DISCUSSION PAPER SERIES

DP14515
(v. 2)

THE PRICE OF REMOTENESS: PRODUCT AVAILABILITY AND LOCAL COST OF LIVING IN ETHIOPIA

Julien Martin, Florian Mayneris and Ewane Theophile

DEVELOPMENT ECONOMICS

INTERNATIONAL TRADE AND REGIONAL ECONOMICS

ISSN 0265-8003

THE PRICE OF REMOTENESS: PRODUCT AVAILABILITY AND LOCAL COST OF LIVING IN ETHIOPIA

Julien Martin, Florian Mayneris and Ewane Theophile

Discussion Paper DP14515
First Published 20 March 2020
This Revision 23 March 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

THE PRICE OF REMOTENESS: PRODUCT AVAILABILITY AND LOCAL COST OF LIVING IN ETHIOPIA

Abstract

This paper uses the microdata underlying the Ethiopian CPI to examine the spatial dispersion in local prices and availability of 400 items across more than 100 cities. We first show that remote cities face higher prices for individual products and have access to fewer products. All else equal, large cities also face higher individual prices but enjoy access to a wider set of products. To assess the welfare implications of these micro patterns, we then examine the impact of remoteness and population size on aggregate cost-of-living indexes that account for product availability. We find the cost of living is higher in remote areas, but not systematically related to population size. We then show spatial differences in the cost of living are a significant determinant of migration flows across Ethiopian regions. The impact of cost of living mainly channels through spatial differences in product availability.

JEL Classification: N/A

Keywords: N/A

Julien Martin - martin.julien@uqam.ca
Université du Québec à Montréal, CEPR and CEPR

Florian Mayneris - mayneris.florian@uqam.ca
Université du Québec à Montréal, CEPR and CEPR

Ewane Theophile - william_theophile.ewane@courrier.uqam.ca
Université du Québec à Montréal

Acknowledgements
We thank Simon Alder, Nate Baum-Snow, Kristian Behrens, Remi Jedwab, François Libois, Marion Mercier, Daniel Pereira Arellano, Tsur Sommerville, Will Strange, Gonzague Vannoorenberghe, and participants at various seminars and conferences for helpful discussions.

The price of remoteness: Product availability and local cost of living in Ethiopia*

Julien Martin†[1,2], Florian Mayneris‡[1,2], and Ewane Theophile§[1]

[1]Université du Québec, Montréal
[2]Centre for Economic Policy Research

March 4, 2020

Abstract

This paper uses the microdata underlying the Ethiopian CPI to examine the spatial dispersion in local prices and availability of 400 items across more than 100 cities. We first show that remote cities face higher prices for individual products and have access to fewer products. All else equal, large cities also face higher individual prices but enjoy access to a wider set of products. To assess the welfare implications of these micro patterns, we then examine the impact of remoteness and population size on aggregate cost-of-living indexes that account for product availability. We find the cost of living is higher in remote areas, but not systematically related to population size. We then show spatial differences in the cost of living are a significant determinant of migration flows across Ethiopian regions. The impact of cost of living mainly channels through spatial differences in product availability.

*We thank Simon Alder, Nate Baum-Snow, Kristian Behrens, Remi Jedwab, François Libois, Marion Mercier, Daniel Pereira Arellano, Tsur Sommerville, Will Strange, Gonzague Vannoorenberghe, and participants at various seminars and conferences for helpful discussions.

†Corresponding author: martin.julien@uqam.ca.

‡mayneris.florian@uqam.ca.

§william_theophile.ewane@courrier.uqam.ca.

1 Introduction

Spatial inequality is pervasive within developing countries, and accounting for cost-of-living differences across places is key to a comprehensive view of its patterns (Ravallion and Van De Walle, 1991).[1] The literature so far has mainly focused on cost-of-living differences between urban and rural areas driven by price differentials of available products and services (see e.g. Ravallion and Van De Walle, 1991; Deaton and Tarozzi, 2000; Muller, 2002).

In this paper, we propose a fresh view on the geography of cost of living within a developing country - Ethiopia. First, we compute a measure of the cost of living that accounts for products and services availability on top of their price, which seems crucial because many products and services are locally missing in developing countries (see, e.g., FAO et al., 2018; WFP and CSA, 2018; FAO et al., 2019).[2] Second, instead of focusing on the urban-rural divide, we provide insights on spatial cost-of-living differences across cities, which allows us to relate the cost of living to two fundamental aspects of economic geography, namely, the population size and remoteness of cities. We find that cost of living is higher in remote places but that population size has no effect on it. Eventually, we show that properly accounting for spatial cost-of-living differences is important to understand the geography of internal migration flows.

The effect of economic geography on the local cost of living is not trivial, because several opposing forces operate. On the one hand, larger cities host higher-income households, which is likely to push prices up. On the other hand, consumers in small and remote cities may face higher trade costs, pushing the price of tradable goods up in those small and remote places.[3] Moreover, an under-studied component of the cost of living is product availability (Feenstra, 1994). Consumers in large and central cities may enjoy direct access to a greater number of products and services, which in the case of consumers' love for variety will push the relative cost of living down in those locations. What forces dominate is a quantitative question, but quite surprisingly, the literature has thus far provided little evidence on the spatial differences in terms of cost of living within developing countries.

We try to bridge this gap by estimating the impact of remoteness and population size on individual prices, product availability, and cost-of-living indexes at the city-level in

[1] See Ferré et al. (2012) and Young (2013) for nominal consumption and nominal income differences across urban and rural areas, and Gollin et al. (2017) for spatial differences in terms of non-monetary amenities.

[2] In our data, about 35% of product-city pairs suffer from unavailability.

[3] The role of transportation costs is likely to be prevalent in a developing country such as Ethiopia where internal transport costs are high (e.g., Atkin and Donaldson, 2015; Rancourt et al., 2014).

Ethiopia. We leverage the microdata underlying the Ethiopian CPI to examine spatial dispersion in local prices and product availability across more than 100 markets (more or less equivalent to cities in our data - we refer to them as cities in the rest of the paper). We find that individual prices increase with city size and remoteness. Product availability, that is the probability that a product is available in a given city, also increases with city size, but it decreases with its degree of remoteness. Whereas remoteness has an unambiguous effect on the local cost of living by increasing prices and reducing product availability, the net benefit of living in a large city is less clear.

To make progress on this front, we aggregate our data and compute local cost-of-living indexes; we then examine how these indexes vary with city size and remoteness. We use two alternative measures of cost of living. The first measure is the spatial version of the Feenstra (1994) CES price index. Although this index was originally built to account for product availability over time, it has been adapted to measure spatial differences in the cost of living (Handbury and Weinstein, 2015). This index rests on the assumption that the price of unavailable varieties tends to infinity. We develop an alternative index that relaxes this assumption by assuming instead that, if a product is not available in a location, a consumer can still purchase it by traveling (thereby incurring travel costs) to the cheapest alternative location where the variety is available.[4]

Armed with these two alternative measures, we show the cost of living is significantly higher in remote locations. However, we do not find systematic evidence that population size affects the cost of living, which suggests the price premium paid for available products in large cities is entirely compensated by access to a wider array of products. These results are robust to computing the cost of living with different methods and different weighting schemes. They also are not sensitive to the introduction of various control variables.

We then use data on income per capita in Ethiopian regions to show that accounting for cost-of-living differences is important to have an accurate view of spatial income inequality. The ranking of regions based on their income per capita adjusted for the price index changes significantly if one includes product availability in the computation of the cost-of-living index.

We eventually investigate how spatial differences in terms of cost of living shape migration flows across Ethiopian regions. We run gravity-like equations and find that cost-of-

[4]In an inter-temporal context, the assumption that a consumer at date t cannot consume goods that disappeared between $t-1$ and t, and that consumers in $t-1$ cannot enjoy varieties that appear at date t is reasonable. However, although consumers cannot travel over time, they can travel over space - and even more so within a country. Therefore, the assumption that consumers in a location cannot benefit from varieties available in other locations may be more debatable when computing a spatial price index.

living differences are a significant determinant of internal migration flows. All else equal, people prefer moving to low-cost-of-living regions. Moreover, accounting for differences in terms of cost of living substantially improves the predictive power of the gravity model (the adjusted R-square of the regression rises by more 20% when cost of living is controlled for). Also note that using a standard price index that neglects product availability does not improve the predictive power of the gravity model. Last, accounting for the cost of living in the destination region completely cancels out the negative and significant impact of remoteness on migration; hence, higher prices and reduced product availability can entirely explain the lower attractiveness of remote regions for migrants. Having an adequate measure of local cost of living that accounts for both prices and product availability is thus crucial to understand internal migration patterns in a developing country such as Ethiopia.

Related literature. We contribute to four strands of the literature. First, we contribute to the literature on spatial differences in terms of cost of living in developing countries (see, e.g., Ravallion and Van De Walle, 1991; Deaton and Tarozzi, 2000; Muller, 2002; Timmins, 2006; Ferré et al., 2012). We see three important contributions of our work with respect to existing studies: (1) Our measure of cost of living accounts for both the availability of products and services and for their price when they are available, whereas existing papers usually focus on the latter; (2) By leveraging the micro-data underlying the consumer price index in Ethiopia, our analysis has a greater spatial and industrial coverage than existing studies that usually focus on a narrower set of products and villages/regions; (3) Instead of emphasizing the urban/rural divide, we analyze cost-of-living differences across cities and relate them to economic geography, namely, population size and geographic remoteness of cities.[5]

Second, a couple of recent papers examine prices and product availability in the context of Ethiopia (Gunning et al., 2018; Krishnan and Zhang, 2018). Gunning et al. (2018) demonstrate that households in remote villages of Ethiopia have access to a lower variety of goods. Our paper differs from these other papers along three important dimensions. First, the focus is different. These papers examine individual prices and product availability in remote

[5]Ethiopia is an ideal laboratory to understand the impact of intra-national isolation on price variations across space in a developing economy context. With 100 million inhabitants, it is the second most populated country in Africa and the largest landlocked country in the world. The population is spread unevenly in the country. While 3% to 5% of the population live in the metropolitan area of Addis-Ababa, about 20 cities spread throughout the country have population with 5-digit figures. Like other countries in the region, Ethiopia has experienced strong GDP growth since the beginning of the 2000s and a sharp improvement in several development indicators.

villages. Our analysis instead covers Ethiopian main cities. Second, our data allow us to compute local price indexes and thus to directly compare the welfare across cities - and relate welfare differences to city size and remoteness.[6] Third, we are the first to study the implications of prices and product availability differentials for migration decisions.

Third, our paper contributes to the literature on the measurement and determinants of price indexes in the presence of missing varieties. Seminal papers in this literature have developed methods to measure the costs and benefits of disappearing and appearing varieties over time (Feenstra, 1994; Hausman, 1996; Diewert and Feenstra, 2019). A few papers apply this method to a spatial context, and explore the link between the cost of living and city size (see, e.g., Handbury and Weinstein, 2015; Feenstra et al., 2019, using U.S. and Chinese data, respectively).[7] We complement this literature by showing that, in a country such as Ethiopia, remoteness has a stronger impact on spatial variations in terms of cost of living than city size. We also propose an alternative spatial price index that relaxes the standard assumption that products unavailable in a city cannot be purchased in surrounding cities by traveling.

Last, our paper contributes to the literature on the determinants of internal migration flows. Different factors have been shown to drive internal migration, including access to domestic and/or foreign markets (Crozet, 2004; Hering and Paillacar, 2016), access to public and private amenities (Bryan and Morten, 2019; Lagakos et al., 2020; Fafchamps and Shilpi, 2013), earnings differences (Young, 2013), conflicts (Kondylis and Mueller, 2014), or differences in housing prices (Monras, 2018; Hsieh and Moretti, 2019). We show that cost-of-living differences (particularly those driven by access to a wider set of products and services) are an important driver of migration flows across Ethiopian regions.[8]

The rest of the paper is organized as follows. Section 2 describes the data used in the empirical analysis. Section 3 presents the results on the impact of remoteness and population density on individual prices and product availability. Section 4 introduces our cost-of-living indexes, and presents our analysis of the determinants of the spatial dispersion in the cost–of-living across Ethiopian cities. Section 5 present the results on the impact of cost-of-living

[6] The literature also examines the impact of remoteness on outcomes other than prices or product availability. For instance, Dercon and Hoddinott (2005) show that better access to market towns allows rural households to buy their inputs at a lower price and to sell their outputs at a higher price in Ethiopia; in the same vein, Aggarwal et al. (2018) find poor market access implies a poor harvest output in rural Tanzania.

[7] See also Matsa (2011) on the link between competition and inventory shortfalls.

[8] Jedwab and Storeygard (2017) show that better market access has spurred population growth in sub-Saharan cities. The effect is more pronounced for small and remote cities. Our results suggest that these differences in population dynamics might partly channel through increased availability and lower prices of products and services in these places.

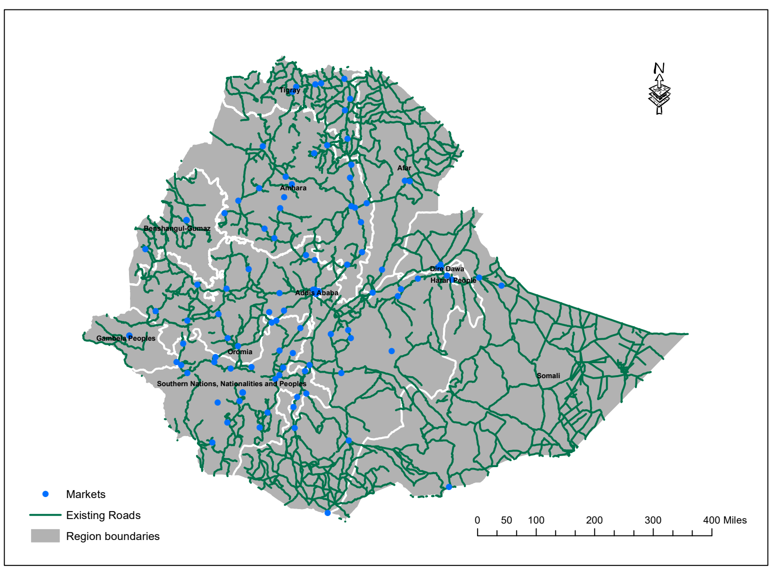
differentials on migration across Ethiopian regions. Section 6 concludes.

2 Data

This section presents the different datasets used in the paper, and discusses their strengths and limitations. Crucial to our analysis are the consumer price data underlying the Ethiopian CPI. These data are described in the next section. We then describe the economic, demographic, and geographic variables used throughout our empirical analysis. The calculation of price indexes requires information on how consumers allocate their expenditure across the various products, and on the elasticity of substitution between these products. We describe these variables at the end of the section.

2.1 Consumer price index data

Figure 1: ***Markets covered by the Ethiopian CPI***



Source. Individual price data are collected by the Ethiopian Central Statistical Agency (CSA) to build the national Consumer Price Index. We work with the data for the year 2015. The price quotes of more than 420 products are collected every month by enumerators in 117 markets. The markets are urban centers (cities or towns) in each woreda (a woreda being the smallest administrative division with local government in Ethiopia). Addis-Ababa, the

Ethiopian capital city, is split into 12 markets. The different markets are mapped in Figure 1. In each market, enumerators survey a pre-determined sample of outlets (1 to 3 per location) every day during the first two weeks of each month. Outlets comprise a representative sample of open markets, kiosks, groceries, butcheries, pharmacies, super markets, and so on. Enumerators are asked to find precisely defined products and to report the product as missing if they cannot find it. When they find the product, they determine its typical price (after bargaining) by interviewing both sellers and consumers (Atkin and Donaldson, 2015).

Extraction. Every month, the CSA releases a 200-page document reporting the price of every product in each of the 117 cities. We extract this information from pdf files for all the months of year 2015. We then manually check the obtained dataset. We change the label of some products or product categories that appear to be obviously wrong. The resulting dataset is available on our personal websites.

Description. The survey covers 427 products and services that can be grouped into 12 major groups and 55 categories.[9] These products and services include food products such as bread and cereals, but also, among others, clothing and footwear products, household equipment products, or hair-cuts and restaurants. The product descriptions range from barcode-like data with brandname products ("Coca-Cola bottle 300c"), to very specific products without a brand (e.g., "bed sheet (Patterned Kombolcha) 1.90m x 2.50m"), to more generic product categories (e.g., "sorghum yellow, kg" or "rice imported, kg").

Outliers and yearly price. Prices are collected every month. In 2015, we thus have a maximum of 12 price quotes for a given product within a given location. We have detected a few outliers in the data. A visual check suggests two main types of outliers exist. First, the price might be abnormally high because, for instance, the marker for decimals was forgotten. Second, enumerators sometimes reported the monthly price of the product at the wrong line. For instance, in Akaki, the product "VCD-Player (Mayato Japan)" is missing every month of 2015 except the month of July. That month, the reported price is 5 Birrs. However, in the file, the product below "VCD-Player (Mayato Japan)" is "VCD Cassette rent". In Akaki, this product is available every month at the price of 5 Birrs, except for the month of July, where it is missing. Moreover, the median price of the "VCD-player (Mayato Japan)"' in markets where it is available is around 1,000 Birrs. Thus, an obvious mistake exists where the price of the "VCD Cassette rent" in Akaki was reported at the wrong line for the month of July.

To systematically detect and delete outliers, we drop monthly observations for which the price is 5 times higher or 5 times lower than the median price of the product across

[9]The list of products is presented in Table A.2 in the Appendix.

markets. We thereby exclude 469 observations out of more than 300,000. Note that with this procedure, we do not automatically create missing varieties, because, as explained below, our definition of "missing" is computed at the annual level.

Because we are interested in price differences across locations (and not over time), and to alleviate potential remaining issues related to misreporting, we take the median price per product and location computed over the months the product is not missing. In the original dataset, Addis-Ababa is split into 12 markets. We merge these markets into a single one by considering the median price across the 12 markets. This approach leaves us with information on 427 products across 106 cities.

Missing products. A missing price in our original dataset means the enumerators were unable to find the product in that city in that month. We follow Atkin and Donaldson (2015) and assume the product was not available if the price is missing. Our definition, however, is stricter than theirs because we require the product to be missing every month of 2015 to consider it unavailable. In our final dataset, a product is considered missing for about 35% of the product-location pairs.

Prices and vertical differentiation. Given our interest in building a local price index that best reflects the cost of living for consumers, we consider that using all the products and services in the dataset (including those with a less precise description) is an upside of our study. Not restricting the sample to barcode products matters because many products in the consumption basket of consumers have no barcode, especially in a developing country such as Ethiopia. This choice comes at the cost that for some products, price differences across locations might partly reflect differences in terms of quality. Actually, we identify in our dataset 31 products (out of 427) whose description clearly mentions a brand and a unit that make them as close as possible to barcode products.[10] As Figure 2 shows, across cities, the coefficient of variation of barcode products tends to be, as expected, lower on average than the one observed for the other products within the same product category. Still, Figure 2 shows the prices of barcode products also exhibit substantial spatial variation, suggesting the latter cannot be reduced to spatial variation in the quality of the available varieties.

Moreover, in most of our specifications, we use use a proxy of local prices purged from spatial differences in terms of quality; more precisely, we regress the price quote of product p in city c p_{pc} on the local per-capita income, and when the coefficient we obtain is statistically significant at the 5% level, we take as a proxy for the price of product p in city c the residual of this regression. The idea is that in the presence of non-homothetic preferences, wealthier consumers are more likely to buy higher-quality varieties; consequently,

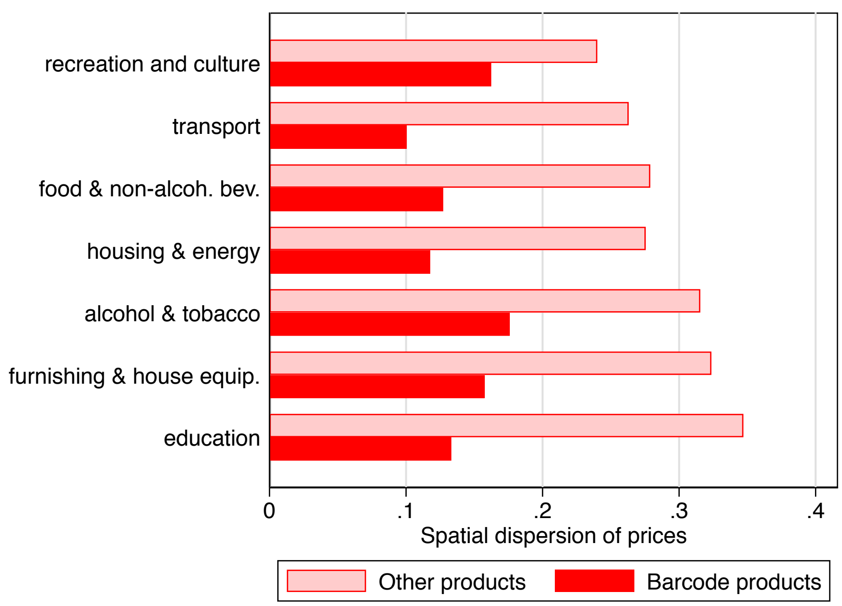
[10] See the list in Table A.3 in the Appendix.

price quotes and per-capita income should be positively correlated at the local level. Note that in doing so, we also control for the possibility of wealthier consumers having a higher willingness to pay than poorer ones. The regression of price quotes on local per capita income is run separately for the 55 product categories and includes product-type fixed effects. The value of the coefficient exhibits a fair amount of variation across product categories (average equal to 0.07, standard-deviation equal to 0.14). The correlation is positive (resp. significantly positive) in 39 (resp. 21) of these categories, and negative (resp. significantly negative) in only 16 (resp. 6) product categories. Hence, the correlation is statistically significant for less than half of the product categories only.

Finally, we are interested in the spatial differences in terms of cost of living and their consequences on the internal migration patterns, which could well be driven by the type of varieties consumers have access to; one could thus argue the composition of the set of available varieties in terms of quality is part of the heterogeneity we want to capture. For this reason, we also present the results with non-purged price quotes.

Overall, our results are similar regardless of the proxy we use for local prices, and thus we believe vertical differentiation is not a first-order issue for our study.

Figure 2: ***Coefficient of variation of prices across space***



Product classification. We test for possible heterogeneity in the macro-determinants of local prices across specific broad categories of products.

More specifically, we manually identify services from the product description. Out of

427 products, 44 are services (10%), including, for instance internet services, haircuts or airplane tickets to Addis-Ababa.

We also use the products' description to identify imported products. For some items, the description indicates the product is imported (e.g., "socks (cotton) imported"). For others, the country of origin is indicated (e.g., "under wear China"). We tag all the products whose description contains a word related to import as an imported product.[11] In our sample, 8% of the products are imported based on this definition (keeping in mind that it might not provide an exhaustive list of imported varieties). Finally, we define as "food products" those products in the major group "Food and non-alcoholic beverages".

Reliability of the price data. Because inflation is a major political concern in Ethiopia, one may worry the price-quote data are manipulated for political reasons. Note that 2015 was not a year of hyper inflation limiting the political motivations for price manipulation. To further assess the reliability of the price quotes, we test whether the Benford law holds in our data. The Benford law states that the frequency distribution of the first digit of numerical data is stable across samples (the frequency of number 1 is 30%, number 2 is 17%, and number 9 appears as the first digit in only 5% of the occurrences). Researchers have used deviations from this law to detect reporting issues in survey data (e.g., Judge and Schechter, 2009; Demir and Javorcik, 2018). The underlying idea of this test is that manipulating the data and still making them fit the Bendford distribution is difficult.

Figure 3 presents the frequency distribution of the first digit of price quotes in our data together with the Benford distribution. The observed frequencies fit with frequencies given by the Benford law.[12] Hence, the distribution of the first digit of price quotes in our data is consistent with the Benford law, suggesting price manipulation is not a major issue here.

2.2 Demographic and economic data

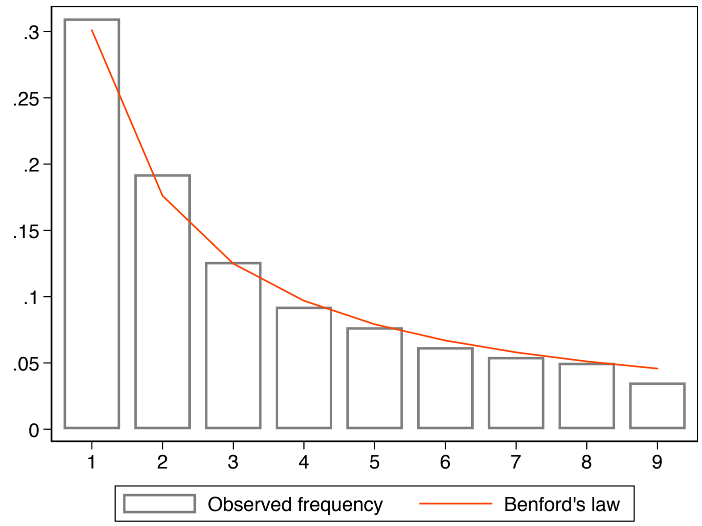
We complement our price data with various variables extracted from different waves of the Ethiopian Census. For some locations, we did not find the required information in the Census, and we had to look for other data sources. We present the details below.

Population. The 2015 population of the urban centers considered in the analysis comes from the population projection figures for 2015 made by the CSA based on the 2007 population and housing census of Ethiopia (see CSA, 2013). For 25% of the urban centers the

[11] More specifically, a product is imported if its description includes at least one word among the following list: imp., import, imported, China, Japan, England, India.

[12] We also did the Kuiper test to assess the deviation of our data from the Benford law. The mean of the Kuiper statistic is 0.0264 which is over the 1% threshold to reject the null hypothesis that the observed distribution deviates from the Benford law.

Figure 3: ***Benford law for price quotes***



Each bar in the diagram represents the observed frequency of a number of the first digit of price quotes. The red curve represents the frequency of a price digit given by the Benford's law.

CSA does not provide any projection. For these locations, we rely on projection figures for 2015 made by the Ethiopian Ministry of Water and Energy in 2011 as part of the urban water-supply universal access plan (see Ministry of Water and Energy, 2011). The size of the cities in our dataset ranges from 764 inhabitants (Deri) to 3,273,000 (Addis-Ababa). In our sample, 63% of cities have less than 30,000 inhabitants. These cities are defined as "rural".

Income. To measure per-capita income, we use data from the 2015/2016 Ethiopian Household Consumption Expenditure survey (HCE) conducted by the CSA (see CSA, 2018a). Per-capita income is measured by region (Ethiopia has 11 regions in total), with two separate measures for urban and rural areas in each region. It is thus available at a more aggregate geographic level than prices. It measures people's total expenditures in Birr and it also includes consumption in kind. It is available by quintiles in each region and each type of area. The average income in the first quintile is 3,900 Birr, whereas it reaches 24,200 Birr for households in the fifth quintile. We use average income per capita as a measure of income in our analysis.

Ethnic diversity. Data on ethnic diversity comes from the 2007 Population and Housing Census of Ethiopia conducted by the CSA (see CSA, 2008). We compute the share of each ethnic group (93 in the country) in the total population of every woreda. Similar to Schiff (2015), our measure of ethnic diversity is the inverse of a Herfindahl index computed with these shares. At the country-level, Oromos are the most represented ethnic group, with 34.6% of the total population.

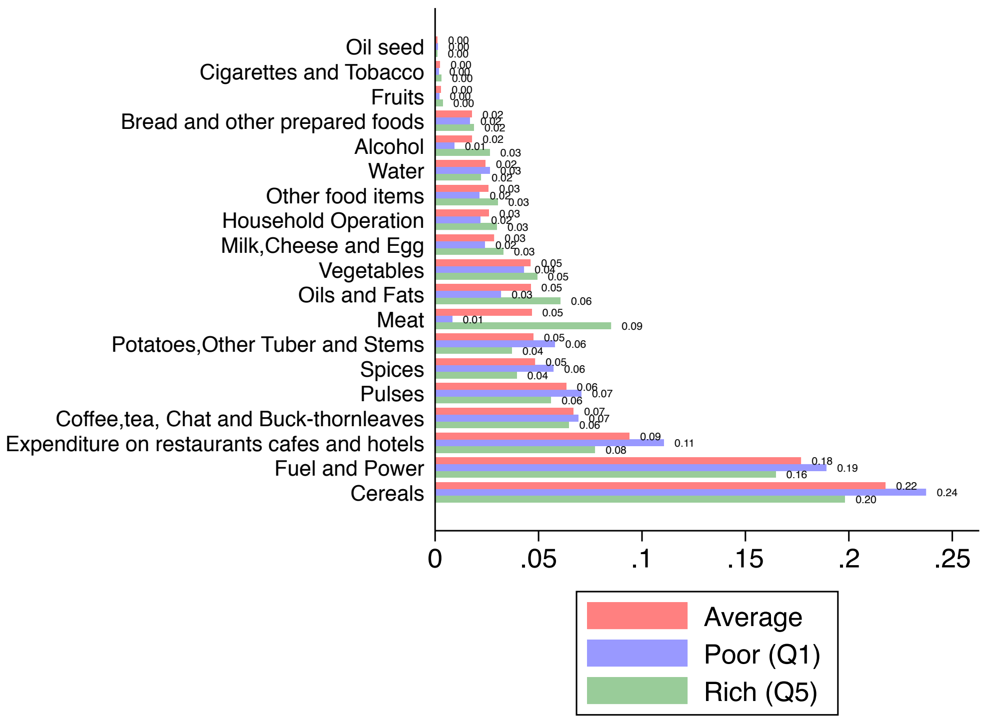
Expenditure weights. Aggregate price indexes require information on expenditure weights across product categories. We use expenditure weights reported in the 2011 Ethiopian Household Consumption Expenditure survey (HCE) conducted by the CSA (see CSA, 2012). We have expenditure weights for 19 semi-aggregated categories. These categories are related to food and daily expenditures. The products for which we do not have weights are thus those that are purchased less frequently. In particular, we do not take into account housing. As compared to the standards in the literature, we do not think this issue is a major one here. Indeed, several papers on the US use housing prices alone as a measure of the local cost of living (Moretti, 2013; Monras, 2018), with housing expenses representing 33% of US households' expenses. Here, the products and services for which we have reliable information account for 62.5% of overall expenses of Ethiopian consumers. Moreover, housing only represents 9.9% of total household expenditures in Ethiopia (see CSA, 2012). We are thus confident the set of products and services for which information on expenditure weights is available allows for a proper measure of local cost of living. Information is available separately for rich and poor households. As depicted in Figure 4, for both poor and rich, Cereals represent the highest expenditure share, and Oil seed represents the lowest.

Elasticities of substitution. To compute the CES price index described in section 4, we will need elasticities of substitution between varieties within semi-aggregated categories. For most of the 19 semi-aggregated categories, we use import demand elasticities estimated by Broda et al. (2017). More precisely, we assign HS codes for which estimated elasticities are available to one of the 19 semi-aggregated categories for which we have expenditure shares, and take the average elasticity. Ethiopia is not available in this dataset. We thus use the elasticities estimated for Togo and Madagascar instead because these countries are the closest to Ethiopia in terms of GDP per capita and linguistic proximity for which information is available. We try two options: consider the elasticities of Togo, or take the average elasticity between Togo and Madagascar. This has little impact on the results. Cigarettes and Tobacco have the highest substitution elasticity (84.83), and Alcohol has the lowest (1.59). These elasticities are for tradable goods. For restaurants, we use the elasticity of 8.8 estimated by Couture (2014) for the US.[13]

Auto-consumption. In some robustness checks, we account for auto-consumption, which could be correlated with both prices and product availability on the one hand, and city size and remoteness, which are our variables of interest, on the other. We compute

[13]Unfortunately, no estimate is available for Ethiopia or for a close developing country. However, this 8.8 for restaurants is close to the average elasticity of substitution measured across the product categories taken into account in the price index, which is equal to 8.84.

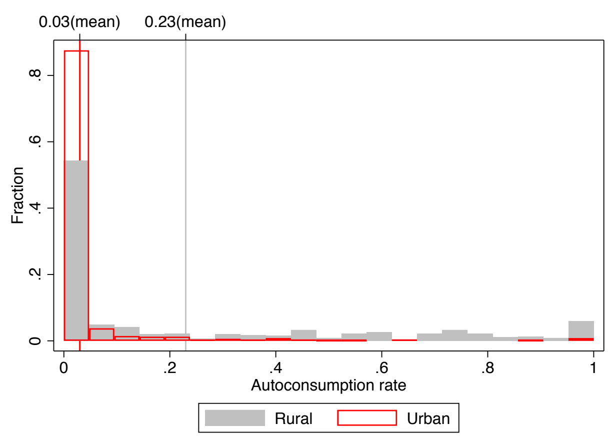
Figure 4: ***Weights in terms of expenditures***



Note: Each bar represents the share of an item in the total expenditures dedicated by a household for the selected products. The blue bars represent the weights for poor consumers, the green bars represent the weights for rich consumers, the red bars represent the average weights.

data on auto-consumption using the third wave (2016) of the Ethiopian socioeconomic survey conducted by the CSA and the World Bank Living Standards Measurement Study-Integrated Surveys on Agriculture (LSMS-ISA) team (see CSA, 2018b). Obviously, auto-consumption is a concern for food products mainly, and information on auto-consumption is thus available only for these products. We define auto-consumption as the share of household consumption (in terms of quantity) that comes from its own production. These data are available for each product-region-area triplet, where area refers to urban versus rural areas. "Dairy products" (e.g., cow milk) and "Cereals unmilled" (e.g., sorghum) are the product categories with the highest share of auto-consumption in the country (average of 38% and 24%, respectively). Not surprisingly, the auto-consumption rate is higher in rural areas with an average of 23% compared to only 3% in urban areas. Note, however, that these averages hide large discrepancies at the household level, as shown in Figure 5. Auto-consumption is below 5% for almost all of the households living in urban districts. For households in rural districts, the distribution of auto-consumption shares is bi-model with a (large) peak around 0 and a (small) peak close to 1.

Figure 5: ***Auto-consumption***



Note: Auto-consumption rates for food products computed from the 2015 household survey (LSMS data). The plain gray bars represent the distribution of auto-consumption shares for households living in rural districts. The red empty bars are for households living in urban districts. The red and grey vertical lines indicate the average auto-consumption rate for urban and rural households, respectively.

Migration data. We use migration data to investigate the effects of spatial differences in terms of cost of living on internal migration flows. We use migration data from the 2013 national labour force survey (see CSA, 2014).

Employment opportunities and surface area of regions. To investigate the determinants of migration flows, we need a measure of regional employment opportunities at the destination. For this purpose, we use 1 minus the regional unemployment rate as a proxy for employment opportunities. We use data on unemployment rates from the 2005 national labor force survey (see CSA, 2006).[14]

The regions in Ethiopia have different sizes. Beyond total population size, surface area (and thus population density) could matter for migration decisions. We use data on the surface area of regions from the 2007 population and housing census of Ethiopia.

2.3 Geographic variables

A key part of our analysis is about the relationship between distance to other cities (or remoteness) and the local cost of living. We build our distance measures as follows.

Distance and travel time between cities. We first need to assess how far each city in our dataset is from the other cities. Distance is computed thanks to the Stata package Geodist. Travel time is computed using the Stata package Georoute (Weber and Péclat, 2016). The travel times to other cities provided by this package are missing for 10 cities. For these cities, we predict travel time based on bilateral distance.[15] We use a similar strategy to compute the distance and travel time to the capital city (Addis-Ababa) and to the main international trade corridor (Kombolcha, through which shipments from and to Djibouti transit).[16]

Remoteness. For each city, we aggregate the information on bilateral distances with the 105 other cities by building a remoteness index, which is the average travel time to the other Ethiopian cities:

$$\text{remote}_c = \frac{1}{105} \sum_{j \neq c} \text{travel time}_{cj},$$

where remote_c is the remoteness index for city c, and travel time_{cj} is the travel time between cities c and j. In robustness checks, we use a population-weighted version of this remoteness index.

Travel costs. Part of our analysis relies on the travel cost paid by a consumer to buy a product that is not available in her own town. To compute travel costs between towns c and

[14]In the migration literature, previous employment rates are often used as determinant of current migration flows (see, e.g., Crozet, 2004). Consequently, we use 2005 unemployment rates to compute regional employment rates.

[15]Travel time is regressed on a polynomial of degree 7 of distance as well as fixed effects for origin-destination pairs of regions. For the 8,730 pairs of cities for which we have both forms of information, the $R-squared$ is close to 90%. Travel time is then predicted for all the city-pairs involving the 10 destinations for which information on travel time is missing.

[16]But in this case, a linear fit with region fixed effects is enough.

j, we use the following formula:

$$\text{travel\_cost}_{cj} = \text{distance}_{cj} \times \text{cost\_per\_km}.$$

The previous paragraph explains the computation of distance between any pair of cities. More difficult to obtain is a reliable measure of the travel cost per kilometer. Estimates of transport costs in Ethiopia are available, but they are based on the cost of transporting products by truck from the location of production or port of entry to a given destination. We are instead interested in the cheapest transport mode for a consumer who wants to travel to a location where the variety is available. Interviews with people from the field suggest the cheapest mode of transport across Ethiopian cities is van taxis and intercity buses. We use the intercity bus fare of 0.1235 birr/km reported in the management of the commercial road transport survey (see Addis Ababa Chamber of Commerce and Sectoral Associations, 2009).[17]

From this figure, we are able to compute the price of consuming in city c a product j from city c' as follows:

$$p_{jcc'} = p_{jc'} + 2\text{travel\_cost}_{cj}.$$

3 Spatial dispersion of individual prices and product availability

In this section, we assess the "macro" determinants of local individual prices and product availability. More specifically, we investigate whether city-size and geographic remoteness affect the price of a product when it is available, and the probability that this product is available in that city. The regression we estimate is thus the following:

$$\text{y}_{pc} = \alpha \text{pop}_c + \beta \text{remote}_c + \omega_p + \epsilon_{pc}, \tag{1}$$

where y_{pc} is either the log median price quote of product p in city c in 2015 or a dummy equal to 1 when p is missing every month of 2015 in city c. City-size is measured by the log population in c. Our measure of remoteness is the log (unweighted) average time to reach all the other Ethiopian cities in the database from c. Finally, ω_p is a product fixed effect, and ϵ_{pc} is the error term.

Before getting to the baseline results, we provide in Table 1 a variance decomposition of local prices and of the dummy identifying product unavailability. The product fixed effects explain most of the dispersion in prices, which simply reflects that different products

[17]Note this cost is not product-specific, because it measures the cost for a consumer to travel to a location where the product is available.

have different average prices. Once we take into account the product fixed effects, the share of the remaining variation explained by city fixed effects is equal to 6.37%. The last row of Table 1 shows once product-region fixed effects are accounted for the R-squared rises to 32.86%. The picture is qualitatively similar (though quantitatively different) for the dummy identifying missing products. Thus, the correlation between local prices/product availability and city size and remoteness might not be the same across products. We discuss this heterogeneity at length in our analysis of the macro-determinants of local prices. This heterogeneity will also be important when aggregating the micro-data to examine the effect of city-size and remoteness on local price indexes.

Table 1: ***Variance decomposition***

	Log p_{pc}	$\mathbb{1}_{\text{missing}_{pc}}$
Product fixed effect	97.4%	35.5%
City fixed effect	6.37%	12.07%
Product-Region fixed effect	32.86%	21.85%

Notes: p_{pc} is the price quote of product p in city c and $\mathbb{1}_{Miss_{pc}}$ is a dummy equal to 1 if product p is missing in city c. The table reports the R^2 of regressions including the fixed effects mentioned in the first column. For city and product-region fixed effects, the dependent variable is the price or the dummy $\mathbb{1}_{Miss_{pc}}$ net of the product fixed effects.

3.1 Baseline results

Table 2 presents our baseline results. As can be seen from regression (1), remoteness and city-size are significant determinants of the local prices of available products. Large and more remote cities are more expensive than the others. When we control for the product category-level correlation between prices and income per capita (which should capture quality differences across cities), the coefficient on our remoteness proxy is even higher (even though not statistically different from the one in the first column), whereas the coefficient on city-level population is reduced but remains positive and highly significant. The pattern of the coefficient on population size is consistent with large cities also being the wealthiest ones. Several explanations could account for this positive association between price and city-size. The distribution costs could be higher in larger cities due to fancier distribution chains or higher land rents for example; products sold in larger cities could also be less likely to be produced locally, especially food products.

Column (3) displays the results on product availability. The probability that a product is missing is significantly higher in remote cities and significantly lower in large cities, with elasticities equal to 10.5% and -7.2%, respectively.

Table 2: ***Macro determinants of micro prices and availability***

	Log p_{pc}		$\mathbb{1}_{\text{missing}_{pc}}$
	(1)	(2)	(3)
Ln Remoteness	0.119^a	0.134^a	0.105^b
	(0.030)	(0.031)	(0.052)
Ln Population	0.040^a	0.028^a	-0.072^a
	(0.004)	(0.005)	(0.009)
Per capita income	No	Yes	No
Observations	28,183	28,183	42,506

Notes: The dependent variables are the log price (column 1), the log price net of its product category-level correlation with the local per-capita income (column 2), and a dummy equal to 1 if the product is unavailable in the city (column 3). These dependent variables are defined in the $product \times city$ dimension. Standard errors clustered at the city level in parentheses, ^{a}, ^{b} and ^{c} respectively denote significance at the 1, 5, and 10% levels.

3.2 Robustness checks

In Table 3, we check the robustness of our baseline empirical regularities. More specifically, we propose four types of checks. First, the Ethiopian population is composed of various ethnic groups, but many cities are characterized by the presence of a (very) majoritarian one. If ethnic groups have very specific tastes, more ethnically diverse cities could be cities where varieties are more likely to be available (e.g., Schiff, 2015), and this diversity could also affect the price at which they are sold. We thus control for the inverse of a Herfindahl index based on the share of the various ethnic groups in the population. Second, remoteness might not only be related to distance to all other cities, but more specifically to distance to the capital city, Addis-Ababa, or to the most important commercial routes, namely, Kombolcha, which is on the corridor to Djibouti. We thus control for distance to these two cities. Third, we use a weighted average distance to all other Ethiopian cities (using population as weights) as an alternative measure of remoteness. Finally, we drop Addis-Ababa from the regression sample.

In a nutshell, controlling for ethnic diversity does not change anything. The distances to Addis-Ababa and to the import corridor are not significant for the price of the available varieties or for the probability that a product is missing, and their introduction does not affect the coefficient on population size and remoteness. The weighted average distance to other cities is positively related to prices but not significantly related to product availability. What matters for product availability is thus how far a city is from the other cities, not from the large ones only. Finally, dropping Addis-Ababa from the sample does not significantly change the results.

Table 3: ***Macro determinants of micro prices and unavailability: Robustness checks***

	Log p_{pc}					$\mathbb{1}_{\text{missing}_{pc}}$				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ln Remoteness	0.134^a	0.136^a			0.141^a	0.105^b	0.103^c			0.116^b
	(0.031)	(0.032)			(0.031)	(0.052)	(0.053)			(0.052)
Ln Population	0.028^a	0.028^a	0.027^a	0.027^a	0.026^a	-0.072^a	-0.072^a	-0.068^a	-0.072^a	-0.075^a
	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)	(0.009)	(0.009)	(0.010)	(0.009)	(0.010)
Ln Ethnic diversity index		-0.008					0.007			
		(0.014)					(0.026)			
Ln Travel time to import corridor			-0.012					0.017		
			(0.007)					(0.010)		
Ln Travel time to Addis-Ababa			0.008					0.005		
			(0.008)					(0.010)		
Ln Remoteness (weighted)				0.054^b					0.058	
				(0.027)					(0.038)	
Per-capita income			Yes					No		
Addis-Ababa	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No
Observations	28,183	28,183	27,697	28,183	27,810	42,506	42,506	41,704	42,506	42,105

Notes: The dependent variables are the log price (columns 1 to 5) net of its product category-level correlation with the local per-capita income and a dummy equal to 1 if the product is unavailable (columns 6 to 10). These dependent variables are defined in the $product \times city$ dimension. Standard errors clustered at the city-level in parentheses, ^{a}, ^{b} and ^{c} respectively, denote significance at the 1, 5, and 10% levels.

In the context of a developing country such as Ethiopia, another important issue that could undermine our benchmark results is auto-consumption. For food products in particular, households sometimes produce their own consumption. The phenomenon is particularly important in rural areas (see section 2.2). If products are produced directly by those who consume them, they might be unavailable on the market without necessarily being unavailable for consumption. For a subset of food products in our database, we have information on the share of auto-consumption in total households' consumption by region and type of area (rural or urban). After reproducing the benchmark results for the subset of observations for which auto-consumption is available, we thus directly introduce auto-consumption in the regression. Table 4 shows auto-consumption is not statistically significantly related to prices or product availability, and controlling for it does not affect the results.

All in all, our benchmark results are thus very robust: large cities are places where products and services are more expensive but also more likely to be available, whereas

Table 4: ***The role of auto-consumption***

	Log p_{pc}			$\mathbb{1}_{\text{missing}_{pc}}$		
	(1)	(2)	(3)	(4)	(5)	(6)
Ln Remoteness	0.134^a	0.104^b	0.102^b	0.105^b	0.175^a	0.176^a
	(0.031)	(0.041)	(0.041)	(0.052)	(0.039)	(0.039)
Ln Population	0.028^a	0.032^a	0.028^a	-0.072^a	-0.047^a	-0.044^a
	(0.005)	(0.007)	(0.008)	(0.009)	(0.006)	(0.007)
Auto-consumption rate			-0.046			0.023
			(0.029)			(0.022)
Per capita income		yes			no	
Observations	28,183	8,672	8,672	42,506	12,370	12,370

Notes: The dependent variables are the log price (columns 1 to 3) net of its product category-level correlation with the local per-capita income, and a dummy equal to 1 if the product is unavailable in the city (columns 4 to 6). These dependent variables are defined in the $product \times city$ dimension. Standard errors clustered at the city-level in parentheses, ^{a}, ^{b} and ^{c} respectively denote significance at the 1, 5, and 10% levels.

remote cities are places where products and services are more expensive and less likely to be available, in line with the idea that these cities are more difficult to cater to.

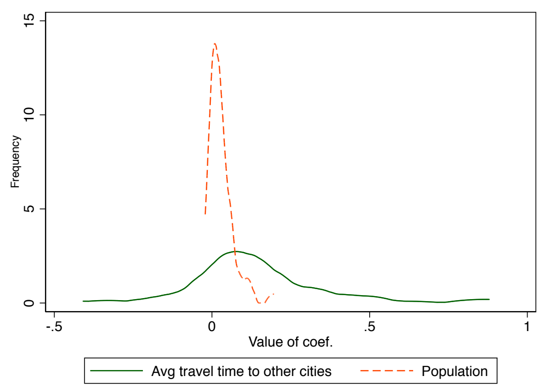
3.3 Heterogeneity

To go beyond the average correlations of our baseline results, we run the regressions of columns (2) and (3) of Table 2 separately for the 55 product categories present in our sample. We summarize the results in Figure 6. Panel (a) shows that for almost all products, prices are higher in larger cities. The coefficient on population is equal to 2.8% on average, with a standard deviation of 3.9%. It is positive for 44 out of 55 products and positive and significant at the 10% level at least for 24 of them. The coefficient on remoteness is equal to 14.5% on average (standard deviation of 22.4%.). It is positive for 47 out of 55 products, and positive and significant for 21 of them.

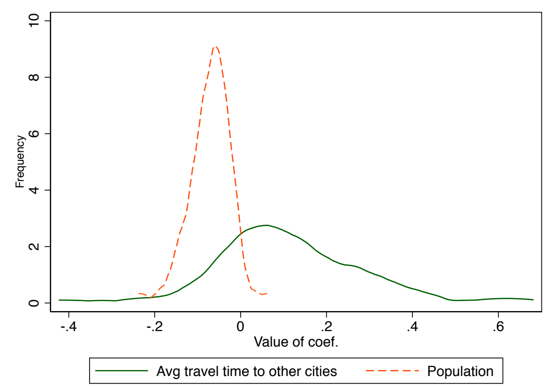
Regarding product availability, panel (b) of Figure 6 shows that for almost all products, the probability that it is missing is lower in larger cities. The average coefficient on population size is equal to -7.1% with a standard deviation of 4.9%. This coefficient is negative for 53 out of 55 product categories, and negative and significant for 46 of them. The coefficient on remoteness is equal to 11.3% on average with a standard deviation of 17.9%. It is positive for 43 out of 55 product categories and positive and significant for 16 of them.

All in all, these figures largely confirm the findings of Table 2: large cities have a greater access to products and services but are also more expensive, whereas remote cities suffer from both lower availability and higher prices. However, they also clearly demonstrate that a significant amount of heterogeneity exists in the relationship between product prices/product availability and the economic geography of cities.

Figure 6: ***Product price, product availability, remoteness, and city size***



(a) Product price



(b) Probability to be missing

Note: Both graphs present the density distributions of the coefficients obtained when running the regressions of columns (2) and (4) of Table 2 separately for the 55 product categories.

To try to dig deeper into this heterogenity, we focus on four specific broad categories: foreign products, services, food products, and barcode products (see section 2.1 for the definitions). We identify the observations corresponding to these types of products thanks to adequate dummies that we interact with our variables of interest.

As Table 5 shows, we detect no significant heterogeneity for imported products. The prices of services increases more in large cities as compared to other products, whereas the probability that they are missing increases less with remoteness. This type of heterogeneity is consistent with services being less tradable than goods. The presence of food products is less intensely related to the city's population size; this finding is well in line with the fact that many food products are produced in rural areas where cities are smaller (see our discussion on auto-consumption). Note the negative effect of remoteness on product availability is stronger for food products, perhaps due to lower transportability of food products, especially fresh ones, in the context of a developing country such as Ethiopia. Finally, the price of barcode products and their availability are on average less related to cities' population size compared to the other products. This finding could be explained by the fact that barcode products are sold by larger firms with country-wide presence. However, remoteness does increase the price of these products.

Table 5: ***Heterogeneity***

	Log p_{pc}					$\mathbb{1}_{\text{missing}_{pc}}$				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ln Remoteness	0.138^a	0.131^a	0.114^a	0.137^a	0.112^a	0.111^b	0.116^b	0.076	0.107^b	0.098
	(0.032)	(0.030)	(0.029)	(0.033)	(0.031)	(0.052)	(0.053)	(0.061)	(0.053)	(0.067)
Ln Population	0.029^a	0.025^a	0.024^a	0.030^a	0.022^a	-0.071^a	-0.069^a	-0.085^a	-0.073^a	-0.084^a
	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)	(0.009)	(0.009)	(0.011)	(0.009)	(0.011)
Ln Remoteness × Imported	-0.052				-0.034	-0.069				-0.065
	(0.046)				(0.041)	(0.042)				(0.040)
Ln Population × Imported	-0.010				-0.005	-0.011				-0.004
	(0.007)				(0.006)	(0.007)				(0.006)
Ln Remoteness × Services		0.033			0.052		-0.102^a			-0.084^c
		(0.041)			(0.046)		(0.036)			(0.043)
Ln Population × Services		0.025^a			0.028^a		-0.029^a			-0.015^c
		(0.006)			(0.007)		(0.007)			(0.008)
Ln Remoteness × Food			0.049		0.054			0.079^b		0.059
			(0.033)		(0.035)			(0.038)		(0.042)
Ln Population × Food			0.010		0.012^c			0.035^a		0.033^a
			(0.006)		(0.006)			(0.006)		(0.006)
Ln Remoteness × Barcode				-0.043	-0.026				-0.020	-0.016
				(0.040)	(0.038)				(0.039)	(0.040)
Ln Population × Barcode				-0.033^a	-0.027^a				0.013^b	0.019^a
				(0.007)	(0.006)				(0.006)	(0.006)
Observations	28,183	28,183	28,183	28,183	28,183	42,506	42,506	42,506	42,506	42,506

Notes: The dependent variables are the log price (columns 1 to 5) net of its product category-level correlation with the local per capita income and a dummy equal to 1 if the product is unavailable in the city (columns 6 to 10). These dependent variables are defined in the $product \times city$ dimension. Standard errors clustered at the city-level in parentheses. ^{a}, ^{b}, and ^{c}, respectively, denote significance at the 1, 5, and 10% levels.

4 Cost of living across Ethiopian cities

We have shown that, on average, prices are higher in large cities and in remote cities, whereas products are more likely to be available in densely populated and/or less remote locations. However, using this micro-analysis to draw any conclusion on the relationship between city size, geography, and the cost of living is hard. Indeed, we have shown that behind these average relationships, a significant amount of heterogeneity exists across products; the overall penalty that consumers in larger and more remote locations pay will then depend on the distribution of their expenditure shares across products. Moreover, no simple way exists to account for both the price of available products and product availability when measuring the cost of living. To make some progress along these two dimensions, one needs to put some structure on the preferences of consumers to derive a measure of the local cost of living that is consistent with these preferences.

We provide the structure in this section by deriving two alternative measures of the local cost of living. We then revisit our empirical micro-evidence using these local price indexes as dependent variables.

4.1 Measurement

We propose two local price indexes.

4.1.1 Feenstra-CES price index.

For the first one, we follow Handbury and Weinstein (2015) and compute the spatial version of the price index proposed by Feenstra (1994). This index assumes consumers have CES preferences nested in a Cobb-Douglas utility function. More precisely, their utility function is as follows:

$$U = \prod_{g \in G} \left(\sum_{j \in J_g} q_j^{\frac{\sigma_g - 1}{\sigma_g}} \right)^{\alpha_g \frac{\sigma_g}{\sigma_g - 1}}, \tag{2}$$

where G denotes the set of product-categories consumed, J_g the set of varieties consumed in product-category g, q_j the quantity of each product j consumed, σ_g the elasticity of substitution between varieties in product-category g, and α_g the Cobb-Douglas preference parameter for product-category g.

With such preferences, and assuming consumers are utility-maximizers, we can show the local exact price index in location c (EPI_c) is a geometric average of category-level indexes, the weights being the (constant) share of each category in households' consumption α_g. Category-level indexes have two parts: (1) an intensive part (SPI_c) that is a standard price index that tracks the price gap across products in location c compared to a location of reference, (2) and extensive part (VA_c) that measures the utility cost of unavailable products in location c compared to the same reference location. Here, we assume the reference location is a fictitious city where all the products of all the product-categories are available at a price equal to the median price observed in 2015 across all Ethiopian cities.

Then, the formulas are as follows:

$$EPI_c = \Pi_{g \in G} (SPI_{gc} \times VA_{gc})^{\alpha_g} \tag{3}$$

$$SPI_{gc} = \Pi_{j \in J_{gc}} \left(\frac{p_{jc}}{p_{jE}} \right)^{w_j} \tag{4}$$

$$VA_{gc} = \left(\frac{\sum_{j \in J_{gc}} x_j}{\sum_{j \in J_g} x_j} \right)^{\frac{1}{1-\sigma_g}}, \tag{5}$$

where p_{jc} is the individual price of good j in city c, and p_{jE} is the median price of good j across Ethiopian cities, J_{gc} is the set of products in category g that are available in city c (which size is also noted J_{gc}), w_j is the log-ideal Sato-Vartia weight (based on the share of product j in consumers's expenditures for product-category g both at the city- and at the national- levels), x_j is the total expenditures for product j of the nationally representative consumer, so that $\left(\frac{\sum_{j \in J_{gc}} x_j}{\sum_{j \in J_g} x_j} \right)$ represents the share of the products of product-category g that are available in c in the representative consumer's overall consumption of product category g.

Data limitations imply some components of the exact price index cannot be directly measured or estimated. We thus make the following assumptions to overcome this issue:

- We do not have the details of the weights for products j within product categories g. We thus assume even weights within each product-category $w_j = 1/J_g$ and $x_j = 1/J_g$. Product categories instead have different weights based on survey data, as explained in section 2.2.

- A key parameter is σ_g, the elasticity of substitution between varieties within a given product-category. We do not have the data to estimate these elasticities, and no estimates are available for Ethiopia. We thus use the elasticity estimated by Broda et al. (2017) for the two countries the most similar to Ethiopia, which are present in their dataset (see section 2).

4.1.2 Travel Price Index.

In any version of a Feenstra (1994) price index, the consumption of unavailable varieties is assumed to be zero, or put differently, the implicit price of missing varieties tends to infinity. This feature makes sense in the context of Feenstra (1994), who aims to measure the evolution of the price index over time: a variety that is not available at time t cannot be consumed in t. In a spatial context, however, consumers may travel around to shop products that are not available in their hometown. In an alternative index, we thus assume that if the product is not available in their hometown, consumers choose to purchase the product from the cheapest location (including travel costs) where it is available. Formally, the price $\tilde{p}_{jc}$ of product j in city c is given by:

$$\tilde{p}_{jc} = \begin{cases} p_{jc} & if\ j \in J_{gc} \\ Min\left(p_{j1} + 2\text{travel\_cost}_{c1}, p_{j2} + 2\text{travel\_cost}_{c2}, ..., p_{jC} + 2\text{travel\_cost}_{cC}\right) & if\ j \notin J_{gc}, \end{cases}$$

where p_{jc} is the price of product j in city c, J_{gc} is the set of products available in city c, and travel_cost$_{cc'}$ is the travel cost between cities c and c'.

The travel price index (TPI) is then given by:

$$TPI_c = \Pi_{g \in G} \left(\Pi_{j \in J_g} \left(\frac{\tilde{p}_{jc}}{p_{jE}} \right)^{w_j} \right)^{\alpha_g} \tag{6}$$

where J_g is the entire set of products included in category g considered in the analysis. As for the CES-Feenstra index, due to data limitations, we assume even weights within each category $w_{jc} = 1/J_g$.

Facts on the price of unavailable products. Before turning to the determinants of the price index, we present in Table 6 an analysis of the determinants of the price of the cheapest accessible variety as defined above. Columns (1) and (2) only include the price of the cheapest variety when the product is missing locally, whereas as a robustness check, columns (3) and (4) include all of the observations. The cheapest price at which a missing variety is available in the other cities is higher for more remote locations, which is not surprising, because those locations face, by definition, higher travel costs to the rest of the country. This price is also higher in large cities. This finding implies that if a product is not available in a large city, consumers do not have access to cheap alternatives in the nearby cities. Instead, they have to pay a large trade costs to get access to the product in another (distant) city. When we include in the sample all of the observations, as expected, results are qualitatively similar to those of columns (1) and (2) of Table 6.

4.2 Results

Baseline results. We first propose a visual inspection of the relationship between the cost of living as measured by our prices indexes (and their various components) and city size and remoteness. The graphs in Figure 7 plot the level of the various components of the price index against the degree of remoteness of the city; the size of the circle is proportional to the population of the city it represents. Panel (a) relates to the "intensive" component of

Table 6: ***Macro-determinants of the price of the cheapest accessible variety.***

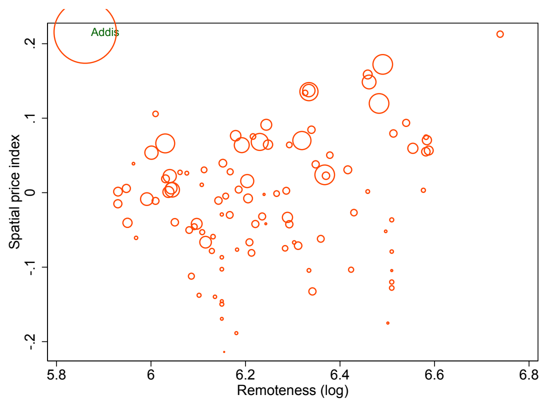
	Log $\tilde{p}_{pc}$			
	(1)	(2)	(3)	(4)
Ln Remoteness	0.581^a	0.644^a	0.331^b	0.356^b
	(0.111)	(0.132)	(0.141)	(0.170)
Ln Population	0.027^b	0.031^b	0.047^a	0.059^a
	(0.011)	(0.015)	(0.015)	(0.019)
Sample	Missing		All	
Purged price	No	Yes	No	Yes
Observations	14,323	14,323	42,506	42,506
R-squared	0.93	0.91	0.93	0.89

Notes: When a product is not available in a given city, the alternative price is the price of the cheapest accessible variety (plus the travel costs).Standard errors in parentheses are clustered in the $region \times market - type$ dimension ($market - type = \{urban, rural\}$). ^{a}, ^{b} and ^{c} respectively denote significance at the 1, 5 and 10% levels.

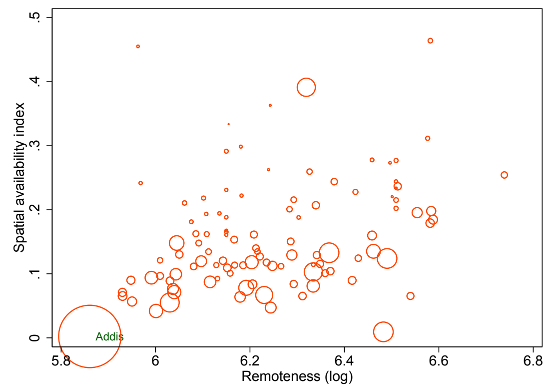
the CES-price index, that is how expensive available varieties are compared to the median price quote observed in Ethiopia. Two main messages emerge: once expenditure shares are accounted for, large cities are still more expensive (big circles are at the top of the graph), and remote cities also still tend to exhibit higher prices (with the notable exception of Addis-Ababa). Regarding the "extensive" component of the price index in panel (b), when expenditure shares and elasticities of substitution are taken into account, large and less remote cities appear to be those in which consumers have access to a greater variety of products, which translates into a lower value of the extensive component of the price index. The combination of the two components gives us the exact price index. It appears on panel (c). The relationship with size is no longer obvious, due to the compensating effects of city size on the level of prices and on product availability. However, remote cities still do have a higher cost of living, due to higher prices and lower product variety. Finally, when using the travel price index in panel (d), the picture is quite similar to the one observed for the Exact Price Index. No clear correlation exists between city size and cost of living, and the correlation with remoteness is positive (but slightly more pronounced).

The graphical observations are largely confirmed by the econometric analysis reported in Table 7. The EPI is the product of the intensive and the extensive components. In large cities, the intensive component is higher (the weighted price index is higher), whereas the availability index is lower, because more products are available. The two effects cancel out

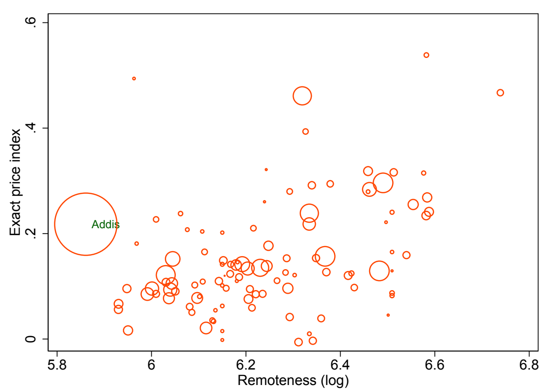
Figure 7: *The geography of prices and product availability*



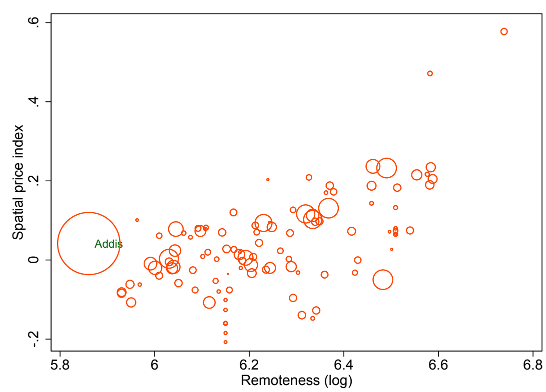
(a) ***SPI***



(b) ***VA***



(c) ***EPI***



(d) ***TPI***

Notes: Each dot is an Ethiopian city. The size of the circle is proportional to the city's population size. The log of the intensive part of the price index is in panel (a). The log of the extensive component is in panel (b). The exact price index in panel (c) is the sum of the indexes presented in panels (a) and (b). The log of the travel price index is in panel (d).

so that city size does not significantly influence the cost of living as measured by the EPI. On the other hand, remote locations exhibit both higher intensive and higher extensive components of the EPI, due to the detrimental effect of remoteness on both the price of available products and on product availability. Results with the TPI are very similar. These results hold in specifications including additional control variables and alternative measures of remoteness as shown in Table 8.[18] More specifically, controlling for the ethnic diversity of the city does not change the correlation between the price index and size and remoteness of the city. Measuring remoteness with a weighted average of bilateral distances between cities does not change the results either. In columns (5) and (6), we find little evidence that the time to an import corridor or the travel time to Addis-Ababa explains spatial differences in terms of the cost of living. The empirical regularities are not driven by Addis-Ababa as shown in columns (7) and (8). Using different weighting schemes (in terms of households' expenditure shares) for urban and rural cities does not affect the results either (columns 9 and 10).

Table 7: ***City-level regressions-local consumer price index***

	Ln SPI	Ln VA	Ln EPI	Ln TPI
	(1)	(2)	(3)	(4)
Ln Population	0.046^a	-0.042^a	0.004	0.015
	(0.007)	(0.007)	(0.011)	(0.012)
Ln Remoteness	0.108^b	0.127^a	0.235^a	0.429^a
	(0.045)	(0.029)	(0.054)	(0.084)
Observations	105	105	105	105
R-squared	0.46	0.46	0.17	0.39

The dependent variable is the log spatial price index (Ln SPI) in columns 1, the log spatial availability index (Ln VA) in columns 2, the log exact price index (Ln EPI=Ln SPI + Ln VA) in column 3, and the alternative spatial price index in column 4. Robust standard errors are in parentheses. ^{a}, ^{b}, and ^{c} respectively, denote significance at the 1%, 5%, and 10levels.

[18] The results are also robust to purging individual prices from their correlation with income per capita as shown in Appendix (Table B.1).

Table 8: ***City-level regressions-local consumer price index : robustness checks***

	EPI	TPI	EPI	TPI	EPI	TPI	EPI	TPI	EPI_UR	TPI_UR
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ln Population	0.004	0.014	0.004	0.016	0.008	0.022	-0.000	0.013	-0.006	0.017
	(0.012)	(0.012)	(0.012)	(0.017)	(0.012)	(0.018)	(0.012)	(0.013)	(0.011)	(0.012)
Ln Remoteness	0.247^a	0.456^a					0.250^a	0.438^a	0.178^a	0.384^a
	(0.048)	(0.075)					(0.055)	(0.086)	(0.044)	(0.080)
Ln of Ethnic diversity index	0.033	0.074^b								
	(0.032)	(0.030)								
Ln Remoteness (weighted)			0.102^c	0.246^a						
			(0.052)	(0.072)						
Ln travel time to import corridor					-0.025	-0.016				
					(0.025)	(0.029)				
Ln travel time to Addis-Ababa					0.024	0.059^c				
					(0.022)	(0.030)				
Observations	105	105	105	105	105	105	104	104	105	105
R-squared	0.18	0.45	0.05	0.21	0.06	0.13	0.18	0.39	0.11	0.35
Addis-Ababa	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes
Weight				common					urban/rural	

Notes: The dependent variable is the exact price index (EPI) in odd columns, the travel price index (TPI) in the even columns, and the price indexes for urban areas (EPI_UR and TPI_UR) are reported in columns 11 and 12. Remoteness is measured as a population-weighted distance of the city to all Ethiopian cities. Standard errors in parentheses are clustered in the $region \times market - type$ dimension ($market - type = \{urban, rural\}$). ^{a}, ^{b} and ^{c} respectively denote significance at the 1%, 5%, and 10% levels.

Rich and poor. We finally compare the impact of remoteness and city size along the income distribution. More specifically, we compute the cost-of-living across Ethiopian cities for households in different quintiles of the income distribution.[19] Doing so requires information on expenditure shares for representative households at different points of the income distribution. We build a series of indexes for rich and poor households and then compute the gap in the cost of living between rich and poor across locations. We present the results in Table 9. The main finding is that the cost of living of poor households deteriorates more strongly with remoteness than the cost of living of richer households. This seems to be driven by the fact that poor households spend a higher fraction of their income on products that are both more expensive when available and more often not directly available in their city.

Table 9: ***Difference in the spatial index of rich and poor***

	gap_SPI		gap_VA	gap_EPI		gap_TPI	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ln average distance to other cities	0.024^{a}	0.025^{a}	0.018^{c}	0.042^{a}	0.043^{a}	0.074^{a}	0.112^{a}
	(0.006)	(0.006)	(0.010)	(0.010)	(0.010)	(0.013)	(0.026)
Ln Population	0.003^{c}	0.001	-0.003^{b}	-0.000	-0.002	-0.000	-0.003
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)	(0.005)
Prices	Raw	Purged	-	Raw	Purged	Raw	Purged
Observations	105	105	105	105	105	105	105
R-squared	0.15	0.12	0.19	0.19	0.20	0.35	0.31

Notes: The dependent variable is the difference in the spatial price index between poor and rich households. (gap_SPI) in column 1, and 2, the difference in the spatial availability index between rich and poor households (gap_VA) in column 3, the difference in the exact price index between rich and poor households (gap_EPI) in column 4, and 5, the difference in the travel spatial price index of poor and rich households (gap_TPI) in column 6, and 7. It is available across regions for urban and rural markets. All these variables are in Log. Robust standard errors are in parentheses. ^{a}, ^{b} and ^{c} respectively denote significance at the 1, 5 and 10% levels.

4.3 Nominal income versus real income

We now ask whether cost-of-living spatial variations across Ethiopian regions have implications for the measurement of income inequality. We have access to a measure of nominal income per capita for 18 regions. These regions are the 10 administrative regions of Ethiopia,

[19]We define poor and rich households as those in the first and fifth quintiles of the income distribution, respectively.

and for eight of them, we have information on income per capita in urban and rural areas (Addis-Ababa and Dire Dawa being only urban).

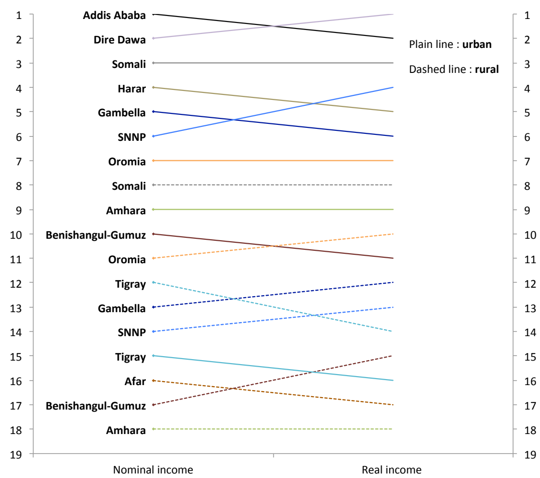
We compare the ranking of regions in terms of nominal income per capita with their ranking in terms of real income per capita. We use three alternative deflators to compute real income, namely, the standard price index (SPI) that does not account for product availability, and the two price indexes (EPI and TPI) that do account for product availability.

We report the exercise in Figure 8, panels (a) to (c). A visual inspection shows that changes in the ranking of regions if we use real income instead of nominal income. The Kendall rank correlation between nominal income and real income ranges from .74 and .90 depending on the index used as a deflator. Accounting for differences in the cost of living is thus essential to get a complete picture of spatial income inequality.

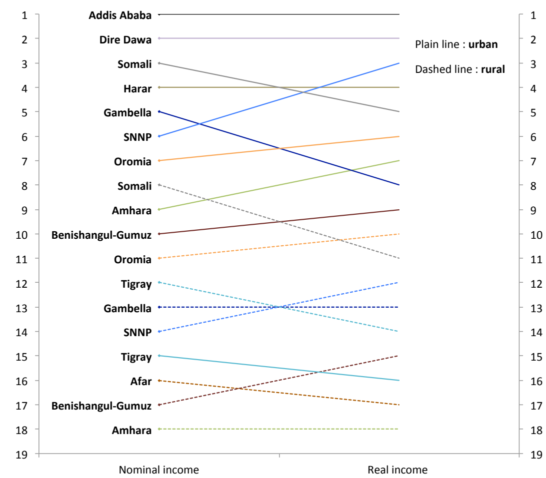
In particular, deflating with a standard price index that does not account for product availability has less of an impact on the ranking than the other two indexes (Kendall rank correlation of .9 in panel (a) vs .82 and .74 in panels (b) and (c)). Interestingly, we see that the ranking of the two wealthiest regions (Addis-Ababa and Dire-Dawa) changes when deflated by the standard price index, whereas it is stable when the deflator is an index accounting for product availability. Therefore, the price level of available products is higher in Addis-Ababa than in Dire Dawa, which more than compensates for the income differential between the two places. However, once the impact of unavailable products is properly measured, the cost of living in Addis-Ababa decreases relative to Dire Dawa, and real income in Addis-Ababa remains higher than in Dire Dawa.

We also see the poorest regions in terms of nominal income are rural regions. These regions are not compensated for the lower nominal income by lower cost of living, although the decision to deflate or not the nominal income influences the ranking among these regions, which highlights the differences in product availability across them.

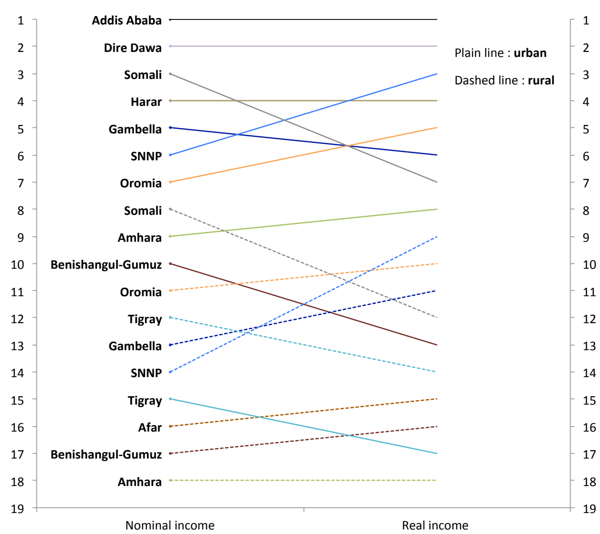
Figure 8: ***Spatial ranking of income: nominal versus real***



(a) ***SPI***



(b) ***EPI***



(c) ***TPI***

Notes: Each line in the graphs links the ranking of a region in terms of nominal income to its ranking in terms of real income. In panel (a) the spatial price index is used to deflate nominal income, the exact price index is used in panel (b), and the travel price index is used in panel (c).

5 Internal migration and cost of living

In this section, we show that properly measuring the local cost of living greatly improves our understanding of internal migration patterns in Ethiopia. A key insight from the economic geography literature is that spatial differences in terms of real income determine migration patterns across regions (Krugman, 1991). Real income in such models featuring love for variety is nominal income deflated by a local ideal price index, which accounts for product availability. Due to the lack of data, the literature has not yet investigated the impact of prices and product availability on internal migration patterns. We here try to fill this gap. We first motivate our empirical model. We then show product availability contributes substantially to the internal movements of people in Ethiopia. We eventually discuss how cost of living intersects with economic geography in determining migration flows.

5.1 Migration equations

We use migration data from the 2013 national labor force survey. Unlike price data, migration flows are available at the region level, not the city level. We thus run our analysis at the region level, and all of the socio-economic explanatory variables included in the analysis are weighted averages of city-level characteristics, using as weights the share of each city in the population of the region where it is located.[20]

In Ethiopia, employment opportunities are the most often cited motivation for migrating. Land scarcity at the origin town and cost of living are also part of the underlying reasons for migrating (Bundervoet, 2018). In addition to these determinants, we control for other standard determinants such as bilateral distance, or income and population size in the destination region (Crozet, 2004). In the end, the empirical model we estimate to uncover the determinants of internal of migration flows is as follows:

$$\begin{aligned} \text{mig}_{od} =& \beta_1 \text{Ln distance}_{od} + \beta_2 \text{Ln emp}_d + \beta_3 \text{Ln pop}_d + \beta_4 \text{Ln area}_d \\ &+ \beta_5 \text{Ln income}_d + \beta_6 \text{ Ln price index}_d + \gamma_o + \epsilon_{od}, \end{aligned} \tag{7}$$

where mig_{od} stands for the share of migrants from region o (origin) who have decided to move to region d (destination), distance_{od} is the distance between region o and region d

[20]For instance, $\text{distance}_{od} = \sum_{i \in o} \sum_{j \in d} \omega_i \times \omega_j \times \text{distance}_{ij}$ where indexes i and j stand for cities in regions o and d respectively, ω_i is the share of city i in the population of region o, ω_j is the share of city j in the population of region d. distance_{ij} is the distance between city i and city j.

and is a standard proxy for migration costs between o and d, emp_d is the employment rate in region d and proxies for employment opportunities at destination, and pop_d and area_d are respectively the total population and the surface area of d and proxy for the size of the destination region. Our main variables of interest are income_d and price index_d and stand for the nominal income and cost of living in d.

Our specification also contains origin fixed effects γ_o, allowing us to abstract from any consideration as to why some regions have migration flows that are more diversified across regions than others. The coefficient associated with each explanatory variable is estimated comparing migration flows from a given region o with each of the possible destinations d. Note that our specification is similar to one where the dependent variable would be the log number of migrants from o to d, instead of the share of migrants from o going to d.

5.2 Results

Table 10 presents our main results. Column (1) corresponds to the estimation of equation 7 excluding the price index from the regressors. Most of the variables have the expected sign. People are less likely to migrate to distant regions, and are more likely to move to larger places in terms of population and surface area. Our proxy for employment opportunities is not significant, which might reflect the difficulty in measuring job opportunities with (the opposite of) the unemployment rate in a country where informality is still pervasive. All else equal, regions with a higher nominal income per capita appear to be more attractive to migrants.

We then add in columns (2) to (4) different measures of price indexes that capture the spatial differences in terms of cost of living. More specifically, when we control for the standard price index (SPI) , which only accounts for the price of available varieties, more expensive regions appear to attract fewer migrants. However, controlling for SPI barely affects the other coefficients, and does little to improve the predictive power of our empirical model. The conclusions are different once we control for the exact price index (EPI), and thus for product availability on top of the price of available varieties. As the results in column (3) of Table 7 show, the coefficient on EPI is strongly negative and significant, and the inclusion of this price index increases the R-squared of the regression by 8 p.p. (i.e. 20%). Measuring local cost of living thanks to TPI instead of EPI yields very similar results. All else equal, the price of consumed goods and services and their availability thus appear to matter both statistically and economically to migrants when they make their location decisions.

To assess the interplay between the cost of living, economic geography, and migration decisions, we reproduce in columns (5) to (8) the first four columns of Table 10, but we now also control for the geographic remoteness of the destination region. As column (5) shows, remote regions attract fewer migrants. Once we control for the price of available varieties in column (6), remoteness remains significantly (and negatively) related to migration, but the coefficient is slightly smaller in absolute value, and the coefficient on population is slightly boosted; these movements in the coefficients are coherent with larger and remote cities/regions being more expensive places to live, so that not controlling for prices tends to downward bias the coefficients on population size and geographic remoteness.

The most interesting results in our view are in column (7). In this specification, we control for both product availability and the price of available varieties thanks to the exact price index (EPI). EPI is, as in column (3), highly negative and significant. Migrants prefer going to places where the price index is lower, that is, with more numerous and less expensive products and services. Moreover, controlling for it causes a sharp increase in the (negative) coefficient associated with remoteness, which becomes statistically indistinguishable from 0, and a decrease in the (positive) coefficient on population, which remains significantly positive. These patterns align with our previous results showing that small and remote places have a narrower set of available products and services. Actually, the negative correlation between migrations flows and remoteness of the destination region vanishes once the exact price index is included, suggesting the correlation is entirely driven by the effect of remoteness on the local cost of living. Using TPI instead of EPI yields qualitatively similar results.

Having an adequate measure of local cost of living that accounts for both prices and product availability is thus crucial to understand internal migration patterns in a developing country such as Ethiopia.

Table 10: ***Migration, income per capita and prices***

	Migration share							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln distance	-0.091^a	-0.094^a	-0.100^a	-0.087^a	-0.079^a	-0.083^a	-0.094^a	-0.090^a
	(0.024)	(0.025)	(0.021)	(0.021)	(0.021)	(0.022)	(0.021)	(0.023)
Ln destination employment rate	0.069	-0.115	-0.265^b	-0.019	0.218^c	0.062	-0.165	-0.075
	(0.108)	(0.113)	(0.109)	(0.097)	(0.124)	(0.113)	(0.131)	(0.130)
Ln destination pop	0.046^a	0.049^a	0.032^a	0.027^a	0.035^a	0.038^a	0.029^a	0.027^a
	(0.010)	(0.010)	(0.008)	(0.008)	(0.008)	(0.009)	(0.007)	(0.008)
Ln destination area	0.031^a	0.030^a	0.034^a	0.036^a	0.030^a	0.029^a	0.034^a	0.037^a
	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Ln income per capita at destination	0.146^b	0.138^b	0.123^b	0.133^b	0.143^b	0.137^b	0.125^b	0.132^b
	(0.062)	(0.061)	(0.057)	(0.055)	(0.058)	(0.058)	(0.057)	(0.055)
Ln destination remoteness					-0.204^a	-0.184^a	-0.081	0.058
					(0.056)	(0.052)	(0.050)	(0.074)
Ln SPI at destination		-0.411^b				-0.317^c		
		(0.182)				(0.163)		
Ln EPI at destination			-0.649^a				-0.571^a	
			(0.125)				(0.117)	
Ln TPI at destination				-0.499^a				-0.576^a
				(0.116)				(0.166)
Observations	110	110	110	110	110	110	110	110
R-squared	0.47	0.49	0.55	0.54	0.51	0.51	0.55	0.54
Adjusted R-squared	0.39	0.40	0.47	0.46	0.42	0.42	0.47	0.46
Origin fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable is the share of migrants from a given region who have decided to move to another region. ^{a}, ^{b}, and ^{c} respectively denote significance at the 1%, 5% and 10% levels. SPI is the spatial price index, EPI the exact price index, and TPI the travel price index.

6 Conclusion

This paper documents geographical discrepancies in the price and availability of products purchased by consumers across Ethiopian cities. We find the cost of living is higher in remote cities because consumers pay higher prices and have access to a narrower set of products and services. All else equal, consumers in large cities pay higher prices but have access to a broader set of products and services. The price effect and the availability effect cancel each other, so that cost of living that accounts for both prices and availability is not related to city size.

One implication of our results is that to have a proper view of income inequality, nominal income should be deflated by local price index, and the choice on the price index matters for the ranking of regions we obtain.

Last, we show that differences in cost of living across Ethiopian regions is also instrumental to understand the destination choice of internal migrants in Ethiopia. Interestingly, differences in cost of living driven by better availability of products and services explain a significant fraction of internal migration patterns. Expanding the availability of products and services might thus be an effective tool for cities to attract migrants or retain their workforce.

References

[1] Addis Ababa Chamber of Commerce and Sectoral Associations, *The Management of Commercial Road Transport in Ethiopia,* May 2009.

[2] Aggarwal, Shilpa, Brian Giera, Dahyeon Jeong, Jonathan Robinson, and Alan Spearot, "Market Access, Trade Costs, and Technology Adoption: Evidence from Northern Tanzania," NBER Working Papers 25253, National Bureau of Economic Research, Inc November 2018.

[3] Atkin, David and Dave Donaldson, "Who's Getting Globalized? The Size and Implications of Intra-national Trade Costs," Technical Report, National Bureau of Economic Research 2015.

[4] Broda, Christian, Joshua Greenfield, and David E. Weinstein, "From groundnuts to globalization: A structural estimate of trade and growth," *Research in Economics*, 2017, *71* (4), 759 – 783. SPECIAL ISSUE IN HONOR OF 40 YEARS OF THE DIXIT-STIGLITZ MODEL.

[5] Bryan, Gharad and Melanie Morten, "The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia," *Journal of Political Economy*, 2019, *127* (5), 2229–2268.

[6] Bundervoet, Tom, *Internal Migration in Ethiopia: Evidence from a Quantitative and Qualitative Research Study*, World Bank, 2018.

[7] Couture, V, "Valuing the consumption benefits of urban density," *Working Paper*, 01 2014, pp. 1–50.

[8] Crozet, Matthieu, "Do migrants follow market potentials? An estimation of a new economic geography model," *Journal of Economic Geography*, 2004, *4* (4), 439–458.

[9] CSA, *Report on the 2005 National Labour Force Survey*, Statistical bulletin, May 2006.

[10] ——, *The 2007 Population and Housing Census of Ethiopia*, Statistical Report at Country Level, 2008.

[11] ——, *Household consumption and expenditure (HCE) survey 2010/11*, Analytical report, October 2012.

[12] ——, *Population Projection of Ethiopia for All Regions At Wereda Level from 2014-2017*, August 2013.

[13] ——, *Analytical report on the 2013 National Labour Force Survey*, Statistical bulletin, June 2014.

[14] ——, *The 2015/16 Ethiopian Household Consumption-Expenditure(HCE) Survey, Results For Country Level*, Statistical Report, January 2018.

[15] ——, *Ethiopia - Socioeconomic Survey 2015-2016, Wave 3*, Report, April 2018.

[16] Deaton, Angus and Alessandro Tarozzi, "Prices and poverty in India," *Princeton, July*, 2000.

[17] Demir, Banu and Beata Javorcik, "Forensics, Elasticities and Benford's Law," CEPR Discussion Papers 12798, C.E.P.R. Discussion Papers 2018.

[18] Dercon, Stefan and John Hoddinott, "Livelihoods, growth, and links to market towns in 15 Ethiopian villages," Technical Report 2005.

[19] Diewert, W. Erwin and Robert C. Feenstra, “Estimating the Benefits of New Products: Some Approximations,” in “Big Data for 21st Century Economic Statistics” NBER Chapters, National Bureau of Economic Research, Inc, March 2019.

[20] Fafchamps, Marcel and Forhad Shilpi, “Determinants of the Choice of Migration Destination,” *Oxford Bulletin of Economics and Statistics*, 2013, *75* (3), 388–409.

[21] FAO, IFAD, WHO WFP, and UNICEF, “The State of Food Insecurity in the World 2018. Building climate resilience for food security and nutrition,” 2018.

[22] ——, ——, and ——, “The State of Food Insecurity in the World 2019. Safeguarding against economic slowdowns and downturns,” 2019.

[23] Feenstra, Robert C, “New Product Varieties and the Measurement of International Prices,” *American Economic Review*, March 1994, *84* (1), 157–177.

[24] Feenstra, Robert C., Mingzhi Xu, and Alexis Antoniades, “What is the Price of Tea in China? Towards the Relative Cost of Living in Chinese and U.S. Cities,” *The Economic Journal*, 2019.

[25] Ferré, Céline, Francisco H.G. Ferreira, and Peter Lanjouw, “Is There a Metropolitan Bias? The relationship between poverty and city size in a selection of developing countries,” *World Bank Economic Review*, 2012, *26* (3), 351–382.

[26] Gollin, Douglas, Martina Kirchberger, and David Lagakos, “In Search of a Spatial Equilibrium in the Developing World,” CEPR Discussion Papers 12114, C.E.P.R. Discussion Papers June 2017.

[27] Gunning, Jan Willem, Pramila Krishnan, Andualem T Mengistu et al., “Fading Choice: Transport Costs and Variety in Consumer Goods,” Technical Report, Centre for the Study of African Economies, University of Oxford 2018.

[28] Handbury, Jessie and David E Weinstein, “Goods prices and availability in cities,” *The Review of Economic Studies*, 2015, *82* (1), 258–296.

[29] Hausman, Jerry, “Valuation of New Goods under Perfect and Imperfect Competition,” in “The Economics of New Goods,” National Bureau of Economic Research, Inc, 1996, pp. 207–248.

[30] Hering, Laura and Rodrigo Paillacar, “Does Access to Foreign Markets Shape Internal Migration? Evidence from Brazil,” *World Bank Economic Review*, 2016, *30* (1), 78–103.

[31] Hsieh, Chang-Tai and Enrico Moretti, "Housing Constraints and Spatial Misallocation," *American Economic Journal: Macroeconomics*, 2019, *11* (2), 1–39.

[32] Jedwab, Remi and Adam Storeygard, "The Average and Heterogeneous Effects of Transportation Investments: Evidence from sub-Saharan Africa 1960-2010," Discussion Papers Series, Department of Economics, Tufts University 0822, Department of Economics, Tufts University 2017.

[33] Judge, George and Laura Schechter, "Detecting Problems in Survey Data Using Benford's Law," *The Journal of Human Resources*, 2009, *44* (1), 1–24.

[34] Kondylis, Florence and Valerie Mueller, "Economic consequences of conflict and environmental displacement," in "International Handbook on Migration and Economic Development," Edward Elgar Publishing, 2014.

[35] Krishnan, Pramila and Peng Zhang, "Restricting Trade and Reducing Variety: Evidence from Ethiopia," Technical Report 2018.

[36] Krugman, Paul, "Increasing Returns and Economic Geography," *Journal of Political Economy*, 1991, *99* (3), 483–499.

[37] Lagakos, David, Mushfiq Mobarak, and Michael E Waugh, "The Welfare Effects of Encouraging Rural-Urban Migration," *Working Paper*, 2020.

[38] Matsa, David A, "Competition and product quality in the supermarket industry," *The Quarterly Journal of Economics*, 2011, *126* (3), 1539–1591.

[39] Ministry of Water and Energy, *Urban Water Supply Access Plan (Part III)*, The Federal Democratic Republic of Ethiopia, December 2011.

[40] Monras, Joan, "Economic Shocks and Internal Migration," CEPR Discussion Papers 12977, Center for Economic Policy Research 2018.

[41] Moretti, Enrico, "Real Wage Inequality," *American Economic Journal: Applied Economics*, January 2013, *5* (1), 65–103.

[42] Muller, Christophe, "Prices and living standards: evidence for Rwanda," *Journal of Development Economics*, 2002, *68* (1), 187–203.

[43] Rancourt, Marie-Ève, François Bellavance, and Jarrod Goentzel, "Market analysis and transportation procurement for food aid in Ethiopia," *Socio-Economic Planning Sciences*, 2014, *48* (3), 198–219.

[44] Ravallion, Martin and Dominique Van De Walle, “Urban-rural cost-of-living differentials in a developing economy,” *Journal of Urban Economics*, 1991, *29* (1), 113–127.

[45] Schiff, Nathan, “Cities and product variety: evidence from restaurants,” *Journal of Economic Geography*, 2015, *15* (6), 1085–1123.

[46] Timmins, Christopher, “Estimating spatial differences in the Brazilian cost of living with household location choices,” *Journal of Development Economics*, 2006, *80* (1), 59–83.

[47] Weber, Sylvain and Martin Péclat, “GEOROUTE: Stata module to calculate travel distance and travel time between two addresses or two geographical points,” Statistical Software Components, Boston College Department of Economics November 2016.

[48] WFP and CSA, “Ethiopia-comprehensive food security and vulnerability analysis (CFSVA),” 2018.

[49] Young, Alwyn, “Inequality, the Urban-Rural Gap, and Migration,” *The Quarterly Journal of Economics*, 2013, *128* (4), 1727–1785.

Appendix material

This set of appendices is structured as follows. Appendix A provides additional tables on the data we use. Appendix B provides additional results.

Appendix A Additional tables on the data

Table A.1: ***List of Product Categories***

Beverages - Alcoholic	Fuel And Power	Vegetables (Fresh)
Beverages - Non Alcoholic	Glass Ware	
Bread And Other Prepared Foods	Household Operation	
Bed Sheet (Non-Patterned Bahir Dar)	Jewellery	
Bed Sheet (Patterned Kombolcha)	Livestock	
Cereals Milled	Meat	
Cereals Unmilled	Medical Care	
Chairs, Tables, Etc.	Metal Ware	
Cigarettes	Miscellaneous Goods And Services	
Clothing	Oil Seeds	
Coffee, Tea, Chat And Buck-Thorn Leaves	Oils And Fats	
Communication	Other Food Items	
Construction Materials	Other Household Equipment	
Cost Of Milling	Other Medical Expenses	
Cost Of Tailoring	Personal Care	
Diary Products And Egg	Personal Effects	
Domestic Service	Plastic Ware (Local Made)	
Draught Animals	Potatoes, Other Tubers And Stems	
Earthen Ware	Prepared Cereal Products	
Education	Pulses Milled Or Split	
Equipment And Accessories	Pulses Unmilled	
Farm Equipment (Hand Made)	Ready-Made / For Adults	
Farm Equipment (Industrial Product)	Ready-Made / For Children	
Fish And Fish Products	Spices	
Food Taken Away From Home	Straw, Bamboo And Others	
Footwear (Men And Women)	Tobacco	
Fruits Fresh	Transport	

Table A.2: ***Product List***

'Katikalla'-Lt	Durrah-Kg	Internet Service-10hrs/month
'Tej' (Mead)-Lt	Hulled Barley-Kg	Mobile Apparatus(Nokia6200)
'Tella'-Lt	Maize (White)-Kg	Mobile call from Tel-Period
Araki (Local)-900cc	Oats -Kg	Telephone Charge (with town)-Period
Beer (Bedele)-330cc	Rice (Imported) -Kg	Telephone Charge to Addia Ababa -Period
Beer (Harar)-330cc	Sorghum Red-Kg	Telephone Line Installation Charge-Once
Beer (Meta Abo)-330cc	Sorghum White-Kg	Bricks (25cm x 12cm x 6cm)
Brandy (Local)-900cc	Sorghum Yellow-Kg	Cement/Bag/(Local)-50Kg
Cognac (Local)-900cc	Wheat Black (Red)-Kg	Chipwood (125cm x 250cm x 8mm)
Gin (Local)-900cc	Wheat Mixed-Kg	Coarse Aggregate Gravel-Meter cube
Saris Wine (Normal)-750cc	Wheat White-Kg	Corrugated Iron Sheet (.2mm)
Ambo Mineral Water-500cc	Book Shelves Wanza (3 Shelves) no Door No	Door made of iron
Coca Cola/Fanta-300cc	Chairs Wanza (Hand Made) Varnished	Floor Board 4m Length
Mineral Water-Lt	Chairs Wanza (Machine Made) Varnished. No	Gutter No 33 -Meter
Pepsi Cola/Mirinda-300cc	Chest of Drawer	Hollow Concrete Block(15x20x40 cm Cube)No
'Dabo' (Traditional Ambasha)-350gm	Cupboard Wanza (2 doors) Varnished	Iron Pipe 6mt. (1/2 inch Wide) Local
'Dabo' (Traditional Sheleto)-350gm	Double Bed Wanza (120cm) Varnished	Key (With hand)
'Enjera' ('Teff' Mixed)-325gm	Sofas (Complete)	Lime-Kg
Biscuits -150gm	Table Wanza (Hand made) Not Varnished. No	Lime/Jeso(Local) -Kg
Bread Wheat (Bakery) -350gm	Table Wanza (Machine made) Varnished	Mega Paints-4Kg
Bed sheet (Non-patterned Bahir Dar)-1.90m x 2.50m	Gissila-Packet	Nail (7cm - 12cm)-Kg
Bed sheet (Patterned Kombolcha)-1.90m x 2.50m	Marlboro-Packet	Nail With Cape-Kg
Bed Cover(Patterned Kombolcha)	Nyala -Packet	Nefas seleke Paints-4Kg
Blanket Woolen(Debre B.)-160cm x 220cm. No	Rothmans/England/ -Packet	Sand -Meter cube
Curtains-meter	Abujedid(Akaki/Bahir Dar)91cm. Meter	Stone for House Construction-Meter cube
Mattress-Sponge (A.A Foam) 120cm	Abujedid(Komb./Arba Min.)150cm Meter	Wall Paints-Super (Fluid) Normal -4Kg
Towel-Local (Kombolcha)	Cotton-Kg	Wall Paints-Super (Fluid) Plastic-4Kg
'Furno Duket' Locally Processed-Kg	Deriya-Meter	Water-Meter cube
'Teff' Black (Red)-Kg	Hisufi-Meter	Water Tanker,Roto (1 meter cube)
'Teff' Mixed-Kg	Jersi-Meter	Window Glass (50cm x 50cm x 3mm)
'Teff' White -Kg	Kashemire-Meter	Wood for House Construction('Atana')
Barley Mixed-Kg	Kefai-Meter	Yewellel Nitaf(Cement Made Tile)
Barley White-Kg	Khaki(Akaki) -Meter	Yewellel Nitaf(Plastic Made Tile)
Durrah-Kg	Khaki(S-10,000 Twil)150cm-Meter	Cereals-100Kg
Maize (White)-Kg	Nylon(Mojo)-Meter	Pepper Whole-100Kg
Oats-Kg	Polyster(Arba Minch/Awasa)-Meter	Pulses-100Kg
Sorghum-Kg	Poplin(Dire Dawa)105cm-Meter	Khaki/Teteron Suit (Boys)
Wheat Mixed-Kg	Poplin(Komb./Arba Minch)150cm-Meter. Meter	Khaki/Teteron Suit (Men)
Wheat White-Kg	Tetron(A.Minch)-Meter	Woolen Suit (Men)
'Teff' Black (Red)-Kg	Wool-England 100%-Meter	Camel Milk-Lt
'Teff' Mixed-Kg	'Chat'-Kg	Cheese Cottage-Kg
'Teff' White-Kg	Buck Thorn Leaves-Kg	Cow Milk (Unpasteurized)-Lt
African Millet-Kg	Coffee Beans-Kg	Cow Milk (pasteurized)-Lt
Barley Black-Kg	Coffee Leaves-Kg	Egg (Traditional)-Dozen
Barley Mixed-Kg	Coffee Whol-Kg	Goat Milk-Lt
Barley White-Kg	Malt-Barley-Kg	Powdered Milk (Me&My)-450gm
Barley for Beer-Kg	Malt-Wheat-Kg	Yoghurt (Traditional)-Lt

Table A.2 (continued).

Maid Servant-Month	Cup of Milk	Neck Laces (6gm 18 carat Local)
Salary for Guard-Month	Cup of Tea	Rings (4gm 18 carat Local)
Unskilled Service (Daily Laborer)-Day	Anbessa Leather Shoes Men(Local)Pair	Seiko-21 Jewels Automatic (Men)
Donkey	Boots for Men Plastic (Local)-Pair	Seiko-21 Jewels Automatic (Women)
Horse	Canvas Shoes (China) Men-Pair	Bull (2-4 Years)
Mule	Canvas Shoes (Local) Men-Pair	Cock (Indigenous)
’Jebena’ Medium Size	Cost of Mending Shoes sole(Men)-Pair	Cow (4 Years and Above)
’Mitad’ (Griddle of Clay)	Leather Shoes Men (Croft)-Pair	Goat (10-15Kg)
Cup for Coffee (China)	Plastic Shoes (Local) Women-Pair	Heifer (2-4 Years)
Local Stove ’(Lakech)’	Plastic shoe(Childern)-Pair	Hen (Indigenous)
Plastic tile-Meter	Sendel Plastic Shoes (Imp.)-Pair	Ox (4 Years and Above)
Plate Clay (imported)	Shoe leather,Children(Imported)-Pair	Sheep (10-15Kg)
Water Pot	Shoe leather,Children(Local)-Pair	Beef-Kg
Ball Point-Bic England	Shoe leather,Women(China)-Pair	Camel Meat-Kg
Day School Fee-Private(Grade 9-10)-Month	Shoe leather,Women(Local)-Pair	Amoxacillin(500mg)-16 caps
Day School Fee-Public(Grade 7 & 8)-Month	Shoe sendel(Plastic) Childern-Pair	Ampicillin(250 mg) Local-56 caps
Exercise Book (50 Leaves) Local	Slippers Sponge Adult (China)-Pair	Asprin (300 mg) Local-20 pills
Night College Fee-Private-Credit/hr	Walking Shoes(Imp.) Non-Leather-Pair	Bactrim(480mg) Local -30 pills
Night School Fee-Government(Gr.9-10)-Month	Avocado-Kg	Chloramphenicol(250mg) Local-56 caps
Night School Fee-Private(Grade9-10)-Month	Banana-Kg	Cough Syrup (Efadykse) Local-125cc
Pencil (China)	Cactus-Kg	Fasider Table-1 pill
School Uniform Fee	Grapes-Kg	Insulin(Lente)-Buttle
Cassette Recorded Original(Local Music)No	Lemon -Kg	Magnesium Oxide -30 Pills
Expense for Photograph(Passport size)4Pho	Mango-Kg	Mezel(250mg) Local-30 caps
Newspaper (Addis Zemen)	Orange-Kg	Paracitamole(500mg) Local -20 Pills
Radio Set Philips 3 Band	Papaya-Kg	Penicillin injection(4 Mu. Local)-Buttle
T.V. Set Philips 21 inch(Colored)	Tangerine-Kg	Tetracycline (250 mg) Local-56 caps
Tape Recorder National (2 Speaker)	Buthane Gas (Shell)-12,5Kg	Vermox(100mg) Local-12 pills
Theater Enterance Fee-Once	Candles	Cooking Pan Medium (Local)
VCD Cassette rent	Charcoal-Kg	Electric ’Mitad’ Aluminium
VCD-Player(Mayato Japan)	Diesel-Lt	Permuze(Japan)
’Digr’	Dung Cake-Kg	Refregrator
’Erfe ’	Electric-Kwatt	Tray (Nickel) Medium N45 Local
’Kember’	Eveready Drycell	Charge for Money Transfer-Once
’Mofer’	Fire Wood-Meter cube	Coffin (Medium Quality)
Plough	Kerosine-Lt	Photocopy-Per page
Sickel	Matches-Box	Wedding Invitation Card-Per page
’Gejera’	Glass for Tea (Duralex)	Castor Beans-Kg
Pick Axe (’Doma’)	Detergent (Omo)-50gm	Ground Nut Shelled-Kg
Sickel	Detergent(Zahira)-50gm	Linseed Red-Kg
Fish Fresh-Kg	Dry Cleaning (Suit Men)	Linseed White-Kg
Sardines (Imported)-125gm	Hard Soap (Imported)-200gm	Niger Seed-Kg
’Fasting Meal Without fish-One Meal	Hard Soap (Local)-200gm	Rape Seed-Kg
’Key Wot Yebeg/Yefyel’-One meal	Incense-Kg	Sesame Seed Red-Kg
’Key Wot Yebere’One meal	Sandal Wood	Sunflower-Kg
’Yebeg Kikil’-One meal	Toilet Paper (Mamko)-Roll	Butter Unrefined-Kg
’Yebeg Tibs’-One meal	Bracelet 20gm (18 carat Local)	Cooking Oil (Imported)-Lt
Cup of Coffee	Earrings (4gm 18 carat Local) -Pair	Cooking Oil (Local)-Lt

Table A.2 (continued).

Vegetable Butter (Sheno & Shady)-Kg	Lentils Split-Kg	Jeans trouser and Jacket
Canned Tomato (Local)-410gm	Mixed Pulses Milled-Kg	Kemise(for childeren)
Dry Yeast(Baking powder)-350gm	Peas Milled-Kg	Shirt Long Sleeved(Imported)Boys
Honey-Kg	Peas Split-Kg	Socks (Imported) Cotton-Pair
Salt-Kg	Peas Split(Roasted)-Kg	Sweater (England) for Girls
Sugar-Kg	Vetch Milled-Kg	Sweater (Local) for Boys
'Kuraz' Small Local Kerosine Lamp	Vetch Split(Roasted)-Kg	T-Shirt
Electric Bulb Philips(40/60 Watt)	Chick Peas-Kg	Basil Dry-Kg
Flash Light	Fenugreek(Green-Kg)	Black Cumin(Local)-Kg
Kerosine Lamp	Haricot Beans-Kg	Black Pepper(Local)-Kg
Bed Charge (Private-Per day	Horse Beans-Kg	Cardamon(Local)-Kg
Bed Charge (gov.)-Per day	Lentils-Kg	Chillies Whole-Kg
Doctor's Fee (Government)-Per visit	Lima Beans-Kg	Cinnamon(Imported)-Kg
Doctor's Fee (Private)-Per visit	Peas Green(dry)-Kg	Cloves(Imported)-Kg
Injection (Service Charge)-Once	Peas Mixed-Kg	Ginger Dry(Local)-Kg
X-Ray(For TB)-Once	Peas White-Kg	Ginger Wet(Local)-Kg
Barbery (Mens Hair Cut)	Soya Beans-Kg	Long Pepper(Local)-Kg
Blade-INDIA	Vetch-Kg	Pepper Whole-Kg
GIV Toilet Soap -90gm	'Gabi'	Tumeric Flour(Local)-Kg
Hair Dressing (Modern)	'Kemisna Netela'	White Cumin(Bishop's Weed) Local -Kg
Modes(Disposable napkins-Packet)	'Netela'	'Sefed'
Parafin Hair Oil -330cc	Geldem	Sack 100Kg Capacity
Perfume-100cc	Jeans Trouser	'Gaya'-Kg
Shaving Machine (medium)	Jogging Suit(sport tuta)	Air Plane (To Addis-Ababa) -Trip
Shoe Polish(Black/Brown)-Once	Khaki Jacket	Animal Transport fare-Trip
Zenith Hair Oil(Liquids Form)-330cc	Khaki Short	Benzene-Lt
Zenith Hair Oil(Non-Liquids Form)-330cc	Leather Jacket	Bus Fare (per km)
Belt (Local) Hand Made	Mekremia	Bus Fare (within Town)-Tarif
Belt (Local) Machine Made	Nylon Dress	Car Washing and Greasing -Trip
Hand Bag (Imported Synthetic)	Pants(for men)	Cart Fare-Trip
Umbrella-Men Medium (Local)	Polyester Suit	Motor Oil (Mobil)-Lt
Umbrella-Women Medium (Imported)	Polyester skirt	Taxi Fare-Trip
Bucket (20 Litres)	Shash (Imported)	Beet Root-Kg
Jerrycan (20 Litres)	Shirts Long Sleeved (Imported)	Cabbage-Kg
'Bula' -Kg	Shirts Long Sleeved (Local)	Carrot
'Kocho' (Unprocessed)-Kg	Shirts Short Sleeved (Imported)	Cauliflower-Kg
Potato-Kg	Singlets (Local) White	Ethiopian Kale-Kg
Sweet Potato-Kg	Socks (Cotton) Imported-Pair	Garlics-Kg
'Dube' Flour-Kg	Sweater (Local) Men	Green Peas-Kg
'Fafa' Flour-Kg	Sweater (Local) Women	Leaks-Kg
Macaroni (Local) Without Egg-Kg	Sweater-Men (Imported)	Lettuce-Kg
Pastini -Kg	Sweater-Women (Imported)	Onions-Kg
Spaghetti (Local) Without Egg-Kg	T-Shirts	Pepper Green-Kg
Chick Peas Milled-Kg	Tetron Trouser	Pumpkin-Kg
Chick Peas Split(Roasted)-Kg	Under Wear China	Spinach-Kg
Fenugreek Milled-Kg	Woolen Suit	Tomatoesv
Horse Beans Milled-Kg	Baby Cloths(Complete)	Tea Leaves(Local)-100gm
Horse Beans Split(Roasted)-Kg	Jeans Trouser	

Table A.3 Barcode products

Category	Product
BEVERAGES - ALCOHOLIC	Beer (Bedele)
BEVERAGES - ALCOHOLIC	Beer (Meta Abo)
BEVERAGES - ALCOHOLIC	Beer (Harar)
BEVERAGES - ALCOHOLIC	Saris Wine (Normal)
CIGARETTES	Gissila
CIGARETTES	Nyala
CIGARETTES	Rothmans/England/
CIGARETTES	Marlboro
COMMUNICATION	Mobile Apparatus(Nokia6200)
EDUCATION	Ball Point-Bic England
BEVERAGES - NON ALCOHOLIC	Ambo Mineral Water
BEVERAGES - NON ALCOHOLIC	Coca Cola/Fanta
BEVERAGES - NON ALCOHOLIC	Pepsi Cola/Mirinda
DIARY PRODUCTS AND EGG	Powdered Milk (Me&My)
PREPARED CEREAL PRODUCTS	'Fafa' Flour
HOUSEHOLD OPERATION	Detergent (Omo)
HOUSEHOLD OPERATION	Detergent(Zahira)
HOUSEHOLD OPERATION	Toilet Paper (Mamko)
OTHER HOUSEHOLD EQUIPMENT	Electric Bulb Philips(40/60 Watt)
CONSTRUCTION MATERIALS	Water Tanker,Roto (1 meter cube)
FUEL AND POWER	Eveready Drycell
FUEL AND POWER	Buthane Gas (Shell)
EQUIPMENT AND ACCESSORIES	Newspaper (Addis Zemen)
EQUIPMENT AND ACCESSORIES	Radio Set Philips 3 Band
EQUIPMENT AND ACCESSORIES	T.V. Set Philips 21 inch(Colored)
JEWELLERY	Seiko-21 Jewels Automatic (Women)
JEWELLERY	Seiko-21 Jewels Automatic (Men)
PERSONAL CARE	GIV Toilet Soap
PERSONAL CARE	Zenith Hair Oil(Liquids Form)
PERSONAL CARE	Zenith Hair Oil(Non-Liquids Form)
TRANSPORT	Motor Oil (Mobil)

Appendix B Additional results

Table B.1: ***City-level regressions-local consumer price index (adjusted prices)***

	Ln SPI	Ln VA	Ln EPI	Ln TPI
	(1)	(2)	(3)	(4)
Ln Population	0.032^a	-0.042^a	-0.010	-0.007
	(0.008)	(0.007)	(0.012)	(0.018)
Ln Remoteness	0.125^b	0.127^a	0.252^a	0.547^a
	(0.052)	(0.029)	(0.063)	(0.115)
Observations	105	105	105	105
R-squared	0.26	0.46	0.18	0.37

Notes: The dependent variable is the log spatial price index (Ln SPI) in columns 1, the log spatial availability index (Ln VA) in columns 2, the log exact price index (Ln EPI=Ln SPI + Ln VA) in column 3, the alternative spatial price index in column 4. Robust standard errors in parentheses with ^{a}, ^{b} and ^{c} respectively denoting significance at the 1, 5 and 10% levels.

Table B.2: ***Migration and real income per capita***

	Migration share							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln destination employment rate	0.069	0.003	-0.013	0.054	0.218[c]	0.147	0.111	0.141
	(0.108)	(0.098)	(0.094)	(0.100)	(0.124)	(0.116)	(0.113)	(0.115)
Ln distance	-0.091[a]	-0.091[a]	-0.090[a]	-0.086[a]	-0.079[a]	-0.080[a]	-0.081[a]	-0.081[a]
	(0.024)	(0.023)	(0.022)	(0.022)	(0.021)	(0.021)	(0.020)	(0.021)
Ln destination area	0.031[a]	0.032[a]	0.037[a]	0.038[a]	0.030[a]	0.031[a]	0.035[a]	0.035[a]
	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Ln destination pop	0.046[a]	0.046[a]	0.036[a]	0.032[a]	0.035[a]	0.035[a]	0.029[a]	0.029[a]
	(0.010)	(0.009)	(0.008)	(0.009)	(0.008)	(0.007)	(0.007)	(0.008)
Ln income per capita at destination	0.146[b]				0.143[b]			
	(0.062)				(0.058)			
Ln destination remoteness					-0.204[a]	-0.194[a]	-0.165[a]	-0.125[b]
					(0.056)	(0.055)	(0.055)	(0.059)
Ln real income per capita at destination (SPI)		0.160[a]				0.152[a]		
		(0.055)				(0.052)		
Ln real income per capita at destination (EPI)			0.197[a]				0.178[a]	
			(0.050)				(0.049)	
Ln real income per capita at destination (TPI)				0.201[a]				0.172[a]
				(0.048)				(0.052)
Observations	110	110	110	110	110	110	110	110
R-squared	0.47	0.48	0.51	0.51	0.51	0.51	0.53	0.52
Adjusted R-squared	0.39	0.40	0.43	0.43	0.42	0.43	0.45	0.44
Origin fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable is the share of migrants from a given region who have decided to move to another region. [a], [b] and [c] respectively denoting significance at the 1, 5 and 10% levels. SPI is the spatial price index, EPI the exact price index and TPI the travel price index.

Table B.3: ***Migration, income per capita and adjusted prices***

	Migration share							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln destination employment rate	0.069	-0.059	-0.245^b	0.023	0.218^c	0.096	-0.103	0.004
	(0.108)	(0.115)	(0.110)	(0.099)	(0.124)	(0.112)	(0.124)	(0.131)
Ln distance	-0.091^a	-0.094^a	-0.101^a	-0.089^a	-0.079^a	-0.082^a	-0.092^a	-0.090^a
	(0.024)	(0.025)	(0.022)	(0.021)	(0.021)	(0.022)	(0.021)	(0.023)
Ln destination area	0.031^a	0.030^a	0.033^a	0.035^a	0.030^a	0.029^a	0.032^a	0.036^a
	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Ln destination pop	0.046^a	0.048^a	0.033^a	0.027^a	0.035^a	0.036^a	0.028^a	0.027^a
	(0.010)	(0.010)	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)	(0.008)
Ln income per capita at destination	0.146^b	0.121^c	0.083	0.088	0.143^b	0.120^c	0.091	0.085
	(0.062)	(0.066)	(0.062)	(0.059)	(0.058)	(0.063)	(0.060)	(0.063)
Ln destination remoteness					-0.204^a	-0.201^a	-0.127^b	0.023
					(0.056)	(0.054)	(0.049)	(0.074)
Ln adjusted SPI at destination		-0.254				-0.236		
		(0.164)				(0.150)		
Ln adjusted EPI at destination			-0.548^a				-0.463^a	
			(0.121)				(0.107)	
Ln adjusted TPI at destination				-0.388^a				-0.410^a
				(0.104)				(0.147)
Observations	110	110	110	110	110	110	110	110
R-squared	0.47	0.48	0.53	0.54	0.51	0.51	0.54	0.54
Adjusted R-squared	0.39	0.39	0.45	0.46	0.42	0.42	0.46	0.45
Origin fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable is the share of migrants from a given region who have decided to move to another region. ^{a}, ^{b} and ^{c} respectively denoting significance at the 1, 5 and 10% levels. SPI is the spatial price index, EPI the exact price index and TPI the travel price index. Incomes are deflated using prices which are adjusted for spatial differences in term of products quality.

Table B.4: ***Migration and real per capita income (adjusted prices)***

	Migration share							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln destination employment rate	0.069	0.000	-0.015	0.068	0.218^c	0.149	0.118	0.148
	(0.108)	(0.099)	(0.094)	(0.101)	(0.124)	(0.116)	(0.113)	(0.116)
Ln distance	-0.091^a	-0.092^a	-0.091^a	-0.086^a	-0.079^a	-0.081^a	-0.082^a	-0.081^a
	(0.024)	(0.023)	(0.022)	(0.022)	(0.021)	(0.021)	(0.020)	(0.021)
Ln destination area	0.031^a	0.031^a	0.036^a	0.038^a	0.030^a	0.030^a	0.034^a	0.035^a
	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)	(0.008)	(0.008)
Ln destination pop	0.046^a	0.046^a	0.037^a	0.030^a	0.035^a	0.035^a	0.029^a	0.028^a
	(0.010)	(0.009)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.008)
Ln income per capita at destination	0.146^b				0.143^b			
	(0.062)				(0.058)			
Ln destination remoteness					-0.204^a	-0.202^a	-0.176^a	-0.117^b
					(0.056)	(0.055)	(0.055)	(0.058)
Ln real income per capita at destination (adjusted SPI)		0.142^a				0.138^a		
		(0.050)				(0.047)		
Ln real income per capita at destination (adjusted EPI)			0.175^a				0.161^a	
			(0.046)				(0.045)	
Ln real income percapita at destination (adjusted TPI)				0.180^a				0.155^a
				(0.040)				(0.043)
Observations	110	110	110	110	110	110	110	110
R-squared	0.47	0.48	0.50	0.51	0.51	0.51	0.52	0.52
Adjusted R-squared	0.39	0.40	0.42	0.44	0.42	0.43	0.44	0.44
Origin fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable is the share of migrants from a given region who have decided to move to another region. ^{a}, ^{b} and ^{c} respectively denoting significance at the 1, 5 and 10% levels. SPI is the spatial price index, EPI the exact price index and TPI the travel price index. Incomes are deflated using prices which are adjusted for spatial differences in term of products quality.