMEO. University of Heidelberg.

Engbom, Niklas. 2018. "Firm and Worker Dynamics in an Aging Labor Market".

Evers, Michael P. 2015. "Fiscal federalism and monetary unions: A quantitative assessment". *Journal of International Economics* 97 (1): 59–75.

Farhi, Emmanuel, Gita Gopinath, and Oleg Itskhoki. 2013. "Fiscal Devaluations". *The Review of Economic Studies* 81, no. 2 (): 725–760.

Farhi, Emmanuel, and Iván Werning. 2017. "Fiscal unions". *American Economic Review* 107 (12): 3788–3834.

Feyrer, James, and Bruce Sacerdote. 2013. "How much would us style fiscal integration buffer european unemployment and income shocks? (A Comparative Empirical Analysis)". *American Economic Review* 103 (3): 125–128.

Foster, Lucia, John C Haltiwanger, and Cornell John Krizan. 2001. "Aggregate productivity growth: lessons from microeconomic evidence". In *New developments in productivity analysis,* 303–372. University of Chicago Press.

Foster, Lucia, John Haltiwanger, and Chad Syverson. 2008. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review* 98 (1): 394–425.

— . 2016. "The Slow Growth of New Plants: Learning about Demand?" *Economica.*

Fuchs-Schündeln, Nicola, Dirk Krueger, and Mathias Sommer. 2010. "Inequality trends for Germany in the last two decades: A tale of two countries". *Review of Economic Dynamics* 13, no. 1 (): 103–132.

Furceri, Davide, and Aleksandra Zdzienicka. 2015. "The Euro Area Crisis: Need for a Supranational Fiscal Risk Sharing Mechanism?" *Open Economies Review* 26 (4): 683–710.

Garicano, Luis, and Esteban Rossi-Hansberg. 2006. "Organization and Inequality in a Knowledge Economy". *The Quarterly Journal of Economics* 121 (4): 1383–1435.

Gourio, François, and Leena Rudanko. 2014. "Customer Capital". *The Review of Economic Studies* 81, no. 3 (): 1102–1136.

Griliches, Zvi. 1969. "Capital-Skill Complementarity". *The Review of Economics and Statistics* 51 (4): 465–468.

Guvenen, Fatih, Greg Kaplan, Jae Song, and Justin Weidner. 2017. *Lifetime Incomes in the United States over Six Decades.* Working Paper, Working Paper Series 23371. National Bureau of Economic Research.

Hagedorn, Marcus, Fatih Karahan, Iourii Manovskii, and Kurt Mitman. 2013. *Unemployment Benefits and Unemployment in the Great Recession: The Role of Macro Effects.* Working Paper, Working Paper Series 19499. National Bureau of Economic Research.

Haltiwanger, John C., Julia I. Lane, and James R. Spletzer. 2007. “Wages, productivity, and the dynamic interaction of businesses and workers”. *Labour Economics* 14 (3): 575–602.

Haltiwanger, John, Ron S. Jarmin, and Javier Miranda. 2013a. “Who Creates Jobs? Small versus Large versus Young”. *Review of Economics and Statistics* 95 (2): 347–361.

— . 2013b. “Who creates jobs? Small versus large versus young”. *Review of Economics and Statistics* 95 (2): 347–361.

Haltiwanger, John, Ron Jarmin, and Javier Miranda. 2011. *Historically Large Decline in Job Creation from Startup and Existing Firms in the 2008-2009 Recession*, Business Dynamics Statistics Briefing 5. Kauffman Foundation.

Hopenhayn, Hugo A. 1992a. “Entry, exit, and firm dynamics in long run equilibrium”. *Econometrica: Journal of the Econometric Society*: 1127–1150.

— . 1992b. “Entry, Exit, and firm Dynamics in Long Run Equilibrium”. *Econometrica* 60 (5): 1127.

Hopenhayn, Hugo A., Julian Neira, and Rish Singhania. 2018. *From Population Growth to Firm Demographics: Implications for Concentration, Entrepreneurship and the Labor Share.* Working Paper, Working Paper Series 25382. National Bureau of Economic Research.

Hopenhayn, Hugo A., and Richard Rogerson. 1993a. “Job Turnover and Policy Evaluation: A General Equilibrium Analysis”. *Journal of Political Economy* 101 (5): 915–938.

Hopenhayn, Hugo, and Richard Rogerson. 1993b. “Job turnover and policy evaluation: A general equilibrium analysis”. *Journal of political Economy* 101 (5): 915–938.

Hosios, Arthur J. 1990. “On the Efficiency of Matching and Related Models of Search and Unemployment”. *Review of Economic Studies* 57 (2): 279–298.

Hosono, Kaoru, Miho Takizawa, and Kenta Yamanouchi. 2020. *Firm Age, Productivity, and Intangible Capital.* Tech. rep. Research Institute of Economy, Trade and Industry (RIETI).

House, Christopher L., Christian Proebsting, and Linda L. Tesar. 2018. *Quantifying the Benefits of Labor Mobility in a Currency Union.* NBER Working Papers 25347. National Bureau of Economic Research, Inc.

Jaimovich, Nir, Seth Pruitt, and Henry E. Siu. 2013. “The demand for youth: Explaining age differences in the volatility of hours”. *American Economic Review* 103 (7): 3022–3044.

Jeong, Hyeok, Yong Kim, and Iourii Manovskii. 2015. “The price of experience”. *American Economic Review.*

Jiang, Helu, and Faisal Sohail. 2019. “Skill-Biased Entrepreneurial Decline”.

Jin, Hehui, and Kenneth L Judd. 2002. *Perturbation methods for general dynamic stochastic models.* Tech. rep.

Johnson, George E. 1997. “Changes in Earnings Inequality: The Role of Demand Shifts”. *The Journal of Economic Perspectives* 11 (2): 41–54.

Jovanovic, Boyan. 1982. "Selection and the Evolution of Industry". *Econometrica: Journal of the Econometric Society*: 649–670.

Juncker, Jean-Claude. 2015. *The Five President's Report: Completing Europe's Economic and Monetary Union.* EUROPEAN COMMISSION BACKGROUND DOCUMENT ON ECONOMIC AND MONETARY UNION.

Jung, Philip, and Keith Kuester. 2015. "Optimal Labor-Market Policy in Recessions". *American Economic Journal: Macroeconomics* 7, no. 2 (): 124–156.

Karahan, Fatih, Benjamin W. Pugsley, and Aysegul Sahin. 2018. "Demographic Origins of the Startup Deficit". *Working Paper.*

Katz, Lawrence F., and K. M. Murphy. 1992. "Changes in Relative Wages, 1963-1987: Supply and Demand Factors". *The Quarterly Journal of Economics.*

Kehrig, Matthias, and Nicolas Vincent. 2018. *The Micro-Level Anatomy of the Labor Share Decline.* Working Paper, Working Paper Series 25275. National Bureau of Economic Research.

Kenen, Peter. 1969. "The Theory of Optimum Currency Areas: An Eclectic View". In *Monetary Problems of the International Economy,* ed. by Robert Mundell and Alexander Swoboda, 41–60. Chicago: University of Chicago Press.

Khan, Aubhik, and Julia K Thomas. 2008. "Idiosyncratic shocks and the role of nonconvexities in plant and aggregate investment dynamics". *Econometrica.*

Kim, Jinill, Sunghyun Kim, Ernst Schaumburg, and Christopher A Sims. 2008. "Calculating and using second-order accurate solutions of discrete time dynamic equilibrium models". *Journal of Economic Dynamics and Control* 32 (11): 3397–3414.

King, Robert G., and Sergio T. Rebelo. 1999. *Resuscitating real business cycles.*

Klette, Tor Jakob, and Samuel Kortum. 2004. "Innovating firms and aggregate innovation". *Journal of political economy* 112 (5): 986–1018.

Klosterhuber, Wolfram, Patrick Lehnert, and Stefan Seth. 2016. *Linked Employer-Employee Data from the IAB. FDZ-Datenreport 05/2016.* Tech. rep. Nuremberg.

Kondo, Illenin, Logan T Lewis, and Andrea Stella. 2018. "On the US Firm and Establishment Size Distributions".

Krueger, Dirk, and Harald Uhlig. 2006. "Competitive risk sharing contracts with one-sided commitment". *Journal of Monetary Economics* 53, no. 7 (): 1661–1691.

Krusell, Per, Lee E. Ohanian, José Víctor Ríos-Rull, and Giovanni L. Violante. 2000. "Capital-skill complementarity and inequality: A macroeconomic analysis". *Econometrica.*

Landais, Camille, Pascal Michaillat, and Emmanuel Saez. 2018. "A Macroeconomic Approach to Optimal Unemployment Insurance: Theory". *American Economic Journal: Economic Policy* 10, no. 2 (): 152–81.

Lentz, Rasmus, and Dale T Mortensen. 2005. “Productivity growth and worker reallocation”. *International Economic Review* 46 (3): 731–749.

— . 2008. “An empirical model of growth through product innovation”. *Econometrica* 76 (6): 1317–1373.

Levintal, Oren. 2017. “Fifth-order perturbation solution to DSGE models”. *Journal of Economic Dynamics and Control* 80:1–16.

Luttmer, Erzo GJ. 2007. “Selection, growth, and the size distribution of firms”. *The Quarterly Journal of Economics* 122 (3): 1103–1144.

— . 2010. “Models of growth and firm heterogeneity”. *Annu. Rev. Econ.* 2 (1): 547–576.

Magnus, Jan R, and Heinz Neudecker. 1999. *Matrix Differential Calculus – with Applications in Statistics and Econometrics.* Wiley.

Marcet, Albert, and Ramon Marimon. 1992. “Communication, commitment, and growth”. *Journal of Economic Theory* 58 (2): 219–249.

Meyer, Bruce D. 1990. “Unemployment Insurance and Unemployment Spells”. *Econometrica* 58 (4): 757–782.

Mitman, Kurt, and Stanislav Rabinovich. 2015. “Optimal unemployment insurance in an equilibrium business-cycle model”. *Journal of Monetary Economics* 71 (C): 99–118.

Moyen, Stephane, Nikolai Stähler, and Fabian Winkler. 2016. *Optimal Unemployment Insurance and International Risk Sharing.* Finance and Economics Discussion Series 2016-054. Board of Governors of the Federal Reserve System (U.S.)

Mundell, Robert A. 1961. “A Theory of Optimum Currency Areas”. *The American Economic Review* 51 (1775): 657–665.

Nekarda, Christopher J., and Valerie A. Ramey. 2011. “Industry evidence on the effects of government spending”. *American Economic Journal: Macroeconomics.*

Oates, Wallace E. 1999. “An Essay on Fiscal Federalism”. *Journal of Economic Literature* 37, no. 3 (): 1120–1149.

Persson, Torsten, and Guido Tabellini. 1996. “Federal Fiscal Constitutions: Risk Sharing and Moral Hazard”. *Econometrica* 64 (3): 623–646.

Peters, Michael. 2019. “Heterogeneous Markups, Growth, and Endogenous Misallocation”.

Petrongolo, Barbara, and Christopher A. Pissarides. 2001. “Looking into the Black Box: A Survey of the Matching Function”. *Journal of Economic Literature* 39 (2): 390–431.

Phelan, Christopher, and Robert M. Townsend. 1991. “Computing Multi-Period, Information-Constrained Optima”. *The Review of Economic Studies* 58 (5): 853–881.

Pugsley, Benjamin W., Petr Sedlacek, and Vincent Sterk. 2018. “The Nature of Firm Growth”. London.

— . 2019. "The nature of firm growth". *Available at SSRN 3086640.*

Queralto, Albert. 2019. "A model of slow recoveries from financial crises". *Journal of Monetary Economics.*

Ravn, Morten O., and Vincent Sterk. Forthcoming. "Macroeconomic Fluctuations with HANK & SAM: An Analytical Approach". *Journal of the European Economic Association.*

Salgado, Sergio. 2019. "Technical Change and Entrepreneurship".

Schmucker, Alexandra, Andreas Ganzer, Jens Stegmaier, and Stefanie Wolter. 2018. *Establishment History Panel 1975-2017. FDZ-Datenreport, 09/2018.* Tech. rep. Nuremberg.

Sedláček, Petr, and Vincent Sterk. 2017. "The growth potential of startups over the business cycle". *American Economic Review* 107 (10): 3182–3210.

Shimer, Robert. 2004. "The Consequences of Rigid Wages in Search Models". *Journal of the European Economic Association* 2 (2-3): 469–479.

Silva, José Ignacio, and Manuel Toledo. 2009. "Labor Turnover Costs and the Cyclical Behavior of Vacancies and Unemployment". *Macroeconomic Dynamics* 13 (S1): 76.

Stepner, Michael. 2013. *BINSCATTER: Stata module to generate binned scatterplots.* Statistical Software Components, Boston College Department of Economics.

Tinbergen, Jan. 1956. "On the Theory of Income Distribution". *Weltwirtschaftliches Archiv* 77:155–175.

Werner, Pierre. 1970. "Report to the Council and the Commission on the realisation by stages of Economic and Monetary Union in the Community – "Werner Report" – (definitive text) [8 October 1970]. Bulletin of the European Communities, Supplement 11/1970". Download from Archive of European Integration at http://aei.pitt.edu/1002/.