ZENTRUM FÜR ENTWICKLUNGSFORSCHUNG (ZEF)

Patterns of Urbanization and Economic Development

Evidence from household surveys in Ethiopia

Dissertation

zur Erlangung des Grades

Doktor der Agrarwissenschaften (Dr. agr.)

der Landwirtschaftlichen Fakultät

der Rheinischen Friedrich-Wilhelms-Universität Bonn

von

Mekdim Dereje Regassa

aus

Nekempt, Äthiopien

Bonn 2022

Referent**:** Prof. Dr. Joachim von Braun

Korreferent**:** Prof. Dr. Jan Börner

Tag der mündlichen Prüfung: 15.02. 2022

Angefertigt mit Genehmigung der Landwirtschaftlichen Fakultät der Universität Bonn

Acknowledgements

Many thanks to my advisor Prof. Dr. Joachim von Braun for the opportunity to pursue my Ph.D. studies, for providing me academic guidance, for being patient and understanding. Thank you! I would like to thank Prof. Dr. Jan Börner, for agreeing to become my second supervisor. My gratitude also goes to my tutor Dr. Zaneta Kubik. Thank you for keeping your office doors open for me and for the wonderful discussions we have had, and for the encouragement throughout my studies.

I am grateful to the Deutscher Akademischer Austauschdienst (DAAD) for the generous financial support for me and my family throughout the Ph.D. program. I also appreciate the financial support from Foundation Fiat Panis for the fieldwork. I massively benefited from the administrative support of the ZEF Doctoral Studies team: Dr. Günther Manske, Ms. Maike Retat-Armin, Max, Anna, and Henrik. Similarly, I highly appreciate the data support from Javier Muro, IT support from Mr. Ludger Hammer, Anja Köhler, and Library support from Mr.Volker Merx.

I could have not made it so far without my wonderful wife, Meron Solomon, and my two children, Milto and Natay. You provided perspective to my work and life and I am thankful. Love you a lot. To my father, Dereje Regassa, and my siblings: Tesfaye (Wonde), Hawi, Derartu, and Abdi thank you! To Haregewoin (Etetu): thank you for all the kindness. To Solomon (RIP), Dawit, Henok, and Dani: thank you for the encouragement and the support. I also like to extend my gratitude to my dear friends: Teshome Tefera, Semere Bekele, Tekalign Sakketa, Ermias Engida, Ibrahim Worku, Muluken Elias, Sinidu Workineh, Solomon Benti, Tirsit Genye, Anteneh Andarge, Fikru Wubishet, Adiyam Debesay, and Girma/Yeshi and all ESSP colleagues. My life is better off with you in it.

To my current and former colleagues: Gashaw Abate, Bart Minten, Kalle Hirvonen, Alemayew Seyoum, Mequanint Melese, and Kibrom Abay: it is my great honor to have known you and work with you. Thank you for the helpful comments on the different versions of my papers and the insights. To Tilman Brück: thank you for your generosity to let me wrap up the thesis and for the encouragement. I needed the final push! I would also like to extend my appreciation to the many colleagues at ZEF including Pallavi Rajkhowa, Evelyn Ewere, Ana Maria Perez, Andre Castro, Rahel Deribe, Emmanuel Rukundo, Chiara Kofol, Essa Mussa, and participants of ZEF Research Colloquiums for your inputs to my work. Finally, I would like to thank Petra Brown, for proofreading all my thesis chapters and the wonderful suggestions, and Barbara Schmidt-Eule for kindly translating the abstract into German.

Above all, Praise is to God, for all the blessings and for getting me to this point (against all odds).

DEDICATION

To my late mother, Ejigayew Asfaw
Sorry I couldn’t wake up before you left. But I am eternally grateful for everything!

May your soul rest in peace!

Abstract

Countries in sub-Saharan Africa (SSA) are urbanizing at an unprecedented fast rate. This trend has the potential to affect the welfare of households by altering the degree of urban proximity as well as the size of the existing urban areas. While ample evidence exists regarding the effect of urban proximity, rigorous empirical evaluation of the heterogeneous effect of different sized urban areas in the region is scant. The absence of research in this dimension is often attributed to the lack of an objective and a disaggregated measure of the level and dynamics of urbanization. Studies presented in this thesis aim at bridging this gap by combining satellite-based nighttime light (NTL) intensity data and standard definitions of urbanization to study the implications of urbanization on households' welfare and livelihood in Ethiopia. The main research questions explored in this thesis are: (i) Does the effect of urbanization on household welfare depend on the degree of urbanization? (ii) What are the heterogeneous effects of urban proximity on nutritional outcomes? (iii) Does the degree of urbanization influence the degree of intergenerational mobility? And (iv) Which interventions are effective to improve the delivery of agricultural extension service in remote areas?

To address the first three questions, three rounds of Ethiopian Living Standards Measurement Study-Integrated Surveys on Agriculture (LSMS-ISA) are geo-spatially linked to NTL data. The first three analytical chapters in the thesis (addressing the first three questions listed above) are organized in such a way as to capture the effect of urbanization on welfare across different generations. **Chapter 2** examines the effect of urbanization on broader indicators of household welfare based on the New Economic Geography (NEG) framework and threshold data analysis technique. The study finds that intermediate towns are more strongly associated with household welfare as compared to large towns, small towns, and the rural hinterland. **Chapter 3** examines the effect of the distance to and the size of the proximate urban areas on children's health and nutrition outcomes. An Instrumental Variables (IV) approach is combined with Inverse Probability Weighting (IPW) to account for endogeneity and self-selection issues in the estimation of the basic model. The study finds a statistically and economically significant positive effect of investment in rural infrastructure on health and nutrition outcomes. It also finds that, for households in intermediate and large towns, diet diversity is higher (by 1.2 percentage points) and child stunting is lower (by about 3 percentage points) compared to households in rural areas. **Chapter 4** uses ordered logistic regression method to assess the intersection between urbanization and intergenerational mobility in occupational status. It finds that intergenerational mobility in occupational status is weaker in large urban areas, and this is largely explained by huge inequality in educational attainment. Once individual education level is accounted for, large urban areas offer better mobility in occupational status.

For the fourth research question, which is addressed in **Chapter 5**, a choice experiment was conducted to elicit the preferences of 761 agricultural Extension Agents (EAs) for job attributes. A novel random parameters logit model (RPL) is used to estimate parameters of interest and to simulate the impact of possible policy interventions. Results show that offering education opportunities is by far the most powerful instrument to attract and retain EAs. It increases the uptake of the extension job in remote locations by 77 percentage points, which is significantly higher than the effect of doubling current salary levels. EAs also expressed strong preferences for work environments with basic amenities, housing, transportation services, and well-equipped Farmer Training Centers (FTCs).

The overarching finding from all the chapters is that while there is a considerable rural-urban gap in living standards, these spatial disparities are underlined by pervasive differences in access to basic public services and employment opportunities. Therefore, policy interventions that target to improve overall welfare as well as reduce the spatial imbalance need to remove the constraints facing isolated households in remote areas as well as the marginalized poor in urban areas. Accordingly, the thesis identified a set of relevant policy recommendations tailored to the different locations along the rural-urban spectrum, based on their degree of urbanization and their level of economic development.

Muster der Urbanisierung und der wirtschaftlichen Entwicklung

Evidenz aus Haushaltsbefragungen in Äthiopien

Zusammenfassung

Die Länder in Afrika südlich der Sahara (SSA) verstädtern in einem noch nie dagewesenen Tempo. Dieser Trend hat das Potenzial, die Wohlergehen der Haushalte zu beeinflussen, indem er den Grad der städtischen Nähe sowie die Größe der bestehenden städtischen Gebiete verändert. Während es zahlreiche Belege für die Auswirkungen der städtischen Nähe gibt, gibt es nur wenige rigorose empirische Untersuchungen zu den heterogenen Auswirkungen der unterschiedlich großen städtischen Gebiete in der Region. Das Fehlen von Forschung in diesem Bereich wird oft auf das Fehlen eines objektiven und disaggregierten Maßes für den Grad und die Dynamik der Urbanisierung zurückgeführt. Die in dieser Arbeit vorgestellten Studien zielen darauf ab, diese Lücke zu schließen, indem sie satellitengestützte Daten zur nächtlichen Lichtintensität (NTL) und Standarddefinitionen der Urbanisierung kombinieren, um die Auswirkungen der Urbanisierung auf das Wohlergehen und die Lebensgrundlage der Haushalte in Äthiopien zu untersuchen. Die wichtigsten Forschungsfragen, die in dieser Arbeit untersucht werden, sind: (i) Hängt der Effekt der Urbanisierung auf die Wohlergehen der Haushalte vom Grad der Verstädterung ab? (ii) Was sind die heterogenen Auswirkungen der Stadtnähe auf die Ernährungssituation? (iii) Beeinflusst der Grad der Verstädterung den Grad der intergenerationalen Mobilität? Und (iv) Welche Interventionen sind effektiv, um die Bereitstellung von landwirtschaftlichen Beratungsdiensten in abgelegenen Gebieten zu verbessern?

Um die ersten drei Fragen zu beantworten, werden drei Runden des Ethiopian Living Standard Measurement Study-Integrated Surveys on Agriculture (LSMS-ISA) geo-räumlich mit den NTL-Daten verknüpft. Die ersten drei analytischen Kapitel der Arbeit (die sich mit den ersten drei oben genannten Fragen befassen) sind so strukturiert, dass sie die Auswirkungen der Urbanisierung auf die Wohlergehen über verschiedene Generationen hinweg erfassen. Kapitel 2 untersucht die Auswirkung der Urbanisierung auf breitere Indikatoren der Wohlergehen der Haushalte auf der Grundlage des New Economic Geography (NEG)-Rahmens und der Technik der Schwellenwertdatenanalyse. Die Studie zeigt, dass Zwischenstädte stärker mit der Wohlergehen der Haushalte assoziiert sind als Großstädte, Kleinstädte und das ländliche Umland. Kapitel 3 untersucht den Effekt der Entfernung zu und der Größe der nahe gelegenen städtischen Gebiete auf die Gesundheits- und Ernährungszustand von Kindern. Ein Instrumentalvariablen (IV)-Ansatz wird mit Inverse Probability Weighting (IPW) kombiniert, um Endogenitäts- und Selbstselektionsprobleme bei der Schätzung des Grundmodells zu berücksichtigen. Die Studie findet einen statistisch und ökonomisch signifikanten positiven Effekt von Investitionen in ländliche Infrastruktur auf die Gesundheits- und Ernährungszustand. Sie stellt außerdem fest, dass für Haushalte in Mittel- und Großstädten die Ernährungsvielfalt höher (um 1,2 Prozentpunkte) und das Stunting von Kindern niedriger (um etwa 3 Prozentpunkte) ist als für Haushalte in ländlichen Gebieten. Kapitel 4 verwendet Methoden der geordneten logistischen Regression, um den Zusammenhang zwischen Urbanisierung und intergenerationaler Mobilität im Berufsstand zu untersuchen. Es zeigt sich, dass die intergenerationale Mobilität im Berufsstand in großen städtischen Gebieten schwächer ist, was größtenteils durch die große Ungleichheit im Bildungsniveau erklärt wird. Sobald das individuelle Bildungsniveau berücksichtigt wird, bieten große städtische Gebiete eine bessere Mobilität im Berufsstand.

Für die vierte Forschungsfrage, die in Kapitel 5 behandelt wird, wurde ein Auswahlexperiment durchgeführt, um die Präferenzen von 761 landwirtschaftlichen Beratern (EAs) für Berufsattribute zu eruieren. Ein neuartiges Random-Parameter-Logit-Modell (RPL) wird verwendet, um die interessierenden Parameter zu schätzen und die Auswirkungen möglicher Interventionen zu simulieren. Die Ergebnisse zeigen, dass das Ausweitung von Bildungsmöglichkeiten das bei weitem stärkste Instrument ist, um EAs anzuziehen und zu halten. Es erhöht die Inanspruchnahme der Job in entlegenen Gebieten um 77 Prozentpunkte,

was deutlich höher ist als der Effekt einer Verdoppelung des aktuellen Gehaltsniveaus. EAs äußerten auch starke Präferenzen für Arbeitsumgebungen mit grundlegenden Annehmlichkeiten, Unterkünften, Transportdienstleistungen und gut ausgestatteten Farmer Training Centres (FTCs).

Die übergreifende Erkenntnis aus allen Kapiteln ist, dass es zwar ein beträchtliches Land-Stadt-Gefälle im Lebensstandard gibt, diese räumlichen Ungleichheiten aber durch tiefgreifende Unterschiede im Zugang zu grundlegenden öffentlichen Dienstleistungen und Beschäftigungsmöglichkeiten unterstrichen werden. Daher müssen Interventionen, die darauf abzielen, die allgemeine Wohlbefinden zu verbessern und das räumliche Ungleichgewicht zu verringern, die Einschränkungen beseitigen, mit denen isolierte Haushalte in abgelegenen Gebieten sowie die marginalisierten Armen in städtischen Gebieten konfrontiert sind. Dementsprechend wurde in dieser Arbeit ein eine Reihe relevanter politischer Empfehlungen identifiziert, die auf die verschiedenen Orte entlang des Land-Stadt-Spektrums zugeschnitten sind, basierend auf ihrem Urbanisierungsgrad und ihrem wirtschaftlichen Entwicklungsniveau.

Contents

List of Figures

List of Tables

List of Abbreviations

AIC	Akaike information criterion
ALMP	Active labour Market Policies
BIC	Bayesian Information criterion
BoA	Woreda Bureau of Agriculture
CEPs	Continuing Education Programs
CI	Confidence Intervals
CL	Conditional logit
CPI	Consumer price index
CSA	Central Statistical Authority of Ethiopia
DCE	Discrete choice experiment
DG	Digital Green
DHS	Demographic and Health Survey
DMSP-OLS	Defense Meteorological Satellite Program Operational Line Scanner
EA	Extension Agents
FAO	The Food and Agriculture Organization of the United Nations
FTCs	Farmer Training Centers.
GDP	Gross Domestic Product
HAZ	Height-for-age
HCT	The Human capital theory
HDDS	Household Diet Diversity Score
HQIC	Hannan-Quinn information criterion
IAIP	Integrated Agro-Industrial Parks
IFPRI	International Food Policy Research Institute
IGM	Intergenerational Mobility
IIA	Independence of irrelevant alternatives
IPW	Inverse Probability Weighting
IV	Instrumental Variables
LPM	Linear Probability Model
LSMS-ISA	Living Standard Measurement Study-Integrated Survey of Agriculture
MLE	Maximum Likelihood Estimation
MoE	Ministry of Education
MoFED	Ministry of Finance and Economic Development
MoUDC	Ministry of Urban Development and Construction

MSEs	Micro and Small Enterprise
MSL	Maximum simulated Log-likelihood
NEG	New Economic Geography
NGDC	the National Geophysical Data Center
NOAA	The National Oceanic and Atmospheric Administration
NPOESS	National Polar-Orbiting Operational Environmental Satellite System
NTL	Nighttime light
OLS	Ordinary Least Squares
PCA	Principal Components Analysis
RPL	Random Parameter Logit model
SACCO	Savings & Credit Co-operative
SD	Standard Deviation
SOL	Sum of Nighttime Light
SSA	Sub-Saharan Africa
TLU	Tropical Livestock Units
UNDP	United Nations Development Programme
UNDSA	United Nations Department of Economic and Social Affairs
UNECA	The United Nations Economic Commission for Africa
VIIRS	Visible Infrared Imager Radiometer Suite
WASH	Water, Sanitation, and Hygiene
WAZ	Weight-for-age
WHO	World Health Organisation
WTP	Willingness To Pay

CHAPTER **1**

1. Introduction

1.1. Background

Most African countries are urbanizing at an unprecedented rate. The current rate of urban population growth in the continent is 3.3 percent per year, which is the highest in the world. At this rate, by 2050, more than half of the continent‘s population is projected to live in the cities and towns, and the total urban population will be 1.2 billion people — almost a quarter of the world‘s urban population (UN Habitat 2014). The urbanization rate of countries in sub-Saharan Africa (SSA) is even faster. Although the current share of the urban population in SSA is 40 percent, which is lower than other developing countries, the region is expected to urbanize rapidly over the coming decades. It is projected that by 2050, 57 percent of the region's population will live in urban areas (UN Habitat 2014).

This rapid rate of urbanization has the potential to create new opportunities but also presents challenges for inclusive and sustainable growth. While urbanization can spur economic growth and create more jobs due to economies of scale and agglomeration (Christiaensen, De Weerdt, and Todo 2013; World Bank 2009), this is not always guaranteed. In Africa, for example, urban areas are often associated with poverty, inequality, and unemployment (Dorosh and Thurlow 2014; Gollin, Jedwab, and Vollrath 2016; World Bank 2013a). Urban households are also more vulnerable to food price hikes and youth unemployment (African Development Bank 2011).

This underperformance of urban areas is partly attributable to the lack of adequate infrastructure and economic systems to support the increasing levels of urbanization. In most African countries, urbanization is not accompanied by industrialization as is the case in developed or middle-income countries at the same stage of development (Gollin et al. 2016; Henderson, Roberst, and Storeygard 2013). Gollin et al. (2016) argue that in most African countries, urbanization is concentrated in "consumption cities". Meaning, urban areas in Africa are dominated by non-tradable personal services and commerce and considerably large shares of the population engage in the informal sector.[1]

In Ethiopia, urban areas share most of these challenges, which are exacerbated by the rapid pace of urbanization in the country. Statistics from the United Nations World Urbanization Prospects 2018 show that Ethiopia is experiencing an annual urban population growth rate of more than 4 percent (UNDESA 2019). However, analogous to other SSA countries, the urbanization process in Ethiopia is accompanied by neither structural transformations nor by integrated planning (Abay et al. 2020). This has limited the ability of urban areas to support both the growing urban population as well as to create the necessary linkages with the surrounding rural areas.

On the positive side, the country is still in the early stages of its urban transformation. This provides a unique opportunity to proactively manage urban development programs to ensure inclusive and sustainable growth (Abay et al. 2020). Nevertheless, there is little rigorous empirical evidence on the impact of this urbanization trend on the welfare of households living in the urban spaces and on the surrounding rural population. In particular, there is a notable gap in research on possible heterogeneity among different-sized urban areas. Recent studies suggest that the implication of

[1] This is in contrast to "Production cities", the type of cities in developed or middle-income countries during the same stage of development, where the majority of workers in urban areas engage in manufacturing sectors (Gollin et al. 2016).

the pattern of urbanization may be at least as important as its *aggregate rate* (Christiaensen, De Weerdt, and Todo 2013; Kanbur, Christiaensen, and De Weerdt 2019). It is noted that the pattern of urbanization — whether the growth rate of large cities is higher or lower than that of small towns — considerably affects the relationship between urbanization and welfare. African urbanization statistics indicate that only 10 percent of the population resides in large cities of more than 5 million inhabitants, while small and medium-sized towns host the majority of the urban population. Moreover, the population in these urban areas has doubled in the last decade and is expected to grow by more than 30 percent in the next decade (UNDESA 2015). This trend especially calls for a disaggregated study of the impact of urbanization across different stages. Notwithstanding, there is little rigorous empirical research on the heterogeneous effect of different-sized urban areas in Africa.

Perhaps, one of the main reasons for this paucity in research is the lack of an objective, robust, and disaggregated measure of the degree and dynamics of urbanization. Conventionally, measurements and definitions of urbanization rely on survey - and census - based aggregate rural-urban indicators. The most common of these indicators is a simple binary urban-rural indicator. However, besides often being subjective, these indicators tend to reflect political and bureaucratic dispositions rather than the services the spaces provide. As a result, they cannot adequately capture the enormous heterogeneities among the rural-urban areas and the rapid dynamics of urbanization (von Braun 2007; Henderson 2010; De Poel, O'donnell, and Doorslaer 2012; Satterthwaite and Tacoli 2003; UNECA 2017). To circumvent this problem and account for the continuum between rural and urban areas, researchers have attempted to use several alternative indicators of urbanization including, *inter alia*, population size, population density, and index of infrastructure and market access (Deichmann, Shilpi, and Vakis 2009; De Poel et al. 2012).

One of the latest and perhaps the most promising indicator of urbanization along this line is the use of satellite-based Nighttime Light (NTL) data.[2] NTL — an indicator of the intensity of light emitted from the earth at night — offers a unique potential for measuring urbanization and urban expansion. Because NTL is a basic urban amenity, its intensity per unit area is a valid indicator of urbanization (Abay et al. 2020; Henderson et al. 2003).

The use of NTL to study urbanization and related socioeconomic development dates back to 1992, the year the NTL database was digitalized. Since then, several studies have used NTL datasets to examine distributional and temporal patterns in key socioeconomic variables such as urban boundaries, population dynamics, built-up area, and electrification (Bennett and Smith 2017). However, the majority of these studies focus on Asia and the US. The use of NTL in empirical studies in the context of sub-Saharan Africa is quite recent and limited. However, in recent years, studies that use NTL as a proxy for urbanization have proliferated (see, for instance, Abay et al. 2020; Amare et al. 2017; Ameye 2018; Binswanger-Mkhize and Savastano 2017; Chen and Nordhaus 2015; Savory et al. 2017).

[2] Another promising development is the use of smartphone-based location traces (Hoteit et al. 2014; Vieira et al. 2010; Williams et al. 2015). Compared to residence-based urbanization measures, this approach has the advantage of collecting spatial-temporal trajectories of individuals' travel information. However, for developing countries where penetration of smart phones is restricted, its applicability might be limited.

In this thesis, this literature is followed to combine satellite-based NTL intensity data and standard definitions of urbanization to study the impact of patterns of urbanization on households' welfare in Ethiopia. More specifically, the objective is to examine whether and how urbanization and the different types of urban areas in Ethiopia improve household welfare. Importantly, the study aims at parsing out the dynamics of household welfare not only between rural and urban areas but also across the different sized urban areas. To this end, throughout this thesis, NTL data is used in two forms[3]. First, the continuous NTL index is used to conduct a micro-level analysis of the welfare implications of urbanization[4]. Second, the NTL index is used to categorize sample households into four groups based on the intensity of nighttime light — the value of NTL — at the place of residence: rural areas, small towns, medium-sized towns (hereafter intermediate towns), and large towns[5].

Compared to the binary rural-urban measure that is commonly used in the literature, the use of the NTL index has several advantages. First, because the NTL index is available at a high spatial resolution, it allows for a continuous assessment of urbanization. That is, instead of considering urban and rural areas as distinct geographic spaces, the NTL index allows the rural-urban space to be examined as a continuum. This facilitates a more disaggregated classification of urban areas, allowing for a micro-level analysis of the patterns and effects of urbanization. In the context of SSA, this is particularly interesting as it facilitates the study of the role of small and intermediate towns, which are mushrooming throughout the region (Satterthwaite and Tacoli 2003).

Second, the use of the NTL index eliminates reliance on the national administrative definition of urban and rural areas. The national urban statistics are often sporadic, unreliable, and lag behind reality, especially in developing countries (Bennett and Smith 2017; von Braun 2014b; Donaldson and Storeygard 2016; Satterthwaite and Tacoli 2003). Administrative definitions also tend to lack comparability across regions and over time as assignments are often based on political dispositions rather than services the spaces provide (Satterthwaite and Tacoli 2003; UNECA 2017). The use of the NTL index mitigates these shortcomings as it is measured with consistent quality, and its availability over a long period of time allows for reliable temporal analysis.

Third, the use of the NTL index allows for the construction of spatially detailed measures of urbanization. Unlike the binary rural-urban classification, this helps to inform policies aimed at promoting place-based development along the rural-urban continuum. In turn, place-based development policies — compared to sector-based policies — tend to create stronger rural-urban linkages, reduce regional inequalities, promote balanced urban systems, and ensure more inclusive growth (OECD/PSI 2020; Satterthwaite and Tacoli 2003)[6].

---

[3] This is notwithstanding several shortcomings of the measure as a robust indicator of urbanization. Section 2.2.1 details the pros and cons of this measure.

[4] To be precise, what is used as a measure of urbanization is the Sum of Light (SOL) — a variable that adds up the NTL within the 10km radius. Details on the description of this variable is presented in the next chapter.

[5] Rural areas are those where the economic activity is predominantly agrarian. The study separates large towns from the other urban areas based on size and because of the increasing concern that large urban areas are becoming too crowded. Such an investigation is particularly imperative as recent studies argue that urbanization in Africa has unfolded differently than the rest of the world (Gollin et al. 2016; Henderson et al. 2013).

[6] Two additional benefits of the use of NTL are related to the possibility of deriving the SOL. The next chapter presents more details on this.

1.2. Patterns of urbanization and economic development

This study examines the spatial distribution of household welfare within the framework of New Economic Geography (NEG). NEG is the latest version of economic geography, a long-standing strand of literature in the field of economics that seeks to answer the questions of what kinds of economic activities occur, where, and why. The pioneer of this approach of “spatial economics” is Johann Heinrich von Thünen (Fujita 2010, 2012; Krugman 1997). According to Fujita (2012), the von Thünen model is the first spatial economy model that precisely describes how agricultural land use around the city is determined based on the distance from the center. This model also laid the foundation for the nascent economic geography models that attempt to describe the spatial pattern in economic activities and welfare, based only on physical and economic factors. One limitation of the von Thünen model is the assumption that the location of the city around which economic activities are organized is exogenously determined (Krugman 1997).

Christaller and Lösch later advanced the central place theory to address the question of why spatial structure would eventually emerge even in a setting where all sites were initially the same (Fujita 2010; Fujita, Krugman, and Venables 2000). Considering dense human settlements as "central places" that serve the population in the surrounding areas, Christaller developed a model that predicts the pattern of settlement locations using geometric shapes and thereby laid the foundation of the hierarchy of central places (Fujita et al. 2000). In this thesis, the study of heterogeneous effects of urban size on welfare is partly inspired by this theory of the hierarchy of urban areas. August Lösch is credited for his work in formalizing the central place theory as well as for his pioneering work on the development of real spatial economic theory based on the hypothesis of general equilibrium (Fujita 2010). The main limitation of central place theory, according to Krugman (1997), is that the model was not fully specified and formalized to describe how urban systems would emerge from the interaction of economic agents.

The NEG was also inspired by regional science, urban systems theory, and international trade theories (Krugman 1997; Lafourcade and Thisse 2008). For example, concepts such as endogenous growth of cities, cumulative causation, agglomeration, increasing returns to scale, monopolistic competition, and transportation costs, which later became the building blocks of NEG, were theorized and elaborated in these fields of study (Fujita 2010; Krugman 1997). However, before NEG, the economics discipline was evolving independently of these and other disciplines. In essence, NEG is a theory that has managed to unify these fragmented developments in order to explain the spatial distribution of economic activities and welfare based solely on the trade-off between increasing returns to production and transportation costs. Fundamentally, NEG argues that the spatial distribution of production sites depends on the balance between agglomeration and dispersion factors and that these factors are underlined by economies of scale, the degree of market competition, and transportation costs.

1.2.1. The core-periphery model of spatial development

The objective of the core-periphery (c-p) model is to demonstrate how the interactions among increasing returns to scale at the level of the firm, transport costs, and factor mobility can cause a spatial economic structure to emerge and change. This part briefly presents the basic derivation as well as the predictions of the model[7].

A. Assumptions of the model

i) In a spatial economy, there exist two sectors: A monopolistically competitive manufacturing sector (M), and a perfectly competitive agricultural sector (A).

ii) Each of these sectors employs a single resource: the manufacturing sector employs workers and the agricultural sector employs farmers.

iii) The supply of workers and farmers is fixed. Let L^A and L^M represent the total available number of farmers and manufacturing workers, respectively.

iv) There are R regions. The geographical distribution of farmers is determined exogenously. In each region, $r \in R$, is endowed with ϕ_r share of the agricultural labor force. On the other hand, the geographical distribution of workers is determined endogenously at any point in time, and workers are mobile. Let λ_r represent the share of region r in total manufacturing labor. Since anyone works either in the agriculture sector or the manufacturing sector, a unit could be chosen such that: $L^M = \mu$, $L^A = 1 - \mu$.

v) Agricultural goods could be transported free of cost. In contrast, manufactured goods are subject to the "iceberg" transport cost. That is, if one unit of a good is shipped from region r to region $s \in R$, only $^1/_{T_{rs}}$ units will arrive. Of course, $T_{rs} > 1$.

vi) Since the agricultural sector exhibits constant returns to scale (due to the competitive market structure assumption), and the shipment of agricultural goods is costless, the wage rate of agricultural workers is the same across all regions. Therefore, the agricultural wage rate is considered a numeraire in the model. That is: $\omega_r^A = 1$.

vii) The wages of manufacturing workers may differ in nominal and in real terms both across regions and over time. Define the nominal and real wage rate of manufacturing workers in region r, respectively by ω_r, and w_r. Workers are mobile and they move toward regions that offer high real wages and away from regions that offer a below-average real wage.

[7] This part heavily relies on Fujita, Krugman, & Venables (2000) and Krugman (1991).

B. The Model

The model determines the geographical distribution of manufacturing firms and manufacturing workers based on the level of income, the price index of manufacturing goods, and the nominal and real wages rate of workers. To set the groundwork, these variables are defined as follows:

i) Income

The total income in region r is the sum of the total income of agricultural laborers (farmers) and manufacturing workers. Since $L^A = 1 - \mu$ and $\omega_r^A = 1$ and the share of region r in L^A is ϕ_r, the total income of agricultural laborers is $(1 - \mu)\ \phi_r$.

On the other hand, since $L^M = \mu$ and the share of region r in L^M is λ_r and the nominal wage rate of manufacturing workers is ω_r, the total income of manufacturing workers is $\mu\lambda_r\omega_r$. Therefore, the aggregate income of region r is:

$$Y_r = \mu\lambda_r\omega_r + \ (1 - \mu)\phi_r \qquad (1.1)$$

ii) Price Index

The price index of manufacturers in each region, derived from the profit optimization condition of each firm is given by:

$$G_r = \left[\sum_{s=1}^{R} \lambda_s(\omega_s T_{sr})^{1-\sigma}\right]^{1/_{1-\sigma}} \qquad (1.2)$$

The level of the index depends on the geographical distribution of the manufacturing firms (λ_s), the size of the transportation cost between regions (T_{sr}), the wage rate (ω_s) and the rate of labor requirement of each firm (σ). What is relevant for this model, is the link between change in the share of manufacturing firms and the price index (G_r). With everything else remaining the same, the shift of manufacturing firms into region r tends to lower the price index in the region, which in turn makes the region more attractive to manufacturing workers (higher real wage). To illustrate this, suppose there are only two regions (hence R=2) and the nominal wage rates are the same across the two regions. That is $\omega_1 = \omega_2 = \omega$. Furthermore, let $\lambda_1 = \lambda, \lambda_2 = 1 - \lambda$ and $T_{12} = T_{21} = T, where\ T > 1$. Hence, equation 1.2 could be written separately for region 1 as:

$$G_1 = [\lambda\omega^{1-\sigma} + (1 - \lambda)(\omega T)^{1-\sigma}]^{1/_{1-\sigma}} \qquad (1.2.1)$$ [8]

Rearranging,

$$G_1 = [(1 - T^{1-\sigma})\lambda\omega^{1-\sigma} + (\omega T)^{1-\sigma}]^{1/_{1-\sigma}} \qquad (1.2.2)$$

[8] Note that the transportation cost is removed from the first term as this refers only to region 1. The second term of the equation accounts for the distance between region 1 and region 2 hence, it consists of the transportation cost, T. The corresponding equation for region 2 is given by $G_2 = [\lambda(\omega T)^{1-\sigma} + (1 - \lambda)\omega^{1-\sigma}]^{1/_{1-\sigma}}$.

Since $T > 1$, then $T^{1-\sigma} < 1$. Therefore, the higher the λ, the lower is G_1. This represents a forward linkage - the positive effect of the concentration of firms on real wage through product prices (Fujita et al. 2000).

iii) **Nominal Wages**

Fujita, Krugman, & Venables (2000) indicate that the nominal wage rate in region r at which each manufacturing firm breaks even is given by:

$$\omega_r = \left[\sum_s Y_s\, T_{rs}^{1-\sigma} G_s^{\sigma-1}\right]^{1/\sigma} \qquad (1.3)$$

Equation 1.3 suggests that the nominal wage firms pay depends on the income level in the region (Y_s), the transportation cost between regions (T_{sr}), and price index (G_s). The study emphasizes the link between income level and wages (ω_r). From the equation (1.3), it is clear that the higher the income level in the region, the higher the nominal wage a firm would pay. This represents the backward linkage — the positive effect of economic density on the income of workers (Fujita et al. 2000).

iv) **Real Wages**

Since the price of agricultural good is normalized to equal one in all regions, and the share of manufacturing goods in total expenditure is determined to be μ, the real wage rate can be defined as:

$$W_r = \omega_r G_r^{-\mu} \qquad (1.4)$$

C. Determination of Equilibrium

The model is said to be at equilibrium when solutions are obtained simultaneously for the income equation, the price indices, the nominal wage equations, and the real wage equations (Fujita et al. 2000). Alternatively, the model is at equilibrium if manufacturing workers are receiving a real wage rate that is at least as high in their current location as in other locations, and hence have no incentive to move. That is, at equilibrium, we require: $W_r = W_s, s = 1, 2, 3, \dots\dots, R \text{ and } s \neq R$. This, however, requires solving four non-linear simultaneous equations, which is not tractable. To address this, Krugman (1991) suggested limiting the number of regions to two (r=1, 2) and to assume that agricultural laborers are evenly distributed between the two regions (i.e., $\phi_{1=}\phi_2 = 1/2$). This special case is known as the *core-periphery* model. The equilibrium in this model is determined recursively by assuming that all manufacturing firms are concentrated at a single point — the 'core' of the economy — and checking whether this state is self-sustaining or not. With these modifications, the model appears as follows:

Since $\phi_{1=}\phi_2 = 1/2$, $\lambda_{1=}\lambda$, and $\lambda_{2=}1-\lambda$, then from equations (1.1-1.4) and for the two regions:

$$Y_1 = \mu\lambda\omega_1 + \frac{(1-\mu)}{2} \qquad (1.5)$$

$$Y_2 = \mu(1-\lambda)\omega_2 + \frac{(1-\mu)}{2} \tag{1.6}$$

$$G_1 = \left[\lambda \omega_1{}^{1-\sigma} + (1-\lambda)(\omega_2 T)^{1-\sigma}\right]^{1/_{1-\sigma}} \tag{1.7}$$

$$G_2 = [\lambda(\omega_1 T)^{1-\sigma} + (1-\lambda)\omega_2{}^{1-\sigma}]^{1/_{1-\sigma}} \tag{1.8}$$

$$\omega_1 = [Y_1 G_1^{\sigma-1} + Y_2 T^{1-\sigma} G_2^{\sigma-1}]^{1/_{\sigma}} \tag{1.9}$$

$$\omega_2 = [Y_1 T^{1-\sigma} G_1^{\sigma-1} + Y_2 G_2^{\sigma-1}]^{1/_{\sigma}} \tag{1.10}$$

$$W_1 = \omega_1 G_1^{-\mu} \tag{1.11}$$

$$W_2 = \omega_2 G_2^{-\mu} \tag{1.12}$$

Now assume that region 1 is the "core", and region 2 is the "periphery". That is $\lambda = 1$. Let us start with $W_1 = 1$. If this represents an equilibrium, then it should be self-sustaining. That is $W_2 \leq 1$.

From equations $1.5 - 1.8$, we get that

$$Y_1 = \frac{(1+\mu)}{2}\ , Y_2 = \frac{(1-\mu)}{2}\ , G_1 = 1, G_2 = T$$

Replacing for the real wage in region 2 in equation 1.12,

$$W_2 = T^{-\mu}\left[\frac{(1+\mu)}{2}T^{1-\sigma} + \frac{(1-\mu)}{2}T^{\sigma-1}\right]^{1/_{\sigma}} \tag{1.13}$$

Rearranging,

$$W_2{}^{\sigma} = \frac{(1+\mu)}{2}T^{1-\sigma-\mu\sigma} + \frac{(1-\mu)}{2}T^{\sigma-1-\mu\sigma} \tag{1.14}$$

Based on equation (1.14), the following are the possible scenarios. **Scenario 1**: transportation cost is higher than one. In this case, the second term in 1.14 becomes arbitrarily high. Therefore, the core-periphery model would not be in equilibrium. **Scenario 2**: $T = 1$ (no transportation cost). This implies, $W_2 = 1$ and hence, location does not matter. **Scenario 3**: a small increase in transport cost at T=1. Totally differentiating 1.14 and evaluating the derivative at $T = 1$ yields:

$$\frac{\partial W_2}{\partial T} = \frac{\mu(1-2\sigma)}{\sigma} < 0 \tag{1.15}$$

Equation 1.15 suggests that at small levels of transportation cost, since $W_2 < W_1 = 1$, the core-periphery model is self-sustainable. Figure 1.1 summarizes this relationship between transportation cost and the long-run spatial equilibrium. It shows that at a sufficiently high transportation cost, higher than T(B), the economy exhibits a unique equilibrium whereby manufacturing is equally divided between the two regions (the share of each region is ½, $\lambda = 1/2$). When the transportation cost declines below a certain threshold level, all manufacturing firms concentrate in region 1- the 'core' of the economy (Fujita et al. 2000; Krugman 1991).

The mechanism that ensures stable equilibrium in this model is what is commonly referred to as the 'home market effect' (Krugman 1991). The process follows the following pattern. **First**, a large market, due to its sheer size, attracts a higher demand for manufactured goods. As a response, firms concentrate at this location, subsequently pushing nominal wages up. **Second**, the local competition among the firms lowers product prices leading to higher real wages. **Third**, the rise in real wages induces the flow of more labor, further encouraging agglomeration. The combination and supporting interaction of these factors lead to the eventual agglomeration of all firms and consumers/workers in a single region — the core of the economy, while the other region forms the periphery.

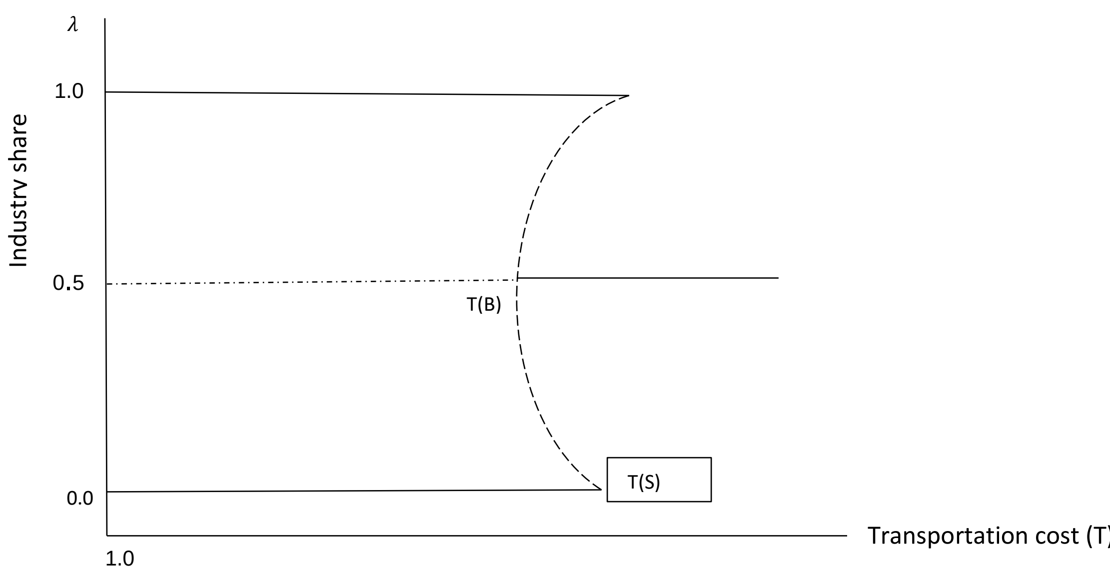


Figure 1.1. Transportation cost and industry share in the core-periphery model
Source: Adapted from (Fujita et al. 2000)
Note: The solid lines indicate stable equilibria and the broken lines indicate unstable equilibria.

The degree of this concentration largely depends on the level of transport cost[9]. If transportation cost is prohibitively high, it is not profitable to conduct transactions over space and each local market would only serve consumers in its respective neighborhood. When transportation cost declines and, in general, economic integration gets deeper, the intensity of the agglomeration force increases whereas the intensity of the dispersion force decreases. This allows firms to exploit their economies of scale more intensively. The deepening of integration also reduces the advantages associated with geographical isolation in the small region where there is less competition. These two effects push toward more agglomeration of the manufacturing sector,

[9] While transportation cost is used for simplicity, it broadly represents transaction cost that hinder spatial integration. This constitutes information costs, transport costs, and tariff and non-tariff barriers.

inferring that, as transport costs reduce, the small region becomes de-industrialized to the benefit of the larger one[10].

1.2.2. The bell-shaped pattern of spatial development

The canonical core-periphery model relies on a set of strong assumptions, the relaxation of which leads to a different pattern of spatial development. The model assumes that the agglomeration of firms and workers at the core continues indefinitely. However, a growing concentration of firms and workers might lead to undesirable consequences including congestion, pollution, and crime. It might also lead to higher housing costs and a longer commute as the concentration of industry intensifies competition for land. At the extreme, the implicit and explicit costs associated with these factors might more than offset the higher real wage workers receive in the agglomerated location. In other words, even when real wages increase with employment density, housing and commuting costs, as well as pollution and crime rates, could make such large agglomerations less attractive (Autor 2020; Glaeser 2020; Lafourcade and Thisse 2008).

The increased importance of the congestion costs at large agglomerations has brought what is known as a bell-shaped pattern of spatial development to the fore (Fujita 2010; Lafourcade and Thisse 2008; Tabuchi and Thisse 2002). The model hypothesizes that as transport cost falls, the spatial economy rather exhibits a bell-shaped pattern. That is, it evolves over three stages: dispersion, agglomeration, and re-dispersion. In the third stage, when transportation cost is sufficiently low and concentration exceeds some threshold level, firms and workers re-disperse away from large agglomerations to alleviate the corresponding congestion costs. At such limits, high commuting costs together with pollution and high crime rate are sufficient to prevent the formation of an oversized primate city and ensure the distribution of economic activities over several small, medium, and large cities (Lafourcade and Thisse 2008; Tabuchi 1998; Thisse 2011). Figure 1.2 demonstrates the bell-shaped pattern of spatial development.

There are additional factors that reinforce the bell-shaped pattern of spatial development. First, labor might not be as mobile as it is assumed in the canonical model. In the model, workers are assumed to be homogenous in their preference and react only to real wage differentials in their migration decision. However, workers might differ in their valuation of non-economic factors affecting the quality of their life including the amenities and social capital, especially once they achieved minimum material welfare. In these situations, the wage premium at the core needs to be substantial to attract additional workers (Tabuchi and Thisse 2002; Thisse 2011). Second, the development of better communication infrastructure might also reduce the association between economies of scale and agglomeration. Once communication technologies are advanced and their costs are sufficiently reduced, firms could cut their transportation cost and increase their market access without proportionate loss in economies of scale. One possible method is through vertical linkage by relocating their production activities to low-wage regions while keeping their strategic functions at the core (Fujita and Thisse 2006; Lafourcade and Thisse 2008; Thisse 2011).

[10] It is important to the note that this core-periphery analysis describes only about the distribution of welfare not about the change in total welfare. That is, agglomeration of economic activity is said to increase the distribution of the national income in favor of the larger region. This, however, does not mean that the periphery gets poorer. In fact, numerous extensions of the NEG model suggests that agglomeration is, overall, Pareto improving (Lafourcade and Thisse 2008; World Bank 2009).

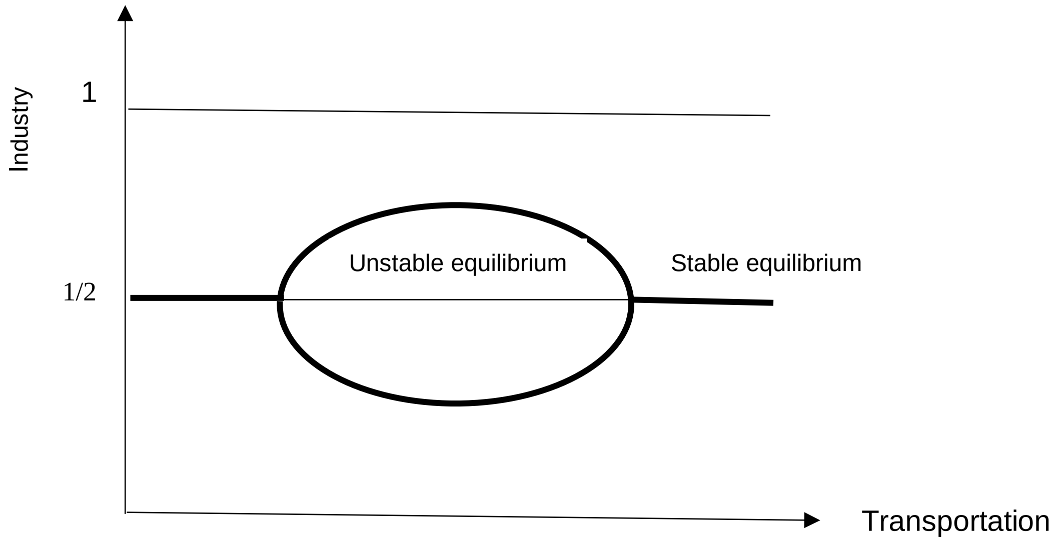


Figure 1.2. The bell-shaped pattern of spatial development
Source: Figure adopted from Lafourcade & Thisse (2008)

Two important factors that are highly likely to affect the spatial pattern of welfare, particularly in developing countries, are left out in the NEG model, canonical, or otherwise. These are **human capital** and **amenities**[11]. The Human Capital Theory (HCT), for instance, argues that the demand for and supply of human capital endowments is the main driver of the spatial wage differential. The proponents of the theory suggest that there is considerable spatial sorting by education and skill, and this explains a significantly large share of the spatial variation in welfare (Combes, Duranton, & Gobillon, 2008). Another strand of literature — the Spatial Equilibrium Model (SEM) — argues that regional differences in wellbeing are explained by differences in natural amenities such as raw materials endowment and climate features (Gollin, Kirchberger, and Lagakos 2017; Roback 1982)[12]. This is particularly important to explain wage inequality in countries where the exploitation of natural resources is a key source of regional income. Broadly defined, SEM is also important to explain the spatial variation in wellbeing in heavily agriculture-dependent countries like Ethiopia where the location of amenities expressed in terms of climate condition, soil fertility, and slope play a considerable role in determining productivity and overall welfare (Christiaensen, Demery, and Kuhl 2011; Haggblade, Hazell, and Reardon 2010; Stifel, Minten, and Dorosh 2003).

This study aims to consolidate these three analytical concepts — the NEG, the HCT, and the SEM — to examine the pattern in and the deriving factors of the spatial economy. The insights generated from this analysis are helpful to inform policies targeted to improve overall household welfare as well as reduce spatial disparity.

[11] Another important explanation for the spatial economy relates to geographical favoritism in government policies. Government policies could be biased towards one geographical location or to a sector at a particular location in terms of taxation, price regulation, and investment/spending. Examples of these policies are urban-biased and industry-focused policies followed by developing countries in 1970s (von Braun 2007) and smallholders-focused rural policy (Collier and Dercon 2014).

[12] This corresponds to the "first nature" geography, according to which some regions are favored to others because they are amenable to human habitation, output production, and the transport of goods (Gallup, Sachs, and Mellinger 1997; Henderson et al. 2017; Venables 2005). In contrast, the underlying factors of spatial pattern in NEG are sometimes called "second nature" geography.

1.2.3. Patterns of urbanization and economic development in Ethiopia

As discussed in the previous sub-section, the bell-shaped pattern of spatial development hypothesizes that the fall in transportation cost sets in motion a chain of self-reinforcing events. **First**, as transportation cost falls, firms agglomerate into a central place to benefit from economies of scale and to have access to a larger market. The concentration of firms increases the nominal wage rate (due to competition in the labor market) and decreases the price of output (due to competition in the product market), leading to a higher real wage rate for the workers at the core. **Second**, while more workers initially move into the core in response to the increased real wage, this lasts only until the benefit from dispersion (low crime, less commuting, better housing) is sufficient enough to compensate for the lower income at less agglomerated locations. At equilibrium, the tension between agglomeration and dispersion forces will lead to the distribution of economic activities over several small, medium, and large cities. This sub-section highlights the pattern of urbanization in Ethiopia in light of the prediction of the model.

Ethiopia has witnessed remarkable growth in road networks over the last three decades and this has resulted in a substantial decline in transportation costs (Shiferaw, Siba, and Alemu 2012; World Bank 2015). In line with the prediction of the NEG, the rate of agglomeration has also intensified over the same period. In 1990, there were only 78 urban areas with more than 10,000 inhabitants in the country, accommodating less than 10 percent of the population. Over the subsequent 25 years, the number of these urban areas increased to 510 and the share of the population residing in such urban areas jumped to 27 percent (see Figure 1.3 and Table A1.1 in the Appendix). It is not only that the number of urban areas has increased, but the size of the existing ones has also expanded. For instance, between 2000 and 2015 alone, the population of the capital city, Addis Ababa, more than doubled from 2.4 million to 4.5 million (Dorosh and Thurlow 2013); see also Table 1.3.

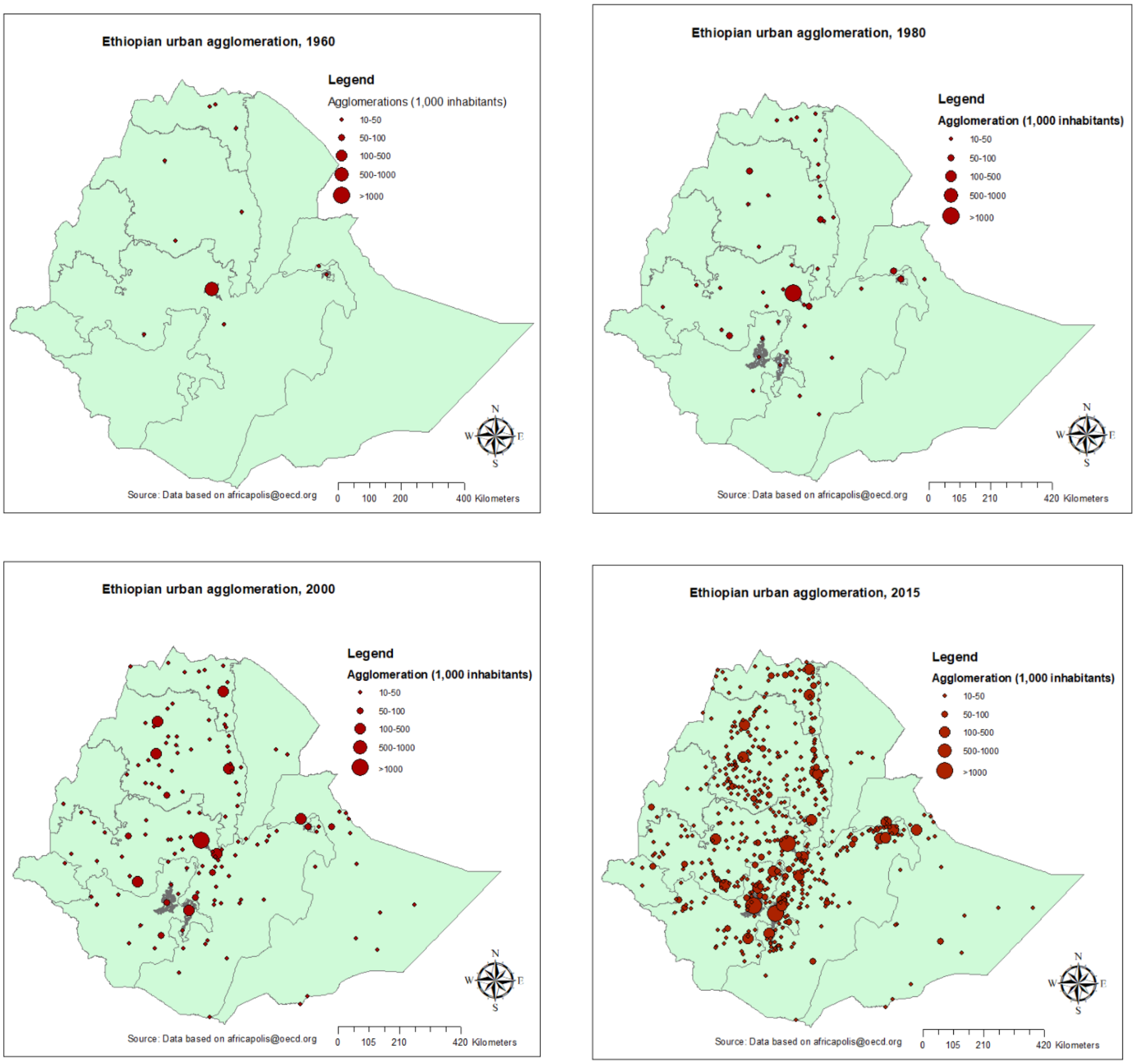


Figure 1.3. The pattern of urbanization in Ethiopia (1960-2015)
Source: Map based on data from africapolis@oecd.org

With urbanization (agglomeration), NEG hypothesizes that firms cluster together, which enhances the welfare of the household through better wages, variety supply, and increasing income (Fujita et al. 2000). In line with this prediction, the distributions of economic development and household welfare also exhibit a spatial pattern in Ethiopia. Table 1.1 and Table 1.2, respectively, present the spatial distribution of enterprises and household welfare in Ethiopia. Table 1.1 shows that as of 2015, Addis Ababa hosts more than 3 times the number of enterprises compared to other locations in the country combined[13]. By and large, this is the case regardless of the type of activities the enterprises engage in or their size (see also Appendix Table A1.2).

[13] This does not include Microenterprises. However, the pattern of the distribution of enterprises that includes Microenterprises is generally the same (OECD/PSI 2020).

Table 1.1. Spatial distribution of Ethiopian enterprises by type of engagement

Administrative Regions	Food	Textile & Garments	Non-Metallic Minerals	Other Manufacturing	Retail	Transport	Other Services	Grand Total
Panel A: Number of enterprises								
Total	1,702	534	334	1,880	9,168	23,807	34,397	71,822
Addis Ababa	322	250	160	1,233	4,909	18,441	28,703	54,018
Amhara	80	12	10	62	553	549	1,398	2,664
Dire Dawa	22	2	2	16	5		26	73
Oromia	581	171	40	221	2,735	4,700	3,036	11,484
SNNPR	551	5	15	37	69	4	220	901
Tigray	146	94	107	311	897	113	1,014	2,682
Panel B: Share of enterprises (%)								
Addis Ababa	18.9	46.8	47.9	65.6	53.5	77.5	83.4	75.2
Amhara	4.7	2.2	3.0	3.3	6.0	2.3	4.1	3.7
Dire Dawa	1.3	0.4	0.6	0.9	0.1	-	0.1	0.1
Oromia	34.1	32.0	12.0	11.8	29.8	19.7	8.8	16.0
SNNPR	32.4	0.9	4.5	2.0	0.8	0.0	0.6	1.3
Tigray	8.6	17.6	32.0	16.5	9.8	0.5	2.9	3.7

Source: World Bank enterprise survey document, 2015.

Note. Regions represent sub-national administrative classification.

Table 1.2 shows the pattern in household welfare at different stages of urbanization. Consistent with the prediction of NEG, welfare improves with agglomeration. It shows that per capita consumption in large urban areas is more than 50 percent larger than in rural areas. Furthermore, households in large urban areas, by comparison, have a more diverse diet, are more food secure, and the share of those at the bottom 40 percent of the income distribution is much lower.

Table 1.2. Patterns in household welfare by urbanization status

Place of residence	Per capita expenditure (in ETB)	Share of Poor (%)	Diet Diversity	Food security gap (%)	Number of food insecure months
Rural	6,209	45.8	5.8	28.8	3.3
Intermediate towns	8,293	32.3	6.8	30.1	3.2
Large towns	9,490	21.9	7.8	10.7	3.9
Total	7,074	40.0	6.3	28.0	3.3

***Source**: Authors' computation based on LSMS-ISA (2014 & 2016)*

***Notes**. Per capita expenditure refers to total household expenditure on food and non-food items per capita per year, after adjusting for inflation and differences in cost of living; Share of poor represents the share of households at the bottom 40% of the income distribution. Diet diversity represents the number of different food groups (out of 12) that households consume. The food security gap represents the share of households that reported to have faced food security issues over 12 months prior to the survey, and the number of food-insecure months is the measure of the severity of the problem.*

Table 1.3 shows the evolution of the settlement pattern of the population across the hierarchy of cities. It indicates that while urbanization has increased, it has increased faster in small and medium-sized towns. The pattern of agglomeration between 1950 and 2015 reveals that up until 1990, Addis Ababa was the only large city with a population number larger than 100,000, however by 2015, 24 other urban centers exceeded this threshold (Table 1.3). The Table also shows that while the capital city accounted for about 72 percent of the total urban population in 1950, this share has continuously been declining to reach only 15.3 percent in 2015. The accelerating growth of the urban population in general and the deceleration in the share of the capital city in the total urban population implies that the growth of other urban areas has been much faster[14]. Figure 1.4 compares the average annual population growth rate of Addis Ababa with the overall growth rate and the growth rate in other urban areas. It shows that the growth rate in Addis Ababa is diverging downwardly from the national figure, particularly since 1990.

Table 1.3. Distribution of Ethiopian population by the size of urban centres (1950-2015)

Size of agglomerations	1950	1960	1970	1980	1990	2000	2010	2015
Panel A: Urban Population (1,000s)								
Total	**503**	**778**	**1,341**	**2,385**	**3,895**	**6,521**	**11,054**	**24,292**
10-100	140	236	545	1,098	1,854	3,038	5,925	11,556
100-500	363	-	-	-	237	1,077	2,144	4,580
500-3000	-	542	796	1,288	1,805	2,406	2,986	4,445
3000+	-		-	-	-	-	-	3,711
Panel B: Number of urban agglomerations								
Total	**6**	**11**	**24**	**45**	**78**	**147**	**288**	**510**
10-100	5	10	23	44	75	138	275	485
100-500	1	-	-	-	2	8	12	22
500-3000	-	1	1	1	1	1	1	2
3000+	-	-	-	-	-	-	-	1
Panel C: Share of population in urban agglomerations (%)								
10-100	27.8	30.4	40.6	46.0	47.6	46.6	53.6	47.6
100-500	72.2	-	-	-	6.1	16.5	19.4	18.9
500-3000	-	69.6	59.4	54.0	46.3	36.9	27.0	18.3
3000+	-	-	-	-	-	-	-	15.3

Source: Authors' Computation based on data from africapolis@oecd.org

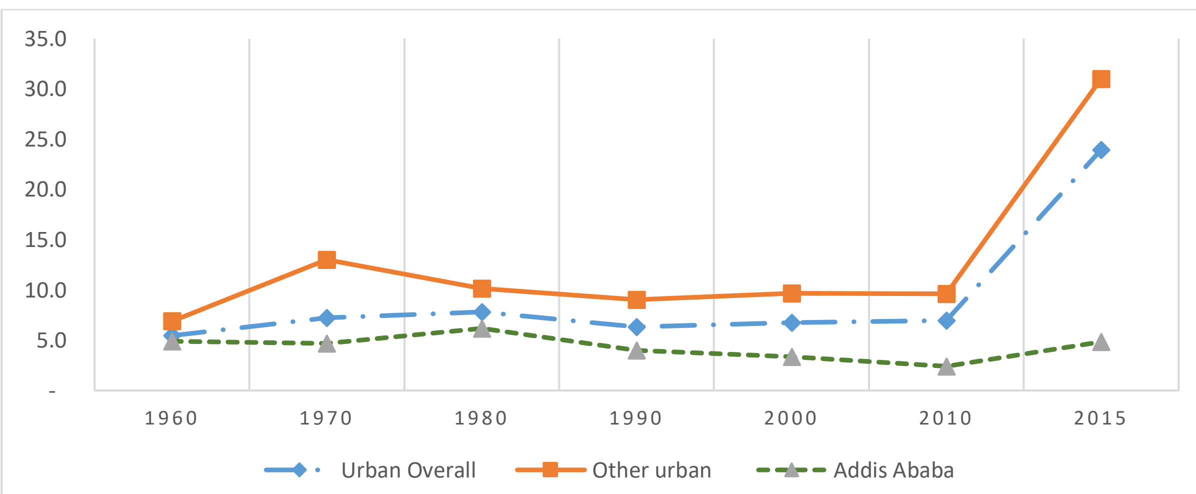


Figure 1.4. The growth rate of urban population by urbanization status (1950-2015)
***Source**: Authors' Computation based on data from africapolis@oecd.org*

[14] This is consistent with the pattern seen across Africa. In Africa the growth rate of the small and medium sized towns is more than twice as large as the growth rate in large towns (UN Habitat 2010, 2014).

According to the bell-shaped pattern of spatial development, the dispersion of workers and firms away from the core is due to congestion at the core and — the main indicators of which include: higher commuting costs, higher housing rent, and a higher rate of crime and pollution. Table 1.4 presents the average rental price of a room, the average expense of a household on transportation, and the proportion of households that reported transportation costs[15]. This Table suggests that the average rental price of a room in the capital city is twice as large as in other large towns and more than 3 times the amount in rural areas and small urban areas. Moreover, and perhaps unsurprisingly, the proportion of households that paid for transportation and their average payment is greater in the capital city than elsewhere.

Table 1.4. House rent, transportation cost, and share of commuters, by urbanization status

Location of residence	Rent per room (in ETB)	Expenditure on Transportation (in ETB)	Commuters (%)
Rural	335.7	26.3	30.8
Small town	475.9	61.8	36.8
Other large towns	528.4	62.3	52.1
Addis Ababa	1,093.7	99.2	77.2
Total	**412.5**	**38.0**	**36.7**

Source*: Author's computation based on LSMS-ISA (2014 &2016)*

Note*. Rent per room is calculated by dividing the total monthly rent by the number of rooms of a residential place. This is based only on households who reported to have rented a residential place. Expenditure on Transportation is computed as the average amount of money spent by households for transportation. The share of commuters is the share of households who spent money on transportation.*

1.3. Objective and outline of the thesis

The primary aim of this thesis is to improve our understanding of the relationship between urbanization and economic development, in order to advance policies that can help harness the potential of the urbanization process to achieve sustainable economic growth, poverty reduction, and balanced spatial development. To achieve this, it employs a novel, continuous measure of urbanization — the sum of nighttime light — to account simultaneously for the continuum between rural and urban areas as well as for the heterogeneity of urban areas. The central focus of the thesis is on the analysis of underlying mechanisms of the spatial economy to inform policies targeted to improve overall welfare while reducing unbalanced spatial development. To that end, the analytical chapters subsequently link one dimension of welfare indicators (e.g. food security, nutrition, labor market outcomes) carefully with the major sources of the underlying mechanisms. The last chapter, as a case study, examines one of the understudied underlying factors of spatial development – public service delivery. Based on the synthesis of the findings from the different chapters, the thesis identifies a menu of potential policy interventions relevant to enhance overall welfare while mitigating the inequality across the rural-urban spectrum.

The thesis is organized in such a way as to capture the effect of urbanization on welfare across different generations. In successive chapters, it examines the effect of urbanization on broader indicators of household welfare (**Chapter 2**), the health and nutritional outcomes of children

[15] Due to lack of data, spatial patterns in pollution and crime rate are not computed

(**Chapter 3**), and its effect on reinforcing intergenerational mobility (**Chapter 4**). The main research questions explored in this thesis are the following:

1. Does the effect of urbanization on household welfare depend on the degree of urbanization? If so, what are the main underlying factors?
2. What are the heterogeneous effects of urban proximity on nutritional outcomes?
3. Does the degree of urbanization influence the extent of intergenerational mobility?
4. Which interventions are effective to improve the delivery of agricultural extension services in remote areas?

The analytical chapters in this thesis address these four interrelated research questions, and each shall be a stand-alone academic paper with substantive contributions to the literature.

Chapter 2 focuses on identifying whether and how the patterns of urbanization are associated with household welfare in Ethiopia as well as exploring the major underlying mechanisms. The data used for this analysis comes from three rounds of Ethiopian LSMS-ISA (2012, 2014, and 2016) geospatially linked to nighttime light data. Based on the New Economic Geography (NEG) framework and threshold data analysis, the findings of this chapter suggest that the effect of the pattern of urbanization is at least as important as the aggregate rate of urbanization. Specifically, the findings indicate that intermediate towns are more strongly associated with household welfare as compared to large towns, small towns, or the rural hinterland. The chapter emphasizes the roles of market access, employment opportunities, and differential access to public services as major underlying mechanisms.

Chapter 3 extends the analysis in Chapter 2 and investigates the effect of proximity to town and the heterogeneous effects of the size of towns on nutritional outcomes. For identification, the study combines an Instrumental Variables (IV) approach with Inverse Probability Weighting (IPW). Using nationally representative LSMS-ISA household and community survey data, the study finds that both the proximity to urban areas as well as the size of the proximate urban areas affect households' nutritional status. More specifically, while proximity to towns has a strong positive effect on nutritional status, households surrounding large towns are better off compared to those around small towns. Reducing the cost of transportation to the nearest town by half leads to a 0.3 percentage point increase in diet diversity and a 0.8 percentage point reduction in child stunting. The results corresponding to the size of towns suggest that while the diet diversity of households in large towns is likely to be higher by 1.2 percentage points, child undernutrition is likely to be lower by about 4 percentage points.

Chapter 4 examines the role of the degree of urbanization of place of residence on the extent to which inequalities in economic and social status are transmitted across generations. Based on the intensity of the nighttime light (NTL) at the place of residence as a marker of urbanization, the chapter presents strong evidence of the interaction between parental characteristics and urbanization. In general, children whose parents are employed in better-paying occupations are more likely to be employed in similarly better-paying occupations themselves, and this intergenerational correlation is more pronounced in large urban areas. Moreover, it shows that the inequality observed in occupational opportunities in large urban areas is explained mainly by differences in educational attainment. Once individual education level is accounted for, large urban areas offer better employment mobility than rural areas and small towns. This suggests that broadening access to and reducing the dropout rates at post-elementary schools and improving the quality of education are the most effective mechanisms to reduce spatial and intergenerational inequality in living standards in Ethiopia.

Chapter 5 assesses the responsiveness of agricultural extension agents (EAs) to potential policy interventions by employing a discrete choice experiment (DCE) design. Using a carefully designed DCE and a novel quantitative approach, the chapter offers several interesting insights. First, there is a general dissatisfaction among the EAs with their current work and living conditions. Second, contrary to popular perception, increasing salaries is not always the strongest incentive for EAs. The findings suggest that offering educational opportunities is by far the most powerful instrument to attract and retain EAs in remote locations. Upward salary adjustment only comes in at a second position, followed by the provision of housing and transportation facilities. EAs are also likely to respond to such incentives as the availability of basic amenities (electricity, drinking water, mobile telephone network) in the villages to which they are posted, as well as the provision of adequate work materials.

CHAPTER **2**

2. Patterns of urbanization and household welfare

Abstract

Countries in sub-Saharan Africa are urbanizing at an unprecedentedly rapid rate. This has intensified interest in the effects of the pattern of urbanization. This study combines parametric and non-parametric regression methods to examine whether and how urbanization and its different stages are associated with household welfare. The main data used in the chapter is from two rounds of LSMS-ISA (2014, 2016) for Ethiopia. The findings suggest that while urbanization positively affects household welfare, there is notable variation across the hierarchy of urban areas in Ethiopia. In general, intermediate towns are more strongly associated with household welfare as compared to large towns, small towns, or the rural hinterland. The disparities in access to markets, employment opportunities, and public services among and within these locations are emphasized as major underlying mechanisms for the differences in welfare. From the analyses throughout, this chapter concludes with the impact such findings have on policy decisions.

JEL Classification: C38, I30, R12, O18

Keywords: Welfare Economics, Threshold Analysis, Size and Spatial Distributions of Regional economic Activity, urbanization, Ethiopia

2.1. Introduction

The distribution of income and other welfare indicators across space suggests that geography is a good predictor of economic development. At the micro-level, empirical evidence indicates that the location where people live is strongly associated with their welfare outcomes[16]. In particular, there is a large and persistent gap in living standards between rural and urban areas (Kraay and McKenzie 2014; Ravallion, Chen, and Sangraula 2007; Sahn and Stifel 2004). In the context of African countries, using cross-sectional data from 24 African countries and wide-ranging welfare indicators, Sahn & Stifel (2004) found that urban households are significantly better off compared to rural households. However, while such an aggregate comparison of rural-urban averages is helpful to indicate levels of spatial disparity in welfare, it is often criticized for being less informative and less useful for two main reasons. First, there is no universal consensus on what constitutes an urban area, even within the same country. Apart from often being subjective, a simple survey and census-based rural-urban binary measure tends to reflect merely political and bureaucratic dispositions, rather than services the spaces provide (Henderson 2010; Satterthwaite and Tacoli 2003; UNECA 2017).

Second, with the rapid urbanization and improvements in information and communication infrastructure, the use of a binary rural-urban classification has proven to be too simplistic to represent the complex reality of urbanization (von Braun 2014b; Muzzini 2008; OECD/PSI 2020)[17]. Rural-urban space is increasingly viewed as a continuum with numerous intermediate stages, ranging from small towns to peri-urban areas (von Braun 2014a). Failure to account for this continuum hinders the micro-level analysis of welfare impacts of urbanization (Christiaensen and Todo 2014; Ingelaere et al. 2018; Satterthwaite and Tacoli 2003).

To address these challenges, researchers have developed several alternative indicators of urbanization including population size, population density, index of infrastructure, and distance to market and urban areas (Deichmann et al. 2009; De Poel et al. 2012). One of the latest and the most promising innovations is the use of nighttime light (NTL) as a measure of urbanization. NTL — an indicator of the intensity of light emitted from the earth at night — offers a unique potential for measuring urbanization and urban expansion. As NTL is a basic urban amenity, its intensity per unit area could be used as a valid indicator of urbanization (Abay et al. 2020; Henderson et al. 2003). NTL not only provides an alternative continuous metric of urbanization, but also has the advantage of consistent quality measurement at a high spatial resolution, and it is available over a long time period to allow reliable temporal analysis (Abay et al. 2020).

The availability of NTL as a continuous variable across the rural-urban space also helps to assess the potential heterogeneity among different-sized urban areas. Recent studies show that a credible estimation of the welfare impact of urbanization should factor in the implications of the *pattern* of urbanization (Christiaensen, De Weerdt, and Todo 2013; Kanbur et al. 2019). That is, the pattern of urbanization — the differences in the growth rates of large, intermediate, and small

[16] At macro level, studies have shown the existence of strong correlation between location of countries and the level of economic development (Bloom et al. 2010; Collier and Gunning 1999; Sachs et al. 2001; World Bank 2009).

[17] Another common approach to measurement of the degree of urbanization is the use of distance to the nearest urban area. As will be discussed in the next chapter, alone, this as well is inadequate to capture the full picture of urbanization.

towns — significantly affects the link between urbanization and welfare. Such a disaggregated study of urbanization across different stages of urban development is particularly useful in the African context where larger cities of 5 million or more inhabitants accounts for only 10 percent of the urban population. Small and medium-sized towns do not only host the majority of the continent's urban population, the population in these urban areas has also doubled in the last decade and is projected to grow by more than 30 percent in the next decade (UNDESA 2015).

Regardless of this, rigorous empirical research on the heterogeneous effect of different-sized urban areas in Africa is scant. This study attempts to bridge this gap. It combines large, nationally representative household- and community-level survey data with satellite-based nighttime light intensity data to examine the effects of the patterns of urbanization on the welfare of households in Ethiopia. The main focus is to identify whether and how urbanization and the different types of urban areas in Ethiopia are related to household welfare, and then to explore the major underlying mechanisms. In doing so, the study aims at parsing out the dynamics of household welfare not only between rural and urban areas but also among the different sized urban areas. This aligns with the recent evidence of the roles of intermediate (secondary) towns in employment generation and overall poverty reduction in developing countries (Christiaensen, De Weerdt, and Kanbur 2013; Dorosh and Thurlow 2013; Kanbur et al. 2019). However, unlike these studies, the analysis in this chapter does not rely on the administrative definition of urban areas. In this regard, it follows the nascent flourishing literature where satellite-based NTL information is being used to proxy urbanization and urban growth (Abay et al. 2020; Amare et al. 2017; Ameye 2018; Henderson 2014; Henderson, Storeygard, and Weil 2009).

The data used in this paper comes from the Living Standard Measurement Study-Integrated Surveys on Agriculture (LSMS-ISA). For Ethiopia, the LSMS-ISA is collected by the Central Statistical Authority of Ethiopia (CSA) in collaboration with the World Bank. The GPS information in this longitudinal dataset is used to merge it with satellite-based secondary datasets, including nighttime light and population density. Non-parametric regression analysis, as well as alternative parametric econometric approaches, are employed to examine the relationship between urbanization and household welfare.

The study finds that urbanization, represented by the sum of nighttime light (SOL), is a robust predictor of household welfare. However, the relationship is non-linear. It rather resembles an s-shaped pattern, whereby welfare measures increase slowly at first and then increase at an increasing rate before flattening out at a more advanced stage of urbanization. In general, the analysis indicates that intermediate-sized towns are associated with better household welfare compared to the rural hinterland, small towns, and large towns. These results have important policy implications for the design of poverty reduction interventions in both rural and urban areas.

The remainder of the chapter is organized as follows. The next section first describes the data sources, then presents measurement issues of urbanization and the outcome variables. This is followed by descriptive statistics. While Section 2.3 presents the estimation result of the spatial pattern in household welfare over the urbanization gradient, Section 2.4 explores the channels through which these spatial patterns may occur. Section 2.5 provides some robustness checks and sensitivity analyses before section 2.6 concludes the chapter.

2.2. Data sources, measurement, and descriptive results

2.2.1. Data sources

For the main part of the study, the Ethiopian Living Standard Measurement Study - Integrated Surveys on Agriculture (LSMS-ISA) is merged with satellite-based NTL data. This sub-section briefly describes these two datasets.

A. Living Standard Measurement Study (LSMS-ISA) data

The LSMS-ISA is a rich geo-referenced and nationally representative longitudinal dataset collected jointly by the Central Statistical Authority (CSA) of Ethiopia and the World Bank over three rounds in 2012, 2014, and 2016[18]. In the first wave, the data was collected only from rural areas and small towns. The subsequent two rounds encompassed larger urban areas while maintaining available samples from the first round. The questionnaires are comparable across waves and include household and Enumeration Area (EA) [19] level surveys. The household survey collected detailed information on, *inter alia,* households' agricultural activities, labor market participation, food security, diet quality, and consumption expenditure. The EA (also called community) survey gathered detailed information on the availability of and distance to public services, employment opportunities, market prices, etc. Very importantly, the survey data is geo-referenced, which enabled the author to exploit satellite-based data on NTL.

For this study, only the 2014 and 2016 rounds of the household and EA level data are used. As the objective of the study is to assess the welfare pattern across the entire rural-urban spectrum, the first round – which *only* covered rural areas and small towns – is excluded from the analysis to maintain comparability over time. In order to reduce bias that might result from endogenous dynamic migration decisions, the data analysis is limited to those households remaining in the village during both survey rounds. Table A2.1 in the Appendix presents the summary statistics of key variables over the two survey rounds.

B. The Nighttime Light (NTL) data

There is a lack of a consistent and universally accepted definition of urbanization and urban area even within one country (Satterthwaite and Tacoli 2003; World Bank 2009). Typically, each country has its own definition of what constitutes an urban area, and researchers and international organizations tend to adopt this definition. This might partly explain both the rarity of a rigorous empirical study of urbanization and the inconsistency in the results of the effect of urbanization among the limited existing empirical studies (Henderson et al. 2003, 2009). In order to address this measurement problem, recent efforts have focused on constructing universally comparable, continuous, and disaggregated indices that capture micro-level variations in urban settlement and urban expansion (Abay et al. 2020; De Poel et al. 2012). Nighttime light (NTL) appeared to be a viable candidate as it helps to capture both spatial as well as intertemporal urban dynamics. Since urban areas generally have higher nighttime light intensities than rural areas, recent empirical

[18] An additional round was collected in 2018/19. However, this is not included as this is a baseline for a new panel, not a follow-up to previous waves.
[19] Enumeration areas (EAs) are equivalent to a village, relatively small, consisting of about 250 households on average.

studies have exploited this satellite-based NTL intensity as a valid marker of urbanization and urban settlements (Abay et al. 2020; Amare et al. 2017; Henderson et al. 2009; Sutton 1997).

This study relies on version 4 NTL time series data from the Defence Meteorological Satellite Program (DMSP), using Operational Line Scanner (OLS)[20]. The available NTL index is a digital number ranging from 0 (no light) to 63 (maximum light) for 1 km^2 pixel. This study makes use of a cleaned and inter-calibrated NTL database made available for Africa for the period 2000-2013 (Savory et al. 2017)[21]. This data was downloaded and merged with the household survey data using the GPS coordinates of the residence place of sampled households[22]. Figure 2.1 presents the geographic distribution of the NTL density of cluster villages in Ethiopia. Within the sample, the NTL ranges from zero in remote rural areas to 61.6 for Addis Ababa[23].

Compared to the aggregate census and survey-based measures of urbanization commonly used in the literature, the NTL data have proved to be a more useful measure since the availability of the data at a high spatial resolution allows construction of spatially detailed measures of urbanization. Furthermore, the index eliminates reliance on the national urban statistics that are often only sporadically available, unreliable, and lag behind reality, especially in developing countries (von Braun 2014b; Satterthwaite and Tacoli 2003). Due to these features, the index has become more popular to delineate urban areas and related human activities (Henderson et al. 2003; Zhang and Seto 2011) and hence promises to hold a huge potential for studying urbanization in sub-Saharan Africa.

However, there are doubts about the robustness of the NTL index as a measure of urbanization. First, NTL is susceptible to measurement problems as it often records positive value where there is actually no light due to a reflection of, for example, lakes or refineries. While one cannot entirely rule out these challenges, their impact is expected to be minimal in Ethiopia as the country is landlocked, and owns no large oil refineries to create flares. Second, the NTL index is measured using pixels that are set to a maximum of 63 and this might result in truncation of pixels at the top – a phenomenon called the *oversaturation of pixels* (Zhang and Seto 2011). For Ethiopia, this is unlikely to be a problem as the maximum NTL density during the survey years is less than 63.

Another concern with using NTL to delineate urban areas is the possibility that it might also pick up local trends in economic activities, electricity, or simply household wealth (Henderson et al. 2009). While the database provided by Savory et al., (2017) has proved to be strongly correlated

[20] The newest set of NTL data that comes from the Visible Infrared Imager Radiometer Suite (VIIRS) Day/Night Band (DNB) of the National Polar- Orbiting Operational Environmental Satellite System (NPOESS) is argued to be superior when compared with DMSP/OLS. But this study uses the DMSP/OLS due to its overlap with the survey data and its availability over long time series (Elvidge et al. 1997, 2017).

[21] This dataset is freely available at https://geodata.globalhealthapp.net/. For the technical aspect of the satellites and the inter-calibration, please refer to Savory et al. (2017). This data is available only for 2000-2013 period. Therefore, for 2015 survey, the data is imputed based on a regression model on the past values, household assets levels and access to infrastructure and electricity. This is similar to poverty mapping in its approach (see Dang, Jolliffe, and Carletto 2019).

[22] To be precise, these are EA level averages. To protect the confidentiality of the sample household, the GPS coordinates in the publicly available version of LSMS-ISA survey data were modified relying on random offset of EA center-point coordinates within a specified range determined by the urban and rural classification. While an offset range of 0-2 km is used for urban areas, 0-5 km offset is used for rural areas. In special circumstances, a maximum of 10 km offset is applied (CSA and World Bank 2017).

[23] That means, the variable is not right-censored and estimation using OLS remains valid.

with other indicators of urbanization, in this study, it is further validated by cross-tabulating the index against the administrative sub-national and urban-rural classification. Table A2.2 in the Appendix shows that the NTL is not only consistent with the census-based classification of urban vs rural areas, but it is also powerful to differentiate among rural areas, small towns, medium towns, and large urban areas. To triangulate the basic results more formally, a detailed sensitivity analysis of the basic result using alternative measures of urbanization (population size and administrative definition) is provided in Section 2.5.

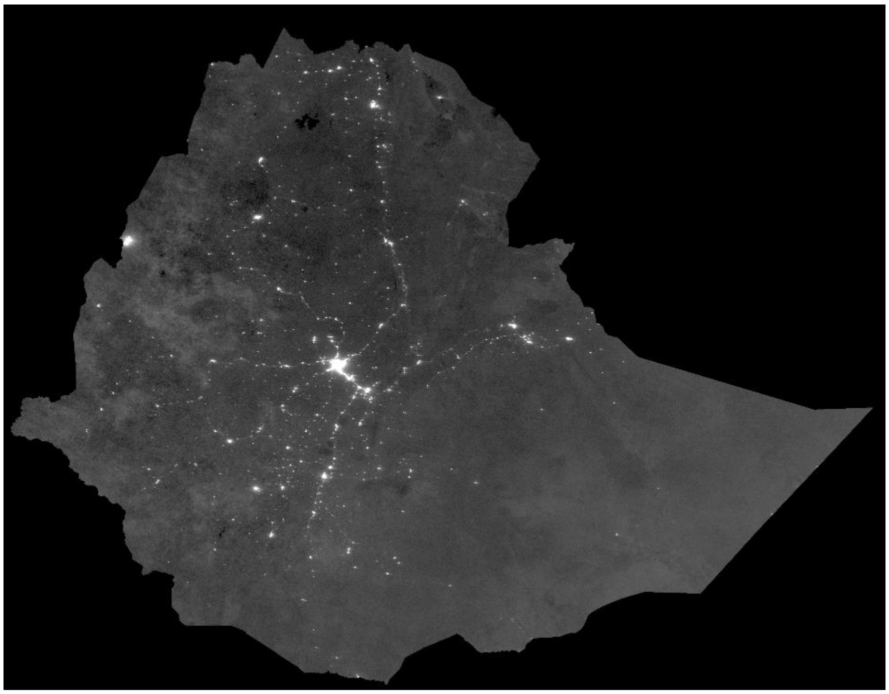


Figure 2.1. Nighttime Light (NTL) intensity of clustered villages in Ethiopia (2015)
Source: Image based on data extracted from Savory et al. (2017)

2.2.2. Measurement of key variables

A. Measurement of urbanization

As pointed out before, the definition of urbanization in this study relies on the NTL dataset, which contains luminous pixels, whose value ranges from 0 to 63 (Figure 2.1). Sample households are clustered at the village level (i.e., Enumeration Areas – EAs) and the average degree of urbanization is determined at the EA level. Depending on the degree of urbanization status of the EAs, the number and intensity of the luminous pixels around the EAs vary considerably. The existence and the size of urban areas around each EA is determined as follows. First, a 10km radius buffer zone is delineated around each EA. Then, the sum of light (SOL) — a variable that adds up the NTL within the 10km radius from the center of the EA — is generated. In this study, urbanization is measured in terms of SOL. Compared to the simple NTL-based approach to urbanization measure, the SOL method commands several advantages. First, since it takes into account both the existence and intensity of lights within the 10km radius buffer zone, it can identify both the existence and the size of urban areas.

Second, it totals the lights from all agglomerations within the delineated buffer zone, and therefore considers the effect of all potential urban centers. This addresses one of the critical shortcomings of the traditional approach where urban influence is measured with respect to the nearest town. Due to this feature, the use of SOL as a measure of urbanization is gaining popularity in empirical research (Gibson et al. 2017; Henderson et al. 2017).

The final advantage of using the SOL approach emanates from the fact that modifications were made to the GPS information when the LSMS-ISA survey data was made available for public use. To ensure the confidentiality of sample households and communities, the World Bank did not provide the original household-level GPS information. Instead, it provided modified coordinates that were cloned from their original levels by applying a random offset of up to 10km[24]. Therefore, the 10km buffer zone created to delineate urban areas in this study eliminates any potential misclassification resulting from the random offsets. Figure 2.2 presents the distribution of SOL in 2014 and 2016.

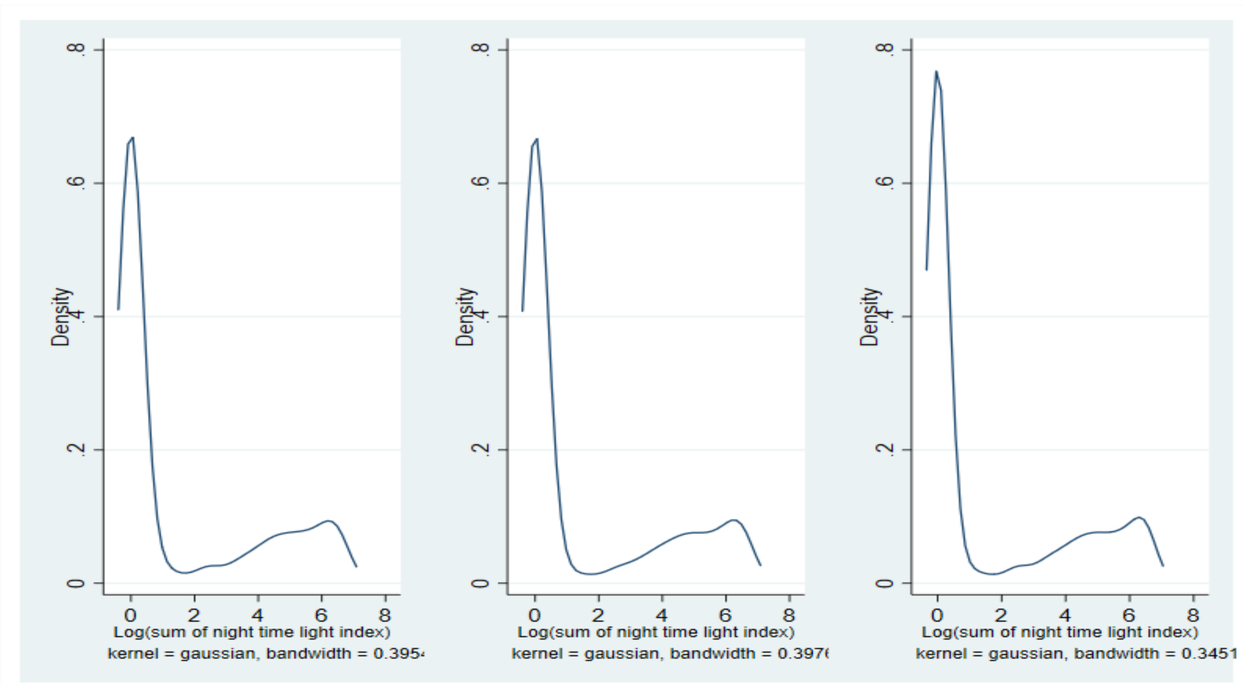


Figure 2.2: Kernel density plots of sum of nighttime lights (Log (SOL))
Note: The graphs represent Kernel density estimates for 2014(left), 2016(center) and pooled (right)

B. Measurement of outcome variables

Three outcome variables — real consumption expenditure per adult equivalent, household dietary diversity index (HDDI), and household food security gap — are used as measures of household welfare.

As a part of the household questionnaire, the LSMS-ISA survey collected detailed information on the consumption of food and non-food items. To minimize recall bias, the information on the consumption of food items was collected on a 7-day recall basis, while information on basic

[24] See https://microdata.worldbank.org/index.php/catalog/2783

household goods (e.g. matches, soap, etc.) and durable assets (e.g. clothing, furniture, etc.) were collected over 1 month and 12 months, respectively. Based on this information, first, the total annual household consumption expenditure (on food and non-food items) is calculated. Then, adjustments were made to account for general price trends over time and differences in cost of living across regions. The general consumer price index (CPI) is used to convert the nominal consumption expenditures into real values. To adjust for differences in the cost of living across different regions, a spatial price index is used[25]. Furthermore, differences in household size and its composition are accounted for by dividing total expenditure by household size based on adult equivalencies. Figure 2.3 presents the distribution of per capita consumption over the two survey rounds. It indicates that real consumption declined between 2014 and 2016. These might partially be attributed to the severe drought in 2015. Some studies estimate that the drought has decreased consumption within affected households by more than 11 percent (Fuje 2018).

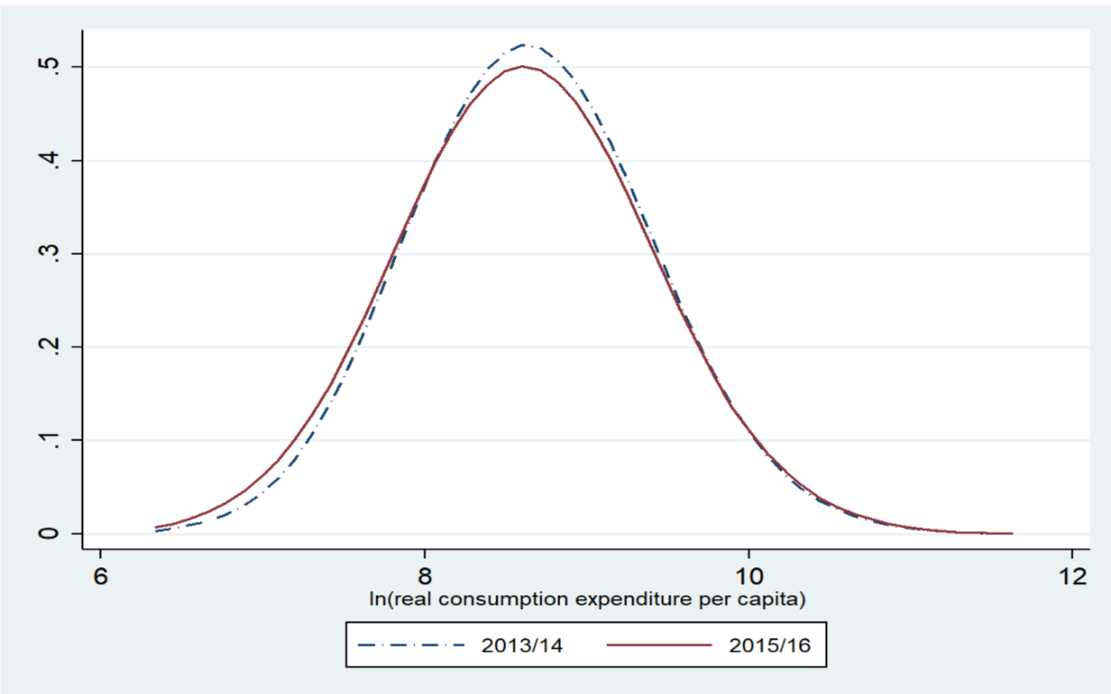


Figure 2.3. Kernel density of real consumption expenditure
***Source**: Author's computation based on LSMS-ISA (2014 & 2016)*
***Note**: Real consumption expenditure per capita represents total annual household expenditure per person (in adult equivalent), after adjusting for general price trends and geographical differences in costs of living.*

The second outcome variable, the household dietary diversity score (HDDS), reflects the economic ability of a household to access diversified foods. Studies have shown that an increase in dietary diversity is a reasonable indicator of household food security and energy availability (FAO 2013; Hoddinott and Yohannes 2002). The household survey collected information on the type and frequency of food items consumed by members of the household. Following FAO (2013) guidelines, these food items are grouped into 12 food groups[26]. An average household in the

[25] The LSMS-ISA data provides the spatial price index computed by Ministry of Finance and Economic Development (MoFED) together with the consumption data. This index captures the difference in the cost of a representative food basket across the administrative regions.
[26] The food groups are Cereals, White tubers and roots, Vegetables, Fruits, Meat, Eggs, Fish and other seafood, Legumes & nuts, Milk and milk products, Oils & fats, Sweets, and Spices &condiments. For the sake of tractability

sample consumes about 4 food groups (see Figure 2.4). There is very little variation in this score both over time and across the survey rounds (see Table A2.1 in the appendix). This result is in line with other studies in the same context (Headey, Hoddinott, and Park 2017). To facilitate interpretation, the diet diversity score is expressed in terms of the proportion of total possible food groups consumed by dividing the number of food security gaps by 12.

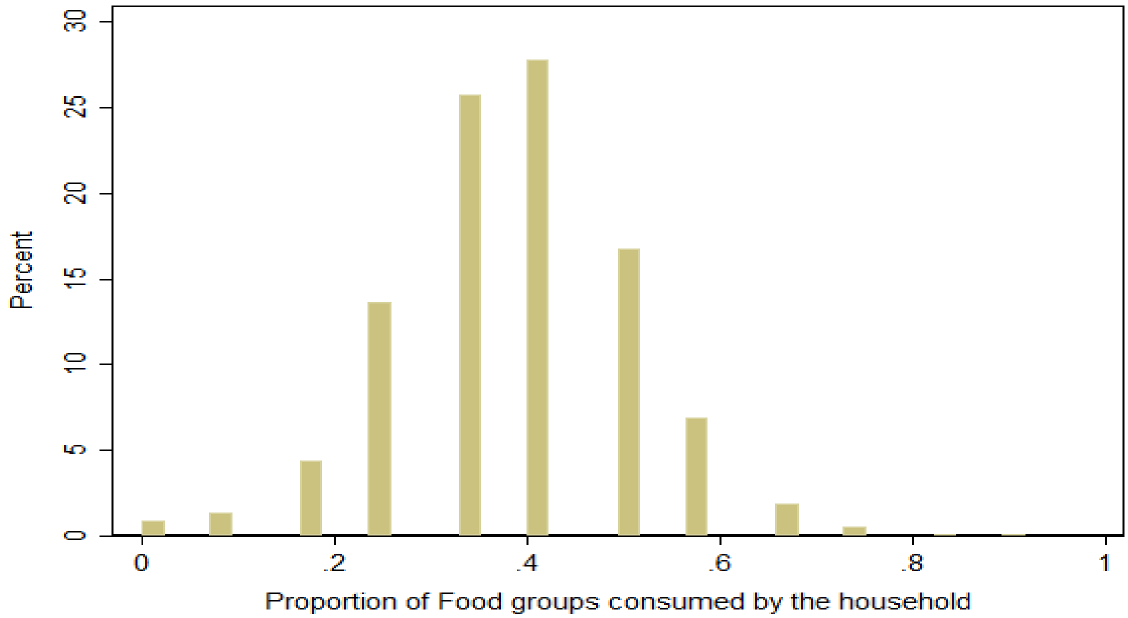


Figure 2.4. Distribution of Household Diet Diversity Index (HDDI)
Source: *Author's computation based on LSMS-ISA (2014 & 2016)*
Note: *HDDI represents the proportion (out of a possible 12) of total food groups consumed by the household.*

The third outcome variable is the household food security gap score. This is measured based on the number of months the households faced food shortages over 12 months prior to the survey. In the food security module of the survey, the respondents were first asked whether they had experienced a food security issue in the previous 12 months, such as a concern that their household would not have enough food. If the response was affirmative, then the number of months in which food shortage occurred was inquired. This information is used as an indicator of the food security level of the household. Depending on the severity of the food security situation, this score could range between 0 and 12. Figure 2.5 shows how the food insecurity score is distributed within the sample. About 73 percent of the households reported being fully food secure. The remaining 27 percent of the sample reported some food insecurity of varying degrees. To facilitate interpretation, the food insecurity score is expressed in terms of the proportion of months by dividing the number of food security gaps by 12.

and for econometric estimation, the resulting total number of food groups consumed by the household is divided by the total number of food groups.

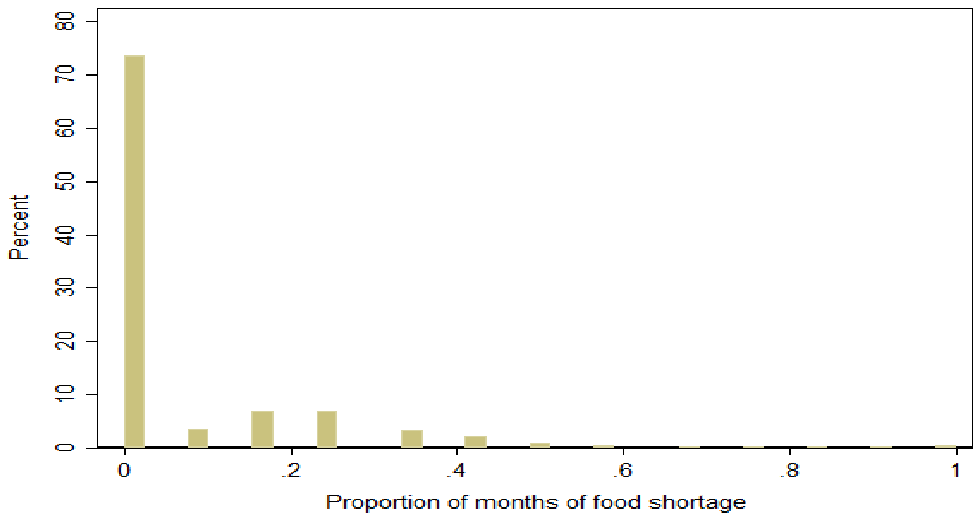


Figure 2.5. Distribution of household food security gap
Source: *Author's computation based on LSMS-ISA (2014 & 2016)*
Note: *food security gap represents the proportion of the number of months per year, over which the household reported food shortage.*

2.2.3. Descriptive results

In order to show the pattern in household welfare at various stages of urbanization, both nonparametric and parametric regression techniques are employed. Non-parametric regressions are used to characterize the relationship between SOL - a proxy of urbanization - and the alternative measures of household welfare. The method is useful to set aside the assumption of a pre-defined parametric relationship and assess the true evolution of the outcome variables over the different stages of urbanization. To this end, the following polynomial regression model is estimated:

$$y_i = \alpha + \beta_i \sum_{1}^{n} z_i{}^n + \varepsilon_i \qquad (2.1)$$

Where y_i represents the outcome variables (consumption per capita, diet diversity score, and food security gap) measured at the household level. The SOL is given by z_i and it is introduced in levels and higher degree polynomial forms to capture possible non-linear patterns. Household-level socioeconomic factors (e.g. demography, wealth, etc.), as well as access to electricity which might be picked up by the SOL, is accounted for by ε_i, the error term. Figure 2.6, Figure 2.7, and Figure 2.8 respectively present nonparametric local polynomial regressions of consumption per capita, household diet diversity score, and household food security gap score on SOL.

These figures suggest that urbanization is strongly and positively associated with both consumption per capita and diet diversity score, while it is negatively associated with the food security gap. However, the relationship between urbanization and household welfare is not linear. On closer examination of the figures, there appears a systematic pattern in welfare across the different stages of urbanization. The welfare measures increase slowly at first and then increase at an increasing rate before they level off at advanced stages of urbanization (s-shape pattern).

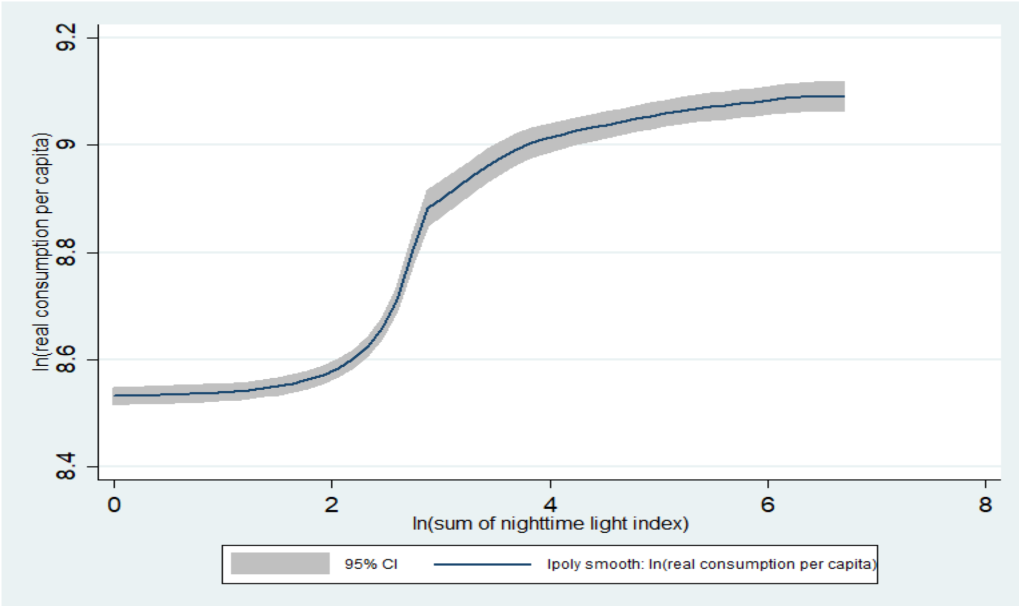


Figure 2.6. Patterns in real consumption expenditure by urbanization status
Source: Author's computation based on LSMS-ISA (2014 & 2016)
***Note**: Real consumption expenditure per capita represents total household expenditure per person (in adult equivalent), after adjusting for general price trend and geographical differences in costs of living*

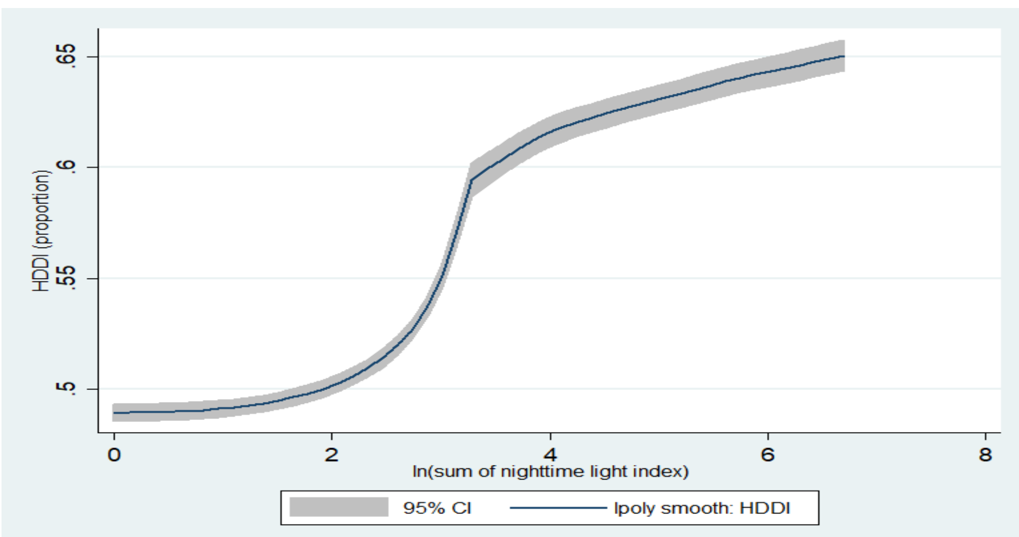


Figure 2.7. Patterns in Household Diet Diversity Index (HDDI) by urbanization status
Source: Author's computation based on LSMS-ISA (2014 & 2016)
Note: HDDI represents the proportion of total food groups consumed by the household.

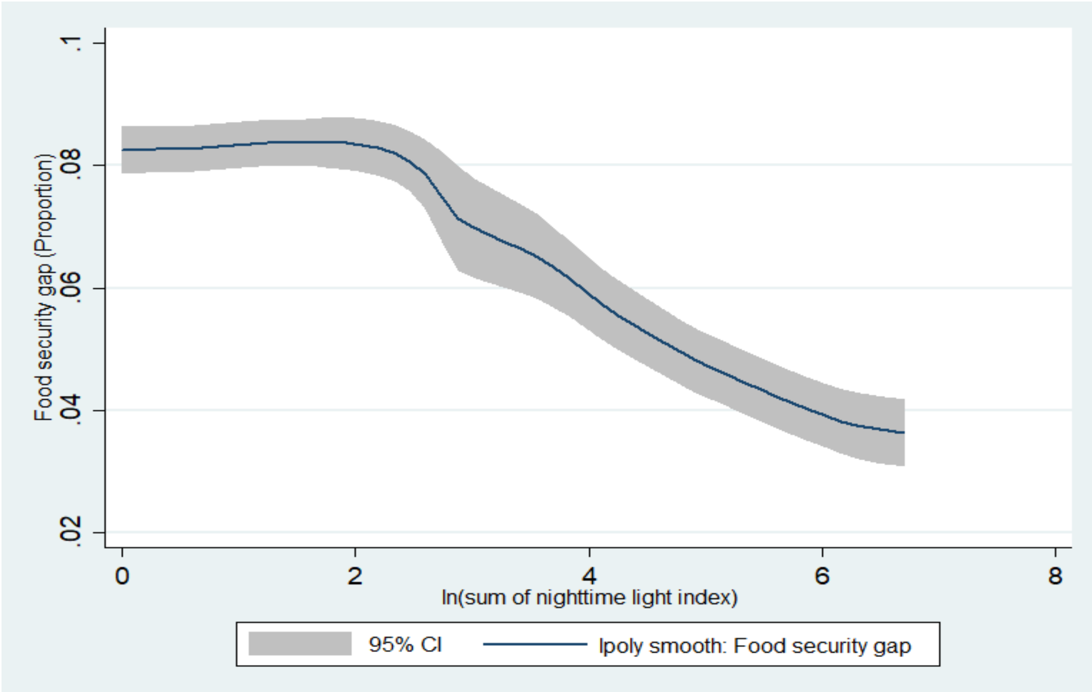


Figure 2.8. Patterns in household Food security gap score by urbanization status

Source: Author's computation based on LSMS-ISA (2014 & 2016)

Note: food security gap represents the proportion of the number of months per year, over which the household reported food shortage.

To understand the welfare dynamics over different stages of urbanization further, the sample households are split into four clusters based on the value of SOL. This is done using the Hansen (2000) threshold estimation technique that extends linear regression to allow coefficients to differ across well-defined clusters based on model fit and information criteria (BIC, AIC, or HQIC). While allowing for disaggregated analysis of urbanization, the approach helps to avoid the use of arbitrary cut-offs. For Ethiopia, during the study period, the logarithm of SOL ranges between 0 and 6.7 (Figure 2.2). The threshold method generated three cut-off points (0.97, 3.46, and 5.82) grouping the sample households into 4 clusters.

Figure 2.9 presents the result of the regression of the outcome variables on SOL over these 4 clusters. In general, it shows that the strength of the relationship between urbanization (as proxied by the SOL) with the outcome variables varies considerably over the different clusters. While the slope is positive only in the third cluster for consumption expenditure and diet diversity score, the slope corresponding to the food security gap is negative only in the fourth cluster. This suggests that the remoteness penalty – the negative effect of isolation on welfare – might depend on the current level of urbanization of the places of residence and hence, policies might need to be tailored to places.

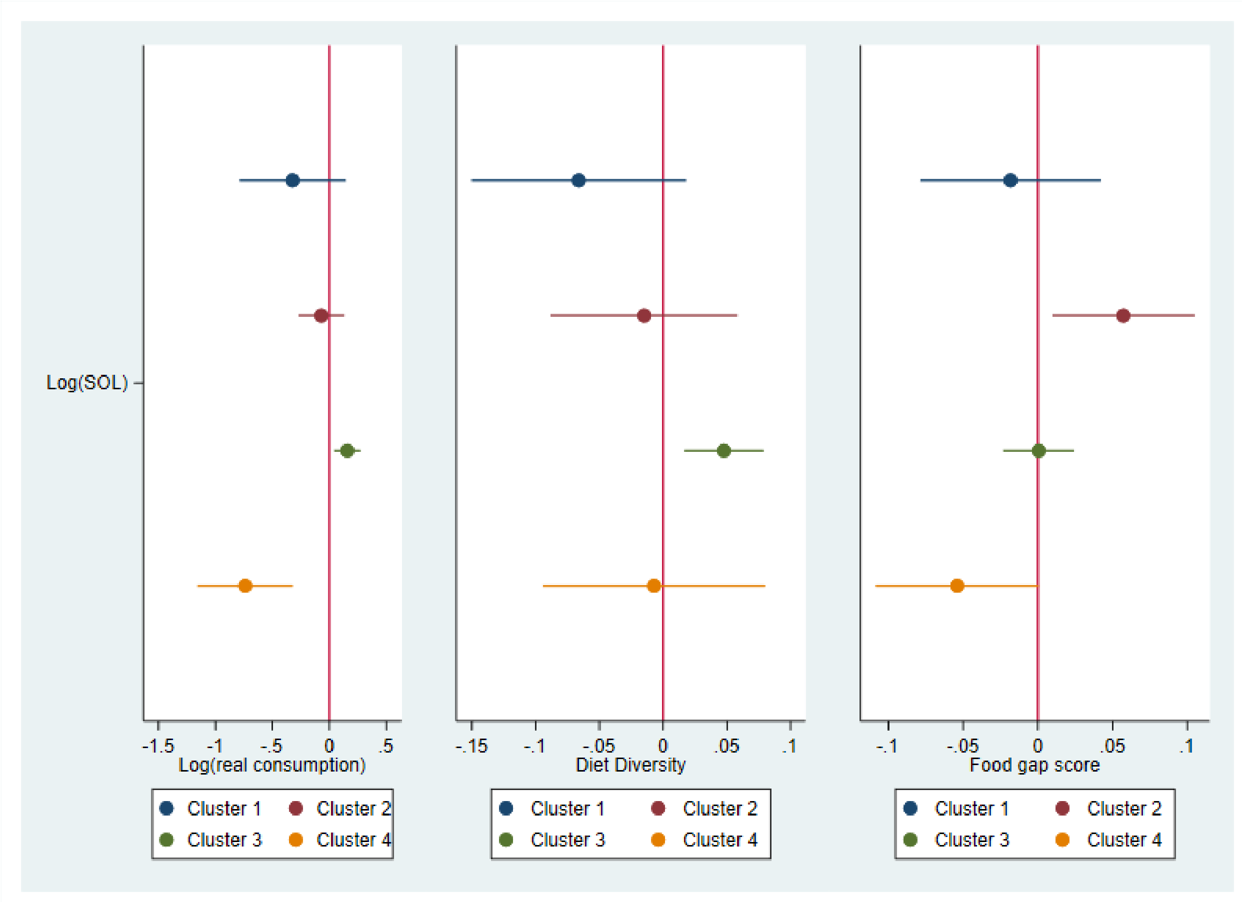


Figure 2.9: Association between urbanization and welfare, threshold estimation

***Source**: Author's calculation based on LSMS-ISA (2014 & 2016)*

***Notes**: "Dots: coefficient from ordinary least square regressions. Bars: 95% confidence intervals". The indicated clusters are generated from the SOL following Hansen (2000).*

Following the above threshold analysis, the sample households are roughly clustered into rural areas (cluster 1), small towns (cluster 2), intermediate towns (cluster 3), and large towns (cluster 4)[27]. Table 2.1 presents descriptive statistics based on these 4 clusters. It suggests that urbanization is generally associated favorably with household welfare. While average consumption and HDDI are higher in urban areas than in rural areas, they improve consistently with the size of the urban areas. For instance, HDDI is 0.49, 0.53, 0.63, and 0.66 in rural areas, small towns, intermediate towns, and large towns, respectively. This pattern is similarly evident for consumption per capita. The food security gap is slightly higher in small towns than in rural areas.

[27] The basic finding endures several sensitivity tests of this classification. Instead of SOL, several alternative measures including population size, population density and administrative classification are also used. See the discussion in section 2.5.

Table 2.1. Descriptive statistics of outcome and covariates by urbanization status

Variables\Urban-rural category [a)]	[A] Rural	[B] Small towns	[C] Intermediate towns	[D] Large town	Mean difference (p-val.) [A] vs [B]	[B] vs [C]	[C] vs [D]
ln(Sum of Nighttime light)	0.02	2.50	4.77	6.35	0.00	0.00	0.00
Outcome variables							
ln (real consumption per capita)	8.53	8.79	9.06	9.08	0.00	0.00	0.43
HDDI (proportion)	0.49	0.53	0.63	0.66	0.00	0.00	0.00
Food security gap (proportion)	0.08	0.11	0.05	0.03	0.00	0.00	0.00
Household characteristics							
ln(Household size)	1.71	1.64	1.50	1.49	0.00	0.00	0.55
ln(Age of household head in years)	3.82	3.80	3.69	3.76	0.23	0.00	0.00
Head is male, yes=1	0.73	0.72	0.64	0.53	0.61	0.00	0.00
Head education, primary=1	0.28	0.30	0.33	0.32	0.26	0.15	0.39
Head education, secondary or higher=1	0.07	0.20	0.37	0.52	0.00	0.00	0.00
Drought shock, yes=1	0.20	0.15	0.07	0.00	0.02	0.00	0.00
Non-drought shock, yes=1	0.44	0.54	0.45	0.44	0.00	0.00	0.64
ln(Land size household owned, ha)	0.63	0.34	0.11	0.00	0.00	0.00	0.00
ln(Livestock owned, in TLU)	1.31	0.86	0.30	0.01	0.00	0.00	0.00
Village characteristics							
ln(Elevation in meters)	7.46	7.47	7.51	7.57	0.50	0.03	0.00
ln(annual rainfall in mm)	7.06	6.98	6.98	6.94	0.00	0.95	0.02
ln(annual rainfall squared)	50.1	49.0	49.0	48.3	0.00	0.99	0.01
ln (mean temperature,^{0}C.)	5.26	5.23	5.24	5.23	0.00	0.14	0.02
Proportion of fertile soil in EA	0.64	0.71	0.65	0.62	0.00	0.01	0.23
EA has electricity, yes=1	0.37	0.68	0.97	1.00	0.00	0.00	0.00

***Source**: Author's calculation based on LSMS-ISA (2014 & 2016)*

***Notes:** The sum of Nighttime light (SOL) represents the sum of NTL intensity around EAs. a) Rural areas, small towns, intermediate towns, and large towns in this table roughly correspond to clusters 1, 2, 3, and 4 in Figure 2.8 respectively.*

It is also worth noting that the patterns in the averages of the welfare indicators in Table 2.1 appear to corroborate the s-shaped curve welfare pattern shown in Figure 2.6, Figure 2.7, and Figure 2.8. Differences between small and intermediate towns are much larger than that between rural areas and small towns as well as between intermediate towns and large towns. That is, while small towns resemble rural areas, intermediate and large towns appear quite similar in terms of average welfare levels, and there is a huge difference between small and intermediate towns.

2.3. Identification strategy, results and discussion

2.3.1. Identification strategy

The estimation based on the nonparametric regression technique and the subsequent threshold analysis implies nonlinear relationships between welfare and urbanization. This section extends this analysis by accounting for household and locational factors. The starting point is to estimate a simple parametric regression of the form specified in equation (2.2). Specifically, the welfare of household *i* in community *j* at time *t* is given by:

$$y_{ijt} = \alpha_j + \beta_1 U_{jt} + \beta_2 H_{ijt} + \beta_3 V_{jt} + \varepsilon_{it} \qquad (2.2)$$

Where y_{ijt} represents the outcome variables (consumption per capita, HDDI, and food security gap) and α_j represents village fixed effects. H_{ijt} and V_{jt}, respectively, represent household and village characteristics. U_{jt} represents the measure of urbanization, the SOL composite index.

Next, as discussed before, the threshold method is used to split the U_{jt} in equation (2.2) into four clusters: rural areas, small towns, intermediate towns, and large towns[28]. With these urbanization categories, equation (2.2) now takes the following form[29]:

$$y_{ijt} = \alpha_i + \beta_1 S_{ji} + \beta_2 M_{ji} + \beta_3 L_{ji} + \beta_4 H_{ijt} + \beta_5 V_{jt} + \varepsilon_{it} \qquad (2.3)$$

With rural areas as a reference, *S*, *M*, and *L* respectively represent small, intermediate, and large towns; everything else is as in (2.2). To quantify the implication of urbanization based on the estimation of the general form of (2.2) or (2.3) involves two main estimation issues. First, there might be selection bias because of systematic differences among households residing at different levels of urbanization. Place of residence, though costly to change over a short time period, is a choice variable. Households determine their location of residence based on their endowments and the available opportunities. Failure to account for this selection bias would obscure the true impact of urbanization on welfare. To partially address this challenge, the sample is restricted to those households remaining in the same village over the 2014-2016 period. This also helps to account for endogenous dynamic migration decisions. It is also important to note that in Ethiopia, the cost of migration is prohibitively high especially for rural households because of absent land markets (Deininger et al. 2003). Ownership of land belongs to the state. Individual farmers have only user rights, and any secured and continuous land use rights are contingent on permanent physical residence in the community. Therefore, it is prohibitively costly for households who seek to enhance their welfare to do so by relocating their place of residence in the short run, reducing the likelihood of selection bias.

The second estimation issue is the potential problem of endogeneity arising from omitted attributes and measurement problems. This is plausible given that most urbanization programs are accompanied by economic growth that can influence the overall livelihood of societies. That is, the SOL variable in (2.2) or the dummy variables representing different sized urban areas in (2.3) may well be correlated with the error term (ε_{it}) owing to omitted variables bias. If this is the case,

[28] The basic finding in this section endures several sensitivity tests of this classification.
[29] This is in line with similar empirical exercises in the literature (Amare et al. 2017; Ameye 2018).

simple OLS estimates of the βs would be biased. To minimize this problem, the panel structure of the data is used to estimate the EA fixed effects model. The EA fixed effect can capture time-invariant differences in welfare across different villages, implying that the parameters associated with SOL estimate the effect of urbanization on household welfare. Moreover, several socio-demographic characteristics of the household that might be associated with urbanization such as education and wealth levels of the households, are controlled for, to address many sources of concern regarding omitted variables bias.

In all regressions, a year dummy is included to account for aggregate shifts in welfare or correlated shifts in the right-hand side variables. Since surveyed households are sampled from stratified village level samples and households from the same village might share common unobservable characteristics, standard errors in all regressions are clustered at the village level.

2.3.2. Results and discussion

Urbanization and welfare

Panel A in Table 2.2 presents the pooled OLS regression result of household welfare. Two regression models are estimated for each of the outcome variables. First, a simple unconditional regression of an outcome variable (e.g. per capita expenditure) is estimated on SOL and survey period dummy. In the subsequent regression, the model is extended by accounting for household and village characteristics as well as the zones of residence. The estimation result shows that urbanization is strongly and positively associated with household welfare, measured in terms of per capita expenditure and diet diversity score. It also shows that urbanization is associated negatively with the food security gap. Specifically, a doubling of the SOL is associated with a 5 percent increase in consumption per capita, a 1.6 percent increase in HDDI, and a 0.4 percent reduction in the food security gap.

Panel B of Table 2.2 presents the estimation result of urbanization on the same outcome variables based on the EA fixed effect. Since these estimators are immune to time-invariant village-level heterogeneities, they identify the causal welfare effect of urbanization (Wooldridge 2002). Although the sizes of the magnitudes are notably larger than the coefficients from the pooled regression, the fixed effect estimators report a qualitatively similar result for consumption and diet diversity. The coefficient of food security score is, however, not statistically significant.

Table 2.2. Association between urbanization and household welfare

	ln(Expenditure)		Diet Diversity score		Food security Gap	
Panel A: Pooled OLS						
ln(Sum of Nighttime light)	0.096***	0.047***	0.028***	0.016***	-0.007***	-0.004**
	(0.007)	(0.008)	(0.002)	(0.002)	(0.001)	(0.002)
HH & village characteristics [a)]	No	Yes	No	Yes	No	Yes
Zone Fixed Effect	No	Yes	No	Yes	No	Yes
Number of observations	9,215	9,210	9,606	9,600	9,606	9,600
R2	0.123	0.375	0.151	0.339	0.013	0.216
Adjusted R2	0.122	0.368	0.151	0.332	0.013	0.207
Panel B: EA fixed effect						
ln(Sum of Nighttime light)	0.151***	0.112***	0.042***	0.031***	0.002	0.005
	(0.031)	(0.025)	(0.007)	(0.006)	(0.016)	(0.012)
HH & village characteristics [a)]	No	Yes	No	Yes	No	Yes
EA Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	9,215	9,210	9,606	9,600	9,606	9,600
R2	0.004	0.156	0.007	0.103	0.000	0.103
Adjusted R2	0.004	0.154	0.007	0.102	-0.000	0.102

***Source**: Author's calculation based on LSMS-ISA (2014 & 2016)*

***Notes**: Village clustered standard error in parentheses: *** p<0.01, ** p<0.05, * p<0.1. The Sum of Nighttime light (SOL) represents the sum of NTL intensity around EAs. [a)] Coefficients omitted to preserve space. For estimation results of the full model, see Tables A2.3 and A2.4 in the appendix.*

Heterogeneity in welfare across city hierarchy

This sub-section employs a parametric regression method to estimate the heterogeneity in the effect of urbanization on household welfare. As discussed before, this proceeds by splitting the sample households into clusters of urbanization based on Hansen's (2000) threshold method. This is tantamount to estimating equation (2.3) with dummy variables representing different levels of urbanization (and a rural household as a reference). The results presented in Figure 2.10 show the coefficient estimates from this model after accounting for household and location characteristics. It shows that all else the same, on average, households in intermediate and large towns consume more per capita than those in rural areas. They also fare better in terms of diet diversity score and food security. Specifically, compared to an average household in rural areas, the per capita consumption is 20 percent higher; diet diversity is higher by 10 percent (approximately by one food group); the food security gap is lower by two weeks for a household in an intermediate- or large- town. The welfare level of a household in a small town is largely comparable to a household in rural areas except for consumption expenditure which is about 10 percent higher in small towns than in rural areas.

In line with the s-shaped welfare pattern from the non-parametric regression, Figure 2.10 also shows that differences between small and intermediate towns are much larger than that between rural areas and small towns as well as between intermediate towns and large towns. That is, while small towns resemble rural areas, intermediate and large towns appear quite similar in terms of average welfare levels, and there is a huge difference between small and intermediate towns. Furthermore, in almost all regressions, the magnitude of the effect is larger for the intermediate towns than for large towns. Moreover, the standard errors corresponding to parameter estimates of large urban areas are larger than those of the intermediate towns, suggesting that inequality in welfare is relatively more prevalent in large urban areas. Table A2.8 in the appendix corroborates this and shows that regardless of the type of measure adopted, welfare inequality is substantially higher in large urban areas. Therefore, while intermediate- and large- towns are comparable in terms of average welfare outcomes, intermediate towns appear to be more inclusive. This result aligns with recent and increasing evidence of intermediate towns having a greater impact on employment generation and overall poverty reduction in developing countries (Christiaensen, De Weerdt, and Todo 2013; Dorosh and Thurlow 2014; Kanbur et al. 2019).

From the comparison of the conditional and unconditional regression coefficients of the full results presented in Table A2.5 and Table A2.6 in the Appendix, it appears that wealth status and human capital endowment of the household head are important drivers of welfare differences. That is, the heterogeneity of the link between welfare and urbanization over different stages is mediated by the spatial distribution of human capital difference, in line with the human capital theory (HCT). However, even after differences in wealth, human capital, and institutional differences across locations are factored in, the spatial disparities in household welfare are considerably minimized but not eliminated. The next section highlights the main underlying factors for this spatial pattern.

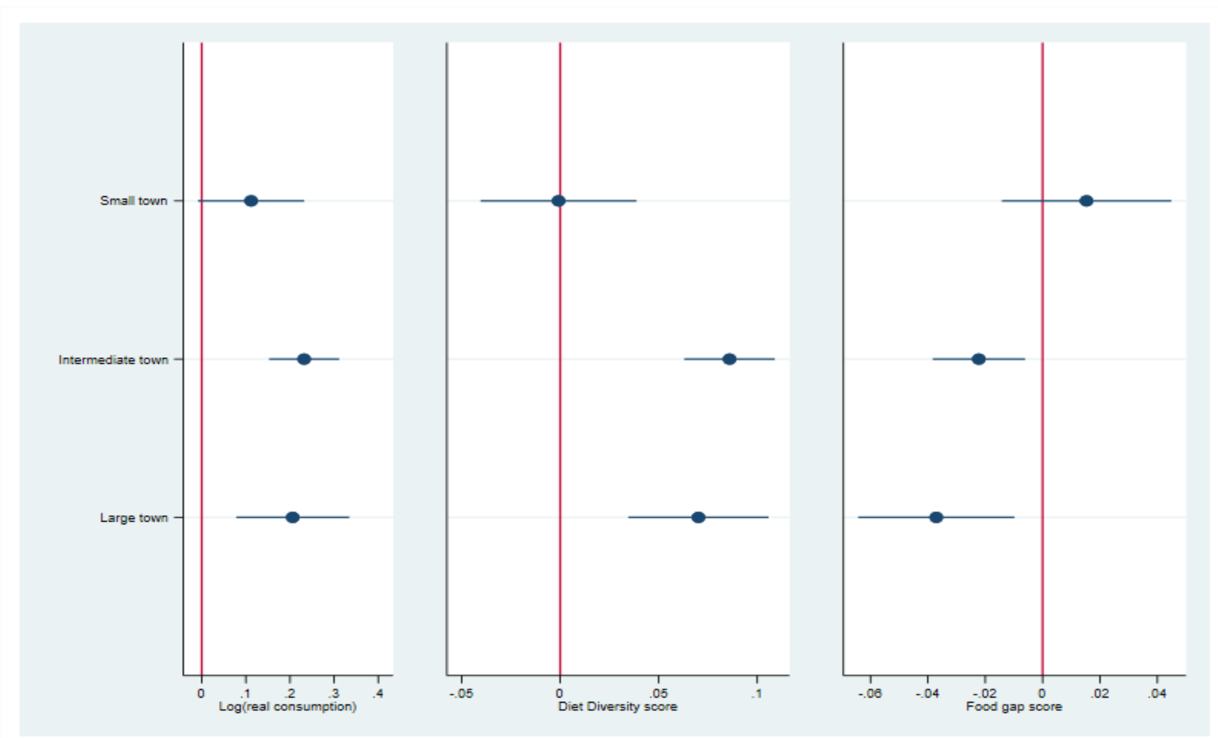


Figure 2.10: Association between stages of urbanization and household welfare
Source*: Author's calculation based on LSMS-ISA (2014 & 2016)*
Notes*: For all regression, standard errors are clustered at the village level. Rural areas, small towns, intermediate towns, and large towns in this figure were generated from the sum of NTL intensity around EA using the Hansen (2000) threshold method, respectively. Other control variables are omitted to preserve space. For estimation results of the full model, see Tables A2.5 and A2.6 in the appendix.*

2.4. Mechanisms

The previous section highlighted the difference in welfare dynamics across the different stages of urbanization. There are several potential mechanisms that explain why household welfare systematically changes over space from rural areas to large urban areas as reflected in this finding. First of all, there is considerable spatial sorting of households by wealth and human capital endowment. Table 2.1, for example, shows that while about 52 percent of household heads in large towns attained secondary or tertiary education level, only seven percent of household heads in rural areas and 20 percent in small towns attained a similar level of education. Since education is a good predictor of wealth, these differences in educational attainments could partially explain the spatial pattern in household welfare.

Table 2.3 presents more direct measures of household wealth ranging from consumption expenditure to housing quality. It shows that both food and non-food real consumption expenditure, ownership of radio, TV, electricity, and mobile phones increases with urbanization. For example, the total expenditure in large towns (after adjusting for differences in the cost of living) is about twice the amount found in rural areas. This is consistent with the pattern observed in the asset-based wealth index. Directly, the differences in ownership of radio, TV, electricity, and mobile phones reflect differences in wealth. Indirectly, they represent households' differential access to information, which proved to be a key predictor of current and future welfare (Hirvonen et al. 2017). A similar pattern is observed in terms of size and quality of housing (roof, floor, wall, type of cooking fuel, and sanitation facility).

Table 2.3. Descriptive statistics of household wealth indicators by urbanization status

Wealth indicators	All Households	Rural	Small towns	Intermediate towns	Large town	F-test p-val.
Real expenditure, ETB	7,449	6,148	8,058	10,313	11,245	0.00
Real food expenditure, ETB	5,445	4,818	5,912	6,882	7,088	0.00
Real non-food expenditure, ETB	1,829	1,267	2,017	3,103	3,449	0.00
Durable assets owned, PCA	0.0	-1.16	-0.13	1.83	4.65	0.00
Ownership of TV, %	23.3	7.3	22.6	52.7	84.6	0.00
Ownership of Radio %	33.5	26.1	32.2	46.5	63.9	0.00
Access to electricity, %	39.3	18.0	49.5	89.2	98.6	0.00
Access to Mobile phone, %	55.9	42.3	64.9	86.1	95.3	0.00
Housing						
Number of rooms	1.87	1.79	1.93	2.05	2.28	0.00
HH has improved roof, %	63.7	50.3	70.6	92.5	98.5	0.00
HH has improved floor, %	20.2	5.6	16.8	43.0	77.3	0.00
HH has improved wall, %	8.5	1.7	2.7	18.4	38.5	0.00
HH uses Improved cooking fuel, %	6.7	0.8	3.1	10.1	41.8	0.00
Type of toilet used						
Improved toilet, %	30.2	23.7	22.9	41.8	64.7	0.00
Less protected toilet, %	40.5	38.1	46.5	47.3	34.2	0.00
No toilet facility, %	29.3	38.2	30.6	10.8	1.2	0.00
Observation	9,606	6,572	513	1,561	960	

Source*: Authors' computation based on LSMS (2014 & 2016).*

Notes*: Rural areas, small towns, intermediate towns, and large towns in this table were generated from the sum of NTL intensity around EA using the Hansen (2000) threshold method.*

Other complementary channels that might underlie the spatial pattern in household welfare include differences in employment opportunities, access to public services, and market access. This section discusses these channels in more detail.

2.4.1. Employment

A major source of consumption risk, particularly for rural households, is the lack of diversified income. In most rural areas of developing countries, population growth is high and there is very limited access to irrigation (McCullough 2017). This implies that rural households' income is seasonal and susceptible to climatic shocks. Urbanization provides an opportunity for these farming households to diversify income sources as it links with higher off-farm employment (Haggblade et al. 2010; Reardon et al. 2006). Table 2.4 shows that between 2014 and 2016, only 6.3 percent of households were wage-employed in rural areas. With urbanization, this increases to reach 47 percent for households in large urban areas. Similarly, ownership of and employment in non-farm businesses also increases with urbanization. During the survey period, the share of households that engaged in non-farm business increased from 18.2 percent in rural areas to 23.9 percent in large urban areas.

Table 2.4. Patterns in type and intensity of employment by urbanization status

	All Households	Rural	Small towns	Intermediate towns	Large town	F-test p-val.
Household Participation in labor market						
Agricultural activities over the last 7 days, %	43.63	56.6	35.5	17.4	1.9	0.00
Non-farm business over the last 7 days, %	21.81	18.2	30.4	32.9	23.9	0.00
Wage employment over the last 7 days, %	15.25	6.3	18.5	32.3	47.0	0.00
Casual employment over the last 7 days, %	7.96	7.7	5.8	10.4	6.8	0.00
Unpaid activities over the last 7 days, %	0.53	0.4	0.0	1.0	0.8	0.00
Off-farm employment over the last 7 days, %	38.79	28.2	48.1	63.2	66.6	0.00
Multiple employment activities over the last 7 days, %	12.63	13.9	13.6	12.0	4.6	0.00
Wage employment over the last 12 months, %	21.21	10.0	26.7	42.2	61.0	0.00
Casual labour work in the last 12 months, %	19.59	22.7	14.8	14.9	8.3	0.00
Non-farm business in the last 12 months, %	35.82	33.1	48.3	44.7	33.0	0.00
Off-farm sector in the last 12 months, %	61.21	52.5	69.2	80.3	85.4	0.00
Households' number of working hours						
Agricultural activities, per week	33.88	44.9	25.9	10.6	0.8	0.00
Non-agricultural activities, per week	10.60	7.3	19.2	18.6	15.9	0.00
Casual activities, per week	2.62	2.4	1.9	3.7	2.5	0.01
Wage, salary, activities, per week	8.91	2.9	10.9	19.1	32.7	0.00
Unpaid activities, per week	0.23	0.2	0.1	0.4	0.4	0.00
Total hours worked, per week	56.24	57.6	57.9	52.3	52.4	0.00
Total hours worked, per week per capita	15.77	15.0	16.1	17.9	17.5	0.00
Primary employment, 12 months	479.4	159.8	602.9	985.3	1,779	0.00
Secondary employment, 12 months	3.72	2.8	3.2	6.9	5.3	0.18
Hours HH spent to fetch water, per day	0.69	0.9	0.7	0.3	0.2	0.00
Hours HH spent to collect firewood, per day	0.78	1.0	0.7	0.3	0.1	0.00
Observation	9,606	6,572	513	1,561	960	

Source*: Author's computation based on LSMS (2014 & 2016)*

Notes*: Rural areas, small towns, intermediate towns, and large towns in this table were generated from the sum of NTL intensity around EA using the Hansen (2000) threshold method.*

Table 2.4 also indicates that the effect of urbanization is not limited to improving the labor supply at the extensive margin — creating employment opportunities. It also improves employment at the intensive margin (i.e. improving the intensity of employment— the number of working hours). Table 2.4 shows that the total number of working hours per capita increases with urbanization. During the survey period, an average adult person works about 3 more hours per week in intermediate and large urban areas than in rural areas. These differences are statistically significant[30].

Why do rural households work fewer hours per capita on average? Two important factors might explain this. First, in the rural areas, underemployment of labor is huge due to the seasonality of farming and agricultural employment, coupled with a lack of private and public job-generating investments. The data from the 2013 labor force survey indicates private and public investments are the major sources of job opportunities in large urban areas. In rural areas, however, these investments barely exist (see Table A2.7 in the Appendix). The second reason might relate to the *less productive time use* in rural areas due to the shortage of public services. The last two rows in Table 2.4 present the number of hours per day that households spend on the collection of drinking water and firewood over geographic space. On average, a rural household spends more than 3 times more time on these activities than an average household in large urban areas. This unproductive time use might reduce the available labor supply for income-generating activities. Therefore, policy objectives that target improvements in the welfare of rural populations, might be achieved by channeling public and private investment to expand households' access to public services and employment opportunities.

Finally, it is also important to note that open unemployment is more prevalent in megacities than in rural and small-sized towns. Table A2.7 in the Appendix shows that 15.4 percent of residents in large urban areas were unemployed in 2013. This is considerably larger than the corresponding figures in rural and intermediate urban areas. The Table also shows that relative to both rural areas and small urban areas, the percentage of the inactive population[31] (28%) and average duration of unemployment (42 months) in large urban areas is notably higher. Labor market conditions, such as these, indicate that large urban areas are riskier. Moreover, given that the labor market is the sole source of livelihood in large urban areas, unemployment tends to correspond with low consumption, food insecurity, and poverty. This calls for targeted intervention in urban areas to enhance the employability of the poor and to create job opportunities tailored to attract the disadvantaged segments of the urban population.

2.4.2. Access to public services

Access to public services such as roads, schools, and health centers are shown to be important determinants of household welfare (Hirvonen et al. 2017; Stifel and Minten 2017; World Bank 2020). A large body of literature also emphasizes the role of credit constraints as a major impediment to labor productivity in Africa (Gine and Klonner 2005; Moser and Barrett 2006).

[30] Note that this difference is for an average adult. If the data is restricted to employed individuals, the difference increases to 16 hours (see Table A2.7 in the appendix).

[31] Inactive population refers to persons that are outside of the labor force (not working, and not looking for work), and includes pre-school children, students, pensioners and housewives or –men.

Therefore, differential access to these factors across the rural-urban spectrum might partly explain the observed spatial pattern in household welfare.

Table 2.5 shows that the distance of households to public services varies systematically and significantly over space, to the disadvantage of rural and small-town households. These differences in access to services have a direct bearing on welfare outcomes, with roads and financial institutions being particularly significant. For instance, while a rural household has to travel 16 kilometers (possibly on an unpaved road) to access a microfinance institution, a household in a large town only needs to drive 0.6 kilometers.[32]

Table 2.5. Patterns in access to public services by urbanization status

Public Service	All Households	Rural	Small towns	Intermediate towns	Large town	F-test p-val.
Household Distance to nearest--						
Major Road, Km	12.6	17.0	9.1	2.0	1.0	0.00
Tar/asphalt road, Km	28.3	39.4	17.1	2.6	0.5	0.00
Daily market, Km	55.0	69.4	45.3	28.8	4.3	0.00
Large weekly market, Km	7.0	9.0	8.3	1.4	1.1	0.00
Primary school, Km	1.7	1.2	0.3	0.4	7.8	0.00
Secondary school, Km	10.5	12.7	4.2	2.0	12.2	0.00
Health post, Km	1.8	1.2	1.9	2.9	3.6	0.00
Hospital, Km	11.3	13.3	7.1	10.9	0.4	0.00
Commercial bank, Km	23.0	24.6	6.0	1.8	0.5	0.00
SACCO, Km	9.0	12.2	4.5	1.8	1.5	0.00
Microfinance institution, Km	12.2	16.6	8.1	1.9	0.6	0.00
Distance to a drinking water source						
<15 min, %	57.1	44.0	57.5	83.4	94.7	0.00
15-30 min, %	25.7	33.5	24.6	10.2	3.4	0.00
30-60 min, %	12.5	16.3	12.1	4.6	1.3	0.00
>60 min, %	4.8	6.2	5.8	1.8	0.6	0.00
Source of drinking water						
Piped water, %	43.7	25.5	56.6	79.9	93.5	0.00
Protected spring/hole, %	26.1	32.9	22.1	14.3	4.0	0.00
Unprotected spring/hole, %	30.2	41.5	21.3	5.8	2.5	0.00
Source of house lightening						
Electricity, %	41.1	18.6	49.8	89.2	98.6	0.00
Improved non-electricity, %	28.4	39.9	14.9	4.2	0.9	0.00
Traditional, %	30.5	41.5	35.3	6.5	0.4	0.00
Observation	9,606	6,572	513	1,561	960	

***Source**: Author's computation based on LSMS (2014 & 2016)*

***Notes**: Rural areas, small towns, Intermediate towns, and large towns in this table were generated from the sum of NTL intensity around EA using the Hansen (2000) threshold method, respectively.*

Similarly, the difference in households' access to drinking water and electricity is rampant across space.

Table 2.5 shows that while piped water and electricity are almost universally accessible in large urban areas, only a quarter of households in rural areas have access to these services. The majority of the remaining households in rural areas rely on unprotected sources of drinking water

[32] This pattern is consistent across most of the selected services except primary schools and health posts. These two public services have been made available in every village by policy.

and traditional sources of house lighting. More than half of the total households also need to travel more than 15 minutes to sources of drinking water. These statistics are unfavorably comparable to large town households. Given that access to these services is a vital input in households' welfare function, these differences may partly explain the spatial difference in welfare outcomes. Hence, policy interventions that target to improve access to these services are likely to be effective to enhance the overall welfare status as well as reducing the spatial disparity[33].

2.4.3. Market access and food prices

Theoretically, the effect of urbanization on food prices is ambiguous. On the one hand, since income is higher in large towns, and food demand/supply is price inelastic, food prices might be higher in large towns compared to rural areas or small towns. On the other hand, since the increased access to the market that follows urbanization tends to attract more producers/suppliers, the competition among suppliers might push the prices down in large towns. The effect might also depend on the type of food item and the size of the market. For non-locally produced goods, the prices in rural areas might be higher than those in large urban areas because of the additional transportation cost as well as the thinness of the market. Table 2.6 seems to confirm the latter hypothesis. It shows that, for commonly imported, seasonal, or localized crops such as onions and potatoes, the prices in rural or small towns are higher than the prices in larger markets.

Table 2.6. Patterns in Food prices by urbanization status

Crop	All locations	Rural	Small towns	Intermediate towns	Large town	F-test p-val.
Teff	16.5	15.7	17.0	17.5	19.3	0.000
Wheat	10.3	9.7	10.6	11.2	12.0	0.000
Barley	9.6	9.1	10.2	9.7	11.7	0.000
Maize	5.8	5.5	6.2	6.1	7.3	0.000
Sorghum	7.2	6.4	8.2	8.0	9.7	0.000
Horse beans	18.7	17.9	19.8	20.3	20.6	0.000
Chickpea	18.4	18.0	19.4	18.8	18.7	0.000
Field pea	20.5	19.4	22.2	21.7	24.3	0.000
Lentils	42.0	40.7	44.1	44.7	42.8	0.000
Haricot beans	10.4	9.5	15.0	10.2	13.6	0.000
Milk	16.9	16.0	18.3	17.4	18.9	0.000
Eggs	2.8	2.6	3.1	3.1	3.4	0.000
Onion	9.9	10.3	9.7	9.3	8.2	0.000
Banana	12.3	11.1	14.2	14.2	15.2	0.000
Potato	8.5	8.6	9.4	8.5	7.6	0.000
Tomato	10.5	10.6	10.5	10.4	9.9	0.000
Orange	19.5	16.6	21.9	23.0	25.4	0.000

Source: *Author's computation based on LSMS (2015/16)*

Notes: *Rural areas, small towns, intermediate towns, and large towns in this table were generated from the sum of NTL intensity around EA using the Hansen (2000) threshold method, respectively. Prices are given in Ethiopian Birr (ETB) per kg.*

[33] A study by UNDP (2006) indicates that Sub-Saharan Africa might save a total $23.5 billion — 5 percent of GDP — if the entire population had access to basic, low-cost water and sanitation technology.

The higher food prices might contribute to lower household welfare in rural areas since, contrary to popular perception, a large share of households in rural areas rely on local food markets for their consumption (Worku et al. 2017).

Table 2.7 shows that while 85.9 percent of households in rural areas finance part or all of their food consumption from their own production, only about 40 percent of food consumption is actually sourced from their own production[34]. For the rest, they rely on the local market. This implies that the lack of sufficient access to a well-functioning market might partly explain the poorer living conditions in rural areas.

Table 2.7. Share of food consumption from own production, by urbanization status

	Rural	Small towns	Intermediate towns	Large town	All Households
Panel A: Share of food consumption financed through own production (%)					
Cereals	63.31	39.64	18.97	0.07	51.50
Pulses & oilseeds	35.34	15.60	6.89	0.18	26.55
Fruits & vegetables	14.24	13.13	2.10	0.01	10.25
Milk & milk products	74.91	79.45	16.97	1.66	62.72
Egg	42.87	46.42	15.90	0.61	29.12
Meat & Fish	13.89	3.89	1.48	-	8.10
Other foods	21.42	12.24	2.47	0.03	15.81
Total	40.26	27.61	8.44	0.15	30.69
Panel B: Households that finance part or all of consumption through own production (%)					
Cereals	71.5	39.2	18.3	0.1	58.3
Pulses & oilseeds	42	18.8	9.4	0.7	33
Fruits & vegetables	28.1	12.5	3.6	0.4	21.9
Milk & milk products	71	53.6	17.3	0.7	56.9
Egg	46.9	55.4	14.7	0.5	34
Meat & Fish	12.6	10.8	1.7	0	8.4
Other foods	40.3	25	7.4	0.4	32.5
Total	85.9	61.4	28.2	1.5	70.1

Source*: Author's computation based on LSMS (2014-2016)*

Notes*: Rural areas, small towns, Intermediate towns, and large towns in this table were generated from the sum of NTL intensity around EA using the Hansen (2000) threshold method, respectively.*

2.5. Sensitivity analysis

As discussed before, NTL as a marker of urbanization has proved to abate several shortcomings of the traditional survey and census-based definitions of urbanization. NTL is particularly appealing to capture micro-level variations in urban expansion as it allows the construction of continuous and disaggregated indices. However, its use to delineate and classify urban areas involves a number of issues. First, as discussed before, NTL might simply represent local trends in economic activity or electricity rather than divulge the existence and degree of urbanization (Henderson et al. 2009). Second, even though its value and use might increase in the future as countries prosper, the applicability of NTL is currently limited in developing countries as it often lacks sufficient variation across geographical space. Third, clustering the NTL index based on a statistical method, as is done in this study to generate the basic result, might not represent actual

[34] Food expenditure accounts about three quarter of the total household expenditure in Ethiopia in 2015 (LSMS-ISA data)

stages of urbanization rendering it less useful to inform policy. Forth, NTL is also not yet an official measure of urbanization. Most countries, including developed countries, use one or a combination of population size, population density, and access to infrastructure to delineate urban areas and classify the urban areas into different clusters based on size.

In Ethiopia, population size is mainly used to delineate urban areas. The Central Statistical Agency (CSA) of Ethiopia defines urban areas as localities with 2,000 or more inhabitants. Within urban areas, “1 million inhabitants” is the cut-off that separates large urban areas from other urban areas. Based on the 2015 LSMS-ISA survey, the total population of Ethiopia was about 97.3 million. Of these 78.4 percent of the population reside in rural areas. While large urban areas accommodate 23 percent of the total urban population, the large majority of the urban population (74%) is concentrated in small and medium towns. A comparison of the 2013 and 2015 surveys shows that the share of small and intermediate towns is increasing fast. On top of serving as a sensitivity test of the result in section 2.3, these statistics suggest why it is beneficial to assess the welfare dynamics based on this taxonomy[35].

For the result presented in Figure 2.11, the definition of the CSA and the World Bank is followed to define rural and small towns. However, distinctions are made between households residing in Addis Ababa from those in other large urban areas. Addis Ababa is a primate city with a population number of more than ten times that of the second-largest city. Figure 2.11 reports the parametric regression result of welfare on indicators of different levels of urbanization. Overall, urban areas appear to be more strongly welfare-enhancing compared to rural areas. However, among households in urban areas, household welfare is better in intermediate towns.

[35] Population density is another common indicator of urbanization and urban growth. Estimates based on this criteria also produces quantitatively similar results.

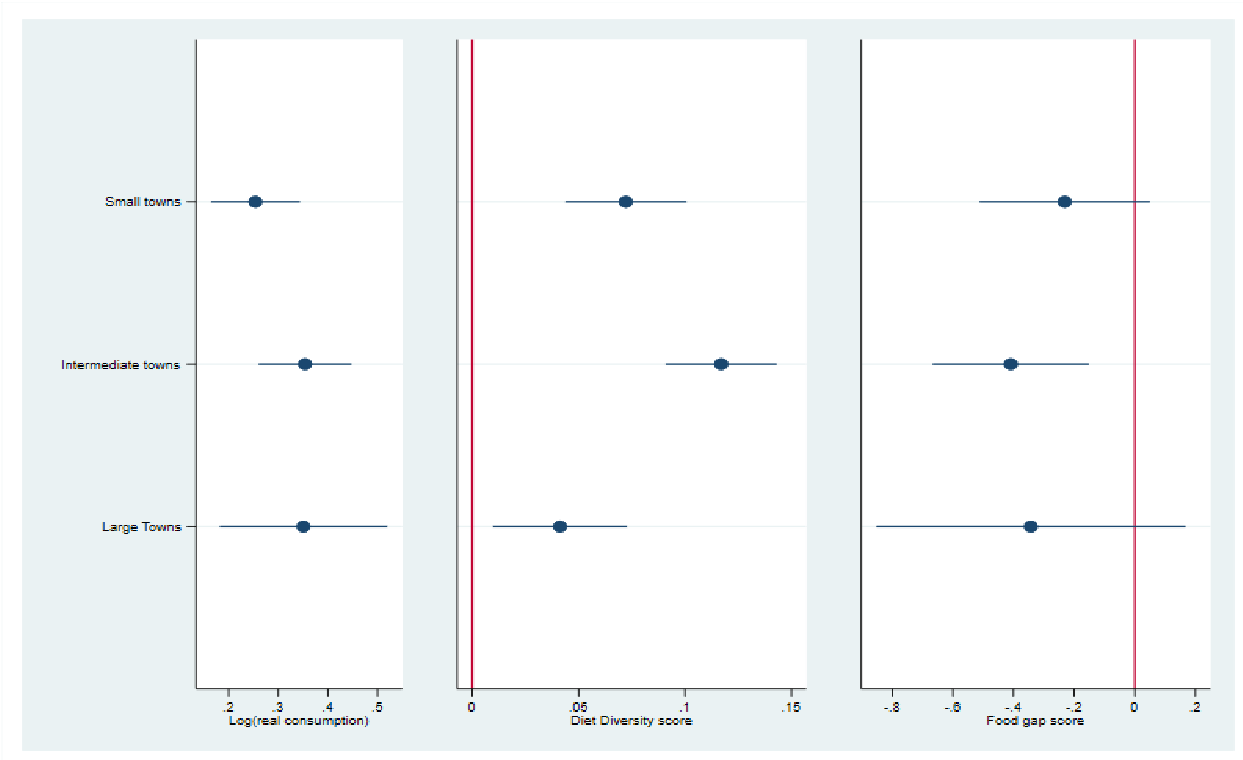


Figure 2.11: Association between stages of urbanization and household welfare

Source*: Author's computation based on LSMS (2014-2016)*

Note*: In all regressions, standard errors are clustered at the village level. The definition of the CSA and the World Bank is adopted to classify sample households into rural areas, small towns, Intermediate towns, and large towns in this table*

2.6. Concluding remarks

Sub-Saharan African countries are urbanizing at an unprecedentedly rapid rate. This has intensified interest in the effects of urbanization. Whether and which – small, intermediate, or large – urban areas lead to more inclusive growth and poverty reduction has remained open to debate. On the one hand, small and intermediate towns might have a stronger effect due to their proximity and linkage with rural areas (Christiaensen and Kanbur 2017; Christiaensen, De Weerdt, and Todo 2013; Dercon and Hoddinott 2005). On the other hand, the agglomeration effect implies that bigger towns might have a stronger effect (Redding 2010; Redding and Venables 2004; Vandercasteelen et al. 2018; Venables 2008; World Bank 2009).

This study contributes to this debate by investigating the relationship between the size of urban areas and household welfare. It focuses on identifying whether and how urbanization and the different stages of urbanization in Ethiopia are associated with household welfare and explores the major underlying mechanisms.

Based on the New Economic Geography (NEG) framework and threshold data analysis, the findings of this chapter suggest that the nature of urbanization is at least as important as the aggregate rate of urbanization. In general, the findings indicate that intermediate and large towns are more strongly associated with household welfare compared to small towns or rural areas. The role of market access, employment opportunities, and differential access to public services are highlighted as the major underlying mechanisms for the spatial disparity in welfare.

The findings have a number of important policy implications. First, it shows that urbanization is welfare-enhancing in Ethiopia since it is associated with improved human capital and asset endowment, diversified employment opportunities, and market access. This is consistent with the New Economic Geography (NEG) literature that describes that high population and economic concentration are particularly important to growth at the early stages when a country has a limited fiscal capacity to finance comprehensive economic infrastructure, and when domestic knowledge accumulation is low (Fujita et al. 2000). However, policies should proactively manage and support the ongoing urbanization process by improving infrastructure, housing, and urban institutions (Henderson 2003; World Bank 2009, 2013b).

Second, the result that points to a significant positive relationship between urbanization and household welfare masks considerable disparity within the intermediate and large urban areas. Table A2.8 in the Appendix reveals that regardless of the type of measures adopted, income inequality is substantially higher in large urban areas than in rural areas, small towns, or intermediate towns. The inequality across large urban areas manifests itself not only in terms of the aggregated welfare indicators but also in terms of the underlying mechanisms. Compared to wealthier households, the poorest households in large urban areas face much lower access to improved water supply, electricity, and sanitation facilities (Muzzini 2008; World Bank 2010, 2020). Therefore, policies should be tailored to the poor to make urban growth more inclusive and facilitate access to basic services. Policy interventions could also target better employment conditions for the poor. Expediting timely access to labor market information, and facilitating access to education, skill training, and small-scale credit facilities, have proved to be effective in this regard (Ali, Deininger, and Duponchel 2014; World Bank 2011). Urban safety net programs are also vital to address the vulnerability and food security issues among the urban poor.

Third, this study presents strong evidence of the effect of intermediate towns on welfare. Together with the result that inequality is much greater in large urban areas, this argues the case for a hierarchical pattern of urban development.

The existence of smaller and intermediate towns in the rural-urban economic space improves access to market, employment opportunities, and urban infrastructure for rural households (Dercon and Hoddinott 2005; Satterthwaite and Tacoli 2003; Vandercasteelen et al. 2018). However, as discussed before, both rural areas and smaller towns feature considerable constraints which limit their potential for more productive interlinkage. A significant share of the population in rural areas and small towns remain poorly connected and lack basic access to critical public services such as roads, health facilities, schools, and markets[36]. Given the importance of these public services to improving household welfare and reduced spatial inequality, policies should focus on resolving issues of accessibility and quality of these services.

Policy interventions are also required to enhance livelihood opportunities in less-connected areas. While rural areas and small towns generally report lower unemployment rates compared to larger

[36] In 2016, for instance, more than half of the total rural population lived more than three kilometers away from an all-weather road (OECD/PSI 2020; Schmidt and Kedir 2009; World Bank 2020).

towns, they face wider issues owing to the lack of diversification of livelihood and underemployment (Kamei and Nakamura 2020; World Bank 2020). This is explained principally by the mismatch between the demand for and supply of labor. Despite the ongoing efforts to create more employment opportunities in these locations through the promotion of private and foreign direct investment, the area has not generated enough jobs to absorb the surplus labor in these locations (Broussar and Tekleselassie 2012; Kamei and Nakamura 2020; OECD/PSI 2020). Governments may need to consider stimulating the economy with a focus on increasing job opportunities, particularly for the low-skilled youth existing in and around small towns. Potential interventions include a combination of active labor market policies (ALMPs), education and training policies, policies for productive agricultural job creation, and large scale public employment programs (von Braun and Kofol 2017). Given the current share of the agriculture sector in total employment and the extent of its linkage with other sectors, increasing productivity in and commercialization of agriculture should be given due attention. Interventions that could enhance land tenure security to farmers and accessibility of labor market information to the landless in the rural areas is worth considering.

CHAPTER **3**

3. Heterogeneous effect of urban proximity on nutritional outcomes

Abstract

African countries are urbanizing at an unprecedentedly rapid rate. While this has led to bigger cities, it mainly unfolded by generating several small towns. Earlier literature has provided evidence of the positive impacts of urban proximity on nutritional outcomes. However, it is less clear whether the size of proximate urban areas also matters. In this paper, we look at heterogeneous effects of urban proximity on the nutritional status of households and hypothesize that, once proximity to urban areas is accounted for, the effects differ for large and small urban areas. We use three rounds of nationally representative household and community datasets from Ethiopia and address the endogeneity of transportation cost and self-selection of households to the place of residence in our econometric specification. Our findings indicate that both the degree of proximity to urban areas as well as the size of proximate urban areas affect households' nutritional status, measured in terms of household dietary diversity and child stunting. Specifically, a reduction in the cost of transportation to the nearest town by half leads to a 0.3 percent increase in diet diversity and a 0.8 percentage point reduction in the probability of child stunting. On the other hand, households located in the proximity of large urban areas are better off compared to those near small towns, as their dietary diversity is higher by 1.2 percent while the probability of child stunting is lower by about 3 percentage points.

JEL Classification: C26, I14, I38, R13, R41, R58.

Keywords: Health and nutrition, Ethiopia, regional planning and policy, IV approach, remoteness

Publication status: Mekdim D. Regassa & Zaneta Kubik (2022): Under review for publication as a journal article.

3.1. Introduction

Countries in Sub-Saharan Africa (SSA) are urbanizing at an unprecedentedly rapid rate. Although the current share of the urban population, at about 40 percent, is lower compared to other developing regions, the rate of urbanization in the region is high and is expected to accelerate over the coming decades. By 2050, 55 percent of the region's population is projected to live in urban areas (UN Habitat 2014; UNDESA 2015). If well-managed and accompanied by structural transformation, urbanization is welfare-enhancing for both city-dwellers and for the population of the surrounding areas (Cali and Menon 2009; Christiaensen and Todo 2014). Since it generally provides better income-earning opportunities and access to markets and services, it also holds a great potential to improve the nutritional status of households and reduce the prevalence of malnutrition among children (Stifel and Minten 2017).

In the context of African countries, while a large body of literature exists on the link between urbanization and nutritional status, the focus has mainly been either on the rural-urban gap (von Braun et al. 1993; Sahn and Stifel 2004; Worku et al. 2017) or the effect of the proximity to urban areas (Headey, Stifel, et al. 2017; Stifel and Minten 2008, 2017). However, while such analyses are informative in and of themselves, they tend to underestimate the true effect of urbanization for two main reasons. First, there exist wide intra-urban and intra-rural spatial disparities in nutritional status as a result of the underlying differences in agro-ecological endowments, access to market and public services. For instance, due to the limited carrying capacity of cities in developing countries, public infrastructure and service provision are not equally distributed within the urban space – their availability and quality are much lower in slums and peri-urban areas, with implications for health outcomes of their dwellers (von Braun et al. 1993; Dorosh and Thurlow 2014). Similarly, in rural areas, there exists a clear spatial pattern across the remoteness gradient. Households and children in remote rural areas have poorer nutritional status compared to those in connected areas owing to their disadvantage in terms of access to markets (Hoddinott, Headey, and Dereje 2015), information (Hirvonen et al. 2017), and public services (Abay and Hirvonen 2017). Therefore, ignoring this apparent heterogeneity of urban and rural areas and using only distance to the nearest urban area misrepresents the true effect of urbanization.

Second, the current rapid rate of urbanization, improvements in infrastructure networks, and developments in information and communication technologies have blurred the distinction between urban and rural areas. This has rendered the use of a binary rural-urban classification or a simple distance to the nearest urban areas too simplistic to represent the complex reality of urbanization (von Braun 2014b; Muzzini 2008). The rapid urbanization in SSA particularly led to the proliferation of small-and- intermediate towns that are distinct from large urban areas in terms of the type of economic activity, amenities and degree of linkage with the surrounding rural areas; which, in turn, translates into heterogeneity of impacts (Christiaensen and Kanbur 2017; Christiaensen, De Weerdt, and Todo 2013; Kanbur et al. 2019; Satterthwaite and Tacoli 2003)[37].

For this reason, to better understand the impact of urbanization on nutritional outcomes in low-income countries, both the degree of urban proximity and the type of the proximate town need to

[37] In the case of African countries, not only has the share of population in small and medium towns doubled in the last decade but that this pattern is also expected to continue. Already now, small and medium towns host the majority of urban population; and over the next decade, their population is expected to grow by more than 30 percent (UNDESA 2015).

be considered jointly. So far, there is little rigorous empirical evidence on the heterogeneous effect of different-sized urban areas on nutritional status in Africa, and the results of the limited existing studies are inconclusive. On the one hand, some studies show that proximity to small towns has a larger positive impact on nutritional status as they tend to have a stronger linkage with the rural hinterlands (Christiaensen and Kanbur 2017; Christiaensen, De Weerdt, and Todo 2013). In particular, production and marketing linkages for agricultural products in small towns are well established due to lower transportation costs and stronger local ties. Proximity to small towns also enables rural households to gain access to specialized services and facilities, input markets, and off-farm employment opportunities (Dercon and Hoddinott 2005; Vandercasteelen et al. 2018). Furthermore, dense social networks and close cultural ties between small towns and surrounding rural areas can facilitate more effective dissemination of new ideas and technologies (Berdegué et al. 2015; Brutzkus 1973). Thus, the growth of small towns could directly benefit the welfare of the households located in surrounding areas by providing increased market access and economic opportunities (Reardon 2016).

On the other hand, larger towns might have a stronger effect on nutritional outcomes compared to small towns due to the agglomeration effect whereby bigger markets provide economies of scale for commerce and concentrate the development of new agricultural technologies and innovations (Redding 2010; Redding and Venables 2004; Venables 2008; World Bank 2009, 2020). According to a strand of economic literature called the New Economic Geography (NEG), sizeable economies of agglomeration in larger towns and cities potentially lead to faster economic growth and off-farm job creation for their inhabitants and those nearby (World Bank 2009). Conversely, the positive effect of town size might dissipate after a threshold where congestion effects start hampering growth in urban areas and surrounding hinterlands (Henderson and Becker 2000; Ingelaere et al. 2018).

These contending arguments suggest that once the distance to the nearest urban area is accounted for, the effect of the size of the nearest urban area on the nutritional status is ambiguous. We contribute to this debate by simultaneously examining the effect of urban proximity - proxied by transportation cost to the nearest town - and the heterogeneous effect of the size of nearest towns on households and children's nutritional status. Moreover, we identify and test the mechanisms that explain these effects.

We use three rounds of nationally representative household and community level surveys from the Living Standard Measurement Study – Integrated Survey of Agriculture (LSMS-ISA). After addressing the potential endogeneity of transportation cost and bias resulting from self-selection to place of residence, we find that both proximity to and the size of the nearest urban area affect households' nutritional status. Specifically, proximity to towns has a strong positive effect on nutritional status, with households close to large towns better off than those close to small towns. A reduction in transportation cost by half leads to a 0.3 percent increase in dietary diversity and a 0.8 percentage point (pp) reduction in child stunting. Moreover, the dietary diversity of households close to large towns is likely to be higher by 1.2 pp points and child stunting is likely to be lower by about 3 pp.

The remainder of the chapter is organized as follows. The next section highlights the conceptual framework that guides the analysis. Section 3.3 describes the data and provides some descriptive results. Section 3.4 presents the econometric approach and the basic result. Section 3.5 presents a sensitivity analysis of the basic results. Section 3.6 discusses the main mechanisms underlying the basic result, and Section 3.7 concludes the chapter.

3.2. Patterns of urbanization and nutritional outcomes

The relationship between urban proximity and nutritional status is underlined by complex household sociodemographic, location, and policy-related factors. Figure 3.1 summarizes these basic relationships. It indicates how proximity to urban areas can lead to higher income and hence improved nutritional status through employment opportunities, market access, public services and remittances. Urban proximity could improve households' access to better employment opportunities and overall income (Fafchamps and Shilpi 2003, 2005). However, in low-income countries, income/wealth alone is not a sufficient condition for optimal nutritional status. Access to well-functioning food markets is also key (Abay and Hirvonen 2017; Hirvonen and Hoddinott 2017; Hoddinott et al. 2015). Access to public services and nutrition knowledge — key determinants of nutritional status — also improve with proximity (Hirvonen et al. 2017).

However, the direction and the degree of the link between urban proximity and nutritional outcomes are conditional on the household and institutional context. In developing countries, proximity to urban areas might be *unfavorable* to the nutritional status of households due to several reasons. First, because most urban centers do not have adequate sanitation systems or waste collection services, households surrounding these urban areas suffer from contaminated water sources and environments, leading to poor health and nutrition outcomes (von Braun et al. 1993). Second, the opportunity cost of home production and the availability of modern grocery stores increases with urban proximity, which leads to unhealthy foods being consumed outside the household (Murphy 2018; Worku et al. 2017). Third, improved employment opportunities for women in proximate urban areas might reduce the duration of breastfeeding (Glick 2002). Fourth, because poor households in and around urban areas rely heavily on daily wages, their income and hence nutritional status could be negatively affected by economic shocks disproportionately, compared to those of remote households (von Braun et al. 1993). Higher prices of locally produced foods that are associated with urbanization also affect the poor in and around urban areas disproportionately as these groups obtain the majority of their food from markets (Minten et al. 2018).

The degree of the links between urban proximity and nutritional outcomes might also depend on the size of the proximate towns. Urban areas in developing countries vary considerably in terms of the type of economic activities carried out, and access to and quality of public services such as roads, water and sanitation, electricity, and mobile and radio network. While these services generally improve with the increasing size of urban areas, this is not always the case. Survey results indicate that a larger percentage of poor households in large urban areas (compared to households in smaller towns) within low-income countries reported poor housing and sanitation issues (von Braun et al. 1993; Muzzini 2008). This partly explains the prevalence of relatively more severe malnutrition issues among poor households in large urban areas (World Bank 2010). Differences in access to and stability of employment, the strength of social capital, and the social and cultural background of resident households might also determine the strength of the link between urban proximity and nutrition outcomes (Kamei and Nakamura 2020). In this paper, we jointly examine the effects of urban proximity and the size of the closest towns to determine the causal impact of urbanization on nutritional outcomes.

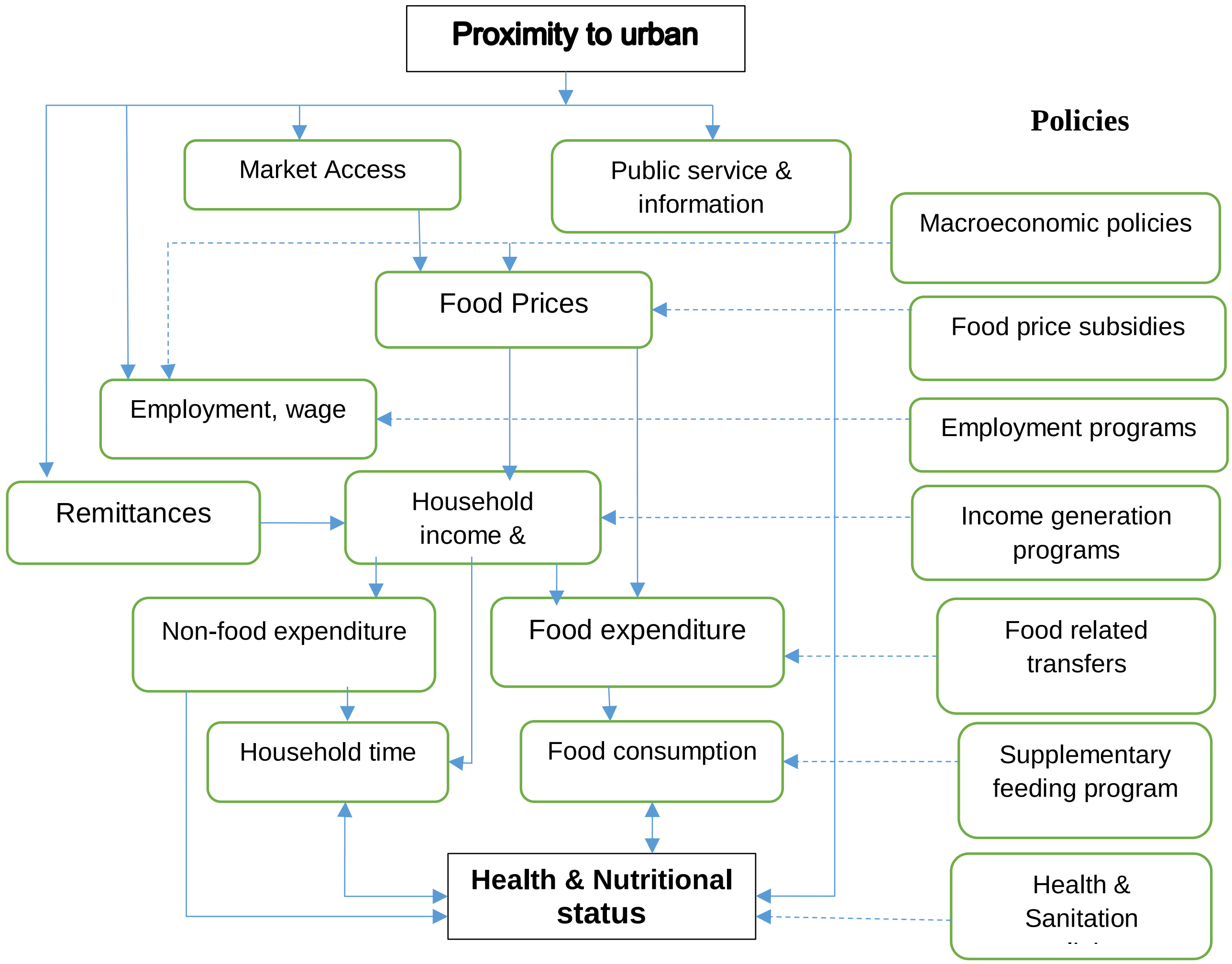


Figure 3.1. Health and nutrition outcomes and urban proximity, conceptual framework
Source*: Modified from* (von Braun et al. 1993)
Note*: Solid lines indicate market and household links; broken lines indicate interventions.*

One challenge in the estimation of the causal impact of urbanization on nutritional status is that both used indicators – urban proximity and size of the nearest urban area – are potentially endogenous in a model explaining nutritional status. For example, the list of policies indicated on the right-hand side of in Figure 3.1 could affect both the measure of proximity as well as the nutritional outcome. At the same time, the place of residence is not exogenously determined because households might systematically self-select to live in a certain place. In this chapter, we use an identification strategy that counters both these problems (see below).

3.3. Data, measurement, and descriptive results

3.3.1. Data

The principal source of data is the Ethiopian LSMS-ISA. This is a nationally representative longitudinal dataset collected jointly by the Central Statistical Agency (CSA) of Ethiopia and the World Bank over three rounds in 2012, 2014, and 2016[38]. The data is panel and covers both rural and urban areas in all administrative regions. The questionnaires are comparable across waves and include surveys at both household and Enumeration Area (EA)[39] levels. The household survey collected detailed information, *inter alia*, on households' demographic characteristics, agricultural activities, food consumption, and labor market participation. The EA (also called *community*) survey gathered detailed information on the availability of and distance to public services, employment opportunities, market prices, etc. Importantly, the surveys also collected GPS coordinates of households' residences as well as detailed information on the type of nearby urban areas, transportation cost, and distance between the place of residence and nearby urban areas. In this study, all three rounds are used.

3.3.2. Measurement of key variables

A. Measurement of explanatory variables

As discussed above, the two key explanatory variables are the size of the nearest urban areas and the measure of the degree of urban proximity. The size of the nearest urban area is determined based on the survey question posed to each respondent as: "is the community in a Woreda town or in a major urban center (regional or zonal capital)?" Based on this information, we classify households into large-town and small-town households. Large town households — households *in and around* large towns — are those for which the nearest town is a regional or zonal capital. *Regional and zonal* capitals represent the first and second-tier cities in the hierarchy of cities in Ethiopia, respectively. These are political and economic centers and serve as hubs that connect different spatial concentrations of human settlement. While there is no universal definition of what constitutes a *large town*, there is a consensus that these towns are likely to accommodate a population size greater than 100,000 (EDRI and GGGI 2015; Roberts 2014b)[40]. This description fits zonal and regional capitals in Ethiopia during the survey period. These account for about 38% of the total households in our sample.

The second group, small-town households – households *in and around* small towns – are those for which the nearest town is either *Woreda* or Kebele capital. Woreda (or district) and Kebele are the third, and the fourth (the lowest) tier administrative units in Ethiopia, respectively. This definition of small towns is in line with the definition from the World Bank and the Ethiopian Ministry

[38] An additional round was collected in 2018/19. However, this is not included as this is a baseline for a new panel, not a follow-up to previous waves.

[39] Enumeration areas (EAs) are equivalent to a village, relatively small, consisting of about 250 households on average.

[40] Many of these centers are growing very rapidly and are projected to accommodate the vast majority of the growing urban population in developing countries (UNDESA 2015). As these centers are also facing enormous urban-development and growth-management problems, their sustainability requires prudent and proactive management.

of Urban Development and Construction (MoUDC) (MoUDC 2012; World Bank and Cities Alliance 2015)[41]. These account for about 62% of the total households in our sample.

Such disaggregation of urban areas by size is vital for building effective institutional and policy frameworks that promise to benefit the economy and society following urbanization (Bloom et al. 2010). It also contributes to the burgeoning debate on the heterogeneous impact of city size on the welfare of its residents and the surrounding population (Christiaensen and Kanbur 2017; Gibson et al. 2017; Ingelaere et al. 2018; Kanbur et al. 2019). From a policy perspective, this is informative as it helps to understand and inform proactive management of the steady urbanization processes of countries in SSA. While it reduces risks associated with congestion, climate effects, and spatial economic inequality, proactively managed urbanization is helpful for optimal allocation of spatial pro-development government resources (Kanbur et al. 2019; Satterthwaite and Tacoli 2003).

Household proximity to urban areas is measured in terms of transportation costs to the nearest urban areas. This information, comprising public transport fares from the center of the village to nearby urban capitals, was collected from village representatives using community surveys. We choose transportation costs rather than the physical distance as a measure of proximity because the former is expected to more accurately reflect the actual cost of remoteness (Stifel and Minten 2008). However, we use the Euclidean distance as one of the instruments in our instrumentation strategy, as explained below.

B. Measurement of outcome variables

We use household dietary diversity index (HDDI) and stunting as indicators of household and child nutritional status, respectively. HDDI reflects the economic ability of a household to access diversified foods. Studies have shown that an increase in dietary diversity is a reasonable indicator of household food security and energy availability (FAO 2013; Hoddinott and Yohannes 2002). The LSMS-ISA household survey collected information on the type and frequency of food items consumed by household members. Following FAO (2013) guidelines, we grouped these food items into 12 categories[42]. An average household consumes 4.5 food groups (Panel A, Table 3.1), with very little variation over time (Table A3.1).

Our second outcome variable - stunting among children - is measured based on the height-for-age (HAZ) score. HAZ score, one of the three common child growth indicators[43], was computed according to the WHO growth standards (Onis et al. 2006; WHO 2006) using children's anthropometric measures collected in the three survey rounds. The first round obtained anthropometric measures (height and weight) of children between 6 and 59 months of age. Subsequent rounds retained children sampled in the first round, including those who became older than 5 years. This explains the increase in children's mean age over time in Table A3.1. Low HAZ is a marker of chronic under-nutrition resulting primarily from prolonged inadequate food intake or infection (WHO 2006). Panel B of Table 3.1 shows that this score is negative for Ethiopia meaning

[41] This classification of urban areas is also consistent with classifications based on population size. In Ethiopia, population centers are distinguished as urban areas if they accommodate a population size of 2,000 or more (MoUDC 2012; World Bank and Cities Alliance 2015).

[42] The food groups are cereals, white tubers and roots, vegetables, fruits, meat, eggs, fish and other seafood, legumes & nuts, milk and milk products, oils & fats, sweets, and spices & condiments.

[43] The other common measures are weight-for-height (WHZ) and weight-for-age (WAZ) score.

that relative to the international reference of well-nourished children, an average Ethiopian child has a lower HAZ score, i.e. he or she is short for his or her age.

In general, children are considered chronically undernourished (stunted) if their HAZ score is below -2. The proportion of stunted children in a population is generally regarded as a good measure of nutritional deprivation and the health status of the population (Pradhan, Sahn, and Younger 2003; Sahn and Stifel 2004). Chronic undernutrition remains widespread in Ethiopia. Nationwide, 38 percent of children under five were reported to be stunted in 2015 — a notable reduction from 58 percent in 2000 (CSA and ICF 2016). Despite this progress, the stunting rate remains very high compared to other developing countries (Headey 2014). In the LSMS-ISA sample, 45 and 35 percent of the children were stunted in 2012 and 2016, respectively (Table A3.1).

Table 3.1. Descriptive statistics of key variables by urbanization status

Variables	Total	Small town	Large town	Mean diff.	Sig.
Panel A: Household-level characteristics					
Number of food groups consumed by HH	4.5	4.4	4.8	-0.39	***
Proportion of food groups consumed by HH	0.38	0.36	0.40	-0.03	***
Transportation cost to the nearest town, ETB	17.5	18.9	15.2	3.65	***
Distance to nearest town, Km	22.5	19.9	26.6	-6.70	***
Mobile phone ownership by HH	49.3	40.1	64.0	23.9	***
Household size, number	4.7	4.9	4.3	0.58	***
Age of household head, years	44.9	45.4	44.0	1.43	***
Male household heads, %	70.6	73.2	66.4	6.8	***
Heads with primary education, %	28.3	28.3	28.4	-0.1	
Heads with secondary education, %	15.7	9.0	26.4	-17.4	***
Household took credit, %	23.3	24.6	21.3	3.3	***
Household runs non-farm enterprise, %	33.3	30.9	37.1	-6.2	***
Livestock owned, TLU [a)]	3.5	4.3	2.3	2.0	***
Durable assets owned, PCA [b)]	0.0	-0.7	1.1	-1.7	***
Observations	14,173	8,722	5,451		
Panel B: Child level characteristics					
Child height-for-age z-score	-1.43	-1.51	-1.29	-0.22	***
Child weight-for-height z-score	-0.44	-0.48	-0.35	-0.13	***
Child weight-for-age z-score	-1.20	-1.27	-1.07	-0.20	***
Prevalence of stunting, %	37.1	39.2	32.9	6.30	***
Prevalence of wasting, %	12.5	12.9	11.7	1.20	*
Prevalence of underweight, %	25.6	27.4	22.2	5.20	***
Child is female, %	48.5	48.3	49.1	-0.80	
Child age in years	4.00	4.01	3.98	0.03	
Observations	12,030	8,044	3,986		

Source*: Author's computation based on LSMS Survey (2012, 2014, and 2016)*

Note*: *, **, and *** represent variables for which the mean difference tests are statistically significant at 10%, 5%, and 1% levels, respectively. [a)] Livestock was measured using tropical livestock units (TLU), which is a common unit used to quantify a wide range of livestock species to a single figure to obtain the total amount of livestock owned by a household. This study employed a tropical livestock unit applicable for SSA [b)]Durable assets owned is an index generated using principal component analysis (PCA) from individual asset items owned by households.*

3.3.3. Descriptive results

Table 3.1 shows the pattern in the outcome and selected covariates by the size of the proximate urban areas. Panel A shows that the size of the proximate urban area is strongly and positively associated with the consumption of more diversified food items. While an average large-town household consumes 4.8 food groups, a small-town household consumes 4.4 food groups. Panel B depicts the same pattern, albeit for child anthropometric scores. While 32.9 percent of children in large towns are stunted, the corresponding figure for children in small-town households is 39.2 percent. The average differences in the outcome variables between small and large town households are statistically significant (p-value<0.05). Similarly, Figure 3.2 shows that both HDDI and stunting exhibit a systematic pattern across the transportation cost gradient. This is consistent with a strand of literature that shows that lower transportation cost is a robust predictor of improved health outcomes (Abay and Hirvonen 2017; Ahmed and Hossain 1990; Headey, Stifel, et al. 2017; Hirvonen et al. 2017; Lokshin and Yemtsov 2005).

However, neither the descriptive results presented in Table 3.1 nor the pattern shown in Figure 3.2 can be used to make causal inferences regarding the effect of city size and transportation cost on nutrition outcomes as they do not account for potential confounding factors (see the discussion below). The next section accounts for these confounding factors to tease out the effect of city size and transportation cost on nutritional outcomes.

Panel A: Household diet diversity score

Panel B: Children's HAZ score

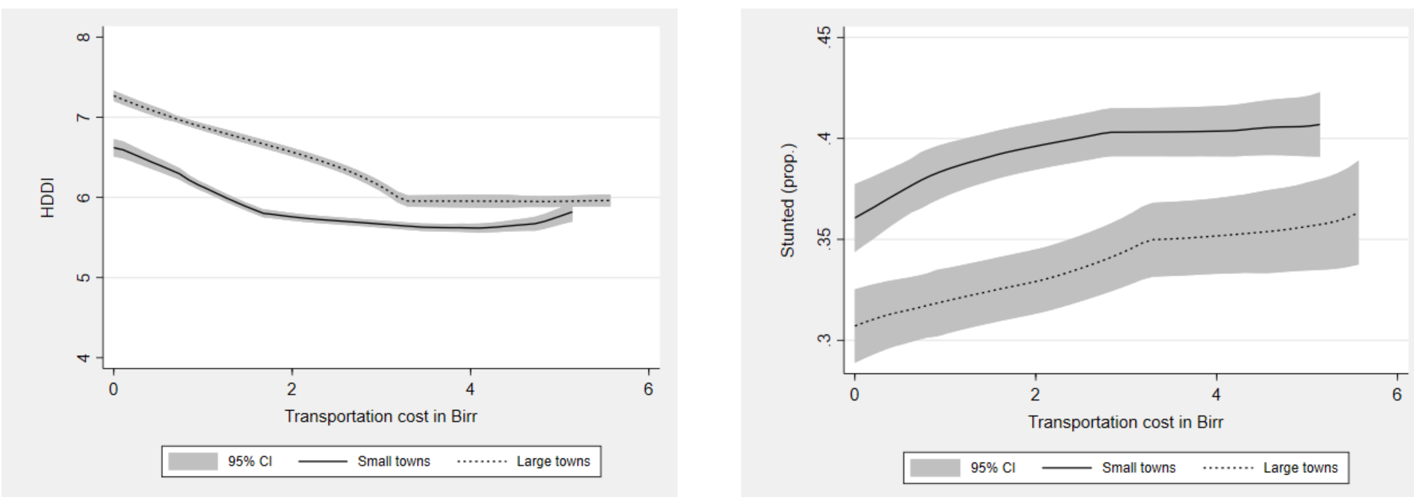


Figure 3.2. Transportation cost and nutritional status

Source: Author's computation based on LSMS Survey (2012, 2014, and 2016)

Table 3.1 further reveals that the two key explanatory variables – transportation cost and size of the proximate town - are closely related. Transportation cost decreases as the size of urban areas increases. It is interesting to note that this difference in transportation cost is not only driven by the physical distance between the households to urban centers, but also by the increasing marginal cost of transportation. Table 3.1 shows that the cost per unit distance - the ratio of transportation cost to distance - is larger for small towns than for large towns. This might partly be explained by non-random placement and the quality of transport infrastructure, as well as the availability and level of competitiveness of transportation services.

From a methodological point of view, an IV method is required to account for the endogeneity of the transport cost. Furthermore, the strong correlation between transportation cost and urban size suggests that it is necessary to simultaneously account for the proximity to and size of proximate urban areas. The next section focuses on these issues[44].

3.4. Econometric approach and basic results

3.4.1. Econometric approach

We model HDDI and stunting ($\mathrm{W_{it}}$) of household/child i at time t as a function of transportation cost to the nearest town ($\mathrm{T_{it}}$), and the size of the nearest town (S_i). The basic econometric model is specified as:

$$W_{it} = \beta_0 + \beta_1 T_{it} + \beta_2 S_i + \beta_3 X_{it} + \varepsilon_{it} \qquad (3.1)$$

In (3.1), X_{it} represents a vector of household and community level characteristics that affect nutrition outcomes. To control for potential confounding factors that may be correlated with both the outcome variables (W_{it}) and the cost of transportation (T_{it}), X_{it} covers several covariates at different levels. Based on the literature, household characteristics that strongly influence nutrition outcomes such as household size, the value of durable assets, and the size of livestock owned are included. Included are also characteristics of the household head including age, gender, and education levels. For stunting, additional controls related to child characteristics such as age and gender are included. Finally, zonal fixed effects are included in all the estimations as observed and unobserved agro-ecological and other location characteristics might influence the nutritional outcomes. The last term in the equation, ε_{it}, is the random error term. Standard errors are clustered at the village level for HDDI and at the household level for stunting. For simplicity, we estimate equation (1) for HDDI and stunting using linear regression models. However, we test the sensitivity of our results by applying alternative econometric models (see section 3.5).

In equation 3.1, β_1 and β_2 capture the main relationships of interest. While β_1 represents the impact of transportation cost, β_2 represents the relative effect of large towns. We hypothesize that for HDDI, β_1<0, β_2>0; and for stunting, $\beta_1 > 0$, $\beta_2 < 0$. That is, urban proximity — measured by lower transportation cost — improves household dietary diversity and reduces the prevalence of child stunting. On the other hand, compared to small towns, large towns improve household/child nutritional status.

As indicated before, transportation cost is likely to be endogenous in equation 3.1, thus rendering the consistency of β_1 estimated using OLS questionable. Transportation costs could be endogenous due to at least three main reasons. First, transport infrastructures are not placed randomly in space and are likely to be influenced by political as well as economic factors. Second, having a wide range of household wealth indicators on the right-hand side may expose this estimation to omitted variables bias, as household wealth is associated with a plethora of unobservable characteristics. Finally, since data on transportation cost was collected in monetary terms, it is susceptible to measurement error.

When transportation cost is endogenous, $corr(T_{it}, \varepsilon_{it}) \neq 0$ and hence β_1 fails to be a consistent estimator (Wooldrigde 2013). To address this concern, we apply an instrumental variable (IV) approach. The first instrument is Euclidian distance from the residence of the household to the nearest town (in kilometers), which we argue to be a valid instrument as it does not follow the

[44] This correlation is, however, not too strong to menace multicollinearity.

potentially endogenous road networks. The second instrument is an interaction between the first instrument, the Euclidian distance, and altitude. All else equal, locations that are on extreme altitudes (very high or very low) are more likely to have higher transportation cost compared to areas at average altitudes (Stifel, Minten, and Koru 2016). While initial construction costs of infrastructure tend to be higher at extreme altitudes, the average fuel cost per passenger is also likely to be higher in these locations[45]. To isolate extreme altitude areas, a dummy variable is generated that takes a value of zero when altitude is within two standard deviations from the mean, one otherwise[46].

The validity of the IV strategy rests on two criteria. The first is the relevance criterion that demands that the instruments should be good predictors of transportation cost. To formally test for this criterion, transportation cost is estimated as a function of the instruments and other relevant household and community characteristics, including several household wealth measures. Table 3.2 shows the first-stage regression results. The first column presents a result of the more parsimonious model where only the instrumental variables and the town size indicator dummy variables are included. From this result, it is evident that the instruments are relevant. That is, the instruments are good predictors of transportation cost. The model's partial F-statistic is larger than 10, the minimum threshold value of the rule of thumb for valid instruments (Staiger and Stock 1997). The second column of Table 3.2 presents the results with more covariates related to household and location characteristics. In this more elaborate model with zonal fixed effects, the coefficients on both instruments are statistically significant and appear with an *a priori* expected sign. Moreover, the additional IV diagnostic tests presented at the bottom of the Table affirm the relevance of the instruments.

[45] One reason for this could be that on extreme altitudes, population density and hence the number of commuters tends be very low. Public transportation facilities often charge a higher fare per person in order to compensate for missing revenues owing to vacant seats.

[46] Rather than 2 standard deviation, the sensitivity of the results with 1 standard deviation and over the whole range of altitude is also tested. The basic results, available from authors upon demand, remained qualitatively the same.

Table 3.2. First Stage regression result: determinants of transportation cost

Explanatory variables	a0	a1
ln(Distance to nearest market town)	0.844***	0.822***
	(0.010)	(0.014)
(Distance to nearest town)*(village is at extreme altitude)	0.045*	0.047*
	(0.025)	(0.026)
Large town, yes=1	-0.238***	-0.218***
	(0.038)	(0.043)
Household and location characteristics	No	Yes
Zonal Fixed Effects	No	Yes
Constant	0.324***	-2.887
	(0.037)	(2.558)
Number of observations	14,139	14,036
R2	0.852	0.889
Adjusted R2	0.852	0.889
F test of excluded instruments:		
F(1, 432)	3,617.4	1,873.0
Prob > F	0.00	0.00
Weak-identification tests:		
Cragg-Donald F-statistic:	3,617	1,873
Kleibergen-Paap rk Wald F statistic:	204	157
---p-value	0.00	0.00
Over-identification test		
Hansen-J	0.01	0.98
---p-value	0.93	0.32

***Source**: Author's calculation based on LSMS-ISA (2012, 2014 & 2016)*
***Note**: Standard errors clustered at the village level in parentheses. Statistical significance indicated by: *** p<0.01, ** p<0.05, * p<0.1; Coefficients on household and location characteristics omitted to preserve space*

The second criterion for good instruments is the *exclusion restriction* which requires that instruments should not affect the outcome variable (i.e. the nutritional status) other than through the transportation cost. One specific concern against this exclusion restriction is that the distance variable might not be exogenous. It might be possible, for example, that households concerned about their welfare may relocate to areas better connected to urban areas. If so, this would violate the exclusion restriction. However, this is not likely to pose a serious threat as the cost of migration is prohibitively high in Ethiopia, especially so for rural households because of absent private land markets (Deininger et al. 2003). Land is owned by the state and individual farmers have only user rights. Securing land use rights is contingent on permanent physical residence in the community. Therefore, it is too costly for households seeking to enhance their welfare to do so by changing their place of residence in the short run, which reduces the likelihood of selection bias. Nevertheless, as a sensitivity test, the model is re-estimated by restricting the sample to those households remaining in the village for the whole period[47]. This last step helps account for endogenous dynamic migration decisions.

Another potential concern with the exclusion restriction is that the altitude indicator variable might directly affect households' nutritional status as it is correlated with agro-ecological factors

[47] This analysis produced similar results and is available from the authors on demand.

(Niermeyer, Mollinedo, and Huicho 2009; Singh et al. 1977). Altitude could also be correlated with other unobserved variables that are correlated with the outcome variables. If either of these two conditions holds, then the exclusion restriction would be violated. To address this issue, the altitude indicator variable is also included as a right-hand-side variable- X_{it}. Furthermore, mean annual temperature and zonal fixed effect are controlled for to ensure that this instrument is not simply picking up differences in agro-ecological factors.

Another concern in the estimation of equation (3.1) is that there might be selection bias because households might systematically self-select to live in and around large towns. If this is the case, β_2 cannot be estimated consistently. Indeed, the descriptive result in Table 1 shows that there might be spatial sorting into large towns based on the human capital endowment. The Table shows that while about 26.4 percent of large-town household heads attained secondary or tertiary level education, only 9 percent of small-town household heads attained a similar level of education. Though not as extreme, one can observe a systematic difference between households residing in the two locations based on asset ownership, housing quality, land size, and access to health and social services.

To address this potential selection bias, we apply a double robust regression approach (Rosenbaum 2012; Rosenbaum and Rubin 1984). This method first involves the estimation of the probability of residing in a large town and then adjusting the regression estimation based on a weight generated from the predicted value of the selection equation. Based on the literature, the selection equation includes variables that are likely to affect the probability of residing in large towns such as household characteristics (age, gender, education, and household size), asset ownership, and housing quality. Table A3.2 in the appendix presents the result from the estimation of the selection equation using a probit model.

3.4.2. Basic results

For reference, we first estimate the outcome variables on transportation costs and the size of the proximate urban area using simple Ordinary Least Square (OLS). The results presented in columns 1 and 3 of Table 3.3 show that proximity to urban areas, as given by lower transportation costs, is associated with improved nutritional status. Reducing the cost of transportation by half is associated with a 0.4 percent increase in dietary diversity and a 0.6 percentage point (pp) reduction in the probability of stunting. The result corresponding to the size of towns indicates that, for households in large towns, dietary diversity is likely to be higher (by 1.2 percent) and stunting is likely to be lower (by about 3 pp), compared to households in small towns.

However, as discussed above, the OLS estimation does not account for the potential endogeneity of transportation costs as well as the possible bias due to self-selection into large towns. To address these issues, we combine an instrumental variable (IV) method with Inverse Probability Weighting (IPW). The results from this estimation is presented in columns 2 and 4 of Table 3.3 for diet diversity and stunting, respectively. The results show that transportation cost has a causally negative effect on household diet diversity and a positive effect on child stunting. Furthermore, the results of the simple OLS regressions and the IV estimations are largely consistent. These results indicate that the pattern of urbanization is as important to household/child nutritional outcomes as the aggregate rate of urbanization. That is, while urbanization is generally important, the growth of large towns is more important for nutritional status than the growth of small towns. This study only explored diversity in food consumption at the household level. More fine-grained geographical distribution of micronutrients across East African countries shows that households in large towns are better off compared to those in megacities and rural areas (Ameye 2018).

Table 3 also reveals that nutritional outcomes are significantly correlated with many other covariates. Consistent with previous studies (Headey, Hoddinott, et al. 2017; Hoddinott et al. 2015), the education level of the household head and wealth indicators — ownership of livestock and durable assets — appear to be important covariates of improved nutritional status.

Table 3.3. Impact of urbanization on nutritional status

	[1]	[2]	[3]	[4]
	Diet Diversity		Stunting	
	OLS	IV-IPW	OLS	IV-IPW
ln(Transportation cost)	-0.008***	-0.005*	0.011**	0.017**
	(0.002)	(0.003)	(0.005)	(0.007)
Large town, yes=1	0.012*	0.012*	-0.030*	-0.037*
	(0.007)	(0.007)	(0.018)	(0.020)
Household size, number	0.006***	0.006***	0.000	-0.005
	(0.001)	(0.001)	(0.003)	(0.004)
ln(Age of household head in years)	-0.010***	-0.009*	-0.024	-0.010
	(0.004)	(0.005)	(0.024)	(0.033)
Head is male, yes=1	0.002	-0.001	0.021	0.025
	(0.003)	(0.003)	(0.016)	(0.022)
Head education, reference=None				
Head education, primary	0.018***	0.020***	-0.058***	-0.032*
	(0.003)	(0.004)	(0.014)	(0.018)
Head education, secondary or higher	0.040***	0.042***	-0.098***	-0.078***
	(0.005)	(0.006)	(0.022)	(0.026)
ln(Livestock owned, in TLU) [ϕ]	0.009***	0.012***	-0.022***	-0.030***
	(0.002)	(0.003)	(0.007)	(0.009)
Durable assets owned, PCA[¶]	0.006***	0.006***	-0.014***	-0.013***
	(0.001)	(0.001)	(0.003)	(0.003)
ln(Village elevation, meter)	-0.010	-0.017	0.097**	0.130***
	(0.014)	(0.016)	(0.039)	(0.048)
ln(Annual Mean Temperature)	-0.066	-0.081	0.334*	0.450*
	(0.074)	(0.081)	(0.202)	(0.246)
Survey round, reference=2012				
Survey round, 2014	0.027***	0.023***	-0.063***	-0.073***
	(0.005)	(0.007)	(0.013)	(0.018)
Survey round, 2016	0.035***	0.033***	-0.060***	-0.071***
	(0.005)	(0.007)	(0.014)	(0.019)
Child Characteristics				
Child is female, yes=1			-0.018*	-0.015
			(0.010)	(0.014)
ln(child age in months)			-0.050***	-0.045***
			(0.008)	(0.010)
Zonal Fixed Effects	Yes	Yes	Yes	Yes
Constant	0.649*		-0.898	-1.761
	(0.392)		(1.097)	(1.324)
Number of observations	14,017	14,005	11,049	11,045
R2	0.224	0.075	0.067	0.071
Adjusted R2	0.218	0.068	0.058	0.063
Weak-identification tests:				
Kleibergen-Paap LM statistic:		96.66		351.11
--- p-value:		0.00		0.00
Over-identification test				
Hansen-J test:		0.76		1.47
--- p-value:		0.38		0.23

Source: Author's calculation based on LSMS-ISA (2012, 2014 & 2016)

*Note: Clustered Standard errors in parentheses. Statistical significance indicated by *** p<0.01, ** p<0.05, * p<0.1; [ϕ]Livestock was measured using tropical livestock units (TLU), which is a common unit used to quantify a wide range of livestock species to a single figure to obtain the total amount of livestock owned by a household. [¶]Durable assets owned is an index generated using principal component analysis (PCA) from individual asset items owned by households.*

3.5. Sensitivity analysis

In this section, we assess the robustness of the basic results in two important ways. First, we re-estimate the basic model by limiting the sample to non-urban households, in contrast to the basic analysis which included both urban and rural households. One could argue that the basic results merely pick up the differences among the urban areas, rather than show the effect of proximity to different sized urban areas. This is potentially a problem, especially if urban areas have weak linkages with the surrounding population.

To examine such a possibility, we limited the analysis to non-urban households - households with a positive distance to town. In the data, urban households account for 17.1 percent of the sample. The result is presented in Table 3.4. While household and community characteristics are controlled for in all regressions, only coefficients associated with the key variables of interest are reported to preserve space[48]. The results in suggest that the exclusion of urban households does not alter the conclusion that households in large urban areas are better off compared to small-town households.

Table 3. 4: Impact of urbanization on nutritional outcomes, excluding urban household

Explanatory variables:	[1]	[2]	[3]	[4]
	Diet Diversity		**Stunting**	
	OLS	**IV-IPW**	**OLS**	**IV-IPW**
ln(Transportation cost)	-0.005***	0.004	0.010*	0.027*
	(0.002)	(0.003)	(0.009)	(0.014)
Large town, yes=1	0.009***	0.007*	-0.055***	-0.050**
	(0.003)	(0.003)	(0.021)	(0.022)
Household & location characteristics	Yes	Yes	Yes	Yes
Child characteristics	No	No	Yes	Yes
Zonal fixed effects	Yes	Yes	Yes	Yes
Constant	0.578***	0.760***	-1.494	-2.388
	(0.177)	(0.238)	(1.189)	(1.472)
Number of observations	10,628	10,628	9,169	9,169
R2	0.237	0.272	0.060	0.068
Adjusted R2	0.232	0.266	0.051	0.059
Weak-identification tests:				
Kleibergen-Paap LM statistic:		631.3		103.2
--- p-value:		0.00		0.00
Over-identification test				
Hansen-J test:		2.87		1.32
--- p-value:		0.09		0.25

***Source**: Author's calculation based on LSMS-ISA (2012, 2014 & 2016)*

***Note**: Standard errors clustered at the village level in parentheses. Statistical significance indicated by: *** p<0.01, ** p<0.05, * p<0.1; Coefficients on child, household, and location characteristics omitted to preserve space.*

Second, the basic results assumed the outcome variables as continuous linear variables. However, since dietary diversity is a count variable and stunting is a binary variable, using a linear model may not be unequivocally appropriate. Linear models are preferable due to their simplicity, interpretability, and because they provide a host of specification tests to assess the validity of the

[48] The results from the full model is available from the authors on demand.

IV strategy (Angrist and Pischke, 2008; Caudill, 1988). However, for limited dependent outcomes, a linear model may be unreliable (Wooldridge 2002). Therefore, we assess the robustness of the basic findings using Poisson regression for dietary diversity and Probit regression for stunting. Table 3.5 reports the result of these regressions along with their respective IV approach to account for the endogeneity of the transportation cost. The results remain robust and do not seem to be driven by the non-linear nature of the outcome variables.

Table 3. 5: Impact of urbanization on nutritional outcomes, alternative econometric models

Explanatory variables:	[1]	[2]	[3]	[4]
	Diet Diversity		**Stunting**	
	Poisson	**IV - Poisson**	**Probit**	**IV - Probit**
ln(Transportation cost)	-0.023***	-0.013*	0.031**	0.047**
	(0.006)	(0.008)	(0.012)	(0.018)
Large town, yes=1	0.032*	0.033*	-0.079*	-0.101**
	(0.019)	(0.018)	(0.042)	(0.046)
Household & location characteristics	Yes	Yes	Yes	Yes
Child characteristics	No	No	Yes	Yes
Zonal fixed effects	Yes	Yes	Yes	Yes
Number of observations	14,035	14,023	11,049	11,045
Pseudo- R2			0.053	

***Source**: Author's calculation based on LSMS-ISA (2012, 2014 & 2016)*

***Note**: Standard errors clustered at the village level in parentheses. Statistical significance indicated by: *** p<0.01, ** p<0.05, * p<0.1; Coefficients on child, household, and location characteristics are omitted to preserve space.*

3.6. Mechanisms

In the previous sections, we showed that an increase in the proximity to urban areas (as measured by lower transportation cost) and the size of the nearest urban areas lead to an improvement in nutritional status. In this section, we highlight the major mechanisms that underpin this basic finding. First, the descriptive result in Table 3.1 shows that, on average, large-town households are wealthier and more educated than small-town households. Furthermore, Figure 3.3 shows that both wealth and level of education decline along the transportation cost gradient. While we control for the direct effect of both of these factors in all the regressions, the indirect effects might still explain the basic result. Empirical studies conducted in low-income countries indicate that differences in household wealth are the single most important explanatory factor of differences in health and nutrition outcomes (Headey et al. 2015; Headey, Hoddinott, et al. 2017).

Panel A: wealth index by transportation cost Panel B: education level by transportation cost

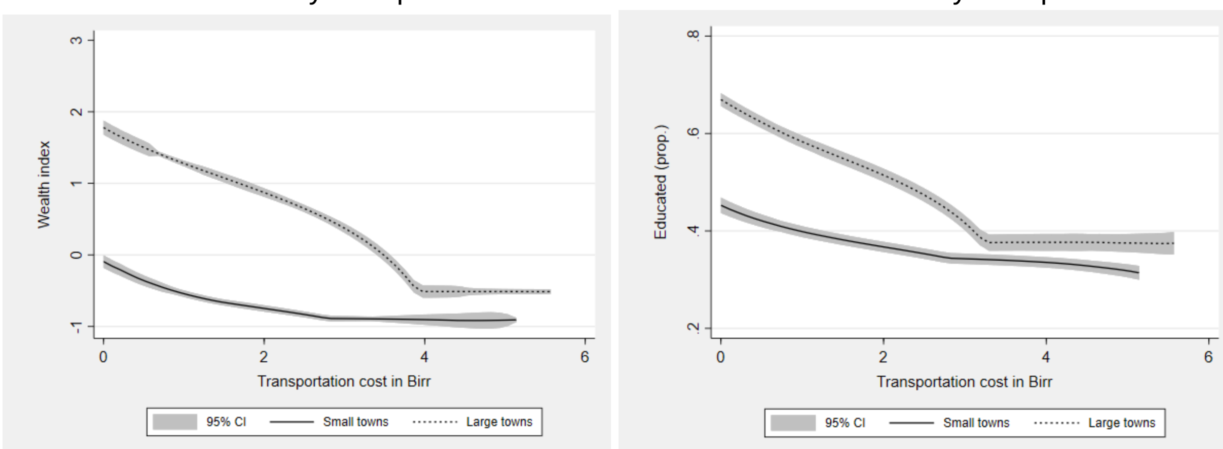


Figure 3.3. Wealth index and education level by transportation cost
Source: Authors' computation based on LSMS Survey (2012, 2014, and 2016)

Other complementary channels that might underlie the spatial pattern in health and nutrition outcomes include differences in access to water and sanitation, public services, employment opportunity, and maternal education and time use. These channels are discussed in more detail below.

3.6.1. Water, Sanitation, and Hygiene (WASH)

Access to clean water, sanitation, and hygiene (WASH) is a fundamental factor to improve health and nutritional status (Humphrey 2009; Spears 2013). Evidence shows that access to WASH is vital, *inter alia*, to improve child and maternal health, reduce water-borne diseases, promote the quality of food hygiene, and reduce inequality based on gender and disability (see Joanna & Oliver, 2016 for review). However, poor access to WASH is widespread across SSA countries. As of the year 2017, less than 30 percent of the region's population had access to basic sanitation (e.g. a clean and safe toilet), and basic handwashing facilities with soap and water. Moreover, 39 percent of the population in the region do not have access to safe drinking water (WHO and UNICEF 2017). This has severe social and economic implications. Estimates show that the lack of access to improved WASH is the second biggest cause of child mortality in Africa. On average, close to 4,000 children under the age of five die every day from WASH-related diseases in the region (UNDP 2006).

While WASH coverage is generally low in the region, there is a significant spatial disparity between and within countries. For instance, Table A3.4 presents the distribution of households' access to water and sanitation facilities by place of residence in Ethiopia[49]. It shows that smaller towns tend to have a larger share of households with substandard housing; i.e., fewer rooms, and inferior quality housing (roof, floor, and wall). Less than 1 percent of small-town households use improved cooking fuel, more than a third resort to open defecation, fetch drinking water from unprotected spring/hole, and use potentially harmful sources of lighting. Half of the households travel more than 15 minutes to a source of drinking water. Table 3.6 shows that even after accounting for household and location characteristics, the quality of housing and WASH systematically vary based on both the degree of urban proximity as well as the size of the proximate urban areas. Given that WASH is a key element of health and nutrition, these observed differences across rural-urban areas may partly explain the spatial difference in health and nutrition outcomes in the country. Therefore, policy interventions that target improvement in WASH are likely to be effective to enhance the overall health and nutrition status as well as reduce the disparity across regions[50].

Table 3. 6: Urbanization and access to clean Water, Sanitation, and Hygiene (WASH)

	Roof	**Floor**	**Toilet**	**Drinking water**
ln(Transportation cost)	-0.067***	-0.025***	-0.034***	-0.078***
	(0.010)	(0.006)	(0.008)	(0.010)
Large town, yes=1	0.052	0.049***	0.085***	0.119***
	(0.033)	(0.019)	(0.027)	(0.036)
Household & location characteristics	Yes	Yes	Yes	Yes
Zonal fixed effects	Yes	Yes	Yes	Yes
Number of observations	14,048	14,048	14,048	14,040
R2	0.383	0.480	0.344	0.422
Adjusted R2	0.378	0.476	0.340	0.418

***Source**: Author's calculation based on LSMS-ISA (2012, 2014 & 2016)*

***Note**: Clustered standard errors in parentheses. Statistical significance indicated by: *** p<0.01, ** p<0.05, * p<0.1; Coefficients, household, and location characteristics are omitted to preserve space.*

3.6.2. Access to public services

Access to public services such as roads, schools, health posts, and communication infrastructure has been shown to be an important determinant of health and nutrition outcomes (Hirvonen et al. 2017; Hoddinott et al. 2015; Stifel and Minten 2017; World Bank 2020). Therefore, the difference in nutritional status between small- and large-town households might be associated with differences in access to these services. Table A3 in the appendix shows that the distance to public services is significantly shorter for large-town households than for small-town households. Relative to large-town households, small-town households live further away from roads, markets,

[49] This pattern is similar across the rest of Sub-Saharan Africa (See WHO and UNICEF, 2017).

[50] A study by UNDP (2006) indicates that Sub-Saharan Africa might save a total $23.5 billion - 5% of GDP- if the entire population had access to basic, low-cost water and sanitation technology.

schools, health posts, and financial services.[51] Furthermore, Table 3.7 shows that large-town households outperform small-town households in terms of access to radio, television, electricity, and mobile phones. These disparities are directly related to differences in wealth. Indirectly, and perhaps more importantly, they represent households' varying access to information, which has been shown to be a key predictor of nutritional outcomes (Hirvonen et al. 2017).

Table 3.7. Urbanization and access to public services and local institutions

	Hospital	Electricity	Mobile phone	Radio	TV
ln(Transportation cost)	-0.077***	-0.097***	-0.041***	-0.027***	-0.031***
	(0.014)	(0.010)	(0.006)	(0.005)	(0.006)
Large town, yes=1	0.087*	0.130***	0.056**	0.015	0.040***
	(0.052)	(0.032)	(0.022)	(0.018)	(0.013)
Household & location characteristics	Yes	Yes	Yes	Yes	Yes
Zonal fixed effects	Yes	Yes	Yes	Yes	Yes
Number of observations	14,048	14,048	14,048	14,048	14,048
R2	0.30	0.58	0.39	0.22	0.67
Adjusted R2	0.30	0.57	0.38	0.22	0.67

***Source**: Author's calculation based on LSMS-ISA (2012, 2014 & 2016)*

***Note**: Clustered standard errors in parentheses. Statistical significance indicated by: *** p<0.01, ** p<0.05, * p<0.1; Coefficients, household, and location characteristics are omitted to preserve space*

3.6.3. Employment opportunity

Another factor to consider is whether opportunities for nonfarm employment differ across urban areas, which could explain the variation in nutritional outcomes. We examine the likelihood of participation and intensity of household employment in off-farm wage employment and non-farm self-employment. Table 3.8 shows that large-town households are more likely to work, and for more hours per week, in wage employment compared to small-town households; but the differences are not significant in the case of non-farm self-employment. Kamei & Nakamura (2020) reported similar results based on spatial analysis of the Ethiopian urban labor market. The results also indicate that households that are better connected to urban areas have better labor-market opportunities related to non-farm activities. This suggests that policy interventions aimed at improving rural infrastructure are likely to improve nutritional outcomes through the labor market.

Table 3.8. Urbanization and patterns in employment status

	Wage employment		Non-farm self-employment	
	Participation	# hours	Participation	# hours
ln(Transportation cost)	-0.014***	-0.056***	-0.023***	-0.094***
	(0.004)	(0.016)	(0.007)	(0.025)
Large town, yes=1	0.050***	0.215***	-0.001	0.005
	(0.013)	(0.049)	(0.021)	(0.086)
Household & location characteristics	Yes	Yes	Yes	Yes
Zonal fixed effects	Yes	Yes	Yes	Yes
Number of observations	14,039	14,039	14,039	14,039
R2	0.229	0.245	0.154	0.168
Adjusted R2	0.224	0.240	0.148	0.162

***Source**: Author's calculation based on LSMS-ISA (2012, 2014 & 2016)*

***Note**: Clustered standard errors in parentheses. Statistical significance indicated by: *** p<0.01, ** p<0.05, * p<0.1; Coefficients, household, and location characteristics are omitted to preserve space.*

[51] This pattern is consistent across most of the selected services except for primary schools and health posts. These two public services are available in every village by government policy.

3.6.4. *Maternal education and time use*

Spatial differences in maternal education, employment, and time use might also explain the heterogeneous impact of urban size on nutritional outcomes. Empirical evidence suggests that maternal education (Alderman and Headey 2017; Emily, Juan, and David 2012; Headey et al. 2015) and maternal productive employment (Alderman et al. 2001) are important determinants of dietary diversity and child health. To test this, we examine the association between maternal education and time use with transportation cost and size of the proximate urban areas in a multivariate regression framework.

Table 3.9 shows that, when compared to mothers in small towns, mothers in large towns are more likely to be educated and more likely to be wage employed. Furthermore, Kamei & Nakamura (2020) reported that, in comparison to small towns, women in large towns are less likely to be unemployed or out of the labor force.

Table 3.9. Urbanization and mothers' education & time use

	Maternal Education	Maternal Employment
ln(transportation cost)	-0.008**	-0.050***
	(0.003)	(0.006)
Large town, yes=1	0.032**	0.064***
	(0.013)	(0.017)
Household & location characteristics	Yes	Yes
Zonal Fixed Effects	Yes	Yes
Number of observations	11,961	11,961
R2	0.440	0.359
Adjusted R2	0.435	0.354

Source: *Author's calculation based on LSMS-ISA (2012, 2014 & 2016)*

Note: *Clustered standard errors in parentheses. Statistical significance indicated by: *** p<0.01, ** p<0.05, * p<0.1; Coefficients, household, and location characteristics are omitted to preserve space.*

3.7. Conclusion

In low-income countries, household nutritional status has a strong spatial character. While the geographical pattern of nutritional status in these countries has been studied to a great extent, the focus has mainly been either on the rural-urban gap or the effect of the proximity to urban areas. However, with the current rate of urbanization and increased use of communication infrastructure, a binary rural-urban classification or a simple approximation of urbanization based on the distance to the nearest urban areas, has become too simplistic to represent the complex reality of urbanization (von Braun 2014b; Muzzini 2008).

In this study, the concept of a rural-urban spectrum is adopted to simultaneously examine the effect of proximity to urban areas and the size of the proximate urban areas on nutritional outcomes. Such spatial differentiation which characterizes the rural-urban space helps understand and inform proactive management of urbanization processes of countries in Sub-Saharan Africa and beyond. As the share of small- and medium-sized towns is increasing amidst accelerating urbanization, this approach is helpful to inform the optimal allocation of spatial pro-development resources.

The findings in this paper suggest that urbanization – both in terms of increasing households' proximity to urban areas and the size of urban areas – has a significant positive effect on nutritional outcomes. Specifically, while the reduction in transportation cost to urban areas has a strong positive effect on nutritional outcomes, large-town households are better off compared to small-town households. Furthermore, the study identifies and discusses the spatial imbalances in the distribution of the sociodemographic and institutional factors that underlie the spatial pattern in nutritional status. The mechanism analysis suggests the availability of a wide range of policy interventions to improve household and child nutritional status, particularly in disadvantaged locations.

CHAPTER **4**

4. Urbanization and intergenerational mobility in Ethiopia

Abstract

In this chapter, we use nationally representative longitudinal data on children and their parents to investigate the extent of intergenerational mobility across rural-urban areas in Ethiopia. The analysis reveals several key results. First, intergenerational mobility is very low in Ethiopia, even when compared to other low-income countries. Second, while the average educational and occupational status improves with urbanization, so does inequality. Third, compared to rural areas or small towns, intergenerational mobility is greater in large urban areas. Fourth, the higher occupational mobility in large urban areas is largely explained by the higher educational level (post-elementary) in this location. These results suggest that making post-elementary schools more accessible and reducing the dropout rates, in addition to improving the quality of education, is one of the most effective mechanisms to reduce spatial and intergenerational socio-economic inequality in Ethiopia.

JEL Classification: J62, J24, O18
Keywords: Intergenerational Mobility, Occupational Choice, Urbanization, Ethiopia

Publication status: Mekdim D. Regassa & Zaneta Kubik (2022): Under review for publication as a journal article.

4.1. Introduction

Mobility in socio-economic status across generations, known as intergenerational mobility (IGM), describes how strongly individuals' position on the socioeconomic ladder is tied to their background, especially the socioeconomic status of their parents (Dearden, Machin, and Reed 1997). The degree of IGM indicates the evolution of living standards across generations and points to how children of more disadvantaged parents fare relative to their more advantaged peers. At an individual level, it reflects the extent to which individuals' chances of success depend on the circumstances in which they were born, such as income, race, gender, and place of birth (Chetty et al. 2014; Narayan et al. 2018). At an aggregate level, IGM signals the degree of equality of economic opportunity in a country.

IGM is particularly important for two reasons: equity and efficiency. Low IGM could be a cause as well as a consequence of overall inequality and has an adverse implication for economic development, social cohesion, and political stability (Narayan et al. 2018; Nybom 2018). By exacerbating inefficiencies in resource allocation, low IGM can also reinforce a vicious cycle of low investment and economic stagnation.[52] Therefore, analyzing IGM and designing policy interventions to improve mobility in the economy is crucial to ensure long-term economic growth and shared prosperity.

Regardless of their importance, however, empirical studies dealing with IGM in sub-Saharan African countries are very scarce[53]. In the Ethiopian context, Haile (2018) examined the intertemporal dynamics in IGM using the Labor Force Survey (LFS) of 2005 and 2013. However, since 2013, Ethiopia has gone through several macroeconomic cycles including (i) an extended period of strong GDP growth; (ii) a sustained increase in urbanization, with the service sector overtaking the agricultural sector as a dominant source of GDP (World Bank, 2015) and (iii) a severe drought in 2015. Pertinent to this study, the share of public spending on education has increased sharply in recent years, reaching 27 percent in 2015, about double the share in 2000.[54] In light of these developments, this study provides important new insights into the extent and drivers of IGM in the country.

In this paper, we focus on one of the main, yet largely neglected, forces driving the pattern of IGM: urbanization. Urbanization holds enormous potential to stimulate economic growth and poverty reduction by facilitating the sectoral reallocation of labor based on productivity differentials and increasing productivity within sectors (World Bank 2009). If managed well, the structural and spatial transformation that accompanies urbanization is essential not only to reduce the underemployment of labor but also to place countries on a long-term prosperous trajectory (Duranton 2015; Kamei and Nakamura 2020) potentially leading to upward economic mobility.

---

[52] In an imperfect capital market setting, an unequal distribution of initial wealth adversely affects the output and the overall development path of an economy by lowering aggregate investment (Galor and Zeira 1993) and negatively influencing the occupational choice of individuals (Banerjee and Newman 1993).

[53] The exceptions in this regard include Alesina et al. (2019); Lloyd and Blanc (1996); and Narayan et al. (2018) that reported IGM of several countries in the region in cross-country setting.

[54] For details, see https://data.worldbank.org/indicator/NY.ADJ.AEDU.GN.ZS?locations=ET

In line with this, several empirical studies indicate that urban areas are positively associated with IGM as they provide better access to public goods, employment opportunities, and information. Moreover, the diversity in economic activities and integration of markets in urban areas are likely to reduce discrimination and the associated inefficiencies in the allocation of talent and skills in the labor market (Chetty et al. 2014; Duranton 2015; Ewing et al. 2016; Narayan et al. 2018; World Bank 2009).

However, there is also a large body of equally compelling empirical evidence indicating that upward social mobility is relatively low in urban areas and even lower in densely populated locations within large urban areas (Autor 2005; Glaeser 2020). Large contemporary inequality in urban areas could limit the opportunities for children from poorer families, perpetuating inequality for future generations. For example, Glaeser (2020) argues that high productivity in urban areas leads to upward mobility only for those at the top of the income distribution. For those at the bottom of the distribution, particularly for the less skilled, urban areas tend to provide limited economic opportunities. With globalization, the return to skilled labor has significantly increased in urban areas further aggravating the disparity between skilled and unskilled workers. This has considerably stifled the prospects of mobility for the poor, especially in more dynamic large urban areas (Autor 2020; Glaeser 2020).

In this study, we explore the extent of intergenerational mobility in Ethiopia and its association with urbanization. First, using non-monetary measures of IGM based on educational and occupational status, we investigate the extent to which inequities in economic and social status are transmitted across generations, both in bivariate and multivariate analysis. Second, we assess whether and how strongly intergenerational mobility interacts with urbanization. In our analysis, we use two rounds of nationally representative household survey data, which we geo-spatially link with the Nighttime Light (NTL) dataset. Based on the intensity of NTL at the place of residence, the sample locations in the datasets are categorized into rural, small towns, and large urban areas.

The data analysis offers several insights. First, there is a strong parent-child correlation in educational and occupational status in Ethiopia. This strong correlation persists even after accounting for a rich set of individual, household, and location characteristics. Second, urbanization reinforces the association between children and parental education. While parental education is generally significantly correlated with the educational attainments of their children, the correlation is stronger in large urban areas than in rural areas or smaller towns. This is particularly the case for post-elementary education. Compared to a child with parents with no education, a child with parents with secondary or tertiary education is roughly twice more likely to attain tertiary education in large urban areas than in rural areas or small towns.

The examination of the interaction between urbanization and IGM in occupational status, similarly reveals several key insights. First, across both rural and urban areas, children whose parents are self-employed are more likely to work, and more likely to work in self-employing jobs than children of employees. More generally, children of self-employed parents are more likely to be employed in better-paying occupations. Second, compared to rural areas and small towns, the correlation between child and parental occupation is weaker in large urban areas. Third, the higher occupational mobility in large urban areas is explained mainly by differences in educational attainment. That is, once individual education level is accounted for, large urban areas offer better mobility in employment opportunities than rural areas or small towns. This suggests that broadening access to and reducing the dropout rates at post-elementary schools, and improving the quality of education, is one of the most effective mechanisms to reduce spatial and intergenerational inequality in welfare in Ethiopia and beyond.

The rest of the chapter is organized as follows. The next section describes the main data sources, definition of key variables, and presents the empirical approach. Sections 4.3 and 4.4 discuss the results from the econometric estimation and section 4.5 concludes the chapter.

4.2. Data, measurement, and methods

4.2.1. Data

Our analysis is based on the Ethiopian Living Standard Measurement Study-Integrated Surveys on Agriculture (LSMS-ISA) which we merge with satellite-based NTL data. The LSMS-ISA is a rich, geo-referenced, and nationally representative household and village level longitudinal dataset, collected jointly by the Central Statistical Agency (CSA) of Ethiopia and the World Bank, every two years, starting from 2012. For this study, we use the 2014 and 2016 rounds of the survey for which all the necessary information is available[55]. The questionnaires are comparable across waves and include household and Enumeration Area (EA)[56] level surveys. The household survey collected detailed individual information, inter alia, on demographic characteristics such as age, gender, educational attainment, and labor market participation. The EA (also called community) survey gathered information on the availability of and distance to public services, employment opportunities, market prices, etc.

The survey also collected two important sets of information. First, for each household member, it provides information on parental education and occupation status; we use this information to examine the extent of IGM. Second, it provides GPS coordinates of the sampled households[57]; we use this information to extract and merge the satellite-based NTL data[58] corresponding to the residential location of each household. Version 4 NTL time series dataset from the Defence Meteorological Satellite Program Operational Line Scanner (DMSP-OLS)[59] is used in this study.

The use of DMSP-OLS NTL datasets to demarcate urban areas involves two commonly cited shortcomings. The first is the lack of intra- and inter-calibration between different satellites from which NTL information is collected. The second is the presence of blooming, overglow, and the oversaturation of pixels (Savory et al. 2017; Zhang and Seto 2011, 2013). For the period between

[55] While the 2012 wave covers only rural areas and small towns, the 2018 round is a baseline for a new panel, not a follow-up to previous waves.

[56] Enumeration areas (EAs) are equivalent to a village, relatively small, and consisting of about 250 households on average.

[57] To be precise, the publicly available versions constituted a modified EA-level coordinates cloned from household level coordinates by applying a random offset of 0-10 km to preserve the confidentiality of sample household and communities (CSA and World Bank 2017).

[58] The National Oceanic and Atmospheric Administration (NOAA) and the National Geophysical Data Centre (NGDC) collaborate to generate the NTL datasets and make them freely available for public use.

[59] The latest version is the Visible Infrared Imager Radiometer Suite (VIIRS) Day/Night Band (DNB) from the National Polar- Orbiting Operational Environmental Satellite System (NPOESS). Although this version has the potential to mitigate the shortcomings of DMSP-OLS version (Zhang and Seto 2013), the dataset is not available to us.

2000 and 2013, Savory et al. (2017) had previously addressed both issues and made the time series data for Africa freely available[60]. The dataset proved to perform well as an indicator of urbanization based on its particularly strong correlations with population and infrastructure density (Savory et al. 2017). It has been used in previous studies that looked at the effect of urbanization on socio-economic development in Africa (Abay et al. 2020; Ameye 2018; Donaldson and Storeygard 2016; Henderson et al. 2009; Henderson, Storeygard, and Weil 2011; Michalopoulos and Papaioannou 2018).

4.2.2. Measurement of variables

A. Outcome variables: Intergenerational Mobility

Commonly, the level of IGM is studied in terms of either monetary indicators such as income, wage, and wealth, or non-monetary measures such as educational and occupational status. While the use of monetary indicators is preferable for its simplicity and interpretability, the LSMS-ISA survey did not collect information on parental income or wages. Besides, the use of these monetary indicators as measures of IGM tends to underestimate the influence of parental characteristics as the transitory variance of measured income might bias estimates (Black and Devereux 2010; Zimmerman 1992).

Therefore, we measure IGM in terms of educational and occupational status. These non-monetary indicators are advantageous as measures of economic mobility for three main reasons. First, unlike income or wealth which are either unavailable or noisy for a large share of the population in developing countries, education and occupation data are mostly available and reliable (Alesina et al. 2019; Porta and Shleifer 2008). Second, measurement error in educational and occupational status is less of a concern relative to monetary indicators (Black and Devereux 2010; Zimmerman 1992). Third, while strongly correlated with income and wealth, education and occupation reflect a broader account of mobility since they have been shown to strongly predict other proxies of well-being including child health and nutrition, aspiration, attitudes towards domestic violence, and proxies of political and civic engagement (Alesina et al. 2019; Haile 2018; Narayan et al. 2018).

For the purpose of our analysis, we group parents' and children's education into four main categories: (i) no schooling; (ii) primary education: grades 1 to 8; (iii) secondary education: grades 9 to 12; and (iv) tertiary education. We define parental education as levels attained by the father or the mother, whichever is the maximum. Similarly, we created four categories for parents' and children's occupation: (i) no or elementary occupation[61]; (ii) unskilled wage employment; (iii) self-employment; and (iv) skilled wage employment. Again, we define parental occupation as the occupation of the father or the mother, whichever is the highest along the occupational ladder. It is important to note, however, that the ordering of these occupational categories is somewhat arbitrary. In the case of wage employment, we differentiate between unskilled and professional wage employment based on educational level, hence the problem is attenuated. On the other

[60] The dataset is available at: https://geodata.globalhealthapp.net/. For the technical aspect of the satellites and the inter-calibration, please refer to Savory et al. (2017). This data is available only for 2000-2013 period. Therefore, for 2015 survey, the data is imputed based on a regression model on the past values, household assets levels and access to infrastructure and electricity. This is similar to poverty mapping in its approach (see Dang, Jolliffe, and Carletto 2019).

[61] This constitutes the unemployed, unpaid family labour, small-scale agriculture.

hand, this is not possible in the case of self-employment. Nevertheless, this classification approach is in line with studies that have examined mobility in education and occupation in other settings (Carmichael 2000; Haile 2018; Nguyen, Haile, and Taylor 2005).

B. Explanatory variables: Indicator of urban areas

As mentioned above, we use the NTL data to determine the urbanization status of EAs where the sampled households reside. The NTL dataset contains luminous pixels that are part of a given light cluster, and these are expressed as digital numbers between 0 (no light) and 63 (maximum light intensity). Depending on the degree of urbanization status of the EAs, the number and intensity of the luminous pixels around the EAs vary considerably. To identify the existence of and determine the size of urban areas, we generate and use a new variable, Sum of Light (SOL), which sums up the NTL within the 10km radius around EAs. Compared to the traditional census-based approach to urbanization measures, the SOL method commands several advantages. First, it allows for continuous assessment of urbanization. That is, rather than considering dichotomous urban and rural distinction, the SOL allows the examination of rural-urban spaces as a continuum. This facilitates a more disaggregated classification of urban areas which in turn enriches the analysis of patterns and effects of urbanization. This is particularly interesting in the SSA setting as it helps to examine the role of small- and intermediate- towns, which are mushrooming up all over the region (Satterthwaite and Tacoli 2003).

Second, the use of SOL eliminates the reliability issues surrounding the national administrative definition of urban areas. Administrative definitions lack comparability and lag behind reality, especially in developing countries. Besides often being subjective, they tend to reflect political and bureaucratic dispositions rather than services a given space provides (von Braun 2014b; Satterthwaite and Tacoli 2003; UNECA 2017). The use of the SOL mitigates these shortcomings as it is measured with consistent quality, and its availability over a long period of time allows reliable temporal analysis (Donaldson and Storeygard 2016; Savory et al. 2017).

Third, the SOL approach adds up the NTL from all agglomerations within the delineated buffer zone. This way, it can identify not only the existence of- but also - the size of urban areas within the buffer zone allowing us to account for the effect of all urban centers. This addresses one of the critical shortcomings of the traditional approaches where urban influence is measured with respect to only the nearest town. Due to this feature, the use of SOL as a measure of urbanization is gaining popularity in empirical research (Abay et al. 2020; Gibson et al. 2017; Henderson et al. 2017).

Finally, the SOL approach appears to be best suited to be used with the LSMS-ISA data. To ensure the confidentiality of sample households and communities, the GPS coordinates in the publicly available version of LSMS-ISA data were modified from their original levels by applying a random offset of up to 10km[62]. Therefore, the 10km buffer zone that is created to delineate urban areas eliminates any potential misclassification resulting from the random offsets.

[62] See https://microdata.worldbank.org/index.php/catalog/2783

4.2.3. Method of data analysis

We start the data analysis by first ranking the educational and occupational status of both children and their parents in a manner illustrated in the previous sub-section. We then analyze these rankings in two different ways. First, following the work of Checchi, Ichino, and Rustichini (1999), Chetty et al. (2014), and Nguyen et al. (2005), we generate transition matrices. These matrices indicate the proportion of children with an educational (occupational) status j whose parents have an educational (occupational) status,i, where P_{ij} represents the entries in the matrix. A comparison of the diagonal and off-diagonal elements of the transition matrices indicates the degree of IGM; the larger the diagonal elements, the lower the IGM. We also use the transitional matrices to generate more intuitive statistics on: (i) the percentage of children with an educational (occupational) status that is lower than their parents; (ii) the percentage of children with the same educational (occupational) status as their parents; and (iii) the percentage of children with an educational (occupational) status that is higher than their parents.

Second, we apply the ordered logit model to estimate the educational (occupational) status of children as a function of parental education (occupation) status. Relative to the analysis based on unconditional transition matrices, the advantage of this approach is that it allows controlling for individual, household, and location characteristics. In light of this, an ordered logit model of the educational (or occupational) status of individual i in family j, y_{ij}^{c}, as a function of parental education (or occupation), y_j^p is specified as follows:

$$y_{i,j}^{c} = \beta_0 + \beta_1 y_j^p + \beta_2 x_{i,j}^{c} + \beta_3 Z_j + \beta_4 U_j + \varepsilon_i^c \qquad (4.1)$$

where $x_{i,j}^{c}$ is a vector of individual-level characteristics that are expected to affect the educational (occupational) status of child i in family j. Based on the human capital literature, $x_{i,j}^{c}$ includes covariates such as age and gender. Children's characteristics are ascribed to influence children's achievements not only through their own preferences and choices but also by influencing the preference and choices of parents in their investment decisions (Becker and Tomes 1986). An important caveat is that we are unable to control for children's intellectual abilities which is likely to bias the estimate of both educational and occupational outcomes (Blackburn and Neumark 1993; Bronars and Oettinger 2006; Cameron and Heckman 1993). One way to address this issue would be by applying panel data methods with fixed effects, assuming that abilities are time-invariant (Wooldrigde 2013). An alternative would be to include test score measures (e.g. Ravens tests, Digit Span test, Stroop test, etc.) as a proxy for intellectual ability. Unfortunately, none of these methods is feasible: too little variation in parental education (occupation) over the survey rounds excludes the fixed effects estimator, and the unavailability of data prevents us from including proxies of abilities. Nevertheless, this is unlikely to pose a serious threat to our estimation since the observed levels of education may absorb most of the variations in measured ability (Cawley, Heckman, and Vytlacil 2001; Zax and Rees 2002).

Similarly, Z_j represents family, and U_j location characteristics. We control for family-related variables including household size, the age and gender of the household head, and a wealth indicator. Finally, zonal fixed effects are included in all the estimations since observed and unobserved location characteristics such as agro-ecology, social capital ties, and locational amenities can influence an individual's achievements regardless of parental characteristics (Becker and Tomes 1986).

In equation (4.1), the main parameter of interest is β_1. It measures the correlation between an individual's educational (or occupational) status and that of his or her parents, hence the degree of intergenerational linkage[63]. The hypothesis is that β_1 is positive: there is an intergenerational correlation between the educational and occupational status.

Our second main objective is to assess whether and how strongly small and large urban areas interplay with intergenerational mobility as compared to rural areas. To that end, we examine the degree of intergenerational mobility based on the urbanization status of the place of residence. Our focus is to test whether and how strongly urban areas foster social mobility relative to rural areas. In this regard, we conduct separate estimations for sub-sample of individuals located in rural areas, small towns, and large towns. An alternative would be to add interaction terms to equation (1) and estimate it for the full sample. However, we assume that not only the slope of the parental effect is different for different categories of rural-urban space, but also the intercept and the remaining parameters vary as well. Hence, we opt for estimation based on sub-samples of individuals. The estimation of model parameters occurs in a probabilistic framework using the maximum likelihood (ML) method.

4.3. Results and discussion

4.3.1. Intergenerational mobility in educational and occupational status

Table 4.1 presents a summary of unconditional probabilities generated from the transition matrix. The detailed transition matrix is available in Table A4.2 in the Appendix. Panel A shows that intergenerational mobility in educational status in Ethiopia is considerably low. Close to 60 percent of children attain the same educational status as their parents, indicating that the degree of intergenerational persistence is high, even when compared to other developing countries[64]. About 28.4 percent of children attain higher educational status compared to their parents, and about half of that, 14.2 percent, attain lower educational status. Table 4.1 also reveals a considerable disparity between daughters and sons. While sons appear to have a similar level of general mobility in educational status as daughters, they show significantly higher upward mobility. Sons are 6.5 percentage points more likely than daughters to achieve higher levels of education and 11.3 percentage points less likely than daughters to achieve lower levels of education than their parents.

Similarly, Panel B of Table 4.1, shows that intergenerational mobility in occupational status is even lower. Approximately 70 percent of children work in the same job category as their parents. On the other hand, although about 20 percent of children attain better occupational status than their parents, 9.7 percent of the children have a worse occupational status. Between sons and

[63] See Carmichael (2000); Checchi (1997); Nguyen et al. (2005); and Zimmerman (1992) for a similar approach.

[64] The corresponding number is 54%, 36%, 29%, 16% for Sudan, Uganda, Egypt and South Africa, respectively (Alesina et al. 2019)

daughters, sons are more likely to attain better occupational status than their parents and are more likely to land a better job (6.4 percentage points) than daughters.

Table 4.1. Mobility in educational and occupational status

	All children	Daughters	Sons
Panel A: Mobility in educational status			
% in lower level than parents [a)]	14.2	19.7	8.4
% in the same level as parents	57.4	55.1	59.9
% in higher level than parents	28.4	25.2	31.7
Panel B: Mobility in occupational status [b)]			
% in lower level than parents	9.7	9.8	9.5
% in the same level as parents	70.6	73.4	67.4
% in higher level than parents	19.7	16.8	23.1

***Source**: Author's calculation based on LSMS-ISA (2014 & 2016)*
***Notes**: [a)] The first row is the percentage of children in a lower educational attainment group than their parents, and similarly for the second and third rows. [b)] The first row is the percentage of children in a lower occupational group than their parents, and similarly for the second and third rows.*

While the descriptive results presented in Table 4.1 are informative, the observed child-parent correlations could also be attributable to several other factors related to children, households, and location characteristics. To tease out the association between child and parental characteristics requires employing a multivariate regression model as in Equation (4.1). Table 4.2 presents the marginal effects from such estimation[65]. The underlying estimation coefficients are reported in Table A3 in the Appendix.[66] The results reveal a strong positive correlation between parents and children's educational levels (Panel A). Children of better-educated parents are more likely to attain better educational status than children of less-educated parents. For instance, figures in column 4 indicate that compared to parents with no education, a child from parents with tertiary education is 54 percent less likely to be uneducated. On the other hand, a child from parents with tertiary education is 34 percent more likely to attain tertiary education than a child from parents without any formal education. Figure 4.1 summarizes these findings. It shows that increased parental education status increases children's chances of attaining higher educational status.

[65] Note: while the full model with full-fledged covariates is estimated, only variables related to urban size are shown here for brevity.

[66] Note that the estimated coefficients in ordered logit model, as in Table A3 and A4, are not directly interpretable. This is because the coefficient estimates of the ordered logit model provide marginal effects on the latent scale, where the true metric is unknown (Wooldridge 2002).

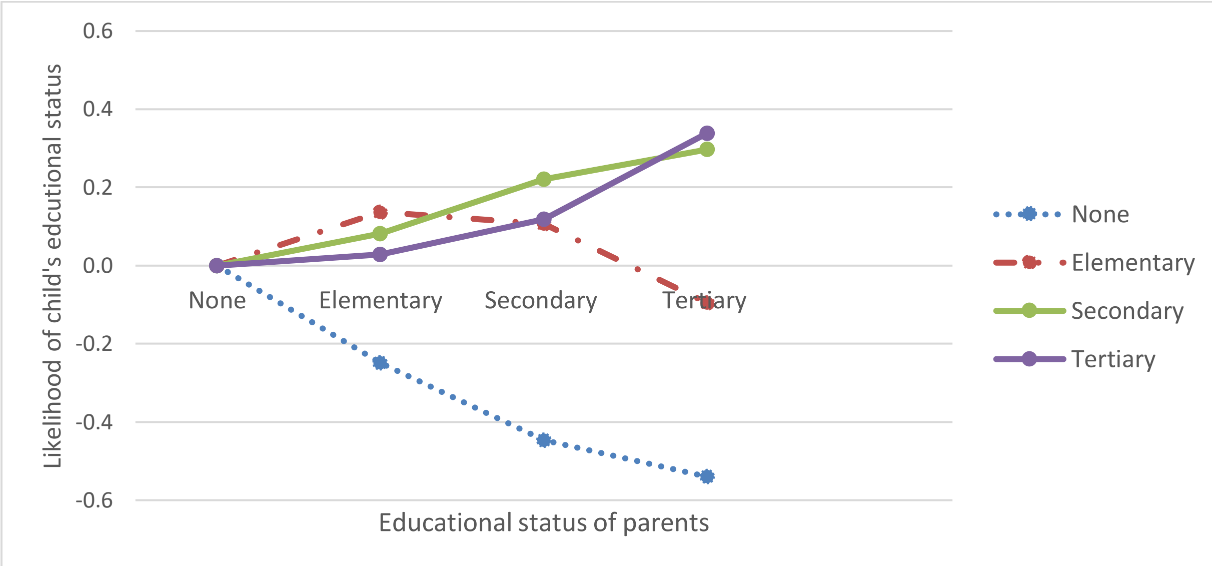


Figure 4.1: Association between child and parental educational status
Source*: Author's computation based on LSMS-ISA (2014 & 2016)*
Note*: The figure shows the pattern in the likelihood of children's educational status, given the educational status of parents. Definition of educational categories: Elementary education (1-8 grade); secondary (9-12 grade) and tertiary (above 12 grade).*

Panels B and C present disaggregated results for daughters and sons, respectively. The results underlie that the strong positive correlation between child and parent education holds regardless of the gender of the child. Moreover, a comparison of Panel B and Panel C suggests that the intergenerational correlation in educational status is stronger for sons than for daughters. For instance, from Panel C, sons from parents with tertiary education are 59 percent more likely to attain tertiary education than those from parents without any formal education. This is 41 percentage points higher than the corresponding figure for daughters (see Panel B).

Table 4.2. Mobility in educational status, marginal effects

Child\Parent Education	(1) No education	(2) Elementary education	(3) Secondary education	(4) Tertiary education
Panel A: All children				
No education	[Reference]	-0.248***	-0.447***	-0.540***
		(0.008)	(0.009)	(0.006)
Elementary education	[Reference]	0.137***	0.107***	-0.0957***
		(0.005)	(0.007)	(0.015)
Secondary education	[Reference]	0.0825***	0.221***	0.297***
		(0.003)	(0.009)	(0.007)
Tertiary education	[Reference]	0.0287***	0.119***	0.339***
		(0.002)	(0.007)	(0.017)
Observations		35,885	35,885	35,885
Panel B: Daughters				
No education	[Reference]	-0.182***	-0.390***	-0.499***
		(0.010)	(0.015)	(0.013)
Elementary education	[Reference]	0.107***	0.152***	0.0957***
		(0.006)	(0.005)	(0.012)
Secondary education	[Reference]	0.0520***	0.148***	0.218***
		(0.004)	(0.009)	(0.011)
Tertiary education	[Reference]	0.0228***	0.0896***	0.185***
		(0.002)	(0.007)	(0.014)
Observations		18,587	18,587	18,587
Panel C: Sons				
No education	[Reference]	-0.308***	-0.468***	-0.514***
		(0.010)	(0.009)	(0.006)
Elementary education	[Reference]	0.151***	-0.0201	-0.338***
		(0.006)	(0.016)	(0.012)
Secondary education	[Reference]	0.128***	0.339***	0.258***
		(0.006)	(0.016)	(0.019)
Tertiary education	[Reference]	0.0293***	0.149***	0.595***
		(0.002)	(0.011)	(0.031)
Observations		17,298	17,298	17,298

Source: *Author's calculation based on LSMS-ISA (2014 & 2016)*

Note*: Standard errors clustered at the household level in parentheses. Statistical significance indicated by: *** p<0.01, ** p<0.05, * p<0.1*

Table 4.3 presents the marginal effects from the ordered logit model of IGM in occupational status. The underlying estimation coefficients are reported in Table A4.4 in the Appendix. Panel A indicates that, compared to parents with elementary occupation, children from self-employed parents are more likely to attain better occupational status. On the other hand, once individual and household characteristics are accounted for, we do not observe a statistically significant correlation between child-parent occupations for wage employment. This result is not surprising given the fact that self-employed parents are more likely to bestow skills of entrepreneurship to their children. This is also consistent with the fact that the employment generation capacity of both private firms and the public sector is very limited in Ethiopia, and the majority of jobs are created by small-scale enterprises that are mainly run by family members (Broussar and Tekleselassie 2012; OECD/PSI 2020).

Table 4.3. Mobility in occupational status, marginal effects

Child\Parent Occupation	(1) Elementary occupation	(2) Unskilled wage	(3) Self-Employment	(4) Skilled Wage
Panel A: All children				
Elementary occupation	[Reference]	-0.023	-0.051***	-0.044
		(0.01)	(0.01)	(0.03)
Unskilled wage	[Reference]	0.008*	0.017***	0.015
		(0.01)	(0.00)	(0.01)
Self-Employment	[Reference]	0.009	0.021***	0.018
		(0.01)	(0.00)	(0.01)
Skilled Wage	[Reference]	0.006	0.013***	0.011
		(0.00)	(0.00)	(0.01)
Observations		28,491	28,491	28,491
Panel B: Daughters				
Elementary occupation	[Reference]	-0.02	-0.052***	-0.022
		(0.02)	(0.01)	(0.04)
Unskilled wage	[Reference]	0.007	0.016***	0.007
		(0.01)	(0.00)	(0.01)
Self-Employment	[Reference]	0.01	0.025***	0.011
		(0.01)	(0.01)	(0.02)
Skilled Wage	[Reference]	0.004	0.011***	0.004
		(0.00)	(0.00)	(0.01)
Observations		15,129	15,129	15,129
Panel C: Sons				
Elementary occupation	[Reference]	-0.024	-0.051***	-0.068
		(0.02)	(0.01)	(0.05)
Unskilled wage	[Reference]	0.009	0.019***	0.025
		(0.01)	(0.01)	(0.02)
Self-Employment	[Reference]	0.008	0.017***	0.023
		(0.01)	(0.01)	(0.02)
Skilled Wage	[Reference]	0.007	0.015***	0.021
		(0.01)	(0.00)	(0.02)
Observations		13,362	13,362	13,362

***Source**: Author's calculation based on LSMS-ISA (2014 & 2016)*

***Note**: Standard errors clustered at the household level in parentheses. Statistical significance indicated by: *** p<0.01, ** p<0.05, * p<0.1*

Panels B and C present disaggregated results for daughters and sons, respectively. In line with the overall result, these results indicate that, regardless of gender, children from self-employed parents are more likely to attain better occupational status. Furthermore, the results show that children whose parents are self-employed are more likely to be self-employed than children of wage-employed parents.

These coefficient estimates in Table A4.4, even though not directly interpretable, also provide interesting insights. Notably, the coefficients of parental occupation remain robust to the inclusion of additional covariates including age, gender, household size, village and zonal characteristics in columns 1, 2, and 3. However in column 4, once the wealth indicator is included, the magnitude of the parameter estimates drops drastically indicating that occupational mobility is relatively lower among the poor. This suggests that the poor remain employed in low-paying jobs over successive generations in line with the theoretical prediction of the human capital theory (Becker and Tomes 1986). This is a classical representation of a poverty trap wherein parental deprivation passes onto the next generation through inadequate schooling and poorly remunerating occupation[67].

[67] Table A6 shows that poor households invest significantly less on education of their children, both in absolute terms and relative to the total household expenditure.

In column 5 of Table A4.4, when children's own education levels are controlled for, the parameter estimates related to parental occupation decline even further and the coefficient of skilled wage becomes statistically insignificant. This suggests that parents employed in better-paying occupations enhance employment opportunities for their children via investment in their education. Indeed, Table A4.9 in the Appendix shows that there is a strong positive correlation between the quality of parental employment and investment into children's education – both in absolute terms and relative to total household expenditure. Lastly, the survey round dummy for 2014 is statistically insignificant. Together with the positive result in Table A4.3 for education, this indicates that while there was an improvement in educational status between 2014 and 2016, this did not translate into better occupational status.

4.3.2. Intergenerational mobility and urbanization

In this section, we explore whether the extent of educational and occupational mobility observed in the previous section interacts with the degree of urbanization of the place of residence. As before, we first report the unconditional probabilities generated from the transition matrix disaggregated into rural areas, small towns, and large towns based on terciles of SOL.

Table 4.4 reveals that there is considerable variation in the degree of educational and occupational mobility across rural-urban areas. Panel A suggests that rural areas and small urban areas offer better upward educational mobility than large urban areas. This might partly be attributable to the nascent expansion of first-cycle education in rural areas across the country.[68] This is also partly attributable to the fact that average educational levels in urban areas are higher, rendering a relatively lower scope for upward mobility (see Table A4.1 in the Appendix). Panel B stands in stark contrast to the pattern in Panel A. It shows that occupational mobility increases consistently from rural to urban areas. While the percentage of children in the same occupational position as their parents is 80 percent in rural areas, the corresponding figure is only 59 percent in large urban areas.

Overall, these results point to two potentially important aspects. First, while there are opportunities for upward educational mobility in rural areas and small towns, these have so far not translated into occupational mobility to the same extent. This suggests that there might be limited labour demand in rural areas and small towns. In large towns, the extent of upward mobility in educational and occupational status are comparable. Second, while large urban areas seem to offer better opportunities for upward mobility, the risk of downward mobility is also higher, which has important implications in terms of rising inequality.

[68] Between 1996 and 2015, the enrolment rate in elementary education in the country rose from only 29% to 86% and the number of elementary schools increased from 11,000 to 32,048. Due to their historical disadvantage, the majority of these changes occurred in rural areas (MoE 2015).

Table 4.4 Mobility in educational and occupational status by urbanization status

Urbanization status [a)]	Total	Rural	Small towns	Large towns
Panel A: Mobility in educational status [b)]				
% in lower level than parents	14.2	12.9	12.6	16.8
% in the same level as parents	57.4	56.9	59.3	57.3
% in higher level than parents	28.4	**30.2**	28.1	25.9
Panel B: Mobility in occupational status [c)]				
% in lower level than parents	8.4	4.3	6.2	14.5
% in the same level as parents	**71.7**	**80.1**	**78.0**	**58.7**
% in higher level than parents	19.9	15.7	15.9	26.8

***Source**: Author's calculation based on LSMS-ISA (2014 & 2016)*

***Notes**: [a)] Sum of the NTL at EA level is used to classify the households from rural (tercile with the smallest light intensity) to large towns (tercile with the highest light intensity). [b)] The first row is the percentage of children in a lower educational attainment group than their parents, and similarly for the second and third rows. [C)] The first row is the percentage of children in a lower occupational group than their parents, and similarly for the second and third rows.*

Next, we re-estimate the basic multivariate model in equation 4.1 by disaggregating the sample households into rural, small towns, and large towns. Table 4.5 and Table 4.6 report extracted marginal effects from the ordered logit model for educational and occupational status, respectively. The results presented in Table 4.5 suggest that regardless of the location of residence, there is considerable persistence in educational status across generations. Compared to parents with no formal education, children from parents with some formal education are significantly more likely to attain higher educational status. However, inequality is much more pervasive in urban areas, particularly in large urban areas, than in rural areas. For instance, in rural areas, a child of parents with tertiary education is 25.4 percent more likely to attain tertiary education than a child of parents with no formal education. The corresponding figure in small and large urban areas is 36.8 percent and 41.2 percent, respectively. Similarly, a child of parents with secondary education is more than twice more likely to attain tertiary education in large urban areas than in rural or small towns. This suggests that urban areas tend to exacerbate rather than abate inequality in access to education. Since wellbeing and educational status are highly correlated, this leads to the further burgeoning of rural-urban as well as intra-urban economic inequality. Unequal schooling attainment and increasing return to the level of education and skills in large urban areas are two key components of income inequality in developing countries (Autor 2020; Glaeser 2020; Binder and Woodruff 2002).

Table 4.5. Mobility in educational status, marginal effects, by urbanization status

Child\Parent Education	(1) No education	(2) Elementary education	(3) Secondary education	(4) Tertiary education
Panel A: Rural Areas [a)]				
No education	[Reference]	-0.240***	-0.444***	-0.559***
		(0.012)	(0.019)	(0.010)
Elementary education	[Reference]	0.155***	0.175***	-0.020
		(0.007)	(0.010)	(0.034)
Secondary education	[Reference]	0.067***	0.194***	0.325***
		(0.005)	(0.017)	(0.016)
Tertiary education	[Reference]	0.018***	0.075***	0.254***
		(0.002)	(0.010)	(0.033)
Observations		17,351	17,351	17,351
Panel B: Small Towns				
No education	[Reference]	-0.252***	-0.496***	-0.596***
		(0.018)	(0.027)	(0.013)
Elementary education	[Reference]	0.162***	0.151***	-0.111**
		(0.011)	(0.024)	(0.049)
Secondary education	[Reference]	0.073***	0.245***	0.339***
		(0.007)	(0.030)	(0.021)
Tertiary education	[Reference]	0.018***	0.099***	0.368***
		(0.003)	(0.020)	(0.064)
Observations		5,770	5,770	5,770
Panel B: Large Towns				
No education	[Reference]	-0.246***	-0.411***	-0.490***
		(0.014)	(0.014)	(0.011)
Elementary education	[Reference]	0.096***	0.022**	-0.175***
		(0.008)	(0.009)	(0.016)
Secondary education	[Reference]	0.104***	0.228***	0.254***
		(0.006)	(0.011)	(0.009)
Tertiary education	[Reference]	0.045***	0.161***	0.412***
		(0.003)	(0.009)	(0.021)
Observations		12,764	12,764	12,764

Source*: Author's calculation based on LSMS-ISA (2014 & 2016)*

Note*: Standard errors clustered at the household level in parentheses. Statistical significance indicated by: *** p<0.01, ** p<0.05, * p<0.1;* [a)] *The sum of the NTL at the EA level is used to classify the households from rural (tercile with the smallest light intensity) to large towns (tercile represents the highest light intensity).*

The multivariate ordered logit regression result for occupation across rural-urban areas is presented in Table A4.5 (coefficients) and Table 4.6 (marginal effects). Table A4.5 shows that, for large urban areas, once educational attainment is accounted for, parental occupation ceases to be significantly associated with individual occupation levels. For rural and small-town sub-samples, only parental occupation in self-employment is associated significantly with individual occupations. This observation indicates that the strong child-parent correlation in occupation, shown in Table 4.4, is being mediated by children's education[69]. That is, parents employed in better-paying occupations enhance employment opportunities for their children via investment in education of their children. Indeed, Table A4.6 in the appendix shows that there is a strong positive correlation between the quality of parental employment and investment into children's education – both in absolute terms and relative to total household expenditure.

[69] Note that individuals' education levels, especially post-secondary level education, appear statistically significant across all levels of urbanization (Table A4.5).

Table 4.6 representing the marginal effect extracted from the basic model, further corroborates the results reported above. Once individual education level is accounted for, large urban areas offer better mobility in occupational status, as compared to rural areas and small towns. This result has huge policy implications. It suggests that policy interventions that effectively address inequality in access to educational opportunities in urban areas might help to address both the inequality in welfare and inefficiency in the labor market. Since the level of education is an important determinant of occupational status (see Table A4.5) and productivity (Barro 2001; Becker 1994), raising the average schooling of disadvantaged individuals and backward regions should indeed reduce inequalities in welfare and inefficiency in the labor market. For similar findings elsewhere, see (World Bank, 2009; World Bank, 2011).

Table 4.6. Mobility in occupational status, marginal effects, by urbanization status

Child\Parent Occupation	(1) Elementary occupation	(2) Unskilled wage	(3) Self-Employment	(4) Skilled Wage
Panel A: Rural Areas [a]				
Elementary occupation	[Reference]	-0.080**	-0.101***	-0.027
		(0.040)	(0.023)	(0.051)
Unskilled wage	[Reference]	0.034**	0.042***	0.012
		(0.016)	(0.009)	(0.022)
Self-Employment	[Reference]	0.035*	0.045***	0.012
		(0.018)	(0.011)	(0.022)
Skilled Wage	[Reference]	0.010*	0.013***	0.003
		(0.006)	(0.003)	(0.006)
Observations		13,217	13,217	13,217
Panel B: Small Towns				
Elementary occupation	[Reference]	0.049	-0.175***	0.035
		(0.041)	(0.032)	(0.098)
Unskilled wage	[Reference]	-0.019	0.053***	-0.013
		(0.016)	(0.009)	(0.039)
Self-Employment	[Reference]	-0.024	0.093***	-0.017
		(0.020)	(0.018)	(0.048)
Skilled Wage	[Reference]	-0.006	0.029***	-0.004
		(0.005)	(0.007)	(0.011)
Observations		4,585	4,585	4,585
Panel B: Large Towns				
Elementary occupation	[Reference]	-0.085***	-0.071***	**-0.152*** **
		(0.020)	(0.013)	(0.047)
Unskilled wage	[Reference]	0.023***	0.020***	0.036***
		(0.005)	(0.004)	(0.008)
Self-Employment	[Reference]	0.031***	0.026***	0.055***
		(0.007)	(0.005)	(0.017)
Skilled Wage	[Reference]	0.031***	0.025***	**0.061*** **
		(0.008)	(0.005)	(0.023)
Observations		10,689	10,689	10,689

Source*: Author's calculation based on LSMS-ISA (2014 & 2016)*

Note*: Standard errors clustered at the household level in parentheses. Statistical significance indicated by: *** p<0.01, ** p<0.05, * p<0.1;* [a] *The sum of the NTL at the EA level is used to classify the households from rural (tercile with the smallest light intensity) to large towns (tercile represents the highest light intensity).*

4.4. Extensions

In this section, we present several extensions of our baseline findings. These extensions are certainly not exhaustive; nevertheless, they provide additional insights into the patterns and drivers of intergenerational mobility. More specifically, we first check to which extent intergenerational mobility interacts with physical mobility; second, we analyze the direction of IGM and focus on upward mobility in particular because of its welfare-enhancing implications.

4.4.1. The linkage between physical and intergenerational mobility

Geography is an important determinant of economic development. The existence of considerable spatial distribution of economic development has led to geographic poverty traps, whereby geographic characteristics of the residence alone can lock people into poverty (Kraay and McKenzie 2014; Ravallion and Wodon 1999). One mechanism to escape such a poverty conundrum is physical mobility.

A large number of authors have looked into the effect of physical mobility on several welfare indicators including income, health and nutrition, education, and labor market outcomes (Beegle, De Weerdt, and Dercon 2011; Lagakos, Mobarak, and Waugh 2018; Nakamura, Sigurdsson, and Steinsson 2016). Most of these studies identified a substantially strong positive effect of physical mobility on intra-generational economic mobility. Perhaps what has been less studied is the evidence of association between physical mobility and intergenerational mobility, particularly in developing country settings. In this sub-section, we examine the difference in intergenerational mobility in educational and occupational status between migrant and non-migrant household members in Ethiopia. A simple comparison of unconditional probabilities in the mobility of migrant and non-migrant members in Table 4.7 shows that physical and economic mobility appear to go hand-in-hand. While 74 percent of non-migrant members in our sample participated in similar occupations as their parents (see Table 4.1), the share was only 46 percent among migrant members. Similarly, while the share of non-migrant members that had the same level of education in the total sample was 59.7 percent, the corresponding percentage was only 45.6 percent for migrant members. It is, however, important to note that while mobility is generally higher for migrants than non-migrant members, this does not necessarily imply upward mobility. Compared to non-migrant members, the risk of downward occupational mobility is higher for migrant members. This result is robust to the inclusion of individual, household, and location characteristics (see Table A4.7, and A4.8 in the Appendix).

Table 4.7. Mobility in educational and occupational status for migrant members

Urbanization status [a)]	Total	Rural	Small towns	Large towns
Panel A: Mobility in educational status [b)]				
% in lower level than parents	18.9	17.5	14.3	23.9
% in the same level as parents	45.6	46.3	48.9	42.7
% in higher level than parents	35.5	36.3	36.8	33.4
Panel B: Mobility in occupational status[c)]				
% in lower level than parents	26.5	20.1	23.7	37.7
% in the same level as parents	45.6	50.0	47.5	38.0
% in higher level than parents	27.9	30.0	28.8	24.3

***Source**: Author's calculation based on LSMS-ISA (2014 & 2016)*

***Notes**: [a)] Sum of the NTL at EA level is used to classify the households from rural (tercile with the smallest light intensity) to large towns (tercile represents the highest light intensity). [b)]The first row is the percentage of children in a lower educational attainment group than their parents, and similarly for the second and third rows. [C)]The first row is the percentage of children in a lower occupational group than their parents, and similarly for the second and third rows.*

4.4.2. From general mobility to the direction of mobility

From a policy perspective, analysis of the covariates of upward/downward mobility might be more informative than general mobility. In this sub-section, we examine the main correlates of upward (and downward) mobility in educational and occupational status. We particularly focus on the effect of places of residence – rural, small town, and large town – on the likelihood of economic mobility. To this end, we specify a logit model where the outcome variable – a binary indicator representing upward (or downward) mobility – is regressed on the individual, household, and community-level characteristics.

For this part, the definition of mobility in educational status has been modified. In the previous sections, upward (downward) mobility in educational status was defined as ascents (descents) in the educational status of a child over the educational status of his parents at any level. According to this definition, two children, one with elementary and the other with tertiary level education, are considered to have achieved (similarly) upward mobility if they were from parents with no formal education. In the context of Ethiopia, however, two major issues stand out with this definition. First, the difference in returns to tertiary and elementary education is disproportionately larger than the differences in the number of years of schooling between the two. As is the case for most developing countries, the return to one additional year of schooling at the tertiary level is significantly larger than that at the elementary level (Binder and Woodruff 2002; Psacharopoulos and Patrinos 2018). Hence, the use of a simple ascent in education level (i.e., a simple comparison of child-parent education level) might bias upward mobility.

Second, the interaction between urbanization and educational mobility is non-linear. Due to the recent expansion of elementary education throughout the rural areas of Ethiopia, upward mobility is higher at a lower educational level in rural areas than in urban areas. However, the completion rate of students is disproportionately lower in rural than in urban areas. In rural areas, significantly large shares of students drop out earlier due mainly to domestic obligations or the lack of financial resources (see Table A4.9 in the Appendix). In 2014 and 2016, for instance, 43 percent of students in rural areas dropped out before completing elementary education. This is 11 and 17 percentage points more than the level in small and large urban areas, respectively (Table A4.9). This implies that the use of a simple ascent in educational status overstates (understates) mobility in rural (urban) areas[70].

To partially account for these two issues, we define upward mobility in education as the likelihood of a child of parents with less than tertiary education to attain a tertiary level education. On the other hand, downward mobility is defined as the likelihood that a child of parents with tertiary education does not attain tertiary education. The first two columns of Table 8 report the marginal effects extracted from the estimation of a logit model with these outcome variables. The last two columns report similar results, albeit for occupational status.

From Table 4.8, several points stand out. First, large urban areas enhance upward mobility in both educational and occupational status. This is consistent throughout the chapter and is in line with similar studies conducted elsewhere (Chetty et al. 2014). Second, compared to rural areas, the risk of downward mobility in occupational status is also more likely in urban areas (both small and large towns). Third, mobility is gender-specific. Compared to daughters, sons are more likely to achieve better educational and occupational status. For similar results, see Haile (2018) and Nguyen et al. (2005). Fourth, occupational mobility increases with children's education level. Fifth, household wealth is associated positively with both upward and downward mobility. While the positive effect of household wealth on upward education and occupation mobility could be explained by parental investment in education and networking, the positive effect on downward mobility is counter-intuitive; but could potentially be explained by the role of wealth on the propensity of risk-taking.

[70] More broadly, this suggests that the decline in inequality in the number of schooling years is not a sufficient condition for reducing income inequality. The stage at which the change occurs might equally be important.

Table 4.8. Covariates of upward and downward mobility, marginal effects

	Education		Occupation	
VARIABLES	Upward	Downward	Upward	Downward
Place of residence(reference=rural)				
Small Town	0.006	-0.008	0.007	0.034***
	(0.004)	(0.006)	(0.009)	(0.009)
Large Town	0.015***	-0.001	0.030***	0.044***
	(0.003)	(0.005)	(0.008)	(0.007)
ln(Age in years)	-0.011***	-0.016***	0.086***	-0.102***
	(0.001)	(0.003)	(0.006)	(0.004)
Male, yes=1	0.006***	-0.019***	0.046***	-0.002
	(0.002)	(0.002)	(0.005)	(0.004)
Education (reference=No education				
Primary			0.117***	0.053***
			(0.006)	(0.005)
Secondary or higher			0.168***	0.081***
			(0.008)	(0.007)
Household size, number	-0.000	-0.004***	-0.014***	-0.007***
	(0.000)	(0.001)	(0.001)	(0.001)
ln(Age of household head in years)	0.022***	-0.032***	-0.129***	0.053***
	(0.004)	(0.005)	(0.009)	(0.008)
Head is male, yes=1	-0.014***	0.010***	-0.001	-0.036***
	(0.003)	(0.003)	(0.007)	(0.006)
Durable assets owned, PCA	0.005***	0.008***	0.012***	0.014***
	(0.000)	(0.001)	(0.001)	(0.001)
ln(Village elevation, m)	0.006	-0.005	-0.022	-0.014
	(0.008)	(0.009)	(0.020)	(0.017)
ln(Annual Temperature, degrees)	0.001	-0.027	-0.013	0.034
	(0.016)	(0.019)	(0.040)	(0.035)
Survey round, 2014	-0.003**	-0.011***	0.001	-0.038***
	(0.001)	(0.002)	(0.004)	(0.004)
Location Fixed Effect?	Yes	Yes	Yes	Yes
Observations	35,885	35,885	28,436	28,436

Source: *Author's calculation based on LSMS-ISA (2014 & 2016)*

Note: *Standard errors clustered at the household level in parentheses. Statistical significance indicated by; *** p<0.01, ** p<0.05, * p<0.1*

4.5. Discussion and conclusion

4.5.1. Discussion

This chapter presents evidence of a substantial dependence of an individual's performance on parental characteristics for Ethiopia. This is in line with studies that indicate that mobility is alarmingly low in sub-Saharan African (SSA) countries. According to a World Bank study, only about 12% of adults in SSA countries have a higher education than their parents. Even when compared to other developing countries, this rate is significantly lower. For example, in parts of East Asia and the Pacific region, over a comparable time period, the share of children with higher education levels than their parents was more than 80 percent (Narayan et al. 2018). For Ethiopia, the statistics are even grimmer. Among the 27 countries they studied, Alesina et al. (2019) found that Ethiopia ranked 23rd in terms of upward mobility in the education of young adults aged between 14 and 18.

This is worrying since higher intergenerational mobility is a desirable economic outcome. An economy with higher mobility promotes a more efficient allocation of resources as educational attainment and occupational placements are based on merit rather than socioeconomic or parental characteristics. The resulting gain in efficiency has the potential to create a virtuous cycle of higher productivity, economic growth, and higher upward mobility in economic status (Binder and Woodruff 2002; Narayan et al. 2018). Therefore, policies that aim at improving socio-economic mobility do not only generate a more equitable distribution of opportunities but also promote long-run growth by mitigating the misallocation of human and financial capital.

Essentially, the analysis in this chapter indicated that a lack of economic mobility results from a combination of differences in parental and location (spatial) characteristics. Therefore, to promote economic mobility, policy interventions could play a proactive role to indemnify disadvantaged individuals for inequalities in parental characteristics and in leveling off the playing field. Table 4.9 summarizes a variety of policy options for achieving this goal.

However, the final policy selection and implementation must be tailored to the needs of the country in question. This is to ensure that the proposed policies are consistent with the country's overall strategy, that they align with the country's existing socioeconomic and macroeconomic policy positions, that the country can implement them, and that a cost-benefit analysis is performed. It is important to note that these suggested policy interventions are neither ranked nor exclusive. They are rather compiled from the literature based on the findings in this chapter.

Table 4.9. Potentially mobility enhancing policies

Sr. No.	Needs assessment (Source of inequality)	Type of policy	Specific policy options	Rationale	Expected outcome
1	Gap in maternal health	Health and Nutrition	Food supplementation programs, nutrition information, prenatal services	Reducing gaps among women of childbearing age can have a positive effect on infant health (Hoddinott et al. 2008)	Equalizing opportunities among women of childbearing age
2	Gap in child health	Health and Nutrition	Food supplementation programs, nutrition information, postnatal services	Nutritional and health improvements in early childhood can yield long-term benefits in education outcomes and wages (Hoddinott et al. 2008)	Equalizing health opportunities among children at an early age
3	Gap in access to and quality of preschool programs	Fiscal policy and education reform	Expand preschool childcare coverage and implement reforms to target specific cognitive and socioemotional development	Cognitive and non-cognitive skills acquired at preschool has a significant effect on education and labour market performance (Heckman 2006)	Equalizing education opportunities among children at an early age
4	Gap in public investment on vital infrastructure	Fiscal policy	Expand access to and quality of public goods targeting disadvantaged locations[71]. (schools, health centres, roads, and subsidized housing)	Lower inequality in access to and quality of public services improves mobility and long-term sustainable growth (Narayan et al. 2018)	Improve the accessibility of public goods in disadvantaged locations
5	Low quality of public services	Public sector reform	Implement rigorous reforms to improve the quality of public services. (education system, extension system)[72]	As in 4 above	Improve the quality of public services accessible in disadvantaged locations
6	Gap in access to and drop out of schools	Fiscal policy (public safety net)	Expand access to and quality of public schools targeting disadvantaged students[73]. E.g. School meal program; secondary school scholarship	As in 4 above	Improve the accessibility of education among disadvantaged students
7	Gap in aspirations window[74]	Cross-cutting	Improve exposure of children and their parents to information,	High aspirations improve educational & labour market	Widen the aspiration window of

[71] Since primary school enrollments in Ethiopia has approach 100 percent, the required expansion in schools might be mainly post elementary school level.

[72] A range of required interventions to improve the quality of learning and reduce inequalities in learning outcomes in developing countries is detailed in World Bank (2018). For spatial distribution of the public services in Ethiopia and the necessary reforms to improve quality service delivery in rural Ethiopia, see Abate et al. (2019)

[73] This could prove to be very effective in Ethiopia as dropout rates at all levels of schools are large (see Table A4.8)

[74] Aspirations window refers to a set of similar (or attainable) individuals whose lives and achievements help form one's future goals (Genicot and Ray 2020)

Sr. No.	Needs assessment (Source of inequality)	Type of policy	Specific policy options	Rationale	Expected outcome
			experiences, and role models to influence aspirations.	outcomes (Genicot and Ray 2020)	disadvantaged individuals
8	Gap in labor force participation	Active labor Market Policies (ALMP)	Parental leave, flexible workplace, childcare service	Directly, labor is one of the major sources of income; Indirectly, the gap in the labor market worsens IGM	Improve labor force participation, especially of women
9	Gap in labor supply	As in 8 above	Facilitate integration of youth into the labor market; offer labor market information and training	As in 8 above	Improve the employability of labor
10	Gap in labor demand	As in 8 above	Wage subsidies to employers; public works targeting disadvantaged individuals and disadvantaged locations	As in 8 above	Stimulate job creation
12	Capital market imperfection	Various	Improve access to credit; otherwise, mitigating the effects of the imperfections through targeted transfers to low-income families (e.g. Unemployment benefit; safety net)	Credit constraints and lack of insurance might limit mobility by reducing investments on education and potentially profitable opportunities (Narayan et al. 2018)	Improved access to capital market

4.5.2. Conclusion

This study examines the extent of intergenerational persistence in social status using data from the Ethiopian Living Standard Measurement Study – Integrated Survey on Agriculture (LSMS-ISA). The findings point to a strong correlation between parental and child educational and occupational status, even after controlling for a wide range of individual, household, and location characteristics.

The results indicate that the extent of economic mobility correlates with urbanization. Urbanization is associated strongly and positively with both the level and inequality in educational status. Low IGM is one of the key channels through which this inequality persists. In large urban areas, parental education significantly influences the educational attainments of their children. This is particularly the case for levels above secondary education. Compared to a child from parents with no education, a child from parents with secondary or tertiary education is about twice more likely to attain tertiary education in large urban areas.

Similarly, while parental self-employment influences their children's occupation status in both rural and urban areas, there is significant persistence in occupation across generations in large urban areas for all employment types. All else the same, the likelihood of employment is significantly different for children from parents with elementary occupations and children from parents with skilled wage employment. In general, in large urban areas, children of parents employed in better-paid occupations are more likely to be employed in similar occupations themselves. The same cannot be said for children from rural areas and small towns.

The study also shows that the inequality observed in occupational opportunities in large urban areas is largely explained by differences in educational attainment. Once an individual's education level is taken into account, large urban areas offer better employment opportunities than small towns or rural areas, regardless of parents' occupational status. This suggests that broadening access to and reducing the dropout rates at post-elementary schools and improving the quality of education, is one of the most effective mechanisms to reduce intra- and inter-generational inequality.

CHAPTER **5**

5. Incentivizing and Retaining Public Servants in Remote Areas: A discrete choice experiment with agricultural extension agents in Ethiopia

Abstract

Increased deployment of agricultural extension agents (EAs) in rural areas is grounded on their importance to spur agricultural productivity and mitigate spatial imbalances in welfare. However, high turnover and low motivation levels of EAs in remote areas pose challenges for equitable service provision and, in some cases, exacerbate geographical welfare disparities. We assess the effectiveness of selected potential policy interventions to incentivize and retain EAs in remote areas of Ethiopia. To this end, we conducted a choice experiment to elicit the preferences of 761 EAs for job attributes. We apply a random parameters logit model to estimate parameters of interest and to simulate the impact of possible policy interventions. Our results show that offering education opportunities is by far the most powerful instrument to attract and retain EAs. It increases the uptake of the extension job in remote locations by 77 percentage points, which is significantly higher than the effect from doubling current salary levels. EAs also expressed a strong preference for work environments with basic amenities, housing, transportation services, and well-equipped Farmer Training Centers (FTCs). Furthermore, the results from sub-sample analyses show that female EAs are less responsive to pecuniary incentives and are more concerned with the availability of infrastructure and amenities. Current salary levels, years of employment, and location of work are also important sources of heterogeneity in the response of EAs to potential policy changes.

JEL Classification: C25, J22, J45, J61, Q16, R50.

Keywords: Agricultural extension agents, choice experiments, Ethiopia, preference heterogeneity, random parameters logit model, remoteness

Publication status: Regassa, M. D., Abate, G. T., Kubik, Z., & Regassa, D. (2021). Incentivizing and retaining public servants in remote areas: A discrete choice experiment with agricultural extension agents in Ethiopia. *Journal of Agricultural Economics*, *72*(3), 878–900. https://doi.org/10.1111/1477-9552.12432

5.1. Introduction

Transformation of the agricultural sector is deemed the fastest and the most effective means to achieve overall poverty reduction and to address the widely prevalent spatial economic disparity in low-income countries (de Janvry and Sadoulet 2009; Mellor 2018; Spielman, Mekonnen, and Alemu 2012). Agricultural extension holds an enormous potential to accelerate the transformation of the agricultural sector through the transfer of knowledge and promotion of technologies and thereby increase agricultural productivity and reduce poverty and spatial inequality (e.g., Stifel and Minten, 2017; Minten et al., 2013; Davis 2008; Dercon et al. 2007). However, empirical evidence of agricultural extension impacts in low-income countries, including Ethiopia, is mixed. Extension systems are widely regarded as ineffective in meeting the needs of poor and remote farm households (Abate et al. 2020; Ragasa et al. 2016; World Bank and IFPRI 2010).

The main problem with public extension services in low-income countries is incentive failure. That is, the extension system is unresponsive to the demand of agricultural Extension Agents (EAs)[75] and smallholder farmers (World Bank and IFPRI 2010). A large body of literature has pointed to extrinsic factors – incentives related to the living and working conditions of EAs – as one of the main reasons for the poor performance of the extension system[76]. This literature has identified a range of factors, such as low salary, poor housing, job insecurity, lack of a transparent remuneration scheme, and inadequate resources to run Farmer Training Centers (FTC) as major constraints on the performance of EAs (Davis 2008; Davis et al. 2010; Kassa et al. 2012; Ragasa et al. 2016).

EAs are expected to play a key intermediary role between smallholder farmers, agricultural researchers, and policymakers. They promote the adoption of improved agricultural technologies and practices generated by the research system and facilitate proper implementation of development policies designed by policymakers (Dercon et al. 2007; Feder et al. 2010; Ragasa et al. 2016). Effective delivery of these mandates requires an adequate number of EAs that are qualified, motivated, and responsive to changes in farmers' demand and in the policy environment. However, observational studies conducted in low-income countries indicate that extension offices are often understaffed and that agents are underqualified and unmotivated. Low morale, absenteeism, and high turnover are commonly reported among agricultural EAs (Davis et al. 2012; Feder et al. 2010; Kassa 2002; Kassa and Abebaw 2004; Ragasa et al. 2016). This problem is particularly severe in remote areas, where quality extension services are much more needed (Birner et al. 2009; Ragasa et al. 2016).[77]

[75] Extension Agents (EAs) in Ethiopia are often named as Development Agent (DAs) to reflect the additional development activities they perform at grass-root level. We used the term extension agents since their primary responsibility is providing agricultural extension and advisory services.

[76] The literature also puts forward intrinsic factors – incentives related to interest in the work, recognition, and assuming responsibility – as complementary reasons for the poor performance of agricultural extension. According to Herzberg (1987), while improvement in intrinsic factors encourage performance, poor extrinsic factors demotivate performance. Qualitative data collected from the EAs surveyed in this study in 2019 shows that extrinsic factors account for the largest share of the factors that hinder effective extension service in Ethiopia (see Table 5.8).

[77] This view is grounded in equity considerations and aims at reducing the existing spatial imbalances. Note, however, that massive deployment of EAs to remote areas might be questionable from a purely economic efficiency perspective if these areas have low agricultural potential.

In the case of Ethiopia, the rapid turnover and general underperformance of EAs pose a serious threat to the extension system and the national goal of achieving food security (Kassa 2002; Kassa and Abebaw 2004). Cognizant of this and given the crucial role of public extension services at the early stages of agricultural transformation, the government of Ethiopia has undertaken several policy measures to improve the attractiveness of jobs in rural extension. These include financial incentives, housing, improved working conditions (e.g., working tools, transportation), and various higher education and training opportunities, among others. The latest agricultural extension strategy of the country shows that additional interventions are being considered to further reinforce the role of EAs (MoANR and ATA 2017).

However, the policy initiatives taken so far do not seem to have abated many of the problems EAs face, as rapid turnover of EAs persists, especially in remote rural areas (Davis et al. 2010; Haile and Abebaw 2012; MoANR and ATA 2017). This could be partly because the design of these interventions has not been grounded in systematic studies of the responsiveness of EAs to the different policy interventions. This suggests the need to study the nature and determinants of EAs' preference for job attributes and their labor supply choices to better design more effective and better-tailored policy interventions that can incentivize and retain EAs, in particular, and public workers in remote areas in general.

Our study fills this research gap by assessing the responsiveness of EAs to potential policy interventions that are closely aligned to the current and prospective plans of the public extension strategy (MoANR and ATA 2014, 2017). Using a carefully designed discrete choice experiment (DCE), we first test the responsiveness of EAs to alternative policy interventions in the realm of EA-specific job incentives, i.e., salary, educational opportunities, equipment, transportation facilities, and housing. We then simulate the impact of different policy changes on the perceived attractiveness of EA jobs. Furthermore, we compute EAs' willingness to pay (WTP) for job attributes related to different living and working conditions. Finally, we explore the heterogeneity in the results based on sociodemographic and current job characteristics of EAs to better tailor interventions. We expect that our findings will offer policy guidance on how to incentivize EAs and, more broadly, any public agent to work in remote rural areas of developing countries.

Our analysis offers several interesting and useful insights. First, there is a general dissatisfaction among the EAs with the status quo, and improvement in any of the proposed EA job aspects is seen as a better option. Second, contrary to popular perception, increasing salaries is not always the strongest incentive for EAs. Our findings suggest that offering educational opportunities is by far the most powerful instrument to attract and retain EAs in remote locations. Upward salary adjustment only comes in at a second position, followed by a provision of housing and transportation facilities. EAs are also likely to respond to such incentives as availability of basic amenities, including improved access to electricity and mobile telephone network in the Kebeles[78] to which they are posted, as well as adequately equipping FTCs with the necessary work materials. Additionally, the analysis of heterogeneity underlying our results demonstrates the importance of accounting for EA sociodemographic characteristics, such as age, gender, or job tenure, when designing policy intervention intended to attract, retain, and motivate EAs. For

[78] Kebele is the lowest administrative unit in Ethiopia. It is a sub-district administrative level that can be loosely equated with a village. There are around 15,000 Kebeles in Ethiopia.

instance, female EAs are more responsive than male EAs to nonpecuniary incentives, such as to the provision or availability of transportation services.

The structure of this article is as follows. Section 5.2 briefly describes the dataset used. Section 5.3 presents the research method, emphasizing the theoretical background and empirical estimation strategy of the DCE. Section 5.4 presents the main results, which are then discussed in Section 5.5. Section 5.6 concludes with the highlights of the main results and some policy implications.

5.2. Data and descriptive statistics

We use a dataset from a survey covering more than 700 EAs in the principal agricultural regions of Ethiopia – Tigray, Amhara, Oromia, and SNNP[79]. The data was collected by the International Food Policy Research Institute (IFPRI) in collaboration with Digital Green (DG) as part of a project that assessed the impacts of video-mediated agricultural extension service provision on farmers' knowledge and the adoption of improved agricultural technologies and practices in Ethiopia (see Abate et al. 2020 for a detailed description of the data). The data were collected in 2017, 2018, and 2019 and covered 896, 781, and 763 EAs, respectively. The dataset contains detailed information on the socio-demographic characteristics of EAs; the extension approaches they use; the incentives they have; their workload, motivation, and knowledge of cereal extension; and information about the Kebeles where they work.

The main part of our analysis is based on a choice experiment module we added to the last round of the IFPRI-DG survey, which was conducted between February and April 2019. Based on a novel discrete choice experiment design, each EA in the survey sample was presented with eight pairwise choices. Each choice set contained two job profiles with varying levels of selected job attributes, as well as an opt-out option. This resulted in 18,264 rows of data that allowed us to elicit information on the preferences and the trade-off EAs made among job attributes.

Table 5.1 presents summary statistics of the socio-demographic characteristics of the EAs in the sample. EAs in the study areas are predominantly male (76 percent), young (less than 30 years of age), and have a college diploma. Average work experience in agricultural extension service provision is six years on average, and most of them came from the same locality in which they are working, i.e., they lived in the same Woreda (district) as a child. This is mainly because the recruitment, placement, and transfer of EAs are primarily done by the Woreda Bureau of Agriculture (BoA), albeit (prospective) EAs have the choice to accept or decline the job placement. About half of the EAs in the sample are computer illiterate.

We also check if spatial inequality in extension services is reflected in our data[80]. The analysis of the profile of EAs disaggregated by the remoteness of their location in columns 3, 4, and 5 of Table 5.1 indicate considerable differences between EAs in more and less remote locations[81].

[79] SNNP refers to Southern Nations, Nationalities, and Peoples' region.

[80] See Abate et al. (2020) for more detailed discussion of the spatial inequality in extension service.

[81] Remoteness is defined based on distance between the center of the Kebele in which an EA is posted and the capital of the local Woreda. 'Nearest tercile' represents Kebeles closest to the Woreda capital.

EAs working in relatively remote locations are younger, less experienced, less educated, and exert less work effort (seemingly due to lack of close supervision) as measured by weekly working hours. These observations are corroborated by the results of locally weighted polynomial regressions of respective outcome variables on the distance from the centre of Kebele in which an EA works to the local Woreda centre (Table A5.1 in the appendix).

Table 5.1. Characteristics of extension agents in study sample, by remoteness tercile

Characteristics	N	All	Nearest tercile	Middle tercile	Farthest tercile	F-test: p-value
Male	2,440	0.76	0.75	0.75	0.76	0.92
Age, years	2,440	28.6	30.7	28.4	26.5	0.00
Number of years working as an EA	2,440	6.47	8.67	6.23	4.44	0.00
Number of years working in current Kebele	2,440	2.31	2.59	2.45	1.89	0.00
Education: Certificate, yes=1	2,440	0.16	0.13	0.15	0.19	0.00
Education: Diploma, yes=1	2,440	0.60	0.64	0.61	0.54	0.00
Education: Degree, yes=1	2,440	0.25	0.23	0.24	0.27	0.10
Computer literate, yes=1	2,440	0.46	0.45	0.48	0.45	0.28
Mobile with internet access, yes=1	2,440	0.48	0.47	0.5	0.48	0.49
Spent childhood: In working Kebele, yes=1	2,439	0.09	0.12	0.11	0.05	0.00
Spent childhood: In working woreda, yes=1	2,439	0.62	0.68	0.61	0.56	0.00
Spent childhood: In working zone, yes=1	2,439	0.85	0.87	0.85	0.83	0.07
Number of EAs in Kebele	2,440	3.42	3.6	3.38	3.28	0.00
Number of farmers' field days organized	2,439	1.91	1.89	1.76	2.08	0.00
Working hours per week: Planting season	2,440	49.0	52.4	47.7	46.9	0.00
Working hours per week: Harvesting season	2,440	36.7	40.2	35.3	34.4	0.00
Working hours per week: Slack season	2,440	23.9	26.9	22.8	22.1	0.00
Working hours per week: Average	2,440	36.5	39.8	35.3	34.4	0.00
Knowledge score: Teff	1,544	70.3	71.1	70.0	69.9	0.32
Knowledge score: Maize	1,544	67.4	68.5	67.2	66.4	0.09
Knowledge score: Wheat	1,544	65.5	65.3	65.4	65.8	0.80
Knowledge score: Average	1,544	67.7	68.3	67.5	67.4	0.35

***Source**: Authors' calculation based on IFPRI-Digital Green's EA survey, 2017, 2018, and 2019.*

***Note**: Knowledge score refers to EAs' work-related knowledge score (out of 100) obtained through quizzes. The knowledge questions (collected only in the last two rounds) focused on the growing practices of Teff, Maize, and Wheat. Remoteness is defined based on the distance between the centre of the Kebele in which an EA is posted and the capital of the local Woreda (district). 'Nearest tercile' represents Kebeles closest to the Woreda capital.*

The descriptive analysis of our main variables of interest in Table 5. 2 indicates that EAs are dissatisfied with their current job and the vast majority believe that their job is worse than other public and private jobs open to candidates with similar education levels. This could be partly because EAs work in relatively remote locations, i.e., far from district capitals and markets, and which lack basic amenities and services, like electricity, water, transportation, and housing. The latter often forces EAs to live outside of their working Kebeles, even though, in principle, they are expected to reside in proximity to the farmers they serve. This could also be due to the limited availability of a performance-based incentive structure. In 2019, only 16 percent of EAs reported having received an award for good performance[82]. Commonly, EAs receive promotions based on

[82] This is commonly expressed through financial rewards, educational opportunities, certificates, and promotions (rank within EA), or transfer to a preferred location.

seniority (number of service years). In Table 5. 2, about 60 percent of EAs stated they have been promoted in the last three years.

The dissatisfaction of EAs with their current job seems to also emanate from inadequate facilities to effectively perform their jobs. As shown in Table 5. 2, about half of Farmer Training Centers (FTC), which are supposed to serve as training and demonstration centers, do not have proper training materials and demonstration plots. The results in Table 5. 2 show that EAs in the most remote locations (farthest tercile) have a relatively poorer work environment compared to those in less remote areas (nearest tercile).

EAs also have limited opportunities to advance their careers through Continuing Education Programs (CEPs). While more than one-third of EAs are enrolled in Continuing Education Programs CEPs, there is a clear disparity on government sponsorship by remoteness. The vast majority of EAs in relatively advanced locations are attending CEPs with government sponsorship (60%) compared to EAs in the most remote locations (23%). This could be because EAs qualify for government-sponsored education after some years of service and at the same time EAs get transferred to more connected areas with increased years of services.

Table 5. 2. Work environment of extension agents in study sample, by remoteness tercile

	N	All	Nearest tercile	Middle tercile	Farthest tercile	F-test: p-value
Perception of EA about their job						
Satisfied with existing incentive structure, yes=1	763	0.15	0.19	0.15	0.12	0.08
Job as compared to other public jobs, worse=1	759	0.62	0.63	0.60	0.62	0.78
Job as compared to private sector jobs, worse=1	747	0.84	0.82	0.87	0.83	0.22
Location characteristics of EAs						
Access to mobile network, yes=1	761	0.99	1.0	0.99	0.98	0.37
Access to electricity, yes=1	760	0.31	0.37	0.31	0.24	0.01
Distance to the nearest market, km	761	6.20	5.4	6.2	7.0	0.01
Distance to the district capital, km	761	18.3	7.2	15.9	32.4	0.00
Housing and transport service						
Access to bicycle or motorcycle, yes=1	763	0.16	0.14	0.19	0.15	0.33
Received housing from the government, yes=1	763	0.22	0.18	0.15	0.33	0.00
If no housing, EA lives outside Kebele, yes=1	599	0.72	0.85	0.73	0.53	0.00
FTC and FTC resources						
Kebele has an FTC, yes=1	761	0.88	0.92	0.89	0.81	0.00
FTC has demonstration plot, yes=1	761	0.81	0.88	0.83	0.7	0.00
FTC has ICT tools, yes=1	761	0.15	0.23	0.13	0.07	0.00
FTC has training materials, yes=1	761	0.52	0.51	0.55	0.50	0.47
FTC has own budget, yes=1	761	0.36	0.41	0.38	0.28	0.01
Educational opportunity						
Enrolled in continuing education (CEP), yes=1	763	0.36	0.33	0.38	0.37	0.39
If CEP enrolled, government sponsored, yes=1	272	0.43	0.60	0.48	0.23	0.00
Education opportunities available, yes=1	762	0.46	0.50	0.46	0.40	0.09
Available incentive structure						
Received any award for performance, yes=1	763	0.16	0.18	0.15	0.14	0.35
Received promotion over the last 3 years, yes=1	763	0.63	0.56	0.69	0.63	0.01

Source: *Authors' calculation based on IFPRI-Digital Green's EA survey, 2019.*

Note: *Remoteness is defined based on the distance between the centre of the Kebele in which an EA is posted and the capital of the local Woreda (district). 'Nearest tercile' represents Kebeles closest to the Woreda capital.*

5.3. Choice experiment design and analytical framework

Empirical studies have adopted various research methods in order to investigate the factors determining the employment choices of EAs. The most common is the use of cross-sectional survey data to evaluate the extent to which outcomes for EAs, such as job satisfaction, motivation, and organizational commitment are correlated with various individual and job characteristics (Davis, Swanson, and Amudavi 2009; Kassa et al. 2012; Kassa and Abebaw 2004). These studies have identified a number of factors that constrain the performance of EAs, such as low salary, poor housing, job insecurity, lack of transparent remuneration scheme, and unavailability of resources to run the FTCs. While this method produces valuable insights, it provides weak evidence on the relative importance of the various constraints EAs face (Lagarde and Blaauw 2009).

An alternative method is to employ longitudinal data on labor supply choices of EAs – a revealed preferences approach. This method, based on panel data, can help provide robust evidence on the relevance of labor market policy interventions. However, to date, it has mainly been used in developed countries (Frijters et al. 2009; Shields 2004). The extensive data it requires is often not available in developing countries. This is particularly the case for EAs due to their high turnover[83].

The Discrete Choice Experiment (DCE) method presents a valid alternative to counter the shortcomings of using cross-sectional and longitudinal data. DCE has become increasingly popular to elicit preferences for attributes of differentiated goods and services. It is particularly helpful to investigate the stated preference of workers for job attributes that are not easily observable in the market (Lagarde and Blaauw 2009). Since it is based on workers' choices from among hypothetical job scenarios, it allows an assessment to be made of the responsiveness of workers to potential policy interventions. An additional advantage of DCE is that it allows analysts to disentangle the contribution of each of the attributes to overall utility. In revealed preference data, it is not possible to isolate the contribution of the attributes due to the strong potential correlation among them, e.g., housing and location of work. In the design of choice experiments, the experimental method allows the researcher to exogenously vary the attributes (Hensher and Greene 2006; Mangham and Hanson 2008; Train 2009).

Although DCE shares the basic features of hedonic wage analysis, contingent valuation methods (CVM), and conjoint rating, it improves on and goes beyond these methods since it allows estimation of a consistent marginal rate of substitution for both existing and prospective traits (Hensher, Rose, and Greene 2005; Louviere, Hensher, and Swait 2010; Train 2009). While the application of choice experiment surveys to elicit preferences is common in the environmental, health, marketing, and transport literature, its application to the evaluation of employees' preferences for job attributes is nascent. The limited existing applications of DCE to job preference are limited to health care occupations (Chomitz et al. 1998; Hanson and Jack 2008; Kolstad 2011;

[83] A credible econometric method is difficult to establish mainly because the use of panel data implies such analysis is based on those that remained employed. This would result in biased estimates as those that remained employed would constitute a systematically different sample from those that left the industry (Angrist and Pischke 2009; Heckman 1979).

Kruk et al. 2010; Scott 2001) and youth employment (Assy et al. 2019). This paper extends the application of DCE to elicit job preferences of rural public agents/servants.

5.3.1. The choice experiment design

The DCE outlines a hypothetical setting in which respondents are asked to repeatedly choose from a limited number of alternatives. Each alternative is described by a number of attributes that take on different levels[84]. Representing job alternatives as bundles of attributes allows assessment of changes in individual choices as one or more of the attributes vary (Lancaster 1966).

In our study, EAs were presented with a series of choice situations, each of which contained a pair of job profiles with six attributes and an opt-out option. The EAs were asked to choose which of the two jobs (or neither) they preferred. The choice of the selected attributes (Table 5.3) is based on an extensive literature review of the factors that are perceived to be important in job choices of EAs in Ethiopia and beyond (Berhane et al. 2018; Dufera et al. 2017; Gebru, Asayehegn, and Kaske 2012; Haile and Abebaw 2012; Kelemu, Sime, and Hailu 2014; Mangham and Hanson 2008; Ragasa et al. 2016). We verified the appropriateness of these attributes and their respective levels based on series of discussions with national and regional extension coordinators, focus group discussions with EAs, and pre-survey piloting.

The number of selected attributes is in line with previous empirical studies. Generally, the attributes and their respective levels need to be realistic enough to provide relevant policy predictions regarding the effect of potential interventions. At the same time, the design does not need to be too complicated in order to minimize fatigue and cognitive burden on the respondents (Kuhfeld 2010; WHO 2012)[85]. The selected attributes alongside their respective levels are shown in Table 5.3. During the interviews, these attributes and their levels were carefully explained to respondents. Explicit information was also included regarding potentially relevant excluded attributes and attribute levels. Respondents were asked to assume that all unstated characteristics of jobs are the same for the two alternatives in a choice set.

[84] For excellent reviews of this method, please see Hensher et al. (2005); Louviere et al. (2010); Train (2009).

[85] In comparable public sector human resource applications, the suggested number of attributes ranges between 2 and 24, with a mode of 6 (De Bekker-Grob et al. 2008; WHO 2012).

Table 5.3. Job attributes and attribute levels used in the choice experiment

Attribute	Definition	Attribute Levels
Location	Whether location of work has reliable mobile coverage, electricity, and piped water (advanced) or not (remote)	(1) Advanced, (2) Remote
Net monthly salary	Net salary at job (reference: current net average salary)	(1) Plus 100%, (2) Plus 50%, (3) Plus 25%, (4) Minus 25%
Provision of housing	Provision of government housing at Kebele of work for residence of the extension agent and her family.	(1) Available, (2) Not available
Extension tools at Farmer Training Centres (FTC)	Adequacy of FTC resources to effectively deliver extension service to farmers (e.g., demonstration plot, adequate budget to run the FTC, adequate teaching materials)	(1) Adequate, (2) Inadequate
Transportation facilities at FTC	Availability of transportation facility at the FTC (bicycle, motorcycle, or horse)	(1) Available, (2) Not available
Education opportunities	Availability of education opportunities after 2 years of service	(1) Available, (2) Not available

Source: Constructed by authors.

In the survey, we presented respondents with a series of pairs of jobs and asked them to choose the one they prefer from each pair or neither. Theoretically, there are 128 (= 2*4*2*2*2*2) distinct jobs characterized by the six attributes, and, therefore, 8,192 (=128*128/2) distinct job pairs. From among these distinct job pairs (called full factorial design), we identified and presented to the respondents 16 different choice sets based on main effects fractional factorial design. This is a D-optimal hypothetical choice design based on the covariance matrix of a multinomial logit model with all the coefficients assumed to be equal to zero. The design offers an efficient combination of orthogonality, level balance, and minimum overlap (Kuhfeld 2010) [86,87]. The 16 choice sets were randomly divided into two blocks in order not to exhaust the respondents. Each respondent thus made eight binary choices with an opt-out option. Table A5.2 in the appendix presents the instructions given to the respondents and an example of the question set-up.

5.3.2. Analytical framework

The analytical framework of the choice experiment data is based on random utility theory, which assumes that a representative individual is rational and, in a given choice situation, selects the alternative that yields the highest level of utility (McFadden 1973). The individual is assumed to know her or his preferences, but a component of these preferences is unobservable to the researcher. Therefore, assuming a linear indirect utility functional form, the utility (U) of an individual i, for alternative j, in choice situation t, is expressed as a sum of a systematic (observable) component V_{ijt}, and a stochastic (unobservable) component, ε_{ijt}.

$$U_{ijt} = V_{ijt} + \varepsilon_{ijt}, \qquad j = 1,2,\dots\dots,m \tag{5.1}$$

[86] Operationalized with SAS analytical software.

[87] In order to generate unlabeled experimental designs suitable for our purpose, SAS choice modelling macros, %MktRuns, %MktEx , %ChoicEff and %MktBlock are used. While there are other popular tools that could be used to generate experimental designs including Stata, Sawtooth, Ngene and R, these SAS macros are also well suited to find good, efficient, and realistic designs (Kuhfeld 1997, 2010).

In line with Lancaster's (1966) theory of demand, which argues that the overall utility an individual generates from a good or service can be decomposed into the sum of separate utilities derived from its constituent characteristics, the systematic part of the utility function can be expressed as: $V_{ijt} = \alpha + x'_{ijt}\beta_i$. After replacing this for V_{ijt}, equation (5.1) becomes:

$$U_{ijt} = \alpha + x'_{ijt}\beta_i + \varepsilon_{ijt} \tag{5.2}$$

where β_i is a vector of individual-specific coefficients, X_{ijt} is a vector of observed attributes relating to individual i, and alternative j, in a choice situation t. In this model, called a random parameter logit model (RPL), ε_{ijt} is a random term that is assumed to be an independently and identically distributed extreme value type I[88]. Consistent with a utility function that is linear in parameters, the probability that an EA i, chooses alternative j, from among m alternatives in a choice situation t, takes a conditional logit specification (McFadden 1973):

$$L_{ij}(\beta_i) = \frac{exp(x'_{ijt}\beta_i)}{\sum_{l=1}^{m} exp(x'_{ilt}\beta_i)} \tag{5.3}$$

The specification in (5.3) assumes that ε_{ijt} is the only source of randomness and that the taste parameter of each EA, β_i, is known to the researcher and fully explained by only using its means. In reality, β_i is unknown to the researcher, and, hence, it is not feasible to condition on β_i (McFadden & Train 2000; Train 2009). Instead, β_i is assumed to be normally distributed with population mean β and covariance Σ_β, and the unconditional probability that an EA will choose alternative j is estimated by integrating the conditional probabilities over all values of each of β weighted by its density function. That is:

$$\begin{aligned} P_{ijt} = Pr[y_i = j] &= \int L_{ij}(\beta_i)\, f(\beta_i|\theta) d\beta_i \\ &= \int \frac{exp(x'_{ijt}\beta_i)}{\sum_{l=1}^{m} exp(x'_{ilt}\beta_i)} f\left(\beta_i\middle|\beta, \Sigma_\beta\right) d\beta_i \end{aligned} \tag{5.4}$$

In equation (5.4), $f(\beta_i|\theta)$ is multivariate normal density for β_i with mean β and covariance Σ_β. The integral is multidimensional with dimension given by the number of components of β_i that are random with non-zero variance (Cameron & Trivedi 2005). For simplicity, we assume that the components are uncorrelated and, hence, the off-diagonal elements of Σ_β are zero. With respect to β and Σ_β, the Maximum Likelihood Estimation (MLE) now maximizes:

$$lnL_N(\theta) = \sum_{i=1}^{N}\sum_{j=1}^{m} y_{ijt} lnP_{ijt} \tag{5.5}$$

Since the integral in (5.4) does not have a closed form, the expression in (5.5) cannot be analytically solved. Instead, simulated probabilities are inserted into the log-likelihood function

[88] When $\beta_i \sim lnN(\beta, \Sigma_\beta)$, for parameters whose sign is known a priori, this model is also known as a mixed logit model (Cameron & Trivedi 2005).

to give a simulated log likelihood (Cameron & Trivedi 2005; Hensher & Greene 2006; Train 2009) of the form:

$$ln\widehat{L_N}(\beta,\Sigma_\beta)=\sum_{i=1}^{N}\sum_{j=1}^{m}y_{ijt}ln\left[\frac{1}{S}\sum_{s=1}^{S}\frac{\exp\left(x_{ijt}'\beta_i^{(s)}\right)}{\sum_{l=1}^{m}\exp(x_{ilt}'\beta_i^{(s)})}\right] \tag{5.6}$$

where y_{ijt}=1 if the EA chooses alternative j in a choice set t, and zero otherwise; and $\beta_i^{(s)}$, with s=1, 2, …, S, are random draws from $f(\beta|\theta)$ [89]. Parameter estimates, $\beta^s\ and\ \Sigma_\beta^{(s)}$, represent the mean and standard deviation generated from equation (5.6) using maximum simulated likelihood (MSL) at r^{th} draw (Cameron & Trivedi 2005; McFadden & Train 2000).

Besides its relevance to capture unobserved heterogeneity, RPL is preferable because it allows possible correlations between the selected alternatives and choice tasks. That is, the model relaxes the strict assumption of independence of irrelevant alternatives (IIA) (Hensher & Greene 2006; McFadden & Train 2000; Train 2009). More importantly, our preferred specification allows estimation of the respondents' marginal rate of substitution for different attributes. When one of the attributes is salary, this produces the willingness to pay (WTP) of EAs for location and different work attributes. For any non-monetary attribute, x^{nm}, the willingness to pay of EA i, could be calculated as:

$$WTP_{ix^{nm}}=\frac{\partial U_i}{\partial X^{nm}}\Big/\frac{\partial U_i}{\partial W}=-\left(\frac{MU_{x^{nm}}}{MU_w}\right) \tag{5.7}$$

where $MU_{x^{nm}}$ and MU_w represent the marginal utility of attribute x^{nm} and salary, respectively. One issue with estimation of the WTP as ratios of the estimated random coefficients of non-monetary attributes to the marginal utility of salary is that it involves dividing distributions on distributions (Hensher & Greene 2006; Train 2009; WHO 2012). Depending on the choice of parameter distributions, this results in WTP distributions which are heavily skewed or distributed with no defined moments (Scarpa, Thiene, & Train 2008; Train & Weeks 2005). Commonly, empirical studies circumvent this problem by assuming that the monetary coefficient is fixed. However, this assumption might be unrealistic as the marginal utility of income tends to vary depending on sociodemographic characteristics (Layard, Nickell, & Mayraz 2008).

In this study, we adopt a novel approach suggested by Train and Weeks (2005) and directly estimate the WTP in a WTP space. This approach, which involves deriving the WTP estimates directly by reformulating the mixed logit model, appears to better fit the data (Scarpa et al. 2008) and produce more realistic WTP estimates (Train & Weeks 2005) than the conventional method. For the sake of illustration, we rewrite the utility function in equation (5.2), differentiating between monetary (W_{ijt}) non-monetary (Z_{ijt}) attributes.

$$U_{ijt}=\eta_i w_{ijt}+z_{ijt}'\varphi_i+\ \varepsilon_{ijt} \tag{5.8}$$

[89] We report results obtained using 100 Halton draws. However, the results remained robust to alternative number of draws.

where η_i and φ_i are individual-specific coefficients for monetary, i.e., salary, and non-monetary attributes of the job and ε_{ijt} is the random term. Dividing both sides of equation 8, we get:

$$U_{ijt} = \eta_i[w_{ijt} + z'_{ijt}\gamma_i] + \varepsilon_{ijt} \tag{5.9}$$

where $\gamma_i = {}^{\varphi_i}/_{\eta_i}$ represents the WTP for the non-monetary attributes. This specification – called model in WTP space – allows direct estimation of the coefficients corresponding to the non-monetary attributes as WTP estimates by using MSL (Train 2009).

5.4. Results

In this section, we discuss estimation results that are based on the elicited preferences of 761 EAs. Initially, the choice experiment included 763 EAs. However, two respondents who provided non-rational choices were excluded from the analysis.[90] We present results based only on the estimation of the RPL model. To assess the pertinence of the RPL model, we initially estimated a conditional logit (CL) model and tested the assumption of independence of irrelevant alternatives (IIA). The Hausman test rejected the IIA assumption, implying that the CL model is not appropriate.[91]

5.4.1. Preferences for job attributes

Table 5.4 presents the simulated likelihood estimates of the RPL model. It is important to note that the coefficients of the parameter estimates do not have an absolute interpretation. This is because the utility represented in the framework merely describes ordinal dependence (Train 2009). However, the sign of the parameter estimates indicates whether the respondents view the attributes positively and their relative magnitude indicates how strongly they do so relative to the alternative attributes.

The findings provide several insights. First, there is a general dissatisfaction with current job characteristics among EAs, as indicated by the constant term of the model. The constant term, which represents EAs' preference for their current job, is negative and statistically significant. This implies that EAs evidently prefer the two hypothetical job choices associated with differing attribute levels than their current job and its respective attributes. This is consistent with other studies that find that the vast majority of EAs are not satisfied with the extant incentive structure within the extension system (Berhane et al. 2018; Davis et al. 2010; Kassa et al. 2012).

Second, even though salary increases appear to be highly valued, as one would expect, they do not come as the top incentive for attracting, motivating, and retaining EAs. Instead, our results suggest that the availability of educational opportunities is by far the most important factor affecting EAs' job choices. Whereas we are not able to interpret the absolute magnitude of the coefficients in column 1 of Table 5.4, the relative magnitude of the parameter estimates clearly point to the value that EAs place on further educational opportunities. To put this in context, the

[90] To assess whether the EAs understood the setup of the questions and were able to make an informed decision between the alternatives, we included a question where one of the choices is superior to the other and asked them to make a choice. Of the 763 EAs, only two EAs selected a strictly dominated alternative. These two EAs were not included in the final sample.

[91] The estimation result of the CL and the conducted test is presented in table A3 in the appendix

availability of educational opportunities comes out as more important than even a 100 percent increase in salary.

Availability of housing, transportation services, adequately equipped FTCs, and access to better infrastructure (access to mobile telephone network, road, and electricity) are also found to be significant factors in the decision process. More generally, the statistical significance of all the selected attributes indicates the availability of a wide range of interventions to policymakers to improve the attractiveness of public service jobs in rural areas of Ethiopia.

Table 5.4. Simulated likelihood estimates of the random parameters' logit model

	(1)	(2)	(3)	(4)
Variables	Structural parameters		SD of the parameter distributions	
	Coefficient	SE	Coefficient	SE
Location is advanced, yes=1	0.848***	0.073	0.953***	0.090
Housing, yes=1	0.498***	0.060	0.399***	0.125
Transport services, yes=1	0.685***	0.061	0.663***	0.096
Adequate FTC, yes=1	0.745***	0.062	0.788***	0.082
Education opportunities, yes=1	2.039***	0.093	1.331***	0.086
Salary (ref: current basic salary)				
Salary increment of 100%, yes=1	1.753***	0.139	1.536***	0.137
Salary increment of 50%, yes=1	0.991***	0.135	-0.049	0.127
Salary increment of 25%, yes=1	0.376***	0.127	0.019	0.160
Salary reduction by 25%, yes=1	-0.727***	0.160	-1.069***	0.131
Constant	-0.101*	0.056		
Number of respondents	761			
Number of observations	18,264			
Log-likelihood	-3,420.7			
LR chi2(9)	457			
McFadden R2	0.48			
Halton draws	100			

Source: Authors' calculation based on IFPRI-Digital Green's EA survey, 2019.

Note: triple (***), double (**), and single (*) represent statistical significance at 1%, 5%, and 10% level, respectively. SD in Columns 3 and 4 represent standard deviations indicating preference heterogeneity.

Column 3 of Table 5.4 presents the standard deviation associated with each of the mean coefficient estimates of the random parameters calculated over the 100 Halton draws. Except for the intermediate salary increments, the standard deviation coefficients are statistically significant, indicating considerable heterogeneity in preferences among EAs. Later in this section, we explore the source of the heterogeneity in preferences for these attributes across EAs based on their socio-demographic and location characteristics.

5.4.2. Willingness to Pay (WTP)

The most important and informative output of the econometric analysis of choice experiment data for policy purposes is the marginal analysis and the related policy impact analysis. The marginal analysis indicates the rate at which EAs are willing to substitute one attribute for another. When the reference attribute is salary, this produces the willingness to pay (WTP) of EAs for a non-pecuniary job attribute. As highlighted in the methods section, we adopt a novel approach suggested by Train and Weeks (2005) and directly estimate the WTP in a WTP space.

The WTP estimates in Table 5.5 provide clear indications about the relative importance that EAs attach to education opportunities.[92] On average, EAs are willing to pay 2,530 ETB in order to obtain education opportunities after two years of service rather than no further educational opportunities. Given that the average salary is about 3,000 ETB per month, this represents an extraordinarily strong preference for continuing education opportunities. EAs require to be paid an additional 1,190 ETB to be willing to work in remote areas rather than locations that are more connected and equipped with basic amenities. This is consistent with the notion of paying extra amounts for people working in difficult conditions. Similarly, the average WTP for transport services and housing is 880 ETB and 690 ETB, respectively.

It is also interesting to note that EAs have a strong intrinsic motivation to deliver quality extension service, as signalled by the large and statistically significant WTP for adequately equipped FTCs. An average EA is willing to sacrifice 830 ETB to work in an FTC that is adequately equipped. Although we are not aware of any other study that computed the WTP estimates for agricultural extension agents with which to compare, our results are comparable to studies of health professionals (e.g., nurses, laboratory technicians, clinical officers) in middle and low-income countries (Hanson and Jack 2008; Kolstad 2011; Mangham and Hanson 2008; Rockers et al. 2012)[93].

Table 5.5. Willingness to Pay (WTP) estimates for job attributes, '000 ETB

Variables	Coefficient	SE	[95% Conf. Interval]	
Mean of estimates				
Location is advanced, yes=1	1.19***	0.08	1.03	1.36
Housing, yes=1	0.69***	0.07	0.56	0.83
Transport service, yes=1	0.88***	0.08	0.73	1.03
Adequate FTC, yes=1	0.83***	0.08	0.68	0.99
Education opportunity, yes=1	2.53***	0.11	2.30	2.75
SD of Estimates				
Location is advanced, yes=1	1.40***	0.10	1.21	1.59
Housing, yes=1	0.93***	0.11	0.71	1.16
Transport service, yes=1	0.75***	0.10	0.56	0.94
Adequate FTC, yes=1	1.06***	0.09	0.89	1.24
Education opportunity, yes=1	1.46***	0.11	1.24	1.68
Number of respondents	761			
Number of observations	18,264			
Chi-squared (df = 6)	711.6			
Log-likelihood	-3,452.3			

Source: Authors' calculation based on IFPRI-Digital Green's EA survey, 2019.

Note: triple (***), double (**), and single (*) represent statistical significance at 1%, 5%, and 10% level, respectively.

Nonetheless, the cost-effectiveness of the alternative interventions should be factored in before concrete policy choices or decisions are made. In this study, we are unable to conduct a thorough assessment of the fiscal cost of these interventions due to a lack of data. Instead, as an example, we conducted a simple *back-of-the-envelope* analysis on education – a strongly preferred job attribute by EAs. The result indicated that EAs WTP is about 2.3 times more than the estimated

[92] Note that these estimates are generated by converting the categorical salary variable to cardinal values based on the median monthly salary of EAs in Ethiopia, which was 2,500 ETB in 2018. In January 2018, 1 USD = 27.34 ETB.

[93] The kernel density plot in Figure A2 in the appendix provides further evidence of the strongly positive, yet notably heterogeneous, preferences for these job attributes.

average cost of education[94]. A similar and thorough cost-effectiveness analysis that combines the cost of the different interventions and computed elasticities will be crucial for determining both the effectiveness of the interventions and their efficacy.

5.4.3. Policy impact

Another salient output of DCE is the policy impact analysis. Calculated by differentiating the probability function with respect to the job attributes, this analysis indicates how effective alternative policy interventions are to improve the attractiveness of an EA job. It shows how the preference for or the probability of taking the baseline job changes due to a change in the level of one of the job attributes (Train 2009). The baseline job is represented by the reference category for all dummy variables, i.e., location is remote; no housing; no transportation service; FTC is inadequate and no educational opportunity, and the current average monthly salary level.[95]

Table 5.6 shows that increasing salary by 25, 50, and 100 percent increases the propensity of EAs to accept a remote job posting by 19, 46, and 70 percentage points, respectively. This reiterates that upward salary adjustment is a powerful tool to improve the attractiveness of an EA job in remote areas. Two additional points are noteworthy. First, our results are consistent with the loss aversion (prospect) theory.[96] While reducing salary by 25 percent reduces the propensity of taking up the baseline job by 35 percentage points, a 25 percent increase in salary only results in a 19-percentage point increase in the probability of employment. Second, the results show that the potency of pecuniary incentives to improve the attractiveness of an EA job diminishes as salary increases. While increasing salary by 50 percent increases the probability of taking up the baseline job by 46 percentage points, increasing it by 100 percent (an additional 50 percent increase from the baseline) increases the probability of take-up only by an additional 24 percentage points.

However, salary regulation alone may not be the most efficient way to retain and incentivize EAs. In line with what was observed in previous sections, offering possibilities for further education after two years of service proves to be more effective than increasing salary by 100 percent. Investing in essential rural infrastructure is another effective policy instrument to make remote areas more attractive to EAs. The results in Table 5.6 show that investment in basic infrastructure increases the probability of taking a remote job by 40 percentage points. Sufficiently equipping FTCs, providing transportation facilities, and providing housing increase the propensity of attracting and retaining an EA in remote areas by 36, 33, and 24 percentage points, respectively.

[94] Based on the cost-sharing regulation of the Ethiopian Ministry of Education (MoE) and reported in Leka & Chalchisa (2012), we estimate that in 2018/19 the average rough monthly cost of public higher education – after adjusting for inflation - is 1,118.2 ETB. This is about 2.3 times more than the WTP of EAs for education.

[95] While this baseline job scenario might appear as unrealistic, it rather closely resembles a typical employment condition of EAs in remote areas in Ethiopia. See, for instance, Davis et al. (2010); and Kassa et al. (2012).

[96] Prospect theory argues that downside changes are far more powerful than upside changes (Ang, Chen, and Xing 2006; Kahneman and Tversky 1979; Schindler and Pfattheicher 2017).

Table 5.6. Simulated preferences under potential policy changes

Variable	Change in probability	Standard Error	[95% Conf. Interval]	
Location is advanced, yes=1	0.40***	0.03	0.34	0.46
Housing, yes=1	0.24***	0.03	0.19	0.30
Transport service, yes=1	0.33***	0.03	0.28	0.38
Adequate FTC, yes=1	0.36***	0.03	0.30	0.41
Education opportunity, yes=1	0.77***	0.02	0.73	0.81
Salary (ref: current basic salary)				
Salary increment of 100%, yes=1	0.70***	0.04	0.64	0.77
Salary increment of 50%, yes=1	0.46***	0.05	0.35	0.56
Salary increment of 25%, yes=1	0.19***	0.06	0.07	0.31
Salary reduction by 25%, yes=1	-0.35***	0.07	-0.49	-0.21

Source: Authors' calculation based on IFPRI-Digital Green's EA survey, 2019.

Note: triple (***), double (**), and single (*) represent statistical significance at 1%, 5%, and 10% level, respectively. Coefficients and related statistics are calculated with the nlcom command in Stata, based on the 'delta method'.

5.4.4. Heterogeneity in preferences

As highlighted in previous sections, the estimated standard deviation of the random parameters suggests that EAs exhibit significant preference heterogeneity for all the job attributes, i.e., not all EAs attach equal weights to the different job attributes. More precisely, the combination of the estimated means and standard deviations of the random parameters provides information about the proportion of the respondent population that has a positive or negative preference for the job attributes (Train 2009; WHO 2012).[97] The result shows that more than three-quarters of the respondents favour well-connected locations, housing, transportation service, adequate FTCs, and education opportunities. In the latter case, an overwhelming 94 percent of the respondents exhibit a strong preference for the availability of education opportunities. Preference is even less homogenous for upward salary adjustment. While 87 percent of the EAs prefer a 100 percent increment in salary, 13 percent prefer a less sizable increment.

In this section, we assess the sources of the preference heterogeneity by re-estimating equation (5.5) for the sub-samples based on gender, work experience, current salary level, and remoteness of place of work. Table 5.7 presents the results. The differences between the subgroups that are statistically significant are indicated in bold. Columns 1 and 2 show that female EAs are less sensitive to pecuniary incentives compared to their male counterparts. That is, increasing salaries over a certain level is less effective in retaining or incentivizing female EAs. On the other hand, female EAs appear to be more responsive to the provision or availability of transport services in the locality. Perhaps, this is related to security and safety issues, as traveling on foot in a sparsely populated area is considered relatively less secure for female than for male EAs.

Educational opportunities remain a powerful instrument to attract, retain, and motivate EAs. This is particularly the case for relatively younger and newly employed EAs. Columns 3 and 4 of Table

[97] The proportion of the respondent population that has a positive preference for the job attribute (%POS) is calculated as:%POS=$\emptyset(\beta/(SD))$, where β and SD represent the estimated means and standard deviations of each of the random taste parameters, respectively and ∅ is the standard normal cumulative distribution function.

5.7 show that, relative to more experienced EAs, younger EAs (those with less than 3 years of experience as EA) have a weaker preference for salary adjustments. Instead, they show a stronger preference for education opportunities, as well as housing and transport services. The existing incentive structure might explain this. Every two years, evaluation of the performance of EAs are conducted by the Woreda Bureau of Agriculture in order to nominate EAs for promotion (Dufera 2018). Younger and more newly employed EAs, thus, might show stronger motivation in early periods of their employment. Gradually, work burdens and frustrations escalate, and they tend to become responsive principally to short term pecuniary incentives. Alternatively, the results might be explained by life-cycle effects, where older workers value monetary returns higher than non-pecuniary ones simply because they need to sustain their families and are generally less flexible in terms of income generation.

Table 5.7. Preferences for job attributes, sub-sample analysis

	1	2	3	4	5	6	7	8
	Gender		**< 3 yrs. as EA[a]**		**> average salary[b]**		**Remote place[c]**	
	Female	**Male**	**No**	**Yes**	**No**	**Yes**	**No**	**Yes**
Location is advanced, yes=1	0.685***	0.884***	0.810***	0.956***	**0.738*** **	**0.867*** **	0.794***	0.987***
	(0.131)	(0.087)	(0.078)	(0.187)	(0.099)	(0.101)	(0.080)	(0.177)
Housing, yes=1	0.404***	0.560***	**0.452*** **	**0.701*** **	0.609***	0.410***	**0.435*** **	**0.812*** **
	(0.112)	(0.072)	(0.066)	(0.160)	(0.087)	(0.082)	(0.066)	(0.155)
Transport services, yes=1	0.694***	0.661***	**0.643*** **	**0.838*** **	0.766***	0.565***	0.690***	0.566***
	(0.116)	(0.070)	(0.064)	(0.157)	(0.085)	(0.080)	(0.066)	(0.137)
Adequate FTC, yes=1	0.619***	0.779***	0.797***	0.737***	0.695***	0.806***	0.764***	0.666***
	(0.107)	(0.073)	(0.069)	(0.163)	(0.089)	(0.086)	(0.069)	(0.133)
Education opportunity, yes=1	**1.805*** **	**2.069*** **	**2.010*** **	**2.528*** **	**2.191*** **	**1.832*** **	**1.955*** **	**2.304*** **
	(0.172)	(0.113)	(0.105)	(0.263)	(0.141)	(0.127)	(0.101)	(0.234)
Salary (ref: current basic salary)								
Salary increment of 100%, yes=1	**1.405*** **	**1.944*** **	**1.905*** **	**1.165*** **	**1.417*** **	**2.113*** **	**2.019*** **	**0.841** **
	(0.255)	(0.169)	(0.154)	(0.339)	(0.198)	(0.197)	(0.158)	(0.336)
Salary increment of 50%, yes=1	**0.789*** **	**1.080*** **	**1.138*** **	**0.243**	**0.598*** **	**1.361*** **	**1.207*** **	**-0.036**
	(0.247)	(0.161)	(0.148)	(0.344)	(0.194)	(0.188)	(0.150)	(0.337)
Salary increment of 25%, yes=1	0.357	0.403***	**0.488*** **	**-0.080**	**0.060**	**0.752*** **	**0.594*** **	**-0.634** **
	(0.231)	(0.153)	(0.140)	(0.324)	(0.185)	(0.173)	(0.143)	(0.312)
Salary reduction by 25%, yes=1	-0.434	-0.811***	**-0.643*** **	**-1.301*** **	-0.902***	-0.502**	-0.46***	-1.93***
	(0.281)	(0.196)	(0.177)	(0.409)	(0.231)	(0.225)	(0.175)	(0.433)
Constant	-0.160	-0.086	-0.134**	-0.014	0.031	-0.229***	-0.17***	0.211
	(0.104)	(0.066)	(0.062)	(0.135)	(0.079)	(0.079)	(0.062)	(0.131)
Number of respondents	191	570	619	142	378	383	626	135
Number of observations	4,584	13,680	14,856	3,408	9,072	9,192	15,024	3,240
Chi-squared (df = 9)	81	358	373	94	203	236	384	60
Log-likelihood	-919.1	-2,504.1	-2,781.4	-628.9	-1,682.6	-1,725.1	-2,805.8	-602.7
Pseudo R2	0.042	0.067	0.063	0.069	0.057	0.064	0.064	0.047

***Source**: Authors' calculation based on IFPRI-Digital Green's EA survey, 2019.*

***Note**: Standard error given in parenthesis; triple (***), double (**), and single (*) represent statistical significance at 1%, 5%, and 10% level, respectively. The differences between the subgroups that are statistically significant are in bold. [a]<3 yrs. as EA: Those that have worked 3 years or less As EAs. [b] Above average salary is defined as EAs whose current salary is above the median. [c] 'Remote place' indicator identifies villages for which the distance between the Kebele and the district capital is larger than the 80th percentile of the distance distribution.*

The responses of EAs also appear to differ significantly based on the current salary level. The higher the current salary level, the more responsive EAs are to salary adjustments. This is intuitive as salary adjustments in the choice experiment are proposed as a percentage of the current salary level. It is well known that the same amount of money will present a different value depending on the baseline income level. Demand for better infrastructure also interacts positively with the current salary level. On the other hand, the availability of government-provided housing and transportation services is not very effective to motivate top-earning EAs. This perhaps emanates from the fact that better earning EAs could afford to rent housing and transportation services on their own. That is, since salaries of EAs are based mainly on work experience and most experienced EAs work in relatively advanced locations close to the district capital, they often have their own housing or the opportunity to rent decent housing in these locations. It is, however, interesting to note that EAs that earn above-average salaries tend to have a weaker preference for further education. This might speak to the high unemployment among graduates and the low expected return to education in the country (Desalegn 2018). Leaving a reasonably well-paying job to pursue further education for which a return is not guaranteed might not be appealing. This might partly be related to the age of EAs: with age (work experience) salary increases and at the same time, the drive for further education diminishes.

To examine the difference in preferences based on the location of work, we introduce an indicator of remoteness. The Kebele where an EA serves is considered remote if the distance between the Kebele and the district capital is larger than the 80th percentile of the distance distribution.[98] Columns 7 and 8 of Table 5.7 show that EAs in remote locations show a stronger preference for government-provided housing as well as educational opportunities. This can be explained by the relatively thin house rental market in remote areas as well as the lack of adequate transportation facilities for daily commuting from workplace to residences located outside of the Kebele. On the other hand, EAs in more connected areas show strong preferences for salary adjustments. This can be explained by the relatively higher cost of living in more connected areas.

We further considered two additional sample splits based on the current level of education and type of motivation[99]. The result is presented in Table A5.4 in the appendix. The disaggregated result by education reveals that less-educated EAs are less satisfied with the terms of their current employment and more sensitive to pecuniary incentives compared to those with advanced education. On the other hand, EAs with advanced education are more sensitive to location amenities, and availability of housing and further education opportunities. The result about education may appear counterintuitive at first sight. However, it is to be noted that government-sponsored education opportunities are relatively more available for those with a Diploma education (to pursue a degree program) than for those with first-degree education (to pursue masters-level education).

Similarly, columns 3 and 4 of Table A5.4 in the appendix compare the responsiveness of intrinsically motivated (motivated by helping farmers) and extrinsically motivated (motivated by factors external to the job) EAs. It shows that intrinsically motivated EAs are more satisfied with the terms of their current employment and less sensitive to location, and availability of educational

[98] See Abate et al. (2020) and Minten, Koru, and Stifel (2013) for similar definition of remote kebeles.

[99] We are grateful to an anonymous reviewer for these insightful suggestions.

opportunities. On the other hand, extrinsically motivated EAs are strongly responsive to the prospect of a downward adjustment in their salaries.

5.5. Discussion and conclusion

5.5.1. Discussion

The results of the choice experiment indicate that the six selected job attributes are statistically significant, implying that EAs are willing to trade pay raises for other job attributes relating to improved living and working conditions. Preferences for further education and infrastructure are particularly strong. Given that improvements in the education of the EAs and in infrastructure are beneficial not only to retain the EAs but also, indirectly, for the productivity of farmers, these appear to be worthwhile investments for the government[100]. These findings are of pragmatic nature in low-income country settings where wage improvement possibilities are limited because governments lack the required fiscal space, especially since salary adjustments for one group of public sector employees are likely to trigger similar demands from other groups[101].

However, to develop concrete policy recommendations, supplementary studies are needed. The role of qualitative information in this regard can be important. To partly address this, we analyse responses from three open-ended questions: (1) the perception of EAs regarding factors that hinder effective extension delivery, (2) factors that motivate EAs in their extension work, and (3) changes EAs suggest to make extension delivery more effective. Table 5.8 asserts that investment in infrastructure, or the lack thereof, is a considerable impediment to effective extension service in rural Ethiopia.

100 One concern is that the education opportunity might not improve the retention of EAs as it might raise their employability outside of agricultural extension (Becker 1994). However, it is important to note that government sponsored educational opportunity are related to agricultural extension. While we do not completely rule out the possibility that the opportunity might increase outside employment options for EAs, this is likely to be lower than for general education (e.g., studying non-agriculture related course).

101 This last remark on government fiscal space and interdependence in salaries of public sector employees emanates from the fact that currently, the vast majority of extension services are provided by the public sector. In the likely and positive future scenario that extension services fall to non-state actors, this remark might not hold any longer but the conclusion about potential synergy from investing on education and infrastructure remains valid.

Table 5.8. Challenges, motivation factors, and suggestions by EAs

Panel A: Challenges faced in conducting extension service	**% EAs**
Poor infrastructure (drinking water, food, electricity, transportation, etc.)	62
Lack of housing	18
Poorly equipped Farmer Training Centre (demonstration plots, ICT tools, budget, etc.)	39
Lack of education opportunities and short-term training	6
Low salary (low basic salary and no or inadequate allowance)	19
Workload (long working hours and long work week)	15
Farmers' resistance (low adoption, low attendance at meetings, etc.)	35
Management approach (poor incentive structure, obsolete extension system, etc.)	45
Extension Agents reporting, no.	722
Panel B: Factors that motivate Extension Agents in conducting extension work	
Nothing	35
On job training	5
Management support (recognition, fair promotion, etc.)	10
Desire to change farmers' lives	49
Active participation of farmers (attentiveness, attendance in meetings, etc.)	14
Interest in agricultural extension (love for profession)	9
Extension Agents reporting, no.	764
Panel C: Suggested changes to make agricultural extension more effective	
Improve infrastructure (improve road quality, provide motorcycles or bicycles, etc.)	65
Provide housing	30
Equip Farmer Training Centres for effective extension	44
Provide educational opportunities, regardless of current level of education	27
Upgrade salary structure to reflect living cost and job market	67
Reduce workload (less frequent reporting, free weekends, etc.)	8
Modify management approach (upgrade extension system, create transparent incentive structure)	46
Extension Agents reporting, no.	718

***Source**: Authors' calculation based on IFPRI-Digital Green's EA survey, 2019.*

***Note**: Since multiple responses were allowed in the survey, the sum of response percentages may exceed 100.*

It is also important to underline that improvements in identified attributes are helpful primarily in reducing EA turnover and job demotivation, but not to spur performance. Factors related to working conditions, i.e., infrastructure, housing, transport service, and FTC materials, and the package of benefits EAs receive, i.e., salary and educational opportunities, are extrinsic to the job. Consequently, their improvement can only partially increase job motivation. According to Herzberg (1987), there are additional factors that lead to work motivation, which are intrinsic to the job. These include love for the profession, desire to make a difference through the job, and desire to advance (through training). Table 5.8 shows that this is indeed the case for EAs in Ethiopia. Close to half (49 percent) of the EAs indicated that the desire to change farmers' living conditions through extension service is their main motivating factor. Other important motivating factors to the EAs include their training and advisory getting accepted by farmers, support from management, the work itself, and access to training (panel B of Table 5.8). This suggests that, in the long run, motivating EAs and, hence, improving their performance requires interventions that are designed and targeted to make extension work more productive and responsive to the demands of farmers. Disengaging EAs from non-extension duties (e.g., tax collection, Kebele administration, promoting political views, etc.) and equipping FTCs with adequate materials (e.g., ICT tools, demonstration plots, training materials, etc.) are examples of changes in the right direction (Table 5.8). This is consistent with findings from other studies (e.g., Berhane et al. 2018; Ragasa et al. 2016).

That said, further experiments that include additional attributes from this qualitative assessment would offer additional insights. The six attributes we considered in the choice experiment represent only a subset of many possible attributes that affect the job choices of EAs. Table 5.8 clearly shows that to most EAs, workload, management practices, and farmers' interest in extension (meetings) are hurdles to effective extension dissemination that require meaningful improvement. While we acknowledge the importance of these remaining attributes and that further research on these attributes would be valuable to inform policy, our basic conclusion regarding the included attributes remains valid. The experimental design we used ensures a valid trade-off among the included attributes, assuming that all relevant excluded attributes remain the same between alternatives (Chomitz et al. 1998; Scott 2001).[102]

Future research can also estimate the interaction effects of attributes beyond the main effects of attributes. Our analysis focused only on the estimation of main effects, i.e., the independent effect of each attribute level on the preference of the EAs. Results from interacting attributes might produce interesting insights regarding synergies in multiple interventions. For instance, it might be the case that preferences for government-provided housing are determined by the location of work. Unfortunately, the estimation of such interaction terms requires a larger number of choice sets to be presented to the respondents (a full factorial design). In this study, we opted to use a fractional factorial design and present eight choice sets to each EA to minimize fatigue and the cognitive burden on respondents. While this design is simple, it is realistic enough to provide relevant policy predictions regarding the effect of policy interventions under the selected attributes (De Bekker-Grob, Ryan, and Gerard 2008; Kuhfeld 2010; WHO 2012). In describing the preferences of employees for job attributes, the main effects are argued to explain most of the variation in preferences (De Bekker-Grob et al. 2008; WHO 2012).

The hypothetical nature of the questions in the choice experiment might be unconvincing as respondents may over or understate their true preferences systematically. This potential bias is often magnified by poor survey administration, including in the selection and training of enumerators, the discussions held with key informants, and pre-testing. While we cannot entirely rule out the possibility of respondent bias, our survey was subjected to a rigorous preparation process to minimize such bias. For studies in which the experimental design was carefully constructed, the stated preferences of respondents were found to closely resemble their revealed preferences (De Bekker-Grob et al. 2008; Lusk and SChroeder 2004; WHO 2012).

[102] In the survey, each respondent was explicitly informed to assume that all unstated characteristics of jobs are the same for the presented alternatives (see Table A2).

5.5.2. *Conclusion*

Low productivity is a considerable hurdle for poverty reduction in rural areas of developing countries and it exacerbates spatial imbalances in welfare (Minten et al. 2013; Stifel and Minten 2017). Agricultural extension agents (EAs) could contribute to reducing this spatial imbalance by promoting the use of modern technologies and production methods, thereby increasing agricultural productivity in remote locations (Dercon et al. 2007; Stifel and Minten 2017). However, studies consistently show that farmers' access to extension services is limited in such locations (Abate et al. 2020; Brinkerhoff, Wetterberg, and Wibbels 2018; World Bank and IFPRI 2010). Extension offices are often understaffed and the quality, motivation, and effort level of EAs are dubious.

In this study, we designed a choice experiment with the objective to inform policy interventions on how to abate the high turnover of EAs and mitigate geographical imbalances in the number and quality of EAs. We employ a random parameter logit model (RPL), which in conjunction with the choice experiment data, allows statistical flexibility and avoids the limitations of using cross-sectional and repeated revealed preference data that are susceptible to endogeneity and selection issues.

We find that offering continuing education opportunities after two years of service is one of the most powerful incentive instruments available to policymakers to attract, motivate, and retain EAs in rural areas. Increasing salaries and offering decent housing and transportation facilities are also effective incentives to EAs, but not as much as offers of further educational opportunities. Good infrastructure, including improved access to electricity and mobile telephone networks in the Kebeles in which EAs are posted, as well as equipping Farmer Training Centres (FTC) are also interventions to which EAs are highly likely to respond.

The sub-sample analysis shows that the preferences of EAs for job attributes vary considerably based on gender, age, current salary level, and place of work. In general, male and experienced EAs, as well as those in more connected areas strongly prefer increased salaries. On the other hand, less experienced EAs and those in remote locations have stronger preferences for further educational opportunities. Overall, these results highlight the importance of accounting for EA sociodemographic factors when designing policy interventions intended to attract, retain, and motivate EAs.

Finally, the cost-effectiveness of the alternative interventions should be considered before concrete policy decisions are made. To this end, future studies might estimate the cost of the different interventions and combine that information with computed elasticities to assess the cost-effectiveness of each intervention. This is crucial for determining both the effectiveness of the interventions and their efficacy.

CHAPTER **6**

6. Summary and contributions to research

6.1. Summarized findings

Broadly, this thesis deals with how urbanization determines the spatial pattern of economic development. Specifically, it investigates if and how proximity to and the size of urban areas influences household welfare in sub-Saharan Africa with data from Ethiopia. Previous empirical studies establish that households in rural areas are, in general, poorer, less productive, and more susceptible to price risks than households in urban areas (Fafchamps and Shilpi 2002; Melesse and Cecchi 2017; Stifel et al. 2003). However, this broad conclusion conceals two key points that have become more evident due to the recent pace and pattern of urbanization in the region. **First**, rural and urban areas are not distinct spaces. Rapid urbanization, improvements in infrastructure networks, and developments in information and communication technologies have blurred the distinction between the two spaces. Now, it is widely acknowledged that rural and urban areas coexist along a continuum with many in-between stages (von Braun 2014b; Satterthwaite and Tacoli 2003).

Second, while urban areas are generally growing, small- and intermediate- urban areas are growing more rapidly in Africa (Dorosh and Thurlow 2013; UNDESA 2015). The current statistics of African urbanization show that about 90 percent of the African urban population resides in cities of less than 5 million inhabitants. Furthermore, the population in these urban areas has doubled in the last decade and is expected to grow by more than 30 percent over the next decade (UNDESA 2015). This pattern has intensified interest into the effect of the nature of urbanization. In particular, a disaggregated study of urbanization over different stages has attracted significant attention. Recently conducted empirical studies along these lines reveal that urban areas are not homogenous and that different urban areas can have different degrees of influence on their surrounding population (Christiaensen and Kanbur 2017; Christiaensen, De Weerdt, and Todo 2013; Vandercasteelen et al. 2018).

Therefore, an empirical study that deals with the effect of urbanization and a rural-urban linkage require an objective measure of the level and dynamics of urbanization. This thesis addresses this issue. It employs a continuous measure of urbanization – Sum of Nighttime Light (SOL) – to account simultaneously for the continuum between rural and urban areas as well as the heterogeneity of urban areas.

One of the central focuses of the thesis is the analysis of underlying mechanisms of the spatial economy. Empirical studies have long-identified a substantial welfare loss associated with remoteness relative to the market or urban areas (Collier and Gunning 1999; David et al. 1998; Kraay and McKenzie 2014; Sachs, Mellinger, and Gallup 2001). However, mechanisms through which this remoteness translates into poorer welfare outcomes have not been explored adequately over the entire rural-urban spectrum. Previous studies from developing countries focused on spatial differences among rural households based on disparities in physical and human capital (Sahn and Stifel 2004; Simler and Dudwick 2009); input use and yield level (Stifel et al. 2003); and access to markets and prices (Chamberlin and Jayne 2013; Melesse and Cecchi 2017). Throughout this thesis, the emphasis is on factors that are Pareto improving - factors that could reduce spatial disparity across the rural-urban spectrum, while also improving the overall welfare of the populations. By examining one of the less studied fundamental factors of spatial development – *public service delivery* - this thesis identifies policy recommendations to address the high turnover and low motivation among agricultural extension agents in remote areas.

To this end, the thesis is organized under the following four Analytical chapters. **Chapter 2**, entitled: "*Patterns of urbanization and household welfare*" focuses on identifying whether and how urbanization and its different stages in Ethiopia are associated with household welfare. The primary data used in this chapter comes from two rounds of LSMS-ISA[103] (2014 and 2016) data which are geo-spatially linked to nightlight data. The findings of this chapter, based on the New Economic Geography (NEG) framework and threshold data analysis, suggest that the implications of the patterns of urbanization are at least as important as the aggregate rate of urbanization. In general, it indicates that intermediate towns are more strongly associated with household welfare as compared to large towns, small towns, or the rural hinterland. The Chapter concludes by emphasizing the roles of market access, employment opportunities, and differential access to public services as major underlying mechanisms.

Chapter 3, entitled: "*Heterogeneous effects of urban proximity on nutritional outcomes*" extends the analysis in Chapter 2 and discusses the effect of the distance to-and the size of - the proximate urban areas on health and nutrition outcomes. Sub-Saharan Africa (SSA) countries are becoming urbanized at an unprecedented fast rate. While this trend has the potential to significantly improve household nutritional status, the underlying mechanisms are not well understood. It is unclear whether and why the effect of proximity to different sized towns on nutrition outcomes varies. This chapter addresses this question by simultaneously examining the effect of proximity to urban areas and the heterogeneous effect of city size on the nutritional status of households in and surrounding urban areas. For identification, an Instrumental Variables (IV) approach is combined with Inverse Probability Weighting (IPW). While the IV approach accounts for the potential endogeneity of transportation costs, the IPW addresses the bias that results from self-selection into the place of residence. Using three rounds of nationally representative LSMS-ISA household and community survey data, the study finds that both the proximity to urban areas and the size of the proximate urban areas affect households' nutritional status. Specifically, while proximity to towns has a strong positive effect on nutritional status, households surrounding intermediate- and large- towns are better off compared to those around small towns.

Furthermore, the chapter identifies several potential mechanisms that may explain why proximity to intermediate and large towns is more likely to improve nutrition outcomes than proximity to small towns. It finds that there is considerable spatial disparity among households in terms of wealth, human capital endowment; access to Water, Sanitation, and Hygiene (WASH) facilities; public services; and employment opportunities. It implies that policy interventions that target improving overall nutritional status as well as reduce the spatial imbalance, need to address access and quality issues in these services in rural areas and smaller towns.

Chapter 4, entitled: "*Urbanization and Intergenerational mobility in Ethiopia*" examines the effect of urbanization on the inequality of opportunities among the current and future generations. Using nationally representative longitudinal data on children and their parents, it investigates the extent of equality of occupational opportunity across rural-urban areas in Ethiopia. The chapter's major

[103] Ethiopian Living Standard Measurement Survey-Integrated Survey of Agriculture (LSMS-ISA)

findings are summarized below. **First**, urbanization is associated strongly and positively with both the quality of and inequality in, occupational status. **Second**, inequality in occupational status transmits across generations due to the strong child-parent correlation in occupation. **Third**, inequality and intergenerational dependence in occupational status are stronger in large urban areas than in rural or small towns. **Fourth**, the inequality observed in occupational opportunities in large urban areas is explained mainly by differences in educational attainment above the elementary school level. Once individual education level is accounted for, large urban areas offer better mobility in employment status. This suggests that expanding access to - and lowering the dropout rates at - post-elementary schools in addition to improving the quality of education, is one of the most effective mechanisms for reducing spatial and intergenerational disparity in welfare. A comprehensive set of potential policy interventions are identified to reduce the inequality in opportunities for the current and future generations.

Chapter 5, entitled "*Incentivizing and Retaining Public Servants in Remote Areas: A discrete choice experiment with agricultural extension agents in Ethiopia*" deals with the geography of public services in Ethiopia and what needs to be done to make it more inclusive. Agricultural extension agents (EAs) are deployed in rural areas to spur agricultural productivity and mitigate spatial imbalances in welfare. However, high turnover and low motivation levels of EAs in remote areas pose challenges for equitable service provision and, in some cases, exacerbate geographical welfare disparities. The chapter assesses the effectiveness of selected potential policy interventions to incentivize and retain EAs in remote areas of Ethiopia. To this end, a choice experiment was conducted to elicit the preferences of 761 EAs for job attributes. A random parameters logit model was then applied to estimate parameters of interest and to simulate the impact of possible policy interventions. The results show that offering education opportunities far exceeds all other job incentives to attract and retain EAs. It increases the job uptake in remote locations by 77 percentage points, which is significantly higher than the effect of doubling current salary levels. EAs also expressed a strong preference for work environments with basic amenities, housing, transportation services, and well-equipped Farmer Training Centres (FTCs). Furthermore, the results from sub-sample analyses show that female EAs are less responsive to pecuniary incentives and are more concerned with the availability of infrastructure and services. Current salary levels, years of employment, and location of work are also important sources of heterogeneity in the response of EAs to potential policy changes.

The overarching principal finding from all the chapters is that while there is a considerable rural-urban gap in living standards, smaller urban centres fare worse across the urban spectrum. Compared to intermediate- and large- towns, rural areas, and small towns are at a disadvantage in terms of consumption per capita and food security (**Chapter 2**); diet diversity and child nourishment (**Chapter 3**); intergenerational mobility (**Chapter 4**); and in public service delivery (**Chapter 5**). These differences persist even after accounting for differences in human capital endowment, wealth, and sociodemographic factors. The studies further demonstrate that these spatial disparities in living standards are underlined by widespread differences in access to basic public services and employment opportunities. From a policy perspective, the broader implication is, interventions that target to improve overall welfare as well as reduce the spatial imbalance need to remove the constraints facing isolated households in rural areas and smaller towns.

Accordingly, this thesis has identified a set of place-based policy recommendations that broadly align with their degree of urbanization and the level of economic development.

In mostly rural areas, the focus of policies should be to continue enhancing the performance of the agricultural sector and to improve the interlinkage between rural and the surrounding urban areas. It is important to acknowledge that even if the contribution of agriculture in national output and employment is bound to decline with urbanization, the roles of agricultural jobs and income remain highly significant in rural and small towns in Ethiopia and beyond (Mellor 2018). Therefore, policy reforms should reflect the importance of the sector while taking into account the increasing importance of the urban areas. To ensure this, one approach is to reform the national agricultural policy so that the focus is to improve the productivity of the entire agricultural value chains rather than merely the yield level of small-scale farmers. In this regard, promoting market linkage and commercialization of agricultural goods is critical for boosting productivity, food security, and employment opportunities for both rural and urban households (von Braun 2007; Collier and Dercon 2014; OECD/PSI 2020) – ***Chapter 5***. Institutional reforms pertaining to land and labour are important to reduce the cost of migration and facilitate social mobility (***Chapter 4***). Policy reforms should also seek to increase investment in infrastructure, and social services - education and health (World Bank 2009) – ***Chapters 2 and 3***.

Small and intermediate urban areas have a huge potential to generate employment opportunities for both rural and urban dwellers, contribute to overall poverty reduction, and help balance the urban system (Satterthwaite and Tacoli 2003). To exploit this potential, the government of Ethiopia (GoE) has been promoting the development of Integrated Agro-Industrial Parks (IAIP) in small and intermediate urban areas[104]. Together with the conducive business environment that these urban spaces offer, and the ever-expanding interregional road network, these policies are generating more jobs than in the capital city (OECD/PSI 2020). However, these additional job opportunities have not been sufficient to absorb the surplus labour from the surrounding areas and the increasing population in these agglomerations. The lack of adequate public services and connective infrastructure is also limiting their potential to function as urban growth poles (***Chapters 2, 3, 4, and 5***).

As shown in ***Chapters 2, 3, and 4***, small and intermediate towns in Ethiopia are facing acute challenges in terms of access to drinking water, electricity, sanitation facilities, health posts, and schools. Unless proper measures are taken, these problems will worsen due to the expected population growth in these locations by threefold over the next decade (OECD/PSI 2020; UNDESA 2015; World Bank and Cities Alliance 2015). To utilize the potential of these locations, mitigate their current challenges, and accommodate for increasing demand in services, the government needs to put in place an extensive institutional reform as well as investment in connective infrastructure. These reforms might include: improving the business climate, reducing conflicts, maintaining macroeconomic stability, along with generous tax incentives. All these potential reforms could attract the desired private investments into these locations.

In large urban areas, the major challenges stem from congestion and economic disparity. While there are encouraging initiatives in Ethiopia to deal with these issues ranging from urban safety

[104] Currently IAIP are located in Mekele, Bahir Dar, Dessie-Kombolcha, Hawassa, Adama, Dire Dawa and Jimma.

net programs to the promotion of Micro and Small Enterprise (MSEs), large urban areas continue to face chronic challenges in relation to unemployment, housing, and social services. Therefore, in addition to institutional reforms and improvements in infrastructure, the government should endorse more aggressive policy interventions that target the most vulnerable residents in these locations to make urbanization more inclusive. These are broad policy suggestions. The choice of specific interventions within these broad categories should be made based on careful evaluation of their effectiveness as well as cost-benefit analysis[105].

6.2. Contributions to research

This thesis contributes to several strands of literature. **First**, it contributes to the literature on the measurement of urbanization and urban influence. The lack of an objective, robust, and disaggregated measure of the level and dynamics of urbanization has limited rigorous empirical evaluation of the impact of urbanization (Satterthwaite and Tacoli 2003). The conventionally used, survey and census-based aggregate rural-urban indicators of urbanization are often only sporadically available, unreliable, and lag behind reality (Bennett and Smith 2017; Donaldson and Storeygard 2016). This study uses an indicator of urbanization that is based on the *nighttime light composite index* (NTL). Since the NTL data is available at a high spatial resolution over a sufficiently long time period, it helps to capture both spatial as well as intertemporal urban dynamics. It enables the creation of a universally comparable, continuous, and disaggregated index representing micro-level variations in urban settlement and urban expansion (Henderson et al. 2009). Due to this, NTL particularly holds a huge potential for sub-Saharan African countries when other, more conventional methods of urbanization measurement fall short, as described earlier.

Second, the thesis expands on the use of NTL by measuring the influence of urban areas based on the sum of Nighttime Lights (SOL) – the sum of NTL within a 10km buffer zone around each village. Compared to the traditional census-based approach to urbanization measure or a simple NTL, the SOL method has a number of advantages. As it considers the effect of all potential urban centres, it addresses one of the critical shortcomings of the traditional approach where urban influence is measured with respect to the nearest town (Gibson et al. 2017; Henderson et al. 2017). The approach also reduces the possibility of incorrectly categorizing villages into rural-urban categories. In order to ensure the confidentiality of sample households and communities, GPS information of sample villages was modified in the publicly available LSMS-ISA data from the original levels by applying a random offset of up to 10km[106]. If the NTL at a single point were used to measure urbanization status, it might lead to misclassification of a large number of villages. The 10km buffer zone that is created to delineate urban areas eliminates any potential misclassification resulting from the random offsets.

Third, the thesis contributes to the use of a statistical method to classify urban spaces into small, intermediate, and large towns. Conventionally, population size or administrative roles are used to

[105] For instance, a World Bank study shows that, compared to consumption subsidies (e.g., for electricity, wheat, kerosene, etc.), targeted social programs are much more effective in terms of both the impact and cost effectiveness to address the needs of the most vulnerable in large urban areas in Ethiopia (World Bank 2016).

[106] See https://microdata.worldbank.org/index.php/catalog/2783

classify urban areas by size (OECD/PSI 2020; Roberts 2014a). While these conventional methods tend to be highly subjective, they often require frequent revision and lack comparability across countries or even across regions within the same country (Satterthwaite and Tacoli 2003; Tacoli 1998). In Chapters 2 and 4, a threshold estimation technique developed by Hansen (2000) is applied to SOL to classify sample villages into rural areas, small towns, intermediate towns, and large towns. While the consistency of this approach with the conventional methods is verified, the method holds huge potential to allow a more rigorous cross-country evaluation of the effect of urbanization and urban concentration.

Fourth, Chapter 5 adds to the use of a choice experiment method to elicit the job preference of public workers in developing countries. While the application of choice experiment surveys to elicit preferences is common in the environmental, health, marketing, and transport literature, its application to the evaluation of employees' preferences for job attributes is nascent (Chomitz et al. 1998; Hanson and Jack 2008; Kolstad 2011; Scott 2001). This study is one of the first, if not the first, to apply a discrete choice experiment (DCE) design to elicit the job preferences of Agricultural Extension Agents (EAs). Moreover, the simulated maximum likelihood method used in the estimation of the preferences for job attributes, and the estimation of the Willingness to Pay (WTP) in the willingness of pay space are novel econometric methods.

Finally, the thesis adds to the scant literature on the microeconomic application of New Economic Geography (NEG) in developing countries. The introductory chapter highlights the basic framework and predictions of the NEG models. It demonstrates how the change in spatial integration cost (i.e. information costs, transport costs, and tariff and non-tariff barriers) leads to a spatial pattern of economic development based on baseline density. It then examines how this basic prediction is altered leading to a bell-shaped pattern of economic development after accounting for congestion and preference heterogeneity (see Chapter 1). The evolution of urban development, the spatial distribution firms, and the geography of household welfare in Ethiopia are examined within this framework.

References

Abate, Gashaw T., Mekdim Dereje, Kalle Hirvonen, and Bart Minten. 2020. “Geography of Public Service Delivery in Rural Ethiopia.” *World Development* 136(July):105133. doi: 10.1016/j.worlddev.2020.105133.

Abay, Kibrewossen, and Kalle Hirvonen. 2017. “Does Market Access Mitigate the Impact of Seasonality on Child Growth? Panel Data Evidence from Northern Ethiopia.” *The Journal of Development Studies* 53(9):1414–29.

Abay, Kibrom A., Luca Tiberti, Tsega G. Mezgebo, and Meron Endale. 2020. *Can Urbanization Improve Household Welfare? Evidence from Ethiopia*.

African Development Bank. 2011. *Transforming Africa’s Cities and Towns into Engines of Economic Growth and Social Development*.

Ahmed, R., and M. Hossain. 1990. *Developmental Impact of Rural Infrastructure in Bangladesh*.

Alderman, Harold, Simon Appleton, Lawrence Haddad, Lina Song, and Yisehac Yohannes. 2001. *Reducing Child Malnutrition: How Far Does Income Growth Take Us?* CREDIT Research Paper 01/05. CREDIT Research Paper 01/05, Nottingham, United Kingdom.

Alderman, Harold, and Derek Headey. 2017. “How Important Is Parental Education for Child Nutrition?” *World Development* 94:448–64.

Alesina, Alberto F., Sebastian Hohmann, Stelios Michalopoulos, and Elias Papaioannou. 2019. *Intergenerational Mobility in Africa*. Vol. 7. Centre for Economic Policy Research DP13497.

Ali, Daniel Ayalew, Klaus Deininger, and Marguerite Duponchel. 2014. *Credit Constraints, Agricultural Productivity, and Rural Nonfarm Participation: Evidence from Rwanda*. Policy Research Working Paper 6769.

Amare, Mulubrhan, Channing Arndt, Kibrom A. Abay, and Todd Benson. 2017. *Urbanization and Child Nutritional Outcomes*. Nigeria Strategy Support Program Working paper 49. International Food Policy Reserach Institute (IFPRI).

Ameye, H. 2018. “Secondary Towns – The Nutritional Sweet Spot. A Study of East Africa.” in *30th International conference of Agricultural economists*.

Ang, Andrew, Joseph Chen, and Yuhang Xing. 2006. “Downside Risk.” *Review of Financial Studies* 19(4):1191–1239.

Angrist, J., and J. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist’s Companion*. New Jersey, USA: Princeton University Press.

Assy, Angela Elzir, Tiago Ribeiro, David Robalino, Furio Rosati, Maria Laura, and Sanchez Puerta. 2019. *The Jobs That Youth Want and the Support They Need to Get Them : Evidence from a Discrete Choice Experiment in Kenya*. IZA DP No. 12864.

Autor, David. 2005. *Work of the Past, Work of the Future*. Vol. 42. NBER Working Paper No. 25588.

Autor, David. 2020. “The Faltering Escalator of Urban Opportunity.” *Aspen Institute. Research Brief.*

Banerjee, Abhijit V. .., and Andrew F. Newman. 1993. “Occupational Choice and the Process of Development.” *Journal of Political Economy* 101(2):274–98.

Barro, Robert J. 2001. “Human Capital and Growth.” *American Economic Review* 105(5):85–88. doi: 10.1257/aer.p20151065.

Becker, Gary S. 1994. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. Third edit. Chicago, USA: The University of Chicago press.

Becker, Gary S., and Nigel Tomes. 1986. “Human Capital and the Rise and Fall of Families.” *Journal of Labor Economics* 4(3):S1–39.

Beegle, Kathleen, Joachim De Weerdt, and Stefan Dercon. 2011. “Migration and Economic Mobility in

Tanzania: Evidence from a Tracking Survey." *Review of Economics and Statistics* 93(3):1010–33.

De Bekker-Grob, Ester W., Mandy Ryan, and Karen Gerard. 2008. "Discrete Choice Experiments in Health Economics: A Review of the Literature." *Health Economics* 1131(2007):1127–31. doi: 10.1002/hec.

Bennett, Mia M., and Laurence C. Smith. 2017. "Advances in Using Multitemporal Night-Time Lights Satellite Imagery to Detect, Estimate, and Monitor Socioeconomic Dynamics." *Remote Sensing of Environment* 192:176–97. doi: 10.1016/j.rse.2017.01.005.

Berdegué, Julio A., Fernando Carriazo, Benjamín Jara, Félix Modrego, and Isidro Soloaga. 2015. "Cities, Territories, and Inclusive Growth: Unraveling Urban-Rural Linkages in Chile, Colombia, and Mexico." *World Development* 73:56–71.

Berhane, Guush, Catherine Ragasa, Gashaw T. Abate, and Thomas Woldu Assefa. 2018. *The State of Agricultural Extension Services in Ethiopia and Their Contribution to Agricultural Productivity*. ESSP Working paper 118. International Food Policy Research Institute (IFPRI).

Binder, Melissa, and Christopher Woodruff. 2002. "Inequality and Intergenerational Mobility in Schooling: The Case of Mexico." *Economic Development and Cultural Change* 50(2):249–67. doi: 10.1086/322882.

Binswanger-Mkhize, Hans P., and Sara Savastano. 2017. "Agricultural Intensification: The Status in Six African Countries." *Food Policy* 67:26–40. doi: 10.1016/j.foodpol.2016.09.021.

Birner, Regina, Kristin Davis, John Pender, Ephraim Nkonya, Ponniah Anandajayasekeram, Javier Ekboir, Adiel Mbabu, David Spielman, Daniela Horna, Samuel Benin, and Marc Cohen. 2009. "From Best Practice to Best Fit: A Framework for Designing and Analyzing Pluralistic Agricultural Advisory Services Worldwide." *The Journal of Agricultural Education and Extension* 15(4):341–55.

Black, Sandra E., and Paul J. Devereux. 2010. *Recent Developments in Intergenerational Mobility*. NBER working paper series 15889. Cambridge, MA.

Blackburn, Mckinley L., and David Neumark. 1993. "Omitted-Ability Bias and the Increase in the Return to Schooling." *Journal of Labor Economics* 11(3):521–44.

Bloom, David E., David Canning, Günther Fink, Tarun Khanna, and Patrick Salyer. 2010. *Urban Settlement: Data, Measures, and Trends*. WIDER Working Paper 2010/12. Helsinki, Finland.

von Braun, J. 2007. "Rural-Urban Linkages for Growth, Employment, and Poverty Reduction." in *Fifth International Conference on the Ethiopian Economy*. International Food Policy Research Institute Washington, D.C., USA.

von Braun, J., John Mccomb, Ben K. Fred-Mensah, and Rajul Pandya-Lorch. 1993. *Urban Food Insecurity and Malnutrition in Developing Countries: Trends, Policies, and Research Implications*. IFPRI, Washington D.C.

von Braun, Joachim. 2014a. "Guiding Urban – Rural Linkages toward Sustainable Development." Pp. 75-95. in *Contributions Towards a Sustainable World In Dialogue with Klaus Töpfer*, edited by F. Schmidt and N. Nuttall. München.

von Braun, Joachim. 2014b. "Urbanization and Decentralization:The Changing Urban-Rural Linkages and Opportunities of Decentralization of Services." in *Regional development & globalization: Best practices*. Center for Development Research (ZEF) University of Bonn, Germany.

von Braun, Joachim, and Chiara Kofol. 2017. *Expanding Youth Employment in the Arab Region and Africa*. ZEF Working Paper 155. Center for Development Research, University of Bonn, Germany.

Brinkerhoff, Derick W., Anna Wetterberg, and Erik Wibbels. 2018. "Distance, Services, and Citizen Perceptions of the State in Rural Africa." *Governance* 31(1):103–24. doi: 10.1111/gove.12271.

Bronars, Stephen G., and Gerald S. Oettinger. 2006. "Estimates of the Return to Schooling and Ability: Evidence from Sibling Data." *Labour Economics* 13(1):19–34. doi: 10.1016/j.labeco.2004.07.003.

Broussar, Nzinga H., and Tsegay Tekleselassie. 2012. *Youth Unemployment: Ethiopia Country Study*. IGC Working paper. International Growth center, Addis Ababa, Ethiopia.

Brutzkus, Eliezer. 1973. “Centralized Versus Decentralized Pattern of Urbanization in Developing Countries: An Attempt To Elucidate a Guideline Principle.” *Economic Development and Cultural Change* 64(1):11–23.

Cali, Massimiliano, and Carlo Menon. 2009. *Does Urbanisation Affect Rural Poverty? Evidence from Indian Districts*. 14.

Cameron, A., and P. Trivedi. 2005. *Microeconometrics: Methods and Applications*. Cambridge: Cambridge University Press.

Cameron, Stephen V, and James J. Heckman. 1993. “The Nonequivalence of High School Equivalents.” *Journal of Labor Economics* 11(1):1–47.

Carmichael, Fiona. 2000. “Intergenerational Mobility and Occupational Status in Britain.” *Applied Economics Letters* 7(6):391–96. doi: 10.1080/135048500351339.

Caudill, Steven B. 1988. “Practitionars Corner: An Advantage of the Linear Probability Model over Probit or Logit.” *Oxford Bulletin of Economics and Statistics* 50(4):425–27.

Cawley, John, James Heckman, and Edward Vytlacil. 2001. “Three Observations on Wages and Measured Cognitive Ability.” *Labour Economics* 8(4):419–42. doi: 10.1016/S0927-5371(01)00039-2.

Chamberlin, Jordan, and T. S. Jayne. 2013. “Unpacking the Meaning of ‘Market Access’: Evidence from Rural Kenya.” *World Development* 41(1):245–64.

Checchi, Daniele. 1997. “Education and Intergenrational Mobility in Occupations: A Comparative Study.” *American Journal of Economics and Sociology* 56(3).

Checchi, Daniele, Andrea Ichino, and Aldo Rustichini. 1999. “More Equal but Less Mobile? Education Financing and Intergenerational Mobility in Italy and in the US.” *Journal of Public Economics* 74(3):351–93. doi: 10.1016/S0047-2727(99)00040-7.

Chen, Xi, and William Nordhaus. 2015. “A Test of the New VIIRS Lights Data Set: Population and Economic Output in Africa.” *Remote Sensing* 7(4):4937–47. doi: 10.3390/rs70404937.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. “Where Is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States.” *The Quarterly Journal of Economics* 129(November):1553–1623. doi: 10.1093/qje/qju022.Advance.

Chomitz, K., G. Setiadi, A. Azwar, N. Ismail, and Widiyarti. 1998. *What Do Doctors Want? Developing Incentives for Doctors to Serve in Indonesia’s Rural and Remote Areas*. Policy Research Working Paper 1888.

Christiaensen, Luc, Lionel Demery, and Jesper Kuhl. 2011. “The (Evolving) Role of Agriculture in Poverty Reduction—An Empirical Perspective.” *Journal of Development Economics* 96(2):239–54. doi: 10.1016/j.jdeveco.2010.10.006.

Christiaensen, Luc, and Ravi Kanbur. 2017. *Secondary Towns and Poverty Reduction: Refocusing the Urbanization Agenda*. Vol. 9. Policy Research Working Paper 7895. World Bank, USA.

Christiaensen, Luc, and Yasuyuki Todo. 2014. “Poverty Reduction during the Rural-Urban Transformation - The Role of the Missing Middle.” *World Development* 63:43–58. doi: 10.1016/j.worlddev.2013.10.002.

Christiaensen, Luc, Joachim De Weerdt, and Ravi Kanbur. 2013. *Urbanization and Poverty Reduction: The Role of Secondary Towns in Tanzania*.

Christiaensen, Luc, Joachim De Weerdt, and Yasuyuki Todo. 2013. “Urbanization and Poverty Reduction: The Role of Rural Diversification and Secondary Towns.” *Agricultural Economics* 44(4–5):435–47.

Collier, Paul, and Stefan Dercon. 2014. “African Agriculture in 50 Years: Smallholders in a Rapidly Changing World?” *World Development* 63(June 2009):92–101.

Collier, Paul, and Jan Willem Gunning. 1999. “Why Has Africa Grown Slowly?” *The Journal of Economic Perspectives,* 13(3):3–22.

Combes, Pierre Philippe, Gilles Duranton, and Laurent Gobillon. 2008. “Spatial Wage Disparities: Sorting

Matters!” *Journal of Urban Economics* 63(2):723–42. doi: 10.1016/j.jue.2007.04.004.

CSA and World Bank. 2017. *Ethiopia Socioeconomic Survey (ESS) Wave Three (2015/2016): Basic Information Document*.

CSA, and ICF. 2016. *Ethiopia Demographic and Health Survey 2016*. Vol. Volume 8. ddis Ababa, Ethiopia, and Rockville, Maryland, USA.

Dang, Hai Anh, Dean Jolliffe, and Calogero Carletto. 2019. “Data Gaps, Data Incomparability, and Data Imputation: A Review of Poverty Measurement Methods for Data-Scarce Environments.” *Journal of Economic Surveys* 33(3):757–97. doi: 10.1111/joes.12307.

David, Bloom, Sachs Jeffrey, Paul Collier, and Christopher Udry. 1998. “Geography, Demography, and Economic Growth in Africa.” *Brookings Papers on Economic Activity* 1998(2):207–95.

Davis, Kristin. 2008. “Extension in Sub-Saharan Africa: Overview and Assessment of Past and Current Models, and Future Prospects.” *Journal of International Agricultural and Extension Education* 15(3):15–28.

Davis, Kristin, E. Nkonya, E. Kato, D. A. Mekonnen, M. Odendo, R. Miiro, and J. Nkuba. 2012. “Impact of Farmer Field Schools on Agricultural Productivity and Poverty in East Africa.” *World Development* 40(2):402–13.

Davis, Kristin, Burton Swanson, and David Amudavi. 2009. *Review and Recommendations for Strengthening the Agricultural Extension System in Ethiopia*. IFPRI Report.

Davis, Kristin, Burton Swanson, David Amudavi, Daniel Ayalew Mekonnen, Aaron Flohrs, Jens Riese, Chloe Lamb, and Elias Zerfu. 2010. *In-Depth Assessment of the Public Agricultural Extension System of Ethiopia and Recommendations for Improvement*. IFPRI Discussion Paper 01041. International Food Policy Research Institute (IFPRI).

Dearden, Lorraine, Stephen Machin, and Howard Reed. 1997. “Intergenerational Mobility in Britain.” *The Economic Journal* 107(440):47–66.

Deichmann, Uwe, Forhad Shilpi, and Renos Vakis. 2009. “Urban Proximity, Agricultural Potential and Rural Non-Farm Employment: Evidence from Bangladesh.” *World Development* 37(3):645–60.

Deininger, Klaus, Songqing Jin, Berhanu Adenew, Gebre-Selassie Samuel, and Demeke Mulat. 2003. *Market and Nonmarket Transfers of Land in Ethiopia: Implications for Efficiency, Equity, and Non-Farm Development*. Policy Research working paper 2992. Policy Research working paper 2992.

Dercon, Stefan, Daneil Gilligan, John Hoddinott, and Tasew Woldehanna. 2007. “The Impact of Roads and Agricultural Extension on Consumption Growth and Poverty in Fifteen Ethiopian Villages.” *American Journal of Agricultural Economics* 91(4):1007. doi: 10.1111/j.l467-8276.2009.01325.x.

Dercon, Stefan, and John Hoddinott. 2005. “Livelihoods, Growth, and Links to Market Towns in Fifteen Ethiopian Villages.” *FCND Brief* 194(July):2.

Desalegn, Yonatan. 2018. “Returns to Education in Ethiopia.” Pp. 434–50 in *Economic Growth and Development in Ethiopia, Perspectives on Development in the Middle East and North Africa (MENA) Region*. Vol. 59, edited by A. Heshmati and H. Yoon. Springer Nature Singapore Pte Ltd.

Donaldson, Dave, and Adam Storeygard. 2016. “The View from above: Applications of Satellite Data in Economics.” *Journal of Economic Perspectives* 30(4):171–98. doi: 10.1257/jep.30.4.171.

Dorosh, Paul, and James Thurlow. 2013. “Agriculture and Small Towns in Africa.” *Agricultural Economics (United Kingdom)* 44(4–5):449–59. doi: 10.1111/agec.12027.

Dorosh, Paul, and James Thurlow. 2014. “Can Cities or Towns Drive African Development? Economywide Analysis for Ethiopia and Uganda.” *World Development* 63:113–23. doi: 10.1016/j.worlddev.2013.10.014.

Dufera, Gerba. 2018. “The Ethiopian Agricultural Extension System and Its Role as a Development Actor : Cases from Southwestern Ethiopia.” Bonn University.

Dufera, Gerba, Girma Kelbero, Till Stellmacher, and Anna Hornidge. 2017. *The Agricultural Extension*

System in Ethiopia: Operational Setup, Challenges and Opportunities. ZEF Working Paper 158. Center for Development Research, University of Bonn, Germany.

Duranton, Gilles. 2015. “Growing through Cities in Developing Countries.” *World Bank Research Observer* 30(1):39–73. doi: 10.1093/wbro/lku006.

EDRI, and GGGI. 2015. *Unlocking the Power of Ethiopia’s Cities: A Report by Ethiopia’s New Climate Economy Partnership*. Addis Ababa, Ethiopia.

Elvidge, Christopher D., Kimberly E. Baugh, Eric A. Kihn, Herbert W. Kroehl, and Ethan R. Davis. 1997. “Mapping City Lights with Nighttime Data from the DMSP Operational Linescan System.” *Photogrammetric Engineering and Remote Sensing* 63(6):727–34.

Elvidge, Christopher D., Kimberly Baugh, Mikhail Zhizhin, Feng Chi Hsu, and Tilottama Ghosh. 2017. “VIIRS Night-Time Lights.” *International Journal of Remote Sensing* 38(21):5860–79. doi: 10.1080/01431161.2017.1342050.

Emily, Greenaway, Leon Juan, and Baker David. 2012. “Understanding the Association Between Maternal Education and Use of Health Services in Ghana: Exploring the Role of Health Knowledge.” *Journal of Biosocial Science* 44(06):733–47.

Ewing, Reid, Shima Hamidi, James B. Grace, and Yehua Dennis Wei. 2016. “Does Urban Sprawl Hold down Upward Mobility?” *Landscape and Urban Planning* 148:80–88. doi: 10.1016/j.landurbplan.2015.11.012.

Fafchamps, Marcel, and Forhad Shilpi. 2002. *Cities and Specilization: Evidence from South Asia*. Policy Research Working paper 48700. Wahsingiton DC.

Fafchamps, Marcel, and Forhad Shilpi. 2003. “The Spatial Division of Labour in Nepal.” *Journal of Development Studies* 39(6):23–66.

Fafchamps, Marcel, and Forhad Shilpi. 2005. “Cities and Specialization: Evidence from South Asia.” *The Economic Journal* 115(503):477–504.

FAO. 2013. *Guidelines for Measuring Household and Individual Dietary Diversity*. Food and Agriculture Organization (FAO).

Feder, Gershon, Jock R. Anderson, Regina Birner, and Klaus Deininger. 2010. “Promises and Realities of Community-Based Agricultural Extension.” Pp. 187–208 in *Community, Market and State in Development*, edited by K. Otsuka and K. Kalirajan. Palgrave Macmillan, a division of Macmillan Publishers Limited.

Frijters, Paul, David W. Johnston, Manisha Shah, and Michael A. Shields. 2009. “To Work or Not to Work?” *American Economic Journal: Applied Economics* 1(3):97–110. doi: 10.1257/app.1.3.97 T4 - Child Development and Maternal Labor Supply M4 - Citavi.

Fuje, Habtamu. 2018. “Welfare Dynamics and Drought in Ethiopia.” in *CSAE Conference, University of Oxford*.

Fujita, Masahisa. 2010. “The Evolution of Spatial Economics: From Thünen to the New Economic Geography.” *Japanese Economic Review* 61(1):1–32. doi: 10.1111/j.1468-5876.2009.00504.x.

Fujita, Masahisa. 2012. “Thünen and the New Economic Geography.” *Regional Science and Urban Economics* 42(6):907–12.

Fujita, Masahisa, Paul Krugman, and Anthony J. Venables. 2000. *The Spatial Economy: Cities, Regions, and International Trade*. Vol. 67. 2nd ed. London, England: The MIT Press.

Fujita, Masahisa, and Jacques-François Thisse. 2006. “Globalization and the Evolution of the Supply Chain : Who Gains and Who Loses?” *International Economic Review* 47(3):811–36.

Gallup, John L., Jeffry D. Sachs, and Andrew D. Mellinger. 1997. “Geography and Economic Development.” *International Regional Science Review* 33(8):179–232.

Galor, Oded, and Joseph Zeira. 1993. “Income Distribution and Macroeconomics.” *Review of Economic Studies* 60(1):35–42. doi: 10.2307/2297811.

Gebru, Gebrehiwot Weldegebrial, Kinfe Asayehegn, and Deribe Kaske. 2012. “Challenges of Development Agents (DAs) Performance in Technology Dissemination: A Case from Southern, Nation, Nationalities and Peoples Regional State (SNNPRS), Ethiopia.” *Scholarly Journal of Agricultural Science* 2(9):208–16.

Genicot, Garance, and Debraj Ray. 2020. “Aspirations and Economic Behavior.” *Annual Review of Economics* 12(1):715–46. doi: 10.1146/annurev-economics-080217-053245.

Gibson, John, Gaurav Datt, Rinku Murgai, and Martin Ravallion. 2017. *For India’s Rural Poor, Growing Towns Matter More Than Growing Cities*. Vol. 98. Policy Research Working Paper 7994.

Gine, Xavier, and Stefan Klonner. 2005. *Credit Constraints as a Barrier to Technology Adoption by the Poor: Lessons from South-Indian Small-Scale Fishery*. World Bank Policy Research Working Paper 3665.

Glaeser, Edward L. 2020. “Urbanization and Its Discontents.” *Eastern Economic Journal* 46(2):191–218. doi: 10.1057/s41302-020-00167-3.

Glick, Peter. 2002. *Women’s Employment and Its Relation to Children’s Health and Schooling in Developing Countries: Conceptual Links, Empirical Evidence, and Policies*. Cornell Food and Nutrition Policy Program Working Paper No. 131.

Gollin, Douglas, Remi Jedwab, and Dietrich Vollrath. 2016. “Urbanization with and without Industrialization.” *Journal of Economic Growth* 21(1):35–70.

Gollin, Douglas, Martina Kirchberger, and David Lagakos. 2017. *In Search of a Spatial Equilibrium in the Developing World*. Working Paper 23916. NBER WORKING PAPER.

Haggblade, Steven, Peter Hazell, and Thomas Reardon. 2010. “The Rural Non-Farm Economy: Prospects for Growth and Poverty Reduction.” *World Development* 38(10):1429–41.

Haile, Getinet. 2018. “Intergenerational Mobility in Socio-Economic Status in Ethiopia.” *Journal of International Development* 30(8):1392–1413. doi: 10.1002/jid.3360.

Haile, MG, and D. Abebaw. 2012. “What Factors Determine the Time Allocation of Agricultural Extension Agents on Farmers’ Agricultural Fields? Evidence Form Rural Ethiopia.” *Journal of Agricultural Extension and Rural Development* 4(10):318–29.

Hansen, Bruce E. 2000. “Sample Splitting and Threshold Estimation.” *Econometrica* 68(3):575–603.

Hanson, Kara, and William Jack. 2008. *Health Worker Preferences for Job Attributes in Ethiopia: Results from a Discrete Choice Experiment*. Health Systems for Outcomes Publication 53122. Wahsingiton DC.

Headey, Derek. 2014. *An Analysis of Trends and Determinants of Child Undernutrition in Ethiopia , 2000-2011*. ESSP working paper 70. IFPRI, Addis Ababa, Ethiopia.

Headey, Derek, John Hoddinott, Disha Ali, Roman Tesfaye, and Mekdim Dereje. 2015. “The Other Asian Enigma: Explaining the Rapid Reduction of Undernutrition in Bangladesh.” *World Development* 66:749–61. doi: 10.1016/j.worlddev.2014.09.022.

Headey, Derek, John Hoddinott, and Seollee Park. 2017. “Accounting for Nutritional Changes in Six Success Stories: A Regression-Decomposition Approach.” *Global Food Security* 13(September 2016):12–20.

Headey, Derek, David Stifel, Liangzhi You, and Zhe Guo. 2017. *Remoteness , Urbanization and Child Nutrition in Sub - Saharan Africa*. IFPRI Discussion Paper 01694. International Food Policy Reserch Institute (IFPRI).

Heckman, James J. 1979. “Sample Selection Bias as a Specification Error.” *Econometrica* 47(1):153–61. doi: 10.2307/1912352.

Heckman, James J. 2006. “Investing in Disadvantaged Children.” *Science* 312(June):2005–7. doi: 10.1016/j.adolescence.2005.09.001.

Henderson, J. Vernon, Tim Squires, Adam Storeygard, and David Weil. 2017. “The Global Distribution of

Economic Activity: Nature, History, and the Role of Trade." *Quarterly Journal of Economics* (November):1–50. doi: 10.1093/qje/qjx030.Advance.

Henderson, M., E. T. Yeh, P. Gong, C. Elvidge, and K. Baugh. 2003. "Validation of Urban Boundaries Derived from Global Night-Time Satellite Imagery." *International Journal of Remote Sensing* 24(3):595–610. doi: 10.1080/01431160304982.

Henderson, Vernon. 2003. "The Urbanization Process and Economic Growth: The so-What Question." *Journal of Economic Growth* 8(1):47–71. doi: 10.1023/A:1022860800744.

Henderson, Vernon. 2010. "Cities and Development." *Journal of Regional Science* 50(1):704. doi: 10.1111/j.1467-9787.2009.00636.x.

Henderson, Vernon. 2014. *Urbanization and the Geography of Development*. Policy Research Working Paper 6877. The World Bank.

Henderson, Vernon, and Randy Becker. 2000. "Political Economy of City Sizes and Formation." *Journal of Urban Economics* 48(3):453–84. doi: 10.1006/juec.2000.2176.

Henderson, Vernon, Mark Roberst, and Adams Storeygard. 2013. *Is Urbanization in Sub-Saharan Africa Different?* Policy Research Working Paper 6481. The World Bank.

Henderson, Vernon, Adam Storeygard, and David N. Weil. 2009. *Measuring Economic Growth From Outer Space*. NBER Working Paper Series 15199. NBER WORKING PAPER.

Henderson, Vernon, Adam Storeygard, and David N. Weil. 2011. "A Bright Idea for Measuring Economic Growth." *American Economic Review* 101(3):194–99. doi: 10.1257/aer.101.3.194.

Hensher, David, and William Greene. 2006. "The Mixed Logit Model: The State of Practice." *Transportation* 133–76.

Hensher, David, John Rose, and William Greene. 2005. *Applied Choice Analysis: A Primier*. New York, USA: Cambrdge University press.

Herzberg, Frederick. 1987. "One More Time: How Do You Motivate Your Employees?" *Hbr* 6(5):76–86.

Hirvonen, Kalle, and John Hoddinott. 2017. "Agricultural Production and Children's Diets: Evidence from Rural Ethiopia." *Agricultural Economics (United Kingdom)* 48(4):469–80. doi: 10.1111/agec.12348.

Hirvonen, Kalle, John Hoddinott, Bart Minten, and David Stifel. 2017. "Children's Diets, Nutrition Knowledge, and Access to Markets." *World Development* 95(February):303–15. doi: 10.1016/j.worlddev.2017.02.031.

Hoddinott, John, Derek Headey, and Mekdim Dereje. 2015. "Cows, Missing Milk Markets, and Nutrition in Rural Ethiopia." *Journal of Development Studies* 51(8):958–75. doi: 10.1080/00220388.2015.1018903.

Hoddinott, John, John A. Maluccio, Jere R. Behrman, Rafael Flores, and Reynaldo Martorell. 2008. "Effect of a Nutrition Intervention during Early Childhood on Economic Productivity in Guatemalan Adults." *The Lancet* 371(9610):411–16. doi: 10.1016/S0140-6736(08)60205-6.

Hoddinott, John, and Yisehac Yohannes. 2002. *Dietary Diversity as a Food Security Indicator*. FCND DISCUSSION PAPER NO. 136. International Food Policy Research Institute (IFPRI).

Hoteit, Sahar, Stefano Secci, Stanislav Sobolevsky, Carlo Ratti, and Guy Pujolle. 2014. "Estimating Human Trajectories and Hotspots through Mobile Phone Data." *Computer Networks* 64:296–307. doi: 10.1016/j.comnet.2014.02.011.

Humphrey, Jean H. 2009. "Child Undernutrition, Tropical Enteropathy, Toilets, and Handwashing." *The Lancet* 374(9694):1032–35. doi: 10.1016/S0140-6736(09)60950-8.

Ingelaere, Bert, Luc Christiaensen, Joachim De Weerdt, and Ravi Kanbur. 2018. "Why Secondary Towns Can Be Important for Poverty Reduction – A Migrant Perspective." *World Development* 105:273–82. doi: 10.1016/j.worlddev.2017.12.025.

de Janvry, A., and E. Sadoulet. 2009. "Agricultural Growth and Poverty Reduction: Additional Evidence." *The World Bank Research Observer* 25(1):1–20. doi: 10.1093/wbro/lkp015.

Joanna, Mills, and Cumming Oliver. 2016. *The Impact of Water, Sanitation and Hygiene on Key Health and Social Outcomes:Review of Evidence*. Vol. 7. SHARE Consortium, UK and UNICEF, USA.

Kahneman, Daniel, and Amos Tversky. 1979. “Prospect Theory: An Analysis of Decision Under Risk.” *Econometrica* 47(2):263–91.

Kamei, Akito, and Shohei Nakamura. 2020. *Urban Agglomerations and Employment Transitions in Ethiopia*. Policy Research Working paper 9184. Policy Research Working Paper 9184.

Kanbur, Ravi, Luc Christiaensen, and Joachim De Weerdt. 2019. “Where to Create Jobs to Reduce Poverty: Cities or Towns?” *Journal of Economic Inequality* (8069). doi: 10.1007/s10888-019-09419-5.

Kassa, Belay. 2002. “Constraints to Agricultural Extension Work in Ethiopia : The Insiders’ View.” *South African Journal of Agricultural Extension* (January 2002).

Kassa, Belay, and Degnet Abebaw. 2004. “Challenges Facing Agricultural Extension Agents: A Case Study from South-Western Ethiopia.” *African Development Review* 16(1):139–68. doi: 10.1111/j.1467-8268.2004.00087.x.

Kassa, Belay, Ranjan S. Karippai, Dawit Alemu, Abera Deressa, and Jemal Yousuf. 2012. *Work Motivation and Job Performance of Development Agents in Ethiopia*. Haramaya University, Dire Dawa, Ethiopia.

Kelemu, Kaleb, Mekonnen Sime, and Mekonnen Hailu. 2014. “Determinants and Levels of Agricultural Development Agents Job Satisfaction : The Case of Kalu Woreda , South Wollo Zone of the Amhara National Regional State.” *Ethiopian Journal of Business and Economics* 4(1):149–75.

Kolstad, Julie Riise. 2011. “How to Make Rural Jobs More Attractive to Health Workers. Findings from a Discrete Choice Experiment in Tanzania.” *Health Economics* 211(January 2010):196–211.

Kraay, Aart, and David McKenzie. 2014. “Do Poverty Traps Exist? Assessing the Evidence.” *Journal of Economic Perspectives* 28(3):127–48. doi: 10.1257/jep.28.3.127.

Krugman, Paul. 1991. “Increasing Returns and Economic Geography.” *Journal of Political Economy* 99(3):483. doi: 10.1086/261763.

Krugman, Paul. 1997. *How the Economy Organizes Itself in Space: A Survey of the New Economic Geography*. SFI WORKING PAPER: 1996-04-021.

Kruk, Margaret, Jennifer Johnson, Mawuli Gyakobo, Peter Agyei-Baffour, Kwesi Asabir, Rani Kotha, Janet Kwansah, Emmanuel Nakua, Rachel Snow, and Mawuli Dzodzomenyo. 2010. “Rural Practice Preferences among Medical Students in Ghana: A Discrete Choice Experiment.” *Bulletin of the World Health Organization* 88(5):333–41. doi: 10.2471/BLT.09.072892.

Kuhfeld, Warren. 1997. “Efficient Experimental Design Using Computerized Searches.” *Sawtooth Software Research Paper Series* 98382(360):0–26.

Kuhfeld, Warren. 2010. *Marketing Research Methods in SAS: Experimental Design, Choice, Conjoint, and Graphical Techniques*. Vol. 13. Cary, NC, USA: SAS Institute Inc.

Lafourcade, Miren, and Jacques-François Thisse. 2008. *New Economic Geography: A Guide to Transport Analysis*. Vol. 33. halshs-00586878.

Lagakos, David, Mushfiq Mobarak, and Michael E. Waugh. 2018. “The Welfare Effects of Encouraging Rural-Urban Migration.” *NBER Working Papers*. doi: 10.3386/w24193.

Lagarde, Mylene, and Duane Blaauw. 2009. “A Review of the Application and Contribution of Discrete Choice Experiments to Inform Human Resources Policy Interventions.” *Human Resources for Health* 10:1–10. doi: 10.1186/1478-4491-7-62.

Lancaster, Kelvin. 1966. “A New Approach to Consumer Theory.” *Journal of Political Economy* 74(2):132–57. doi: 10.1086/226550.

Layard, R., S. Nickell, and G. Mayraz. 2008. “The Marginal Utility of Income.” *Journal of Public Economics* 92(8–9):1846–57. doi: 10.1016/j.jpubeco.2008.01.007.

Leka, Wanna, and Desalegn Chalchisa. 2012. *Cost Sharing in Public Higher Education Institutions in Ethiopia with Special Emphasis on Addis Ababa and Adama Universities*. Addis Ababa, Ethiopia: Forum for Social Studies (FSS).

Lloyd, Cynthia B., and Ann Blanc. 1996. "Children ' s Schooling in Sub-Saharan Africa: The Role of Fathers, Mothers, and Others." *Population and Development Review* 22(2):265–98.

Lokshin, Michael, and Ruslan Yemtsov. 2005. "Has Rural Infrastructure Rehabilitation in Georgia Helped the Poor?" *World Bank Economic Review* 19(2):311–33. doi: 10.1093/wber/lhi007.

Louviere, Jordan J., David A. Hensher, and Joffre D. Swait. 2010. *Stated Choice Methods: Analysis and Application*. Cambridge, United Kingdom: Cambrdge University press.

Lusk, Jayson L., and Ted C. SChroeder. 2004. "Are Choice Experiments Incentive Compatible? A Test with Quality Differentiated Beef Steaks." *American Journal of Agricultural Economics* 86(May):467–82.

Mangham, Lindsay J., and Kara Hanson. 2008. "Employment Preferences of Public Sector Nurses in Malawi: Results from a Discrete Choice Experiment." *Tropical Medicine and International Health* 13(12):1433–41. doi: 10.1111/j.1365-3156.2008.02167.x.

McCullough, Ellen B. 2017. "Labor Productivity and Employment Gaps in Sub-Saharan Africa." *Food Policy* 67:133–52.

McFadden, Daniel. 1973. "Conditional Logit Analysis of Qualitative Choice Behavior." Pp. 105–42 in *Frontiers in Econometrics*, edited by P. Zarembka. New York: Academic Press.

McFadden, Daniel, and Kenneth Train. 2000. "Mixed MNL Models for Discrete Response." *Journal of Applied Econometrics* 15: 447-47(May).

Melesse, Mequanint B., and Francesco Cecchi. 2017. "Does Market Experience Attenuate Risk Aversion? Evidence from Landed Farm Households in Ethiopia." *World Development* 98(453):447–66. doi: 10.1016/j.worlddev.2017.05.011.

Mellor, John W. 2018. "Measuring the Impact of Agricultural Growth on Economic Transformation." Pp. 29–46 in *Agricultural Development and Economic Transformation*. Palgrave Studies in Agricultural Economics and Food Policy.

Michalopoulos, Stelios, and Elias Papaioannou. 2018. "Spatial Patterns of Development: A Meso Approach." *Annual Review of Economics* 10(May):383–410. doi: 10.1146/annurev-economics-080217-053355.

Minten, Bart, Mekdim Dereje, Fantu Bachewe, and Seneshaw Tamru. 2018. *Evolving Food Systems in Ethiopia: Past, Present and Future*. ESSP Working paper 117. Addis Ababa, Ethiopia.

Minten, Bart, Bethlehem Koru, and David Stifel. 2013. "The Last Mile(s) in Modern Input Distribution: Pricing, Profitability, and Adoption." *Agricultural Economics (United Kingdom)* 44(6):629–46. doi: 10.1111/agec.12078.

MoANR and ATA. 2014. *National Strategy for Ethiopia's Agricultural Extension System:Vision, Systemic Bottlenecks and Priority Interventions*. Minstry of Agriculture and natural resources (MoANR) and Agricultural tranformation agency (ATA), Addis Ababa, Ethiopia.

MoANR and ATA. 2017. *Ethiopia's Agricultural Extension Strategy: Vision, Systemic Bottlenecks and Priority Interventions*. Minstry of Agriculture and natural resources (MoANR) and Agricultural tranformation agency (ATA), Addis Ababa, Ethiopia.

MoE. 2015. *Education Sector Development Programme V (ESDP V) 2015/16 - 2019/20: Program Action Plan*. Ministry of Education (MoE), Addis Ababa, Ethiopia.

Moser, Christine M., and Christopher B. Barrett. 2006. "The Complex Dynamics of Smallholder Technology Adoption: The Case of SRI in Madagascar." *Agricultural Economics* 35(3):373–88. doi: 10.1111/j.1574-0862.2006.00169.x.

MoUDC. 2012. *Structure Plan Manual (Revised Version)*. Ministry of urban development and construction (MoUDC), Addis Ababa, Ethiopia.

Murphy, Daniel. 2018. “Home Production, Expenditure, and Economic Geography.” *Regional Science and Urban Economics* 70:112–26.

Muzzini, Elisa. 2008. *Urban Poverty in Ethiopia:A Multi-Faceted and Spatial Perspective*. URBAN PAPERS, UP-4, World Bank, Washington D.C.

Nakamura, Emi, Jósef Sigurdsson, and Jón Steinsson. 2016. “The Gift of Moving: Intergenerational Consequences of a Mobility Shock.” *NBER Working Paper* (22392):1–45. doi: 10.3386/w22392.

Narayan, Ambar, Roy Van der Weide, Alexandru Cojocaru, Lakner Christoph, Silvia Redaelli, Daniel Mahler, Rakesh Ramasubbaiah, and Stefan Thewissen. 2018. *Fair Progress? Economic Mobility across Generations around the World*. Washington DC: The world Bank.

Nguyen, Anh, Getinet Haile, and Jim Taylor. 2005. “Ethnic and Gender Differences in Intergenerational Mobility: A Study Of 26-Year-Olds in the USA.” *Scottish Journal of Political Economy* 52(4):544–64. doi: 10.1111/j.1467-9485.2005.00355.x.

Niermeyer, Susan, P. Andrade Mollinedo, and L. Huicho. 2009. “Child Health and Living at High Altitude.” *Archives of Disease in Childhood* 94(10):806–11. doi: 10.1136/adc.2008.141838.

Nybom, Martin. 2018. *Intergenerational Mobility: A Dream Deferred?* ILO future of work research paper series 7. Geneva, Switzerland.

OECD/PSI. 2020. *Rural Development Strategy Review of Ethiopia: Reaping the Benefits of Urbanisation*. Paris: OECD Development Pathways, OECD Publishing.

Onis, Mercedes De, Adelheid W. Onyango, Elaine Borghi, and Cutberto Garza. 2006. “Comparison of the World Health Organization (WHO) Child Growth Standards and the National Center for Health Statistics / WHO International Growth Reference : Implications for Child Health Programmes.” 9(7):942–47. doi: 10.1017/PHN20062005.

De Poel, Van De, O. O’donnell, and E. Van Doorslaer. 2012. “Is There a Health Penalty of China’s Rapid Urbanization?” *Health Economics* 21(2007):367– 385. doi: 10.1002/hec.

Porta, Rafael La, and Andrei Shleifer. 2008. “The Unofficial Economy and Economic Development.” *Brookings Papers on Economic Activity* 2008(2):275–352. doi: 10.2139/ssrn.1304760.

Pradhan, Menno, David E. Sahn, and Stephen D. Younger. 2003. “Decomposing World Health Inequality.” *Journal of Health Economics* 22(2):271–93.

Psacharopoulos, George, and Harry Antony Patrinos. 2018. *Returns to Investment in Education: A Decennial Review of the Global Literature*. Policy Research Working Paper 8402. The World Bank.

Ragasa, Catherine, John Ulimwengu, Josee Randriamamonjy, and Thaddee Badibanga. 2016. “Factors Affecting Performance of Agricultural Extension: Evidence from Democratic Republic of Congo.” *Journal of Agricultural Education and Extension* 22(2):113–43.

Ravallion, Martin, Shaohua Chen, and Prem Sangraula. 2007. “New Evidence on the Urbanization of Global Poverty.” *Population and Development Review* 33(4):667–701. doi: 10.1111/j.1728-4457.2007.00193.x.

Ravallion, Martin, and Quentin Wodon. 1999. “Poor Areas, or Only Poor People?” *Journal of Regional Science* 39(4):689–711. doi: 10.1111/0022-4146.00156.

Reardon, Thomas A. 2016. *Growing Food for Growing Cities: Transformation of Food Systems in an Urbanizing World*. Chicago Council on Global Affairs, Chicago, USA.

Reardon, Thomas, Julio Berdegué, Christopher B. Barrett, and Kostas Stamoulis. 2006. “Household Income Diversification into Rural Nonfarm Activities.” Pp. 115–40 in *Transforming the rural nonfarm economy. Opportunities and threads in the developing world*, edited by S. Haggblade, P. Hazell, and T. Reardon. Johns Hopkins University Press.

Redding, Stephen. 2010. “The Empirics of New Economic Geography.” *Journal of Regional Science* 50(1):297–311.

Redding, Stephen, and Anthony J. Venables. 2004. “Geography and Export Performance: External

Market Access and Internal Supply Capacity." in *Challenges to Globalization: Analyzing the Economics*, edited by R. E. Baldwin and W. L. Alan. University of Chicago Press.

Roback, Jennifer. 1982. "Wages, Rents, and the Quality of Life." *Journal of Political Economy* 90(6):1257–78.

Roberts, Brian. 2014a. *Managing Systems of Secondary Cities: Policy Responses in International Development*. Cities Alliance/UNOPS, Brussels.

Roberts, Brian. 2014b. *The Systems of Secondary Cities: The Neglected Drivers of Urbanising Economies*. The CIVIS series, citiesalliance.org.

Rockers, Peter C., Wanda Jaskiewicz, Laura Wurts, Margaret E. Kruk, George S. Mgomella, Francis Ntalazi, and Kate Tulenko. 2012. "Preferences for Working in Rural Clinics among Trainee Health Professionals in Uganda: A Discrete Choice Experiment." *BMC Health Services Res.* 12:12–212. doi: 10.1186/1472-6963-12-212.

Rosenbaum, Paul R. 2012. "Model-Based Direct Adjustment Model-Based Direct Adjustment." *Journal of the American Statistical Association* (March 2015):37–41. doi: 10.1080/01621459.1987.10478441.

Rosenbaum, Paul R., and Donald B. Rubin. 1984. "Reducing Bias in Observational Studies Using Subolassification on the Propensity Score." *Journal of the American Statistical Association* 79(387):516–24.

Sachs, Jeffrey, Andrew Mellinger, and JL Gallup. 2001. "The Geography of Poverty." *Scientific American*.

Sahn, D. E., and David Stifel. 2004. *Urban-Rural Inequality in Living Standards in Africa*. Vol. 12. UN-WIDER Research Paper 2004/4. UN-WIDER Research Paper 2004/4, United Nations University.

Satterthwaite, David, and Cecilia Tacoli. 2003. *The Urban Part of Rural Development: The Role of Small and Intermediate Urban Centres in Rural and Regional Development and Poverty Reduction*. iied WORKING PAPER 9.

Savory, David J., Ricardo Andrade-Pacheco, Peter W. Gething, Alemayehu Midekisa, Adam Bennett, and Hugh J. W. Sturrock. 2017. "Intercalibration and Gaussian Process Modeling of Nighttime Lights Imagery for Measuring Urbanization Trends in Africa 2000-2013." *Remote Sensing* 9(7). doi: 10.3390/rs9070713.

Scarpa, Riccardo, Mara Thiene, and Kenneth Train. 2008. "Utility in Willingness to Pay Space: A Tool to Address Confounding Random Scale Effects in Destination Choice to the Alps." *American Journal of Agricultural Economics* 90(4):994–1010. doi: 10.1111/j.1467-8276.2008.01155.x.

Schindler, Simon, and Stefan Pfattheicher. 2017. "The Frame of the Game: Loss-Framing Increases Dishonest Behavior." *Journal of Experimental Social Psychology* 69:172–77. doi: 10.1016/j.jesp.2016.09.009.

Schmidt, Emily, and Mekamu Kedir. 2009. *Urbanization and Spatial Connectivity in Ethiopia: Urban Growth Analysis Using GIS*. Vol. 3. ESSP Working Paper 003. International Food Policy Research Institute (IFPRI).

Scott, Anthony. 2001. "Eliciting GPs' Preferences for Pecuniary and Non-Pecuniary Job Characteristics." *Journal of Health Economics* 20(3):329–47. doi: 10.1016/S0167-6296(00)00083-7.

Shields, Michael A. 2004. "Addressing Nurse Shortages: What Can Policy Makers Learn from the Econometric Evidence on Nurse Labour Supply?" *The Economic Journal* 114(499):464–98.

Shiferaw, Admasu, Eyerusalem Siba, and Getnet Alemu. 2012. *Road Infrastructure and Enterprise Development in Ethiopia*. IGC Working Paper C-32001-ETH-1. International Growth center, Addis Ababa, Ethiopia.

Simler, Ken, and Nora Dudwick. 2009. *Urbanization and Rural-Urban Welfare Inequalities *DRAFT FOR DISCUSSION*. World Bank, USA.

Singh, I., I. S. Chohan, M. Lal, P. K. Khanna, M. C. Srivastava, R. B. Nanda, J. S. Lamba, and M. S. Malhotra. 1977. "Effects of High Altitude Stay on the Incidence of Common Diseases in Man." *International Journal of Biometeorology* 21(2):93–122.

Spears, Dean. 2013. “How Much International Variation in Child Height Can Sanitation Explain?” *World Bank Policy Research Working Paper* (February):1–53.

Spielman, David J., Dawit Kelemework Mekonnen, and Dawit Alemu. 2012. “Seed, Fertilizer, and Agricultural Extension in Ethiopia.” Pp. 84–122 in *Food and Agriculture in Ethiopia: Progress and Policy Challenges*, edited by P. A. DOROSH and S. RASHID. University of Pennsylvania Press.

Staiger, B. Y. Douglas, and James H. Stock. 1997. “Instrumental Variables Regression with Weak Instruments.” *Econometrica* 65(3):557–86.

Stifel, David, and Bart Minten. 2008. “Isolation and Agricultural Productivity.” *Agricultural Economics* 39(1):1–15. doi: 10.1111/j.1574-0862.2008.00310.x.

Stifel, David, and Bart Minten. 2017. “Market Access, Well-Being, and Nutrition: Evidence from Ethiopia.” *World Development* 90:229–41.

Stifel, David, Bart Minten, and Paul Dorosh. 2003. *Transaction Cost and Agricultural Productivity: Implications of Isolation for Rural Poverty in Madagascar*. MSSD DISCUSSION PAPER NO. 56. Washington D.C.

Stifel, David, Bart Minten, and Bethlehem Koru. 2016. “Economic Benefits of Rural Feeder Roads: Evidence from Ethiopia.” *Journal of Development Studies* 52(9):1335–56.

Sutton, Paul. 1997. “Modeling Population Density with Night-Time Satellite Imagery and GIS.” *Computers, Environment and Urban Systems* 21(3–4):227–44. doi: 10.1016/S0198-9715(97)01005-3.

Tabuchi, Takatoshi. 1998. “Urban Agglomeration and Dispersion: A Synthesis of Alonso and Krugman.” *Journal of Urban Economics* 44(3):333–51. doi: 10.1006/juec.1997.2074.

Tabuchi, Takatoshi, and Jacques François Thisse. 2002. “Taste Heterogeneity, Labor Mobility and Economic Geography.” *Journal of Development Economics* 69(1):155–77. doi: 10.1016/S0304-3878(02)00057-3.

Tacoli, Cecilia. 1998. “Rural-Urban Interactions: A Guide to the Literature.” *Environment and Urbanization* 10(1):147–66.

Thisse, Jacques François. 2011. “Geographical Economics: A Historical Perspective.” *Recherches Economiques de Louvain* 77(2–3):141–68. doi: 10.3917/rel.772.0141.

Train, Kenneth. 2009. *Discrete Choice Methods with Simulation*. Vol. 9780521816. 2nd ed. New York, USA: Cambrdge University press.

Train, Kenneth, and Melvyn Weeks. 2005. “Discrete Choice Models in Preference Space and Willingness to Pay Space.” Pp. 1–16 in *Applications of Simulation Methods in Environmental and Resource Economics*, edited by R. Scarpa and A. Alberini.

UN Habitat. 2010. *The State of African Cities 2010:Governance, Inequality and Urban Land Market*. Vol. 93. Nairobi, Kenya.

UN Habitat. 2014. *The State of African Cities 2014:Re-Imagining Sustainable Urban Transitions*. Nairobi, Kenya.

UNDESA. 2015. *World Urbanization Prospects: The 2014 Revisions, (ST/ESA/SER.A/366)*. United Nations, Department of Economic and Social Affairs, Population Division (UNDESA), New York, United Nations.

UNDESA. 2019. *World Urbanization Prospects: The 2018 Revision, (ST/ESA/SER.A/420*. Vol. 12. United Nations, Department of Economic and Social Affairs, Population Division (UNDESA), New York, United Nations.: United Nations.

UNDP. 2006. *Human Development Report 2006. Beyond Scarcity: Power, Poverty and the Global Water Crisis*. New York, USA: Palgrave Macmillan.

UNECA. 2017. *Assessment of Urbanization Data in Africa*. United Nations, Economic Commission for Africa (UNECA), Addis Ababa, Ethiopia.

Vandercasteelen, Joachim, Seneshaw Tambru Beyene, Bart Minten, and Johan Swinnen. 2018. “Big

Cities, Small Towns, and Poor Farmers: Evidence from Ethiopia." *World Development* 106:393–406.

Venables, Aj. 2008. "New Economic Geography." *The New Palgrave Dictionary of Economics* 1–12.

Venables, Anthony J. 2005. "Spatial Disparities in Developing Countries: Cities, Regions, and International Trade." *Journal of Economic Geography* 5(1):3–21. doi: 10.1093/jnlecg/lbh051.

Vieira, Marcos R., Vanessa Frías-Martínez, Nuria Oliver, and Enrique Frías-Martínez. 2010. "Characterizing Dense Urban Areas from Mobile Phone-Call Data: Discovery and Social Dynamics." *Proceedings - SocialCom 2010: 2nd IEEE International Conference on Social Computing, PASSAT 2010: 2nd IEEE International Conference on Privacy, Security, Risk and Trust* 241–48. doi: 10.1109/SocialCom.2010.41.

WHO. 2006. *WHO Child Growth Standards: Length/Height-for-Age, Weight-for-Age, Weight-for-Length, Weight-for- Height and Body Mass Index-for-Age : Methos and Development*.

WHO. 2012. *How to Conduct a Discrete Choice Experiment for Health Workforce Recruitment and Retention in Remote and Rural Areas: A User Guide with Case Studies*. Geneva, Switzerland.

WHO and UNICEF. 2017. *A Snapshot of Drinking Water , Sanitation and Hygiene in Africa:2017 Update and SDG Baselines*.

Williams, Nathalie E., Timothy A. Thomas, Matthew Dunbar, Nathan Eagle, and Adrian Dobra. 2015. "Measures of Human Mobility Using Mobile Phone Records Enhanced with GIS Data." *PLoS ONE* 10(7):1–16. doi: 10.1371/journal.pone.0133630.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. MIT Press.

Wooldrigde, Jeffrey M. 2013. *Introductory Econometrics: A Modern Approach*. 5th Editio. MIT Press.

Word Bank. 2009. *Spatial Disparities and Development Policy*. edited by G. Kochendörfer-Lucius and B. Pleskovic. Wahsingiton DC.

Worku, Ibrahim Hassen, Mekdim Dereje, Bart Minten, and Kalle Hirvonen. 2017. "Diet Transformation in Africa: The Case of Ethiopia." *Agricultural Economics (United Kingdom)* 48:73–86. doi: 10.1111/agec.12387.

World Bank. 2009. *World Development Report: Reshaping Economic Geography*. New York, USA: The world Bank.

World Bank. 2010. *Ethiopia : Re-Igniting Poverty Reduction in Urban Ethiopia through Inclusive Growth*. REPORT NO. 52848-ET. The World Bank.

World Bank. 2011. *From Farm to Firm: Rural-Urban Transition in Developing Countries*. edited by N. Dudwick, K. Hull, R. Katayama, F. Shilpi, and K. Simler. Washington D.C.

World Bank. 2013a. *Global Monitoring Report 2013: Rural-Urban Dynamics and the Millennium Development Goals*. The World Bank.

World Bank. 2013b. *Harnessing Urbanization to End Poverty and Boost Prosperity in Africa*. Sustainable Development Series 81546, The World Bank.

World Bank. 2015. *Ethiopia's Great Run: The Growth Acceleration and How to Pace It*. Report No.: 99399-ET, Addis Ababa, Ethiopia. The world Bank.

World Bank. 2016. *Ethiopia: Priorities for Ending Extreme Poverty and Promoting Shared Prosperity*. Report No: 100592-ET. The World Bank.

World Bank. 2018. "World Development Report 2018: Learning to Realize Education's Promise."

World Bank. 2020. *Ethiopia Poverty Assessment:Harnessing Continued Growth for Accelerated Poverty Reduction*. Washington DC. World Bank.

World Bank and Cities Alliance. 2015. *Ethiopia Urbanization Review:Urban Institutions for a Middle-Income Ethiopia*.

World Bank and IFPRI. 2010. "Gender and Governance in Rural Services:Insights from India, Ghana, and Ethiopia." in *Community, Market and State in Development*. Washington D.C.: The world Bank and

International Food Policy Research Institute (IFPRI).

Zax, Jeffrey S., and Daniel I. Rees. 2002. “IQ, Academic Performance, Environment, and Earnings.” *The Review of Economics and Statistics* 84(November):600–616.

Zhang, Qian, and Karen C. Seto. 2011. “Mapping Urbanization Dynamics at Regional and Global Scales Using Multi-Temporal DMSP/OLS Nighttime Light Data.” *Remote Sensing of Environment* 115(9):2320–29. doi: 10.1016/j.rse.2011.04.032.

Zhang, Qian, and Karen C. Seto. 2013. “Can Night-Time Light Data Identify Typologies of Urbanization? A Global Assessment of Successes and Failures.” *Remote Sensing* 5(7):3476–94. doi: 10.3390/rs5073476.

Zimmerman, By David J. 1992. “Regression Toward Mediocrity in Economic Stature.” *The American Economic Review* 82(3):409–29.

Appendices

Supplementary materials to Chapter One

Table A1.1. Patterns of urbanization in Ethiopia

Year	Urban population (1,000s)	Urban population (% total)	No. of agglomerations	Av. Distance between agglomerations
1950	503	3	6	174
1960	778	4	11	123
1970	1,341	5	24	72
1980	2,385	7	45	56
1990	3,895	8	78	47
2000	6,521	11	147	37
2010	11,064	14	289	24
2015	24,292	27	510	19

Source: Author's computation based on data from africapolis@oecd.org

Table A1.2. Distribution of Enterprises by size, 2015

Location\size of Enterprise	Small	Medium	Large	Unknown	Total
Panel A: Number of Enterprises					
Total	29,827	13,562	1,979	26,454	71,822
Addis Ababa	21,656	9,102	895	22,365	54,018
Amhara	1,517	318	32	797	2,664
Dire Dawa	3	29	3	38	73
Oromia	4,979	2,672	844	2,989	11,484
SNNPR	49	795	57		901
Tigray	1,623	646	148	265	2,682
Panel B: Share of Enterprises (%)					
Addis Ababa	72.6	67.1	45.2	84.5	75.2
Amhara	5.1	2.3	1.6	3.0	3.7
Dire Dawa	0.0	0.2	0.2	0.1	0.1
Oromia	16.7	19.7	42.6	11.3	16.0
SNNPR	0.2	5.9	2.9	-	1.3
Tigray	5.4	4.8	7.5	1.0	3.7

Source: World Bank enterprise survey document, 2015.

Supplementary materials to Chapter Two

Table A2.1. Descriptive statistics of key variables by survey years

Variable	Total	2014	2016	**Mean difference test (p-value)**
ln (real consumption per capita)	8.68	**8.71**	**8.66**	0.00
Proportion of food groups consumed by HH	0.53	**0.52**	**0.54**	0.00
Proportion of months food gap is reported by HH	0.07	0.07	0.07	0.59
ln(Nighttime light Index)	1.56	1.55	1.57	0.72
ln(Household size)	1.65	1.64	1.65	0.20
ln(Age of household head in years)	3.79	3.77	3.81	0.00
Head is male, yes=1	0.69	0.70	0.69	0.57
Head education, primary=1	0.29	0.29	0.29	0.96
Head education, secondary or higher=1	0.17	0.17	0.17	0.62
Household faced drought shock, yes=1	0.15	0.07	0.24	0.00
Household faced non-drought shock, yes=1	0.45	0.34	**0.55**	0.00
ln(Land size household owned, ha)	0.47	0.49	0.45	0.00
ln(Livestock owned, in TLU)	0.99	0.98	1.00	0.19
ln(Elevation in meters)	7.48	7.48	7.48	0.98
ln(total annual rainfall in mm)	7.03	**7.36**	**6.70**	0.00
ln(total annual rainfall squared)	49.66	**54.2**	**45.2**	0.00
ln (mean annual temperature, cent.)	5.25	5.25	5.25	0.93
Proportion of fertile soil in EA	0.64	0.64	0.64	0.74
EA has electricity, yes=1	0.55	0.54	0.55	0.29

Source: Authors' computation based on LSMS (2014 & 2016)

Table A2.2. Cross-tabulation of administrative urbanization indicators and NTL

Urban-rural indicator	Nighttime light statistics		
	Mean	Median	Std.Dev
Rural	5.0	0.0	26.6
Small towns	5.9	0.0	19.0
Intermediate towns	232.9	161.3	196.5
Large town	612.0	736.1	232.7
Total	83.18	0.0	185.98

Source: Authors' calculations based on Savory et.al and LSMS (2014 & 2016)

Table A2.3. Association between urbanization and welfare, pooled OLS

	ln(Expenditure)		Diet Diversity score		Food security Gap	
ln(Nighttime light Index)	0.096***	0.047***	0.028***	0.016***	-0.007***	-0.004**
	(0.007)	(0.008)	(0.002)	(0.002)	(0.001)	(0.002)
Survey round, 2015	-0.06***	-0.14***	0.019***	0.001	0.002	-0.03***
	(0.018)	(0.041)	(0.005)	(0.012)	(0.006)	(0.009)
ln(Household size)		-0.44***		0.070***		0.002
		(0.020)		(0.005)		(0.004)
ln(Age of household head in years)		-0.024		-0.001		0.010*
		(0.023)		(0.006)		(0.005)
Head is male, yes=1		0.003		-0.007*		-0.013***
		(0.016)		(0.004)		(0.004)
Head education, primary=1		0.154***		0.055***		-0.018***
		(0.018)		(0.005)		(0.005)
Head education, secondary or higher=1		0.423***		0.118***		-0.048***
		(0.028)		(0.008)		(0.006)
Household faced drought shock, yes=1		-0.08***		-0.03***		0.092***
		(0.029)		(0.009)		(0.012)
Household faced non-drought shock, yes=1		-0.018		-0.004		0.060***
		(0.016)		(0.005)		(0.005)
ln(Land size household owned, ha)		0.091***		0.000		-0.025***
		(0.025)		(0.007)		(0.006)
ln(Livestock owned, in TLU)		0.071***		0.006		-0.012***
		(0.014)		(0.004)		(0.004)
ln(Elevation in meters)		-0.127		-0.039		-0.024
		(0.082)		(0.024)		(0.028)
ln(total annual rainfall in mm)		2.012***		0.103		-0.275
		(0.770)		(0.224)		(0.254)
ln(total annual rainfall squared)		-0.16***		-0.010		0.020
		(0.059)		(0.017)		(0.019)
ln(mean annual temperature, degrees)		-0.249		-0.033		-0.045
		(0.162)		(0.046)		(0.052)
Proportion of fertile soil in EA		0.074**		0.001		-0.020**
		(0.037)		(0.010)		(0.008)
EA has electricity, yes=1		0.183***		0.055***		-0.032***
		(0.030)		(0.008)		(0.008)
Zone Fixed Effect	No	Yes	No	Yes	No	Yes
Constant	8.557***	4.717	0.478***	0.492	0.084***	1.432
	(0.022)	(2.918)	(0.006)	(0.798)	(0.005)	(1.101)
Number of observations	9,215	9,210	9,606	9,600	9,606	9,600
R2	0.123	0.375	0.151	0.339	0.013	0.216
Adjusted R2	0.122	0.368	0.151	0.332	0.013	0.207

*Note: Village clustered standard error in parentheses: *** p<0.01, ** p<0.05, * p<0.1*

Table A2.4. Impact of urbanization on welfare, village/EA FE

	ln(Expenditure)		Diet Diversity score		Food security Gap	
ln(Nighttime light Index)	0.151***	0.112***	0.042***	0.031***	0.002	0.005
	(0.031)	(0.025)	(0.007)	(0.006)	(0.016)	(0.012)
Survey round, 2015	-0.05***	-0.088**	0.018***	0.007	0.002	-0.031***
	(0.018)	(0.041)	(0.005)	(0.011)	(0.006)	(0.011)
ln(Household size)		-0.464***		0.061***		0.001
		(0.020)		(0.005)		(0.004)
ln(Age of household head in years)		-0.042**		-0.011**		0.011**
		(0.021)		(0.005)		(0.005)
Head is male, yes=1		0.020		-0.002		-0.014***
		(0.015)		(0.004)		(0.004)
Head education, primary=1		0.126***		0.035***		-0.014***
		(0.016)		(0.004)		(0.004)
Head education, secondary or higher=1		0.342***		0.087***		-0.039***
		(0.025)		(0.007)		(0.006)
Household faced drought shock, yes=1		-0.086***		-0.015*		0.077***
		(0.032)		(0.009)		(0.011)
Household faced non-drought shock, yes=1		-0.034**		-0.008*		0.057***
		(0.015)		(0.004)		(0.005)
ln(Land size household owned, ha)		0.146***		0.006		-0.031***
		(0.021)		(0.005)		(0.007)
ln(Livestock owned, in TLU)		0.089***		0.016***		-0.016***
		(0.011)		(0.003)		(0.003)
ln(Elevation in meters)		-0.342		-0.020		0.058
		(0.275)		(0.069)		(0.060)
ln(total annual rainfall in mm)		1.594*		-0.064		-0.199
		(0.824)		(0.239)		(0.275)
ln(total annual rainfall squared)		-0.124**		0.003		0.014
		(0.063)		(0.018)		(0.021)
ln(mean annual temperature, degrees)		-0.663*		0.026		-0.038
		(0.382)		(0.100)		(0.096)
Proportion of fertile soil in EA		-0.019		-0.029*		0.003
		(0.066)		(0.017)		(0.012)
EA has electricity, yes=1		-0.030		-0.004		-0.018
		(0.080)		(0.013)		(0.022)
EA Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Constant	8.471***	10.241**	0.456***	0.722	0.069***	0.539
	(0.050)	(4.634)	(0.011)	(1.240)	(0.025)	(1.243)
Number of observations	9,215	9,210	9,606	9,600	9,606	9,600
R2	0.004	0.156	0.007	0.103	0.000	0.103
Adjusted R2	0.004	0.154	0.007	0.102	-0.000	0.102

*Note: Village clustered standard error in parentheses: *** p<0.01, ** p<0.05, * p<0.1*

Table A2.5. Association between stages of urbanization and welfare, pooled OLS

	ln(Expenditure)		Diet Diversity score		Food security Gap	
Small town	0.261***	0.112*	0.038	-0.001	0.023	0.015
	(0.086)	(0.061)	(0.024)	(0.020)	(0.021)	(0.015)
Intermediate town	0.531***	0.232***	0.144***	0.086***	-0.03***	-0.022***
	(0.043)	(0.040)	(0.012)	(0.012)	(0.008)	(0.008)
Large town	0.552***	0.206***	0.169***	0.070***	-0.06***	-0.037***
	(0.062)	(0.065)	(0.012)	(0.018)	(0.007)	(0.014)
Survey round, 2015	-0.056***	-0.135***	0.018***	0.002	0.002	-0.032***
	(0.018)	(0.041)	(0.005)	(0.012)	(0.006)	(0.009)
ln(Household size)		-0.439***		0.070***		0.002
		(0.020)		(0.005)		(0.004)
ln(Age of household head in years)		-0.021		0.001		0.009
		(0.023)		(0.006)		(0.005)
Head is male, yes=1		0.002		-0.008*		-0.013***
		(0.016)		(0.004)		(0.004)
Head education, primary=1		0.156***		0.056***		-0.018***
		(0.018)		(0.005)		(0.005)
Head education, secondary or higher=1		0.427***		0.119***		-0.048***
		(0.028)		(0.008)		(0.006)
Household faced drought shock, yes=1		-0.083***		-0.034***		0.093***
		(0.029)		(0.009)		(0.012)
Household faced non-drought shock, yes=1		-0.018		-0.003		0.060***
		(0.016)		(0.005)		(0.005)
ln(Land size household owned, ha)		0.093***		0.001		-0.025***
		(0.025)		(0.007)		(0.006)
ln(Livestock owned, in TLU)		0.070***		0.005		-0.012***
		(0.014)		(0.004)		(0.004)
ln(Elevation in meters)		-0.117		-0.041*		-0.021
		(0.083)		(0.025)		(0.028)
ln(total annual rainfall in mm)		2.060***		0.102		-0.265
		(0.771)		(0.224)		(0.252)
ln(total annual rainfall squared)		-0.162***		-0.010		0.019
		(0.059)		(0.017)		(0.019)
ln(mean annual temperature, degrees)		-0.216		-0.032		-0.039
		(0.164)		(0.048)		(0.053)
Proportion of fertile soil in EA		0.073**		0.001		-0.020**
		(0.036)		(0.010)		(0.008)
EA has electricity, yes=1		0.178***		0.053***		-0.032***
		(0.030)		(0.008)		(0.007)
Zone Fixed Effect	No	Yes	No	Yes	No	Yes
Constant	8.553***	4.404	0.479***	0.522	0.082***	1.367
	(0.022)	(2.936)	(0.006)	(0.796)	(0.005)	(1.093)
Number of observations	9,215	9,210	9,606	9,600	9,606	9,600
R2	0.127	0.375	0.152	0.342	0.018	0.217
Adjusted R2	0.127	0.368	0.152	0.335	0.018	0.209

*Note: Village clustered standard error in parentheses: *** p<0.01, ** p<0.05, * p<0.1*

Table A2.6. Impacts of stages of urbanization on welfare, village/EA FE

	ln(Expenditure)		Diet Diversity score		Food security Gap	
Small town	0.653**	0.651***	0.115**	0.075*	-0.052	-0.044
	(0.259)	(0.177)	(0.049)	(0.043)	(0.048)	(0.040)
Intermediate town	0.444*	0.499**	0.205***	0.157***	-0.143*	-0.155**
	(0.265)	(0.201)	(0.054)	(0.050)	(0.076)	(0.064)
Large town	0.632**	0.511**	0.165***	0.116**	-0.054	-0.039
	(0.269)	(0.202)	(0.059)	(0.052)	(0.069)	(0.056)
Survey round, 2015	-0.05***	-0.092**	0.019***	0.006	0.002	-0.031***
	(0.018)	(0.041)	(0.005)	(0.011)	(0.006)	(0.011)
ln(Household size)		-0.46***		0.061***		0.001
		(0.020)		(0.005)		(0.004)
ln(Age of household head in years)		-0.042**		-0.011**		0.011**
		(0.021)		(0.005)		(0.005)
Head is male, yes=1		0.020		-0.002		-0.014***
		(0.015)		(0.004)		(0.004)
Head education, primary=1		0.127***		0.035***		-0.014***
		(0.016)		(0.004)		(0.004)
Head education, secondary or higher=1		0.343***		0.088***		-0.038***
		(0.025)		(0.007)		(0.006)
Household faced drought shock, yes=1		-0.09***		-0.02*		0.077***
		(0.032)		(0.009)		(0.011)
Household faced non-drought shock, yes=1		-0.034**		-0.008*		0.057***
		(0.015)		(0.004)		(0.005)
ln(Land size household owned, ha)		0.145***		0.006		-0.031***
		(0.021)		(0.005)		(0.007)
ln(Livestock owned, in TLU)		0.089***		0.016***		-0.016***
		(0.011)		(0.003)		(0.003)
ln(Elevation in meters)		-0.348		-0.022		0.058
		(0.275)		(0.069)		(0.060)
ln(total annual rainfall in mm)		1.771**		-0.015		-0.276
		(0.827)		(0.246)		(0.281)
ln(total annual rainfall squared)		-0.137**		-0.001		0.019
		(0.063)		(0.019)		(0.021)
ln(mean annual temperature, degrees)		-0.645*		0.032		-0.033
		(0.381)		(0.099)		(0.096)
Proportion of fertile soil in EA		-0.023		-0.030*		0.001
		(0.066)		(0.017)		(0.012)
EA has electricity, yes=1		-0.023		-0.007		-0.014
		(0.080)		(0.013)		(0.020)
EA Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Constant	8.535***	9.623**	0.465***	0.554	0.104***	0.814
	(0.079)	(4.652)	(0.014)	(1.247)	(0.017)	(1.257)
Number of observations	9,215	9,210	9,606	9,600	9,606	9,600
R2	0.004	0.156	0.007	0.104	0.002	0.106
Adjusted R2	0.004	0.155	0.007	0.102	0.002	0.104

*Note: Village clustered standard error in parentheses: *** p<0.01, ** p<0.05, * p<0.1*

Table A2.7. Patterns in status and type of employment by urbanization status

	Rural	Small town	Medium town	Large town
Proportion of employed (%)	85.3	71.4	69	56.4
Proportion of unemployed (%)	1.1	9.1	8.5	15.4
Proportion of inactive (%)	13.7	19.5	22.5	28.3
Average duration of unemployment (mean)	25.1	38.9	28	42.5
Average duration of unemployment (median)	8	13	12	24
Proportion of working additional hours (%)	28.4	17.4	8.7	1.6
Working hours (for employed)				
Average working hours per week on main activity	28.82	36.97	38.52	48.49
Average working hours per week on additional activity	9.28	8.23	8.1	7.46
Average total working hours per week	32.53	39.58	39.93	48.88
Employer (for employed)				
Government	1.8	22.6	22	22.4
Private/NGO	1.66	8.48	15	36.8
Domestic	0.86	4.09	4.13	10.7
Self-employment	48.86	47.99	46.25	25.9
Unpaid employment	46.19	14.43	10.43	0.67
Others	0.62	2.41	2.21	3.53
Type of employment (for employed)				
Managers	0.49	1.68	2.24	3.57
Professionals	0.63	7.35	9.21	13.66
Technicians and associate professionals	0.66	5.96	5.87	7.46
Clerical support workers	0.14	3.78	2.56	4.59
Service and sales workers	4.08	26.87	28.08	23.87
Skilled agricultural, forestry and fishery	66.01	18.71	13.32	1.21
Craft and related trades workers	2.09	8.57	10.32	9.86
Plant and machine operators	0.36	1.36	3.46	8.15
Elementary occupations	25.5	24.71	24.39	26.91
Others	0.03	1.02	0.54	0.73

Source: Ethiopian National Labour Force survey, 2013

Table A2.8. Patterns in inequality in consumption expenditure by urbanization status

Urban category	GE(0)	GE(1)	GE(2)	Gini
Rural	0.20	0.21	0.27	0.34
Small towns	0.21	0.20	0.23	0.35
Intermediate towns	0.19	0.19	0.23	0.33
Large town	0.25	0.32	0.90	0.38
Total	0.23	0.25	0.46	0.37

Source: Authors' computation based on LSMS (2014 & 2016)

Note: GE (0) is the mean logarithmic deviation; GE (1) is the Theil index; GE (2) is half the square of the coefficient of variation

Supplementary materials to Chapter Three

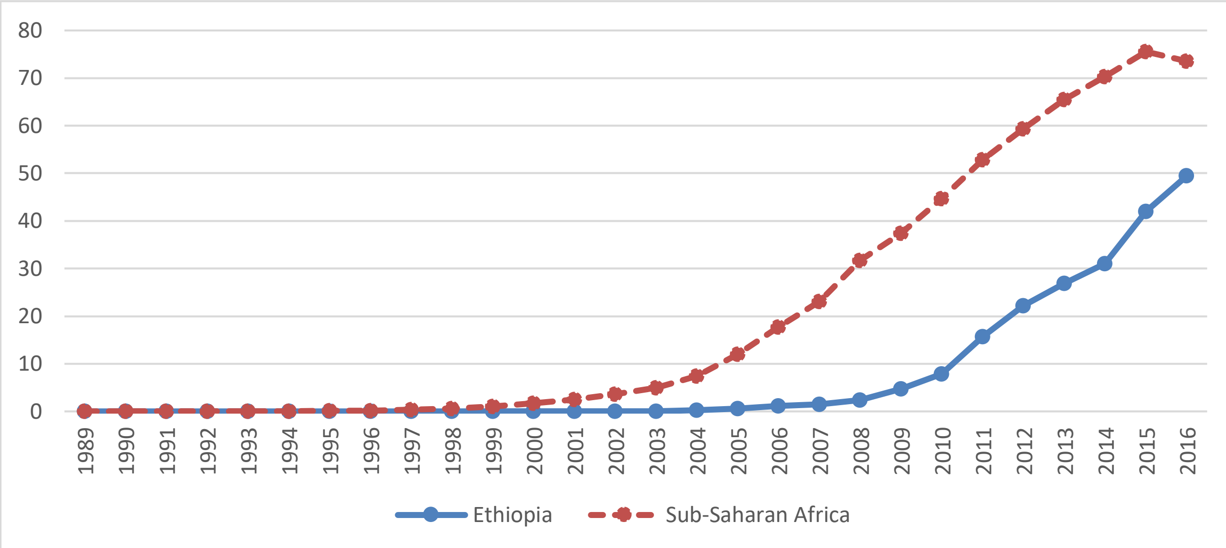


Figure A3.1. Evolution of mobile telephone subscriptions (%)

Source*: World Development Indicators, World Bank, accessed from: https://data.worldbank.org/indicator/IT.CEL.SETS.P2?contextual=region&locations=ET*

Note*: The indicator includes the number of post-paid subscriptions, and the number of active prepaid accounts (i.e. that have been used during the last three months). It applies to all mobile cellular subscriptions that offer voice communications.*

Table A3.1. Descriptive statistics of key variable, over survey rounds

Variables	2012	2014	2016	Total
Panel A: Household-level characteristics				
Number of food groups consumed by HH	4.172	4.605	4.686	4.513
Proportion of food groups consumed by HH	0.348	0.384	0.39	0.376
Transportation cost to the nearest town, ETB	21.595	14.69	17.08	17.452
Distance to the nearest town, Km	28.04	19.30	21.55	22.52
Household size, number	4.754	4.52	4.722	4.656
Age of household head, years	43.988	44.131	46.332	44.864
Male household heads, %	0.746	0.691	0.691	0.706
Heads with primary education, %	0.275	0.285	0.288	0.283
Heads with secondary education, %	0.081	0.194	0.178	0.157
Household took credit, %	0.227	0.251	0.22	0.233
HH has access to electricity, %	0.172	0.418	0.396	0.341
Household runs non-farm enterprise, %	0.27	0.338	0.377	0.333
HH owns Radio, %	0.338	0.359	0.317	0.338
Access to Mobile phone, %	0.298	0.552	0.586	0.493
Livestock owned, TLU	3.814	3.273	3.595	3.537
Durable assets owned, PCA	-0.003	0	0	-0.001
HH has improved roof, %	0.422	0.626	0.649	0.577
HH has improved floor, %	0.048	0.196	0.209	0.159
HH has improved wall, %	0.014	0.08	0.091	0.065
HH distance to the nearest hospital, Km	12.624	12.073	10.192	11.569
HH distance to the nearest commercial bank, Km	31.209	19.345	16.186	21.553
HH distance to the nearest SACCO, Km	.	.	9.227	9.227
HH distance to nearest microfinance institution, Km	14.064	12.266	12.144	12.726
HH distance to the nearest primary school, Km	0.976	2.813	0.733	1.565
HH distance to the nearest secondary school, Km	13.991	12.651	8.576	11.596
Average village elevation, meters	1821.4	1869.9	1872.7	1857.3
Annual Mean Temperature, degrees	195.6	192.7	192.6	193.5
Observations				
Panel B: Child level characteristics				
Child height-for-age z-score	-1.697	-1.42	-1.327	-1.434
Child weight-for-height z-score	-0.401	-0.495	-0.409	-0.439
Child weight-for-age z-score	-1.238	-1.168	-1.204	-1.199
Prevalence of stunting, %	0.449	0.359	0.345	0.371
Prevalence of wasting, %	0.124	0.116	0.134	0.125
Prevalence of underweight, %	0.271	0.241	0.261	0.256
Child is female, %	0.479	0.489	0.486	0.485
Child age in years	2.671	3.719	4.85	3.998
Observations	2,516	4,215	5,301	12,032

Source: LSMS Survey (2012, 2014, 2016)

Table A3.2. Estimation of the propensity score.

Explanatory variables:	Large town, yes=1	
	Coefficient	**Std. Err**
Household size in adult equivalents, number	-0.125**	0.050
Head is male, yes=1	-0.164***	0.046
ln(Age of household head in years)	0.218***	0.063
Head education, primary=1	0.061	0.049
Head education, secondary or higher=1	0.489***	0.067
Primary school in village, yes=1	-0.093**	0.046
Secondary school in village, yes=2	-0.461***	0.055
House has improved roof, yes=1	0.177***	0.046
House has improved wall, yes=1	1.165***	0.093
Household has access to electricity, yes=1	1.352***	0.052
Household owns Radio, yes=1	0.097**	0.043
Household took credit, yes=1	-0.073	0.046
HH affected by health shock, yes=1	0.010	0.049
Constant	-1.625***	0.247
Number of observations	14,113	
Pseudo R2	0.129	

*Note: Robust standard errors in parentheses. Statistical significance indicated by: *** p<0.01, ** p<0.05, * p<0.1; **

Table A3.3. Patterns of household wealth indicators by urbanization status

Variables	Total	Small town	Large town	Mean diff.	Sig.
Real expenditure, ETB	7,775	7,079	8,878	-1,799	***
Real food expenditure, ETB	5,802	5,555	6,193	-638.5	***
Real non-food expenditure, ETB	1,801	1,433	2,384	-951.8	***
Regional spatial price index	1.01	0.97	1.09	-0.1	***
Durable assets owned, PCA	0.0	-0.7	1.1	-1.7	***
Ownership of TV, %	18.9	8.3	35.9	-27.7	***
Ownership of Radio %	33.8	29.3	41.1	-11.8	***
Access to electricity, %	34.2	20.0	56.7	-36.7	***
Access to Mobile phone, %	46.3	38.7	67.2	-28.5	***
Observation	14,173	8,722	5,451		

*Note: Robust standard errors in parentheses. Statistical significance indicated by *** p<0.01, ** p<0.05, * p<0.1; **

Table A3.4. Association between urbanization and participation in the labour market

Outcome:	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
	wage employment		**Non-farm business**		**Off-farm employment**		**Multiple jobs holding**	
ln(Transportation cost)	-0.055***	-0.014***	-0.030***	-0.023***	-0.074***	-0.032***	0.007	0.000
	(0.005)	(0.004)	(0.006)	(0.007)	(0.007)	(0.007)	(0.004)	(0.005)
Large town, yes=1	0.120***	0.050***	-0.038**	-0.001	0.064***	0.039*	-0.074***	-0.019
	(0.015)	(0.013)	(0.018)	(0.021)	(0.022)	(0.023)	(0.013)	(0.014)
Household & location characteristics	No	Yes	No	Yes	No	Yes	No	Yes
Zonal Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes
Constant	0.206***	1.264**	0.353***	-0.025	0.552***	0.398	0.185***	-0.734
	(0.013)	(0.628)	(0.020)	(1.106)	(0.021)	(1.320)	(0.013)	(0.805)
Number of observations	14,087	14,039	14,087	14,039	14,087	14,039	14,087	14,039
R2	0.100	0.229	0.009	0.154	0.055	0.195	0.011	0.102
Adjusted R2	0.100	0.224	0.009	0.148	0.055	0.190	0.011	0.095

*Note: Standard errors clustered at the village level in parentheses. Statistical significance indicated by: *** p<0.01, ** p<0.05, * p<0.1; Coefficients on child, household, and location characteristics omitted to preserve space.*

Table A3.5. Association between urbanization and intensity of employment

Outcome:	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
	wage employment		**Non-farm business**		**Total work hours**		**Total work hours per capita**	
ln(Transportation cost)	-0.224***	-0.056***	-0.136***	-0.094***	0.089***	0.006	0.040**	0.005
	(0.021)	(0.016)	(0.023)	(0.025)	(0.022)	(0.023)	(0.017)	(0.019)
Large town, yes=1	0.495***	0.215***	-0.106	0.005	-0.192***	0.006	-0.070	0.002
	(0.058)	(0.049)	(0.068)	(0.086)	(0.071)	(0.078)	(0.054)	(0.064)
Household & location characteristics	No	Yes	No	Yes	No	Yes	No	Yes
Zonal Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes
Constant	0.800***	4.663*	1.290***	1.221	3.197***	2.304	2.279***	2.477
	(0.053)	(2.422)	(0.077)	(4.256)	(0.068)	(4.401)	(0.052)	(3.712)
Number of observations	14,087	14,039	14,087	14,039	14,087	14,039	14,087	14,039
R2	0.109	0.245	0.013	0.168	0.011	0.217	0.003	0.129
Adjusted R2	0.109	0.240	0.012	0.162	0.010	0.212	0.003	0.123

*Note: Standard errors clustered at the village level in parentheses. Statistical significance indicated by: *** p<0.01, ** p<0.05, * p<0.1; Coefficients on child, household, and location characteristics omitted to preserve space.*

Table A3.6. Patterns in proportion of weaning children by age and urbanization status

Age category	Urban	Rural	Total
<2 months	2.8	1.6	1.7
2-4 months	2.8	1.9	2.0
4-6 months	4.4	3.3	3.5
6-12 months	3.7	2.9	3.0
12-24 months	20.0	16.9	17.3
>24 months	84.2	81.7	82.1
Observation	1,986	9,668	11,654

Source: Ethiopian Demographic and Health Survey, DHS, 2010

Table A3.7. Descriptive statistics of key variable by urbanization status

Variables	Total	Small town	Large town	Addis Ababa
Number of food groups consumed by HH	4.5	4.4	4.7	4.8
Proportion of food groups consumed by HH	0.4	0.4	0.4	0.4
Transportation cost to the nearest town, ETB	17.5	18.9	16.9	-
Distance to the nearest town, Km	22.5	19.9	29.6	-
Household size, number	4.7	4.9	4.3	4.1
Age of household head, years	44.9	45.4	43.7	46.3
Male household heads, %	70.6	73.2	67.8	53.9
Heads with primary education, %	28.3	28.3	28.4	28.5
Heads with secondary education, %	15.7	9	23.5	52.9
Household took credit, %	23.3	24.6	22.6	9.7
HH has access to electricity, %	34.1	20	52.2	97.6
Household runs non-farm enterprise, %	33.3	30.9	38.2	27.4
HH owns Radio, %	33.8	29.3	38	69.1
Access to Mobile phone, %	49.3	40.1	60.6	95
Livestock owned, TLU	3.54	4.29	2.59	0.00
Durable assets owned, PCA	0.00	-0.66	0.59	5.20
HH has improved roof, %	57.7	48.7	69.2	98.2
HH has improved floor, %	15.9	6.2	26.9	72.2
HH has improved wall, %	6.5	1.9	13.1	21
HH distance to the nearest hospital, Km	11.6	13.8	9.0	-
HH distance to the nearest commercial bank, Km	21.6	27.9	11.1	13.9
HH distance to the nearest SACCO, Km	9.2	12.5	5.5	0.6
HH distance to the nearest microfinance institution, Km	12.7	15.7	8.8	0.1
HH distance to nearest primary school, Km	1.6	1.4	2.0	0.5
HH distance to nearest secondary school, Km	11.6	13.8	6.3	23.4
Average village elevation, meters	1,857	1,881	1,757	2,382
Annual Mean Temperature, degrees	193	192	199	162
Observations	14,173	8,722	4,907	544
Child height-for-age z-score	-1.434	-1.508	-1.303	-1.02
Child weight-for-height z-score	-0.439	-0.482	-0.39	0.259
Child weight-for-age z-score	-1.199	-1.265	-1.098	-0.584
Prevalence of stunting, %	37.1	39.2	33.4	25.6
Prevalence of wasting, %	12.5	12.9	12	7
Prevalence of underweight, %	25.6	27.4	22.8	12.6
Child is female, %	48.5	48.3	49	49.4
Child age in years	3.998	4.01	3.979	3.902
Observations	12,035	8,045	3,739	251

Source: LSMS Survey (2012, 2014, 2016)

Supplementary Materials to Chapter Four

Table A4.1. Descriptive statistics of key variables by urbanization status

Variables	N	Rural	Small towns	Large towns	F-test: p-val
Education of Individuals					
No education	44,103	0.46	0.47	0.30	0.00
Primary	44,103	0.46	0.44	0.44	0.00
Secondary	44,103	0.06	0.07	0.15	0.00
Tertiary	44,103	0.02	0.03	0.11	0.00
Parental Education					
No education	44,833	0.585	0.563	0.352	0.00
Primary	44,833	0.331	0.353	0.343	0.00
Secondary	44,833	0.044	0.044	0.146	0.00
Tertiary	44,833	0.04	0.04	0.159	0.00
Occupation of Individual					
None	31,867	0.416	0.419	0.437	0.00
Skilled Agriculture	31,867	0.334	0.339	0.18	0.00
Unskilled Agriculture	31,867	0.112	0.094	0.165	0.00
Self-employment	31,867	0.065	0.082	0.107	0.00
Skilled wage	31,867	0.017	0.02	0.072	0.00
Parental Education					
Skilled Agriculture	29,466	0.93	0.907	0.744	0.00
Unskilled Agriculture	29,466	0.009	0.014	0.066	0.00
Self-employment	29,466	0.057	0.075	0.162	0.00
Skilled wage	29,466	0.004	0.004	0.028	0.00
Individual & Household characteristics					
Age in years	44,542	26.9	26.6	27.7	0.00
Male, Yes=1	44,760	0.49	0.49	0.47	0.00
Household size	44,836	5.7	5.8	5.3	0.00
Age in years (head)	44,815	48.1	47.4	46.8	0.00
Male, Yes=1 (head)	44,836	0.76	0.806	0.712	0.00
Asset Index	44,719	-1.01	-1.06	1.52	0.00

Source: LSMS Survey (2014 & 2016)

Notes: Sum of the nighttime light at EA level is used to classify the households from rural (tercile with the smallest light intensity) to large towns (tercile with the highest light intensity). The labour variables represent extensive margin (i.e., participation in wage, non-farm business or agriculture), but individuals do not necessarily work only on one activity.

Table A4.2. Transition matrix of educational and occupational status

Panel A: Transition Matrix for Education					
Child\Parent	No Education	Elementary education	Secondary education	Tertiary education	Total
No education	0.61	0.26	0.12	0.07	0.40
Elementary education	0.34	0.65	0.40	0.32	0.45
Secondary education	0.05	0.07	0.42	0.17	0.10
Tertiary education	0.01	0.02	0.06	0.44	0.05
Panel B: Transition matrix for Occupation					
Child\Parent	Elementary Occupation	Unskilled wage	Self-employment	Skilled wage	Total
Elementary Occupation	0.79	0.51	0.53	0.49	0.75
Unskilled wage	0.10	0.20	0.22	0.13	0.12
Self-employment	0.08	0.10	0.15	0.07	0.09
Skilled wage	0.02	0.19	0.10	0.31	0.04

Source: LSMS Survey (2014 & 2016)

Table A4.3. Association between child and parental education, ordered logit model

	[1]	**[2]**	**[3]**
Parental Education			
No Education	[reference]	[reference]	[reference]
Elementary	1.078***	1.240***	1.229***
	(0.034)	(0.041)	(0.041)
Secondary	2.805***	2.575***	2.594***
	(0.067)	(0.078)	(0.080)
Tertiary	4.359***	4.056***	4.075***
	(0.092)	(0.115)	(0.115)
ln(Age in years)		-1.208***	-1.227***
		(0.033)	(0.033)
Male, yes=1		0.807***	0.810***
		(0.029)	(0.029)
Household size, number		-0.035***	-0.031***
		(0.007)	(0.007)
ln(Age of household head in years)		0.696***	0.697***
		(0.060)	(0.060)
Head is male, yes=1		-0.435***	-0.426***
		(0.042)	(0.042)
Durable assets owned, PCA		0.220***	0.214***
		(0.008)	(0.009)
ln(Village elevation, m)		0.254***	0.273**
		(0.091)	(0.126)
ln(Annual Mean Temperature, degrees)		-0.000	-0.067
		(0.198)	(0.238)
Survey round, 2014		-0.026*	-0.027*
		(0.015)	(0.015)
Location Fixed Effect?	No	No	Yes
Number of observations	36,026	35,885	35,885
Pseudo R2	0.133	0.217	0.220

Source: LSMS Survey (2014 & 2016)

*Note: Standard errors clustered at the household level in parentheses. Statistical significance indicated by: *** p<0.01, ** p<0.05, * p<0.1. Reported coefficients are from ordered logit model.*

Table A4.4. Association between child and parental occupation, ordered logit model

	[1]	[2]	[3]
Parental Occupation			
Elementary occupation	[reference]	[reference]	[reference]
Unskilled wage	1.392***	0.397***	0.468***
	(0.083)	(0.089)	(0.091)
Self-Employment	1.196***	0.552***	0.571***
	(0.049)	(0.058)	(0.058)
Skilled wage	1.756***	0.652***	0.677***
	(0.187)	(0.201)	(0.206)
ln(Age in years)		0.339***	0.342***
		(0.033)	(0.034)
Male, yes=1		0.582***	0.589***
		(0.033)	(0.033)
Household size, number		-0.150***	-0.147***
		(0.010)	(0.010)
ln(Age of household head in years)		-1.066***	-1.054***
		(0.063)	(0.063)
Head is male, yes=1		-0.076	-0.097**
		(0.047)	(0.046)
Durable assets owned, PCA		0.223***	0.239***
		(0.007)	(0.008)
ln(Village elevation, m)		0.098	0.031
		(0.102)	(0.130)
ln(Annual Mean Temperature, degrees)		0.510**	0.179
		(0.246)	(0.270)
Survey round, 2014		-0.047	-0.039
		(0.029)	(0.029)
Location Fixed Effect?	No	No	Yes
Number of observations	28,605	28,491	28,491
Pseudo R2	0.030	0.109	0.111

Source: LSMS Survey (2014 & 2016)

*Note: Standard errors clustered at the household level in parentheses. Statistical significance indicated by: *** p<0.01, ** p<0.05, * p<0.1. Reported coefficients are from ordered logit model.*

Table A4.5. Association between child and parental occupation, by urbanization status

	Rural	Small Town	Large Town
Parental Occupation			
Elementary occupation	[reference]	[reference]	[reference]
Unskilled wage	0.355	-0.527	0.115
	(0.245)	(0.401)	(0.099)
Self-Employment	0.495***	0.950***	0.108
	(0.137)	(0.164)	(0.075)
Skilled wage	-0.103	-0.480	0.365
	(0.374)	(0.869)	(0.223)
ln(Age in years)	0.417***	0.493***	0.669***
	(0.055)	(0.095)	(0.063)
Male, yes=1	0.257***	0.254***	0.538***
	(0.054)	(0.096)	(0.052)
Education level attained			
No Education	[reference]	[reference]	[reference]
Primary	0.836***	0.735***	0.795***
	(0.064)	(0.105)	(0.071)
Secondary or higher	1.711***	1.509***	1.622***
	(0.099)	(0.170)	(0.091)
Household size, number	-0.115***	-0.123***	-0.126***
	(0.016)	(0.030)	(0.014)
ln(Age of household head in years)	-0.630***	-0.932***	-1.011***
	(0.104)	(0.177)	(0.094)
Head is male, yes=1	-0.179**	-0.262**	0.026
	(0.075)	(0.128)	(0.065)
Durable assets owned, PCA	0.287***	0.241***	0.108***
	(0.025)	(0.039)	(0.010)
ln(Village elevation, meter)	0.032	-1.080	0.442
	(0.195)	(0.719)	(0.361)
ln(Annual Mean Temperature, degrees)	0.292	-1.315	1.241**
	(0.409)	(1.022)	(0.606)
Survey round, 2014	-0.015	-0.226***	0.068
	(0.052)	(0.086)	(0.042)
Location Fixed Effect?	Yes	Yes	Yes
Number of observations	13,181	4,578	10,677
Adjusted R2	0.121	0.120	0.134

Source: LSMS Survey (2014 & 2016)

*Note: Standard errors clustered at the household level in parentheses. Statistical significance indicated by *** p<0.01, ** p<0.05, * p<0.1*

Table A4.6. Education expenditure by wealth category and parental occupation

Categories	Expenditure in Birr			Share of education (%)	
	Education	Non-food	Total	Non-food	Total
Panel A: Wealth category					
Poorest group	292.1	3,252.3	20,699.2	8.98	1.41
Middle-income group	401.5	4,944.0	25,390.9	8.12	1.58
Richest group	2,450.4	11,469.8	41,968.3	21.36	5.84
Panel B: Parental occupation					
Elementary occupation	691	5,616	27,113	12.3	2.5
Unskilled wage	2,558	9,568	35,944	26.7	7.1
Self-employment	2,375	10,618	39,803	22.4	6.0
Skilled wage	4,277	13,692	46,901	31.2	9.1

Source: LSMS Survey (2014 & 2016)

Note: Wealth category is generated from individual asset items owned by households using principal component analysis (PCA).

Table A4.7. Mobility in education, coefficients from ordered logit model, by migration status

Variables	**Migrant**	**Non-migrant**
Parental Education		
No Education	[reference]	[reference]
Elementary	0.602***	1.480***
	(0.086)	(0.046)
Secondary	0.961***	3.112***
	(0.171)	(0.084)
Tertiary	1.328***	4.961***
	(0.207)	(0.122)
ln(Age in years)	-0.268***	-1.455***
	(0.084)	(0.037)
Male, yes=1	0.265***	0.938***
	(0.063)	(0.032)
Household size, number	0.007	-0.042***
	(0.017)	(0.008)
ln(Age of household head in years)	0.719***	0.836***
	(0.130)	(0.062)
Head is male, yes=1	-0.198**	-0.527***
	(0.085)	(0.046)
Durable assets owned, PCA	0.178***	0.221***
	(0.022)	(0.008)
ln(Village elevation, m)	0.212	0.251*
	(0.261)	(0.129)
ln(Annual Mean Temperature)	-0.285	-0.093
	(0.509)	(0.241)
Survey round, 2014	-0.155***	-0.035*
	(0.043)	(0.020)
Location Fixed Effect?		
Number of observations	5,723	30,162
Pseudo R2	0.073	0.266

Source: LSMS Survey (2014 & 2016)

*Note: Standard errors clustered at the household level in parentheses. Statistical significance indicated by *** p<0.01, ** p<0.05, * p<0.1*

Table A4.8. Mobility in occupation, coefficients from ordered logit model, by migration status

Variables	**Migrant**	**Non-migrant**
Elementary occupation		
Unskilled wage	0.854	0.151
	(0.531)	(0.095)
Self-Employment	0.017	0.323***
	(0.119)	(0.074)
Skilled wage	-0.149	0.510**
	(0.415)	(0.212)
ln(Age in years)	1.381***	0.868***
	(0.126)	(0.046)
Male, yes=1	0.688***	0.362***
	(0.081)	(0.040)
Education level attained (REF:No Education)		
Primary education	0.106	0.809***
	(0.092)	(0.049)
Secondary or higher	0.606***	1.736***
	(0.151)	(0.066)
Household size, number	0.000	-0.113***
	(0.021)	(0.012)
ln(Age of household head in years)	0.030	-1.352***
	(0.154)	(0.078)
Head is male, yes=1	-0.046	-0.117**
	(0.109)	(0.055)
Durable assets owned, PCA	-0.023	0.184***
	(0.026)	(0.009)
ln(Village elevation, m)	0.318	-0.042
	(0.352)	(0.152)
ln(Annual Mean Temperature, deg C)	0.133	0.363
	(0.642)	(0.326)
Survey round, 2014	-0.124	0.093***
	(0.106)	(0.033)
Location Fixed Effect?		
Number of observations	3,236	25,200
Pseudo R2	0.075	0.169

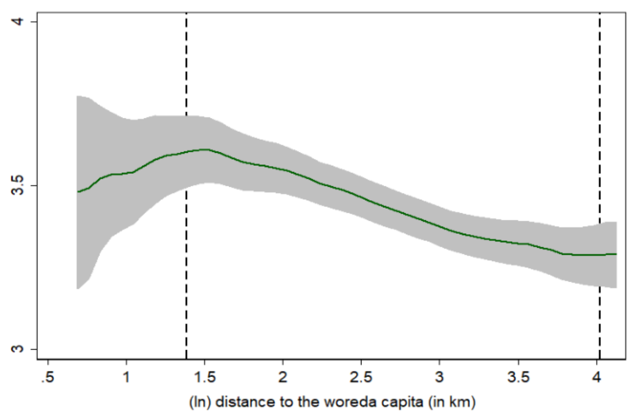
Source: LSMS Survey (2014 & 2016)
*Note: Standard errors clustered at the household level in parentheses. Statistical significance indicated by *** p<0.01, ** p<0.05, * p<0.1*

Table A4.9. Major reasons for school dropout

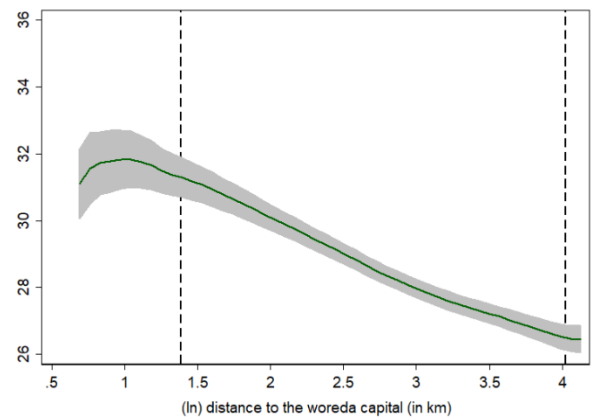
Panel A: Dropout rate by place of residence			
Location	Elementary school	High school	Observation
Rural	43.3	6.7	15,893
Small Town	32.3	16.7	2,727
Large Town	26.0	20.7	7,730
Total	37.1	11.8	26,350

Panel B: Reason for dropout		
Reasons	Freq.	Percent
No time / no interest	1,689	29.98
Marital obligation	1,092	19.39
Domestic obligation	593	10.53
Lack of money	544	9.66
Too old to attend	500	8.88
Had enough schooling	479	8.5
Other Specify	367	6.52
Sickness	183	3.25
Death of parents	74	1.31
Awaiting admission	49	0.87
Separation of parents	39	0.69
No school / lack of teachers	15	0.27
Disability	9	0.16
Total	5,633	100

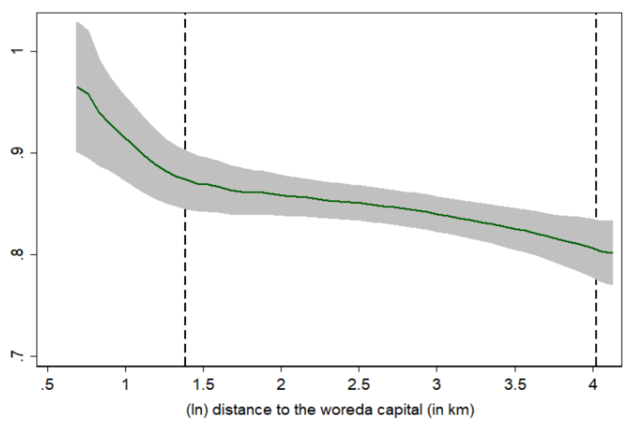
Supplementary materials to Chapter Five



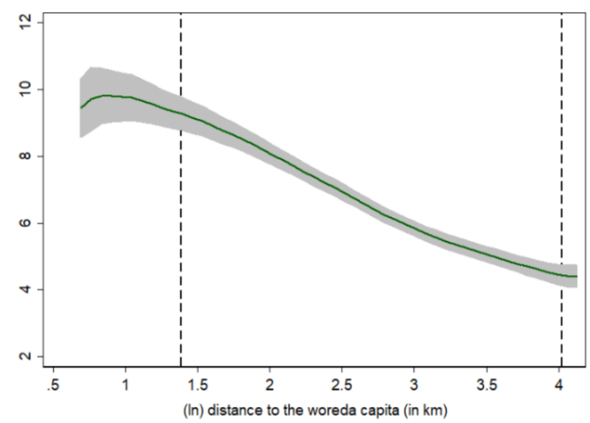
a) Number of EAs in a Kebele



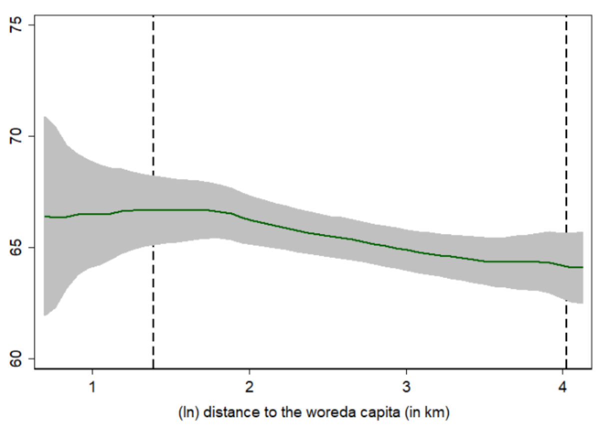
b) Average age of EAs in Kebele in years



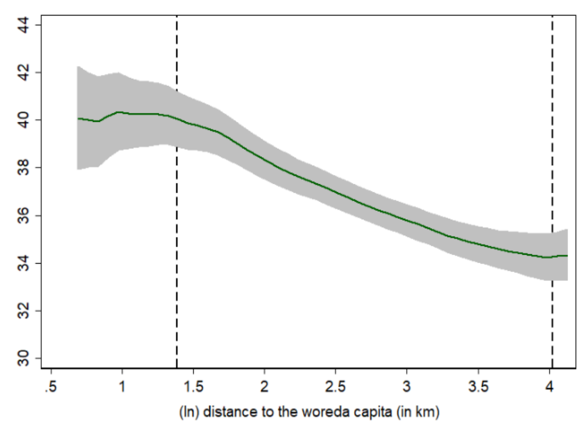
c) Education level of EAs (Diploma or higher)



d) Average years of work experience of EA



e) Technical knowledge of agri. practices



f) Average number of working hours per week

Figure A5.1. Profile and effort levels of Extension Agents (EAs) by remoteness

Source: *Authors' calculation based on IFPRI-Digital Green's EA survey, 2017, 2018 and 2019.*

Note: *The locally weighted polynomial regressions of EAs characteristics on the distance from the Kebele to the district capital. The area between the dashed lines indicates the 90% of the distance distribution.*

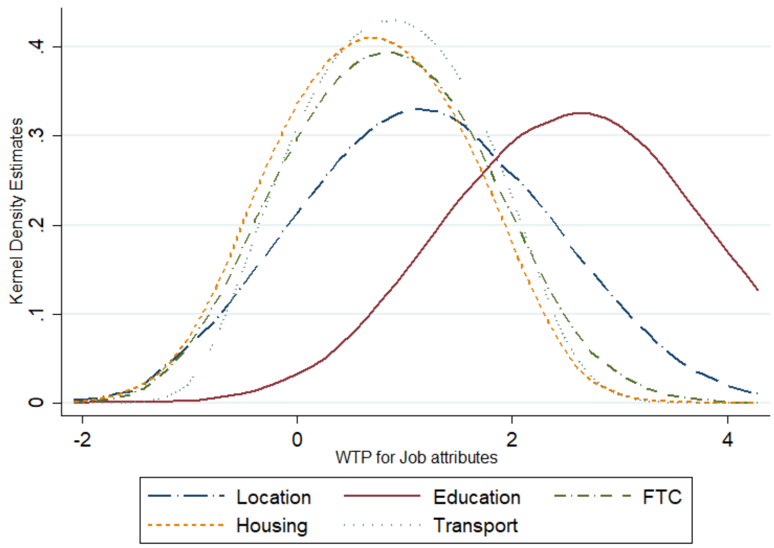


Figure A5.2. Kernel density of willingness to pay for attribute parameters

Source: Authors' calculation based on IFPRI-Digital Green's EA survey, 2019.

Table A5.1. Characteristics of extension agents in study sample, over survey rounds

Characteristics	2017	2018	2019	All
Male	0.75	0.77	0.75	0.76
Age, years	27.8	28.8	29.3	28.6
Number of years working as an EA	5.59	6.72	7.27	6.48
Number of years working in current Kebele	1.88	2.53	2.6	2.31
Education: Certificate, yes=1	0.25	0.17	0.03	0.16
Education: Diploma, yes=1	0.54	0.55	0.7	0.6
Education: Degree, yes=1	0.2	0.28	0.27	0.25
Computer literate, yes=1	0.46	0.41	0.5	0.46
Mobile with internet access, yes=1	0.37	0.48	0.62	0.48
Spent childhood: In working Kebele, yes=1	0.09	0.1	0.09	0.09
Spent childhood: In working Woreda, yes=1	0.6	0.61	0.64	0.61
Spent childhood: In working zone, yes=1	0.83	0.85	0.87	0.85
Number of EAs in Kebele	3.47	3.53	3.27	3.42
Number of Farmers' field days organized	2.02	1.86	1.83	1.91
Working hours per week: Planting season	51	49.6	46.1	49
Working hours per week: Harvesting season	40	36	33.5	36.7
Working hours per week: Slack season	26.1	22.4	23	23.9
Working hours per week: Average	39	36	34.2	36.5
Knowledge score: Teff		68.4	72.3	70.3
Knowledge score: Maize		65.1	69.7	67.4
Knowledge score: Wheat		62.1	68.9	65.5
Knowledge score: Average		65.2	70.3	67.7
Observations	896	781	763	2,440

***Source**: Authors' calculation based on IFPRI-Digital Green's EA survey, 2017, 2018, and 2019.*

***Note**: Knowledge score refers to EAs' work-related knowledge score (out of 100) obtained through quizzes. The knowledge questions focused on growing practices of Teff, maize, and wheat.*

Table A5.2. Description of the choice experiment task

Instruction for choice experiment

[Interviewer: Please read the following aloud to the respondent]

Below I will present you with a number of jobs with features similar to current working conditions in rural Ethiopia. However, these jobs are hypothetical and do not necessarily reflect the working condition of any particular job. In each round, I would like to ask you to make a choice between two jobs based only on the information given corresponding to the jobs. Since these jobs are hypothetical, you are not supposed to evaluate whether the jobs are realistic or not. Based only on the information given, you need only to choose which one of the two jobs (Job 1, Job 2, or neither) you prefer. Assume that all unstated characteristics of jobs are the same for the two jobs.

Description of the Attributes of the Jobs:

Location: Refers to the quality of services at the location of work and distance from quality services and takes **two values**.

1. **Advanced location**: Location with reliable mobile telephone coverage, electricity, and piped water.
2. **Remote location**: Location with no or unreliable mobile telephone coverage, electricity, and piped water.

Pay (salary): This assumes **four values**.

1. **Minus 25%**: 25 percent less than the current net salary of the responding EA. That is, the current basic salary minus 25 percent of the current basic salary;
2. **Plus 25%**: Current basic salary plus 25 percent of current basic salary.;
3. **Plus 50%**: Current basic salary plus 50 percent of current basic salary;
4. **Plus 100%**: Twice current basic salary.

Housing: This refers to the Provision of government housing at Kebele of work for residence of the extension agent and his/her family. It takes two levels:

1. **No**: No housing;
2. **Yes**: Housing with basics enough for survival.

Farmer Training Centres (FTC): This stands for extension tools at FTC and assumes **two levels**:

1. **Inadequate**: Not enough resources to effectively deliver extension service to farmers, e.g., no demonstration plot, inadequate budget to run the FTC, inadequate teaching materials)
2. **Adequate**: Enough resources to effectively run the FTC as well as deliver extension service to farmers.

Transportation: This assumes **two levels:**

1. **No**: No transportation facility at the FTC (no bicycle, motorcycle, or horse)
2. **Yes**: FTC has its own transportation facility (e.g. bicycle, motorcycle, or horses depending on the availability and quality of roads).

Educational opportunity: this refers to government-sponsored continuing education opportunity and takes **two levels:**

1. **No**: No educational opportunities

2. **Yes**: Educational opportunity offered after 2 years of service

Do you have any questions? Is everything clear?
Now, let us take the following as an example.
If you are given the opportunity to choose between **J1 and J2**, which job would you choose?

Block-1; Question -1

Job profiles	Job 1	Job 2
Location	Remote	Advanced
Salary	Plus25%	Plus50%
Housing availability	Yes	No
Transport access	No	Yes
FTC tools/equipment's	Inadequate	Adequate
Education opportunity	No	Yes
Answer (choice)		
1. Job 1		
2. Job 2		
3. Neither job 1 nor job 2		

Table A5.3. Testing for the assumption of independence of irrelevant alternatives (IIA)

	Full (b)	Deducted (B)	Difference (b-B)	S.E. (b-B)
Housing, yes=1	0.291	0.134	0.157	0.016
Transport services, yes=1	0.418	0.334	0.084	0.008
Adequate FTC, yes=1	0.442	0.409	0.032	0.009
Education opportunities, yes=1	1.234	1.110	0.124	0.016
Salary (ref: current basic salary)				
Salary increment of 100%, yes=1	1.843	2.413	-0.570	0.035
Salary increment of 50%, yes=1	1.364	2.072	-0.708	0.053
Salary increment of 25%, yes=1	0.949	1.532	-0.583	0.048
Salary reduction by 25%, yes=1	0.337	1.069	-0.732	0.057
Location is advanced, yes=1	0.578	.	.	.
constant	-0.036	-0.283	0.2475	0.013
Number of respondents	761	761		
Number of observations	18,264	18,264		
Log-likelihood	-3649.2	-3767.5		
Hausman test:				
chi-square (6):	139.45			
-----p-val:	0.000			

Source: Authors' calculation based on IFPRI-Digital Green's EA survey, 2019.
Note: Hausman's test statistics shows that the assumption of independence of irrelevant alternatives (IIA) is rejected at 1% level.

Table A5.4. Preferences for job attributes, sub-sample analysis

	1	2	3	4
	Advanced education [a]		**Intrinsically motivated [b]**	
	No	**Yes**	**No**	**Yes**
Location is advanced, yes=1	**0.799*** **	**0.919*** **	**1.025*** **	**0.603*** **
	(0.082)	(0.153)	(0.112)	(0.087)
Housing, yes=1	**0.435*** **	**0.707*** **	0.477***	0.483***
	(0.070)	(0.129)	(0.089)	(0.079)
Transport services, yes=1	0.688***	0.645***	**0.590*** **	**0.740*** **
	(0.069)	(0.123)	(0.089)	(0.081)
Adequate FTC, yes=1	**0.635*** **	**1.081*** **	0.854***	0.638***
	(0.070)	(0.134)	(0.098)	(0.074)
Education opportunity, yes=1	**2.010*** **	**2.178*** **	**2.195*** **	**1.844*** **
	(0.109)	(0.194)	(0.143)	(0.123)
Salary (ref: current basic salary)				
Salary increment of 100%, yes=1	**1.867*** **	**1.698*** **	**1.277*** **	**2.778*** **
	(0.164)	(0.287)	(0.194)	(0.240)
Salary increment of 50%, yes=1	**1.197*** **	**0.376**	**0.600*** **	**2.013*** **
	(0.156)	(0.284)	(0.184)	(0.235)
Salary increment of 25%, yes=1	**0.511*** **	**0.019**	**-0.223**	**1.614*** **
	(0.148)	(0.251)	(0.172)	(0.229)
Salary reduction by 25%, yes=1	**-0.483*** **	**-1.534*** **	**-1.706*** **	**0.857*** **
	(0.184)	(0.339)	(0.235)	(0.253)
Constant	**-0.136** **	**-0.045**	**-0.215** **	**-0.014**
	(0.065)	(0.115)	(0.084)	(0.072)
Number of respondents	558	203	384	377
Number of observations	13,392	4,872	9,216	9,048
Chi-squared (df = 9)	326	139	313	146
Log-likelihood	-2,519	-885	-1,813	-1,547
Pseudo R2	0.061	0.07	0.08	0.04

Source: Authors' calculation based on IFPRI-Digital Green's EA survey, 2019.
Note: Standard error given in parenthesis; triple (***), double (**), and single (*) represent statistical significance at 1%, 5%, and 10% level, respectively. . The differences between the subgroups that are statistically significant are in bold [a] Advanced education: Those that have first degree. Intrinsically motivated: Those that reported "helping others" as the main motivation for being EAs.