CENTRE FOR THE STUDY OF AFRICAN ECONOMIES

CSAE Working Paper WPS/2018-05-2

Fading Choice: Transport Costs and Variety in Consumer Goods*

Jan Willem Gunning[†], Pramila Krishnan[‡] and Andualem T. Mengistu[§]

September 19, 2018

Abstract

We examine the spatial variation in variety of manufactured goods to study how choice fades with distance. We model monopolistically-competitive trade between market town and village and show how transport costs reduce consumer welfare through reduced variety. We use data from a purpose-designed survey of shops and consumers in villages in Ethiopia and prices of matched source and destination goods to estimate similar tastes for variety across space. Our estimates suggest an average mark-up of 10-15% and welfare costs of falling variety at 19% on average of expenditures on manufactures, in contrast to the effect of prices at an average of 1.75%. The cost of lower variety in remote places is substantial.

1 Introduction

Transport costs reduce consumer welfare not only through lower incomes and higher prices but also through reduced variety: choice fades with distance. Studies of market development

*We would like to thank David Bevan, Stefano Caria, Stefan Dercon, Douglas Gollin, Bart Minten, Kate Orkin, Simon Quinn, Natalie Quinn, Chris Woodruff and seminar participants in Addis Ababa, Freiburg, Gothenberg, Trinity College Dublin, London, Oxford and Nottingham for their valuable comments on earlier versions of this paper. Grants from the International Growth Centre number 89218 and the African Capacity Building Fund are gratefully acknowledged. The views expressed in this article are those of the authors.

†Vrije Universiteit Amsterdam (The Netherlands)

‡University of Oxford and CEPR(UK)

§Ethiopian Development Research Institute(Ethiopia)

UNIVERSITY OF OXFORD

Centre for the Study of African Economies

Department of Economics · University of Oxford · Manor Road Building · Oxford OX1 3UQ

T: +44 (0)1865 271084 · F: +44 (0)1865 281447 · E: csae.enquiries@economics.ox.ac.uk · W: www.csae.ox.ac.uk

typically focus on the spatial integration of markets as measured by price wedges for goods traded across space (see Fackler and Goodwin (2001) for a review). In contrast, the literature on spatial variation in product variety is sparse. Our focus is on the choice of varieties in domestic markets in contrast to most studies of the extensive margin that focus on across countries. We use data from a purpose-designed survey of shops and consumers in rural villages in Ethiopia to understand how remoteness reduces the variety of consumer goods available and the costs to consumers of fading choice.

Why might remoteness affect product variety? First, in remote areas, previous research (Jacoby and Minten (2009), Jacoby (2000), Minten and Kyle (1999), Minten et al.(2013), Stifel and Minten (2008), Dercon et al. (2009) and Khandker et al. (2009)) has shown that productivity is low and poverty is more intense. In the case of Ethiopia for instance, the costs imposed on farm households due to remoteness alone lower incentives to use potentially profitable inputs: Minten et al.(2013) show that incentives for fertiliser use decline quickly over space because output to input price ratios drop by half. Dercon et al. (2009) report that access to all-weather roads increases consumption growth by 16 percentage points and, reduces the incidence of poverty by 7 percentage points. Thus, if the demand for variety has a positive income elasticity[1], then remoteness will be also be associated with a reduction in variety. Second, high transport costs imply that individual varieties will be more costly in rural areas. With any fixed cost of varietal marketing, this will give rise to a reduction in the set of varieties. And third, the pricing power of retailers in remote areas, where market size may be too small to support much competition, may imply that shopkeepers prefer to restrict the set of varieties in order to focus on products with high margins. We investigate the effects of transport costs, incomes and market size on the variety of products available across space in what follows below but before we do so we turn to a brief summary of the related literature.

While the literature on product variety within countries is thin, the effects of fixed costs and scale, incomes and market size have been extensively examined in trade across countries. The theoretical literature on the effect of scale economies in international trade (e.g. Krugman (1991)) has inspired a large empirical literature on the extensive margin or variety in trade. Hummels and Klenow (2005), for example, find that larger and wealthier countries trade more, and that 60 percent of the difference in aggregate trade flows comes from differences in the

[1] Li (2011) examines the Engel curvey for variety in India, with positive estimates of the elasticity of variety in food consumption.

number of goods traded[2]. Smaller countries will have less variety in consumption as a result of fixed costs, both in production and in international trade. This literature suggests that there is a strong response of the extensive margin to changes in trade barriers or country size (Broda and Weinstein (2006), Eaton et al. (2011), Chaney (2008)). Relatively little is known, however, about the variety in domestic trade. For the U.S.A., Handbury and Weinstein (2014) examine detailed barcode data on purchase transactions by households in 49 U.S. cities and find that biases in spatial price measurement due to heterogeneity in quality and the availability of variety are large: correcting for these, they find food prices lower in bigger cities. Handbury (2012) finds systematic variation in the living costs faced by consumers in different income classes and demonstrates that these differences are driven by cross-city variation in product variety rather than prices. Conventional price indices compare only the prices of goods common to locations and ignore differences in variety across locations thus fail to account for such differences in living costs.

Indeed, as far as we are aware, this is one of the few papers to document the spatial variation in variety and to examine the role of local demand and transport costs in explaining it[3]. We assess the welfare effects of transport costs through the channel of reduced variety, allowing for heterogeneity across households in income. We model the number of varieties available by location to understand the local retail environment and use data on local availability of varieties to understand the variation in varieties across villages. Li (2011) offers an excellent complement to the discussion in this paper: using data on variety in food consumption across India, he documents an Engel curve for variety. In contrast to our approach, he takes the local retail environment as exogenous. He discovers substantial welfare effects in food consumption over time with welfare gains accruing largely to rich and urban households and points to the association between consumption diversity, economic development and urbanisation. Two other fascinating and complementary pieces examine the role of reducing trade and transport costs via the expansion of e-commerce in China. Fan et al. (2016) documents the fall in domestic trade costs through the rise of e-commerce in China: using data from the e-retailer Alibaba, they examine the implications of falling costs in reducing spatial consumption inequality. Staying in China, Couture et al. (2018) combine an experiment with survey and administrative microdata to provide evidence that e-commerce leads to a significant reduction in cost of living via lower

[2] Broda and Weinstein (2006) estimate welfare gains from variety growth in imports alone as 2.8 percent of GDP.

[3] Hillberry and Hummels (2008) show that the spatial frictions strongly affect the extensive margin of trade within the U.S.A. Unlike the patterns documented in this paper, they show that spatial frictions affect the trade in intermediate inputs rather than manufactures.

prices and increased product variety.

Differences in prices of goods traded within developing countries have attracted more attention than variety[4]. Atkin and Donaldson (2015) use spatial price differences from official price surveys in Nigeria and Ethiopia as a proxy for trade costs within developing countries. Van Leemput (2016) quantifies the size of internal versus external trade barriers in India and finds that reducing internal trade barriers across states within India offers twice the welfare gains compared to reducing international barriers in terms of price integration.

There are two reasons why the literature on the differences in variety across space is sparse. The first is that household and other surveys typically do not obtain information on availability at retail outlets. The second reason is the difficulty in identifying the impact of transport costs. The causal identification of the impact of roads using cross-sectional data is near impossible given that roads are usually constructed for some reason, often with economic returns in mind, that is likely to be conflated with other characteristics that also affect outcomes of interest (Adam et al. (2016)). Plausible studies of the impact of roads have relied on randomised roll-out of road improvements (e.g.Casaburi et al. (2013) Gonzalez-Navarro and Quintana-Domeque, (2016)) or instrumental variables (Banerjee et al. (2012), Storeygard (2016) and Donaldson (2016)). An alternative strategy used by Jacoby and Minten (2009) and Minten et al.(2013) relies on an innovative survey of households in a small, relatively homogeneous regions of rural Madagascar and Ethiopia, over which transport costs to the same market vary tremendously, due not to the placement of roads but to the terrain.

We implement a similar strategy as in Jacoby and Minten (2009) to identify the effect of remoteness on varieties available, using data from a purposive survey of retail outlets and local traders in Ethiopia. There are two key concerns that we seek to address. The first is that road infrastructure is not randomly built across districts and villages; we needed to account for this non-random placement effect. By sampling up to three villages in the same district, and focusing on within district variation, we can account for placement at the level of the district. In addition, we find no real evidence of placement effects within districts on observable factors but perhaps more telling is that district fixed effects account for the type of road in over half of the villages surveyed. The second issue is the matching of goods at destination with their origin in order to correctly assess the impact of transport costs. First, accounting for the variation

[4]For developed countries there are also some estimates, e.g. Handbury and Weinstein (2014) and Handbury (2012) for the US. Handbury (2012) finds that price differences for different income groups are driven entirely by variety differences across cities.

across districts allows us to concentrate on the relationship between the local (within district) supply of variety in goods and local transportation costs, local (average) incomes and local heterogeneity in incomes. The second concern, the matching of origin and destination in the supply of goods required that we do so by designing the survey so that the local market town is the only supplier of consumer goods and can thus serve as a benchmark for local availability. This is explained in detail later in this paper.

We conducted a survey of 296 villages in 100 districts across the four main regions of Ethiopia. We focus on manufactured consumer goods: this covers about 20 item groups from processed foods, drinks, garments, footwear, cosmetics to kitchen ware, hardware and small electronics. Within these groups there are specific items such as pasta, beers, soaps, plastic tableware, linens, notebooks, and batteries and we can further disaggregate many of these by specific brands. Data were collected in shops in the market towns closest to the villages and in shops and periodic markets (served by traders) in villages. Village officials and six households in each village were surveyed. These data allow us to examine how transport costs, incomes, the income distribution and size of the local market affect the fraction of items available in a remote village compared to its nearest market town. We construct a theoretical framework of monopolistic competition among traders in manufactured goods to examine the relationship between remoteness, incomes and market size; we also use it to estimate the welfare impacts of living in remote areas.

Our main contribution, arising from our results on the availability of variety across space, is to show that the loss of variety can lead to a substantial welfare loss for households who live far from market towns. In particular, we complement the discussion on the effect of variety in domestic trade which has focused on the impact on prices across space(Atkin and Donaldson (2015) Handbury and Weinstein (2014), Handbury (2012)and Glaeser et al. (2001)) and extend this to the impact of space on variety.

The paper is structured as follows. The next section discusses the theoretical framework in two parts. We begin with a simple model of trade and transport, with homogeneous agents. We then generalise this to account for heterogeneity in incomes since it is plausible that the distribution of income might have a separate effect on variety supplied locally. We follow this in Section 3 with a description of the purposive survey and the data obtained and describe the rationale for the design. Section 4 offers results on how choice varies with travel times together with estimates of the price mark-ups across space. Section 5 uses these results to estimate the welfare costs of remoteness. Section 6 concludes.

2 Model of Transport Costs and Availability of Consumer Goods

2.1 The basic model

We begin with a very simple model of trade and transport to investigate how consumers are affected by transport costs, not only through changes in prices, but also through changes in the set of goods they can buy.

We assume Dixit and Stiglitz (1977) preferences:

$$u = \sum_{i=1}^{n} c_i^{\theta} \qquad 0 < \theta < 1 \tag{1}$$

where c_i denotes consumption of good i, the consumer takes the number of available goods n as given, and $1 - \theta$ is a measure of the consumer's taste for variety and y denotes income. For $\theta \to 1$ the consumer treats different goods as close to perfect substitutes, for $\theta \to 0$ as complements.

The consumer solves:

$$\max_{c_1,..,c_n} \sum_{i=1}^{n} c_i^{\theta} \text{ subject to } \sum_{i=1}^{n} p_i c_i = y. \tag{2}$$

The first-order condition gives:

$$c_i = \left(\frac{\theta}{\lambda p_i}\right)^{\frac{1}{1-\theta}} = \frac{y}{n p_i}. \tag{3}$$

Here λ denotes the Lagrange multiplier of the budget constraint. An immediate implication of (3) is that the price elasticity of consumer demand decreases with the taste for variety: the more the consumer cares about variety the more inelastic demand and hence the greater the monopoly power of the trader.

We consider transport and trade to J villages ($j = 1, .., J$), each connected by its own road to a market town where all n goods are available at a given price. The distance between village j and the town is s_j. Each village has m consumers whose income is derived from selling a crop. This income is fixed at y^* at $s = 0$ and declines with the distance from the market town, reflecting iceberg transportation costs:

$$y_j = \frac{y^*}{\gamma s_j}. \tag{4}$$

(The assumptions that villages have the same number of consumers and that incomes are equal within a village will be relaxed in the next section.)

Since goods are modelled symmetrically we can drop the index i. Each trader deals in a single good which he buys in the market town at a given price $\overline{p}$ and transports to a subset of the villages where the good is sold at the price p_j^*. The cost of transporting a quantity q over a distance s is $(\alpha + \beta q)s$. Hence the trader's profits on sales at j are given by:

$$\pi_j = [p_j^* - (\overline{p} + \beta s_j)](m.c) - \alpha s_j. \tag{5}$$

Traders are engaged in monopolistic competition. Hence each trader sets a profit maximising price taking into account the demand curve (3). From (3) and (4) this gives for locations that are served:

$$p_j^* = \frac{\overline{p} + \beta s_j}{\theta}. \tag{6}$$

This shows that each trader charges a markup over marginal costs $\overline{p} + \beta s_j$ and that this markup is increasing in the taste for variety $1 - \theta$.[5][6]

Free entry drives profits to zero and this determines the number goods available at a particular location. From (3)-(5) and the budget constraint this implies:[7]

$$n_j = \frac{(1-\theta)my^*}{\alpha\gamma s_j^2}. \tag{7}$$

[5] Note that for $\theta \to 1$ the markup vanishes: if consumers have no taste for variety then marginal cost pricing is, of course, optimal.

[6] Benassy (1996) points out that this formulation locks the taste for variety to the elasticity of demand and suggests an alternative formulation that would separate the two parameters. While this is theoretically appealing, it is clear that the taste for variety parameter thus separated affects only unobservable variables, namely welfare and the number of optimal varieties (see Section 6 in Benassy (1996)) and hence is unidentifiable in a fundamental sense.

[7] For simplicity we treat n as a continuous variable.

Equation (7) indicates that at a distance s_j from the market town the number of goods available is increasing in the taste for variety and in the size of the market (my_j), measured in terms of fixed transport costs, $my_j/(\alpha s_j) = my^*/(\alpha\gamma s_j^2)$.

In this model a reduction in transport costs (a lower value of α, β, or γ) increases consumer welfare in three ways. First, from (4), when γ falls the value of income increases. Secondly, from (6), when a fall in β reduces the price of consumer goods so that (for given n_j) consumption of each of the available goods increases. Finally, a fall in α increases the number of goods available in a given location. This reduces consumption of each good but raises welfare because of the consumer's taste for variety. We call these three effects the income effect, the price effect and the variety effect respectively.

One of the key assumptions here is the existence of *fixed costs*. This is particularly salient in this setting as we discovered upon interviewing traders. These costs include licences (by either items or groups of items), payment to village authorities[8], and the costs of own transport and accommodation for itinerant traders as well as inventory and storage costs. Table 1 offers a summary of why traders specialise in items: 38% of traders quote licences required as the reason for not trading other items, 33% quote lack of capital while 20% claim that the lack of demand in remoter areas dissuades them from carrying more items.

In the Dixit-Stiglitz specification goods are modelled symmetrically so that in equilibrium prices and quantities are the same for all goods consumed. A corollary is that the equilibrium is affected only by market size (total income in a village), not by the way income is distributed. This is a serious limitation. In the next section we therefore drop the symmetry assumption. As a result the number of goods available in a specific location will depend on the distribution of income.

2.2 Market size and Inequality

For the time being we maintain the assumption that consumers are identical in terms of income. However, we change the utility function by grafting onto the Dixit-Stiglitz specification the two

[8]The existence of fixed lump sum fees per trader is attributed to a combination of land regulation and tax policy. Rozelle et al. (2003) describe similar difficulties facing itinerant traders in rural China: "Local officials clearly understand how difficult it is to collect taxes from itinerant peddlers. As a consequence, officials spend little time trying to collect value-added taxes, relying instead on simple taxation methods such as collecting stall fees or negotiating lump sum fee payments."

key features of the Murphy et al. (1989) specification. First, beyond a certain level, $\overline{c}$, the consumer can raise his utility only by consuming new goods rather than consuming more of the same goods. Secondly, marginal utility declines not only as consumption of the same good increases (as in the Dixit-Stiglitz specification) but also as more goods are consumed. Marginal utility for good i is now given by:

$$u_i = \begin{cases} \frac{\theta v_i}{c_i^{1-\theta}} & \text{if } 0 < c_i \leq \overline{c}. \\ 0 & \text{if } c_i > \overline{c}. \end{cases} \tag{8}$$

where the parameters v satisfy

$$v_i > v_{i+1} > 0 \quad i = 1, 2, ..$$

If n goods are available in the location utility maximization gives:

$$c_i = \min(\left(\frac{\theta v_i}{\lambda p_i}\right)^{\frac{1}{1-\theta}}, \overline{c}) \quad i = 1, 2, .., n. \tag{9}$$

Prices p_i are equal across goods and equal to a village specific price level p_j^*. The trader's profit maximisation gives (as before):

$$p_j^* = \frac{\overline{p} + \beta s_j}{\theta}.$$

Note that the parameters v affect consumption levels but not prices. Since new goods have lower marginal utility c_i decreases with i. If the consumer's income is sufficiently low $c_1 < \overline{c}$ so that

$$\overline{c} > c_1 > c_2 > .. > c_n \tag{10}$$

while at higher income levels consumption of the first k goods ($n > k \geq 1$) will be at the bound:[9]

$$c_i = \overline{c} \quad i = 1, ..k; \ \overline{c} > c_{k+1} > .. > c_n. \tag{11}$$

[9] We ignore the extreme case $k = n$. In that case part of the consumer's income would be useless: spending it would not add to utility. This is an empirical point: in Ethiopia, the "excess" income is likely to be held in livestock assets.

As income rises k will successively take the values $1, 2, ...$

We now relax the assumption that the consumers at a given location are identical. Instead there now are m_{jH} rich consumers with income y^*_{jH} in the village and m_{jL} poor consumers with income y^*_{jL}.[10] We assume that y^*_{jL} is sufficiently low for (10) to hold and y^*_{jH} sufficiently high for (11) to hold.[11]

The number of goods available, n_j, is determined by the zero profit condition for the last good:

$$[p^*_j - (\overline{p} + \beta s_j)](m_{jL}c_{n_jL} + m_{jH}c_{n_jH}) = \alpha s_j.$$

or

$$m_{jL}c_{n_jL} + m_{jH}c_{n_jH} = \frac{\alpha s_j}{\overline{p} + \beta s_j}\frac{\theta}{1-\theta}$$

or, using the two budget constraints:

$$m_{jL}\frac{y^*_{jL}}{\gamma s_j}\frac{v^*_{n_j}}{v^*_1 + .. + v^*_{n_j}} + m_{jH}(\frac{y^*_{jH}}{\gamma s_j} - p^*_j k\overline{c})\frac{v^*_{n_j}}{v^*_{k_j+1} + .. + v^*_{n_j}} = \frac{\alpha s_j}{1-\theta} \tag{12}$$

where

$$v^*_i = v_i^{1/(1-\theta)}.$$

The equilibrium condition in equation (12) determines the number of goods available, n_j.The left hand side (LHS) of this equation measures the value of total consumption for the last good, n_j, in the village[12]. Note that the LHS is decreasing in n_j. Hence any change that increases the LHS must be offset by an increase in n_j. It follows that n_j is increasing in the size of the two groups (m_{jL} and m_{jH}) and in their incomes (y^*_{jH} and y^*_{jL}). Availability is also increasing in market size, measured by total village income $m_{jL}y^*_{jL} + m_{jH}y^*_{jH}$. Less obviously, availability increases with income inequality: a mean preserving spread in the distribution of income increases n_j. Intuitively this is because the poor spend an increase in income on all goods, the rich only on those goods for which their consumption has not yet reached the level

[10]We assume throughout that $y^*_L < y^*_H$.

[11]If (10) would hold for both groups then the rich would consume more of each good in proportion to their income:

$$c_{iH} = c_{iL}(y^*_H/y^*_L)$$

and in this case inequality would not matter: the number of goods available would be determined by total market size (as in the Dixit-Stiglitz model), irrespective of the distribution of income.

[12]Since n_j is integer (12) will not hold as an equality. Instead in equilibrium the $LHS \geq RHS$ (where RHS denotes the right hand side) and $LHS < RHS$ if n_j is replaced by $n_j + 1$. Hence the last trader may make a profit but there is no incentive for a new trader to enter the market with a new good.

$\bar{c}$. A mean preserving spread will therefore raise total demand for the marginal good and this induces an increase in variety.

More precisely, the effect of a mean preserving spread (an increase in y^*_{jH} offset by a reduction in y^*_{jL} so as to keep total income $m_{jL}y^*_{jL} + m_{jH}y^*_{jH}$ and hence mean income constant) is given by the partial derivative

$$\frac{\partial LHS}{\partial y^*_{jL}} = \frac{m_{jH}}{p^*_j}\left[\frac{v^*_{n_j}}{v^*_{k_j+1} + .. + v^*_{n_j}} - \frac{v^*_{n_j}}{v^*_1 + .. + v^*_{n_j}}\right] > 0$$

where n_j is kept constant. This increase of the LHS calls for an offsetting increase in n_j. Hence an increase in inequality unambiguously improves variety, n_j. This is an important testable implication of the model.

It should be noted that in this framework both the rich and the poor consume the same set of varieties, with the poor consuming lower quantities of the varieties. As in Li (2011), the gains in welfare from increased variety occur not because of heterogeneity in tastes in this setting but from being able to counteract diminishing returns to consuming increasing quantities of the same variety.

The model describes a spatial equilibrium where households are fixed in space, which captures a key feature of the Ethiopian context. Farmers in Ethiopia have only user rights to land, which they must relinquish if they migrate to town. Settlement patterns across Ethiopia have remained unchanged for decades and migration rates are amongst the lowest in sub-Saharan Africa, both of which are consistent with low urbanisation rates as well. And finally, it should be stressed that the loss of variety with remoteness in this setting pertains to rather basic goods in the consumption basket making a tradeoff with the intangible benefits of remoteness less plausible. More generally, this picture is consistent with Gollin et al (2017) who argue that there is little evidence that the higher consumption levels of urban areas are offset by lower non-monetary amenities in sub-Saharan Africa.

3 Data and Survey Design

Ethiopia offers a useful setting for examining the role of transport costs and other wedges in market development. It is landlocked which affects external trade while internal trade costs are strongly affected by its particular physical geography. It has a mean elevation of over 1000 metres and the bulk of the population lives on the high plateau, a terrain bisected by mountains. The terrain has also meant that Ethiopia has one of the lowest road densities in the world. While this potential bottleneck to market development has been recognised and resulted in substantial investment in new roads over the last decade[13], it is still the case that vast swathes of rural Ethiopia are dependent on travel to market using mules on country tracks or on foot. Road density in Ethiopia has risen from 0.46 km per 1000 people to 0.57 km, which compares very poorly with the average in sub-Saharan Africa of 3.9 km/1000. With low urbanisation rates at 17 percent (compared to a sub-Saharan average of 33 percent) both physical and human geography in Ethiopia mean that remoteness from markets is fundamental to describing market access.

The data used here come from a purposeful survey of 295 villages in 100 districts across the four main regions of Ethiopia. The survey was designed to address two potential concerns in taking the theory to data. First, it was to ensure that the local market town was the only supplier of consumer goods and could thus serve as a benchmark for local availability. The villages were chosen such that they were linked directly only to one market town within the district. We also interviewed district officials and questioned them on the list of villages within districts and picked a subset of the remoter villages which had only one road or transport connection to the nearest market town. The aim was to ensure that the supply of goods to the village was only possible via the market town and hence the availability of goods in the market town would serve as the benchmark for the range available in the village albeit by separate roads. Each market town is linked to the village by a single road or track (and no other point of natural entry or exit); there are an average of three villages linked to the same market town. Figure 1 provides an illustrative map, based on two of the four sampled regions. Market towns are indicated in shades of blue, denoting differences in travel time to the farthest of the three villages they are connected to, while the black dots denote the villages. Table 2 examines our sampling frame: about 73% of villages within districts are connected to just one market town directly. Our sampling strategy thus produces villages that are not unusually situated given the hilly terrain. Secondly, by sampling three villages on average in the same district, and focusing on within district variation, we can account for placement at the level of the district.

[13]The Ethiopian Government embarked on a major programme of investment in roads with a 10-year Road Sector Development Program in 1997 (RSDP 1997-2007). The first phase of the RSDP (1997-2002) focused on the rehabilitation of the main road network and since then has worked on an investment programme for new roads. See The Ethiopian Road Authority Manual (2015).

(We find no evidence of linear placement effects within districts as discussed below in Table 3). Accounting for the variation across districts allows us to concentrate on the relationship between local supply of variety in goods and local transportation costs, local (average) incomes and local heterogeneity in incomes. Data were collected in shops in the market towns closest to the villages and in shops and periodic markets (served by traders) in villages. Key village officials were interviewed about village characteristics, amenities and endowments. In addition, in each village, six households were interviewed, three chosen from the lower and three from the upper end of the wealth distribution, again with local consultation. The households were interviewed on their basic characteristics, consumption of consumer goods in terms of both variety and expenditures and their incomes (both from agriculture and outside agriculture if relevant).

We obtained data on the infrastructure, population, public goods, local incomes and production and most important, detailed data on transport infrastructure, travel times and quality of roads in each village. The data were collected both from district (woreda) officials and local village (kebele) representatives.

The main part of the survey involved the survey of shops in the nearest local market town on a full list of potential consumer goods across different categories of consumption from processed foods, household goods, toiletries to clothing and shoes[14]. The next stage involved the collection of data on consumer goods available in the village. Depending on the village, there were small fixed local shops and periodic markets where traders bring consumer goods to the village[15]. We surveyed both kinds of outlets on the variety (including brands) of goods and their prices, with additional questions on whether items from the list of goods available in the local market town were usually available even if they were not on sale on the day we visited them.

The key issue in the design of the survey as explained above was to ensure that the relationship between travel time and the availability of goods was not contaminated by the possibility that roads (or improved roads) are more likely to be placed in villages that are also wealthier and more likely to attract a wider range of goods. The Ethiopian Road Authority supposedly uses five main criteria[16] during the preliminary selection of new road projects which unsuprisingly

[14]A full list of items is available upon request.

[15]In a series of articles, Skinner (1964) describes the role of periodic markets in rural China where itinerant traders bring goods to villages. This is a similar setting and as in China Rozelle et al. (2003), fixed stores co-exist with periodic markets.

[16]These are:
i) Roads providing access to areas with economic development potential (20%)
ii) Roads leading to areas with surplus food and cash crop production (20%)

target potentially more productive and populated regions. However, the targeting is effectively in two stages: the first, at the regional level and the second, within regions, at the district level - vitiating any concerns about targeting at the level of the village. Shiferaw et al. (2012) (p. 11) discuss these criteria and argue that while they suggest a degree of targeting to the district, it is not borne out in the data. They conclude that "Regressing our road infrastructure variables on district-level *(or woreda)* control variables, we find that most of the variation in road accessibility is captured by the year dummies and the region fixed effects." Our data yield the same conclusion as Table 3 demonstrates [17]. In this table, we examine the relationship between the type of village road, whether a surfaced road (tarmac or stone) or a dirt track and village-level variables that might capture placement, such as population density (measured by te number of households per km^2), median income[18], the share of households in the village that receive support from the government's safety net programme and the number of nearby hamlets[19] that use the local periodic market, while controlling for district-level fixed effects. It is striking that for 151 villages, or half the sample, controlling for the district fixed effect completely explains the type of road in the village. Thus, over half the villages have their quality of road entirely unrelated to village-level variables, while for the remaining 134[20] villages, there is no evidence of a relationship between plausible village level variables (such as size and incomes) and the type of road. In the analysis that follows, we thus control for district fixed effects and base our identification of the relationship between variety available at the village level and travel time and local demand, entirely on the variation between villages within a district.

We present two sets of descriptive statistics. Table 4 describes households' perceptions of the effects of travel costs on consumer goods variety. We show this separately for households in the bottom and top of the village income distribution. The households were identified thus by

iii) Roads that link existing major roads (20%)

iv) Roads providing access to large and isolated population centers (30%)

v) Roads that bring balanced development amongst the regions in the country and that provide access to emerging regions (10%).

[17] We present a basic set of covariates here, in line with the variables used later to examine the variety of goods available in villages relative to their nearest market town. Variations on this specification do not affect the conclusion that these variables do not explain the quality of village roads and thus there is little evidence that they are explicitly â€œplacedâ€ï¿½ within villages.

[18] Our measure of village median income is obtained from the interviews with the six housesholds, three of which were deemed poor, while the other three were identified as rich by village officials. We also have measures of total grain and other production by village which proved far noisier.

[19] A hamlet refers to a small number of households who do not live in the main village. A hamlet is connected to the village by a walking path or track. The only exit to the road to town is through the village. The average village has 7 hamlets in its neighbourhood.

[20] The total sample size used is 285 because there was missing information on eleven villages for some village level variables needed to measure population density and village incomes.

village officials and it is useful to note that their incomes and assets tally with the description[21]. Poor and rich households differ sharply in their perceptions of the constraints posed by travel costs and distance. Just over half the poor but over 80% of rich households faced constraints in the set of manufactured goods available locally. Similar percentages said they would produce more for market and would travel more frequently if travel were easier. However, both groups were equally inclined to say that the rationale for travelling to town was to both sell produce and buy goods - even if the transactions of the rich are likely to be larger than the poor, as the first set of answers suggest.

What are the items that do not make the last mile into the village? Here there is consensus across households and this is reassuring because it describes a general lack of choice in particular item categories. Clearly, both rich and poor households face a lack of variety in similar item categories even if the rich are expected to spend more within them. The main sets of goods with items unavailable locally are clothing and linens followed by processed foods.

Table 5 describes the main characteristics of the villages in the survey. We begin with a discussion of the extensive margin of variety available: on average, villages have about half the items available in town and within these items, a third of all brands available in town. Seventy percent of towns are connected by a road made of tarmac or of stone to town (labelled a "good road"), while the remainder are served by a track or worse. The average distance to town is about 22 kilometres while the average time to travel to town by the most common form of transport is about an hour and 45 minutes. If the village hosts a periodic market, about 7 local hamlets are likely to participate; this is also our proxy for market size in the regressions below[22]. The villages host about 65 households per square kilometre on average, which is higher than the Ethiopian average since it does not include pastoralist areas, but is about average for East Africa. We also use a proxy measure of whether the village is relatively poor compared to other villages in the sample by checking whether the village is covered by the Productive Safety Net Program (PSNP). This program is targeted to the poorer villages within poor districts [23]. If the village is part of the PSNP, we checked village records to discover how many households in the village receive such support . Forty-two percent of the sampled

[21] We obtained data on monthly expenditures, agricultural and off-farm incomes and values of livestock. Uniformly, for poor households these values are at most a third of those of the rich households within villages. Note that the distribution of income across villages varies substantially.

[22] The market size is at least 1 by construction since we count the main village as participating in the periodic market. Only 4% of villages have no surrounding hamlets. Alternative measures for market size include the population and village officials' estimate of the average number of shoppers at the periodic market.

[23] Forty percent of the country's 710 districts (*woredas*) are covered by the PSNP and the programme supports about 8 million people or 10% of the population (IFPRI (2013)).

villages are in the government's safety net: furthermore in half of these villages, over a third of households are covered by the programme. The share of households in each village who are in the PSNP is thus used as proxy for capturing the level of poverty in a village, which in turn allows us to compare relative poverty across villages. Finally, in terms of amenities in the village, most villages have a health centre and a primary school but access to pharmacies and secondary schools is low. About a quarter of the sample have reliable electricity (which means a reliable supply for 3 days a week or more) and 40% report a reliable cellphone connection in the same vein [24].

4 Results

We now turn to the empirical analysis. We first investigate to what extent these factors such as transportation costs, market size and income distribution affect the distribution of variety across space. We then use data on prices both from our survey for small market towns and the villages and separately from the National Price Surveys that focus on the 118 larger market towns and urban centres across Ethiopia, to estimate the taste for variety captured by $1 - \theta$. Finally, we use these estimates to examine the implications of remoteness on the welfare costs of falling variety. We describe these costs across space in our sample and also for the median urban consumer if faced by less variety, akin to being moved to a remote location.

To measure the extent of variety in manufactured consumer goods available we use a simple count of all items and brands available locally in village markets, across 10 different categories of items. Figure 2 presents a picture of the fall in availability of items across space, while Figure 3 displays the share of items (and brands) available relative to the nearest market town and demonstrates the decline in average shares as travel times to town increase. Clearly, travel time alone does not determine variety: we would expect that population and market size, median incomes and other village attributes might affect this as well. To understand these effects, we turn to our empirical specification.

We model the probability that a village i with characteristics vector, X_i, (which includes transport times to the nearest market town, local market size, income distribution captured as a mean-preserving spread, and village-level amenities), has available exactly Y_i items. We

[24] This may seem at odds with the reliability of the electricity connection; the mobile telephones are often deposited with a shopkeeper in town to be charged, or with the local shop if they have better access to electricity since households have poor access in general. The median access to electricity (and cellphones) is 0.

use a count data specification, with a generalised negative binomial (GNB) model below. The advantage of the GNB specification is that, in contrast to the Poisson, it allows for the variance to be different from the mean, thus accounting for the overdispersion[25] in the data which is captured here through the introduction of regional fixed effects.

The main specification[26] includes the time to travel to town using the most common form of transport[27] and a measure of market size, proxied by the number of more remote hamlets that shop in the local sampled village[28] or the periodic market located there.[29] As an additional measure of market size we use (household) population density: the number of households per square kilometre. Finally, it is clear that some measure of local incomes is required. As explained above, we also surveyed six households in each village, where three households were drawn from the bottom of the distribution and the remainder from the top, as identified privately by local village officials. The households were interviewed on their average monthly consumption on manufactured goods, their income from both agriculture and other sources and their own assessment of their income, within three categories of rich, comfortable and poor. The measure of incomes and expenditures is noisy, given the light nature of the questionnaire and the usual difficulties of obtaining reliable estimates in one short interview. Instead of relying on any single measure, we construct a weighted average of all these measures, as a standardised normal variate, to proxy the distribution of income within villages. We use the bottom 20% and the top 20% (effectively, the income of the poorest and the richest households sampled), to capture the range of incomes or the mean-preserving spread, with the mean being standardised at zero[30]. However, the distribution of incomes thus measured only captures variation within a village, and the proportions of poor and rich are likely to differ across villages. To anchor this income distribution, we use the share of households that receive support from the government's Public Safety Net Programme, which is only targeted to villages deemed poor.

[25]We conducted tests for over dispersion, contrasting the Generalised Negative Binomial with the Poisson and the Negative Binomial distribution. The goodness-of-fit tests strongly reject the hypothesis of constant variances: Deviance goodness-of-fit =1728, with Prob $> \chi^2(201)= 0.00$ and the Pearson goodness-of-fit = 1759.24 with Prob $> \chi^2(201) = 0.00$. Introducing non-linearities does not affect the outcome and the Generalised Negative Binomial seemed the most appropriate alternative.

[26]The number of items/brands available in the nearest market town is treated as the exposure variable, since the counts of availability are better understood as a fraction of the availability in town.

[27]An alternative is to use the distance between village and town, together with controls for quality of road, by season. The results do not differ if these were included instead.

[28]Rozelle et al. (2003) describe similar periodic markets in China.

[29]An alternative measure is the community officials' estimate of the averge number of people who shop at the local periodic market. These are strongly correlated but the number of hamlets is the less noisy measure and thus preferred.

[30]We could also use estimates of village-level incomes as estimated by village officials. These were far noisier (even if correlated with the estimates obtained from the households), hence our reliance on the measures based on the survey of the households.

Table 6 presents the first set of estimates, using the total number of items available in the first 3 columns and then brands (across all item categories) in columns 4-6 as the dependent variable. Availability is strongly affected by travel times and the size of the local market. We report the coefficients in the table: the coefficients on the (ln) variables can be read as elasticities[31]: a 1% increase in travel time, results in a fall of 0.07 - 0.09% in item varieties available locally. Thus, for an average village, a fall of an hour in travel time is associated with an increase of about 9 items or 18 brands on average. Market size matters too: we measure market size as the number of hamlets (comprising the village and any smaller clusters situated further down the track) using the local periodic market and so an extra hamlet added to the size of the local market (a 10% increase) would be associated with 12 extra items or 18 extra brands being available. Columns 2 and 3 (and 5-6) include measures of the measure of inequality (mean-preserving spread) and amenities (captured by reliable electricity and cellphone access). Inequality also plays a significant role here,and an increase in the spread from the median to the highest decile would raise availability by 3 items or 5 brands; while the introduction of reliable access to amenities raises availability by 7 items or 9 brands. Note that apart from transport costs and market size, we are describing an association here rather than causal effects. It is more useful to examine the impact on the coefficients of these key variables when controlling for amenities and it is reassuring that those effects remain similar in size and significance.

Figure 4 demonstrates the heterogeneity of the relationship between variety available across space and the types of items. Heavier, bulkier goods are more likely not to make the last mile: processed foods (which include bottled drinks here), clothing and footwear exhibit sharp declines within 30 minutes of travel time from the nearest market town. The results in Table 5 hide this heterogeneity and disaggregating this relationship by item groups suggests much sharper effects of remoteness for some groups compared to others.

4.1 Prices, Markups and Remoteness

We now turn to examining the difference in prices between town and village. The data we gathered were meant to capture the extensive margin, i.e. whether a good is available in a market or not. As a result, the data have limitations in terms of its use for analysing the impact of distance on prices. First, given our focus on availability, we include a large number of generic brands - for instance, as long as a pair of jeans branded as "made in China" is available,

[31] We use the ln(travel time) given the concave relationship between availability and travel time.

and sold as different from another pair of jeans, it is treated as a different "item". However, this is a generic brand in that we cannot distinguish it as a specific brand like Levi's or Gap. Second, even when the brands are clearly defined (for instance, as a particular brand of flour), the unit of measurement used at point of sale could vary across areas, making comparison between origin and destination difficult in terms of price. We deal with these difficulties in two ways. First, we examine the difference in prices between source town and village for branded items that can be compared in identical weight or volume units. We use 1380 observations on such matched price-brand pairs. This includes largely processed foods (sugar, pasta), beverages (bottled drinks and beer), ready-made clothing of particular brands, cosmetic and hygiene (soaps, detergents, toilet paper), household items (matches, batteries) and educational items (branded notebooks and pens). Second, we carried out a separate second survey of prices alone in half the sample (115 villages and 43 market towns) for specific item-brand pairs obtained from this survey to control for any measurement error in the first survey. This is also a more restricted set of items that does not include clothing or footwear. Note that restricting the matched item-brand pairs across villages in the first survey results in a smaller set of observations because in many cases units were missing or were different. We report the regression of prices paid in the village on prices paid in town and travel time in Table 7 using both sets of data. The regression examines the variation of prices across space for these matched pairs and also offers an estimate of the taste for variety, $1 - \theta$, as in Equation 6, which gives us two estimates, ranging from 0.09 to 0.13, which in turn give us estimates of a markup between 10.3% and 14.6%, where the markup is defined as $\frac{1}{\theta}$. In the next section, we use these estimates to construct the welfare costs of the loss of variety with remoteness. As we will see, these costs are very sensitive to the estimated taste for variety and the consequent estimates of elasticities. Our estimates of θ are obtained from the variation across brands of similar items and hence we would expect the taste for variety to be lower and the consequent elasticity of substitution estimated here to be higher than across goods of different types. Our estimates thus provide an upper bound on the elasticity of substitution. This implies that our estimated welfare costs are lower bounds on the costs of losing variety across space.

The estimate of the taste for variety, θ, we obtain is similar whether we use the data on price differentials over relatively short distances, as in our surveys as reported above or between source town and 118 market towns, in the National Price Surveys collected by the Central Statistical Authority. We use their data on prices of a set of 13 branded products collected monthly since 2010, where the 13 products are chosen such that we can map the source and destination prices accurately. We also obtain the travel time in minutes between the source and destination towns, using data from IFPRI. for 2010 and relying on the fact that until 2015, there was little new investment in these roads. The last column of Table 7 provides the estimate of the taste for variety using these data and this provides similar estimate as above, of a taste for variety of 0.12 and a markup of 14.4%.

The fact that we obtain similar estimates of the taste for variety over both remote locations and their small market towns, as well as the larger urban centres and their source towns (like the capital Addis Ababa), suggests that the taste for variety is similar across space. But the price wedge also depends on transport costs: the estimates in Table 7 suggest that the marginal transport cost (the estimate of β) is far lower between say, Addis Ababa and an urban centre than it is between a small market town and a remote village. The rapid investment in regional road infrastructure over the past twenty years has paid off for the urban consumer.

5 Welfare Costs of Remoteness

To estimate the welfare cost of remoteness we use a simplified two-level CES utility function. We depart from the simpler, one-level framework in order to capture the description offered in Figure 4, where we observed that the fall in variety varies substantially by item categories. Unsurpisingly, bulkier and heavier items are alos more likely to vanish faster across space, where the costs of transportation bite strongly.

$$u = \left(\sum_{i=1}^{n} (\widetilde{c}_i)^{\theta} \right)^{1/\theta} \tag{13}$$

$$\widetilde{c}_i = \left(\sum_{j=1}^{n_i} (c_{ij})^{\theta_i} \right)^{1/\theta_i} \tag{14}$$

where c_{ij} denotes consumption of item j within group i, n_i the number of items of group i that are available, n the number of groups that are available, $0 < \theta < 1, \sigma = 1/(1-\theta), 0 < \theta_i < 1$ and $\sigma_i = 1/(1-\theta_i)$. Note that the v coefficients have disappeared; this makes the utility function homothetic.[32] The indirect utility function for this two-level CES function is:

[32]This assumption makes the welfare analysis tractable but is obviously restrictive. Our aim is to get a sense of the order of magnitude of the welfare effects and by allowing for different elasticities of substitution, we are

$$u = y\left[\sum_{i=1}^{n}\left[\left[\sum_{j=1}^{n_i} p_{ij}^{\frac{\theta_i}{\theta_i-1}}\right]^{\frac{\theta_i-1}{\theta_i}}\right]^{\frac{\theta}{\theta-1}}\right]^{\frac{1-\theta}{\theta}} = y\widetilde{u}(p_{ij}, n, n_i)$$

To calculate the availability effect we solve

$$y\widetilde{u}(p_{ij}, n, n_i) = y\xi\widetilde{u}(p_{ij}, n, n_i') \tag{15}$$

for ξ. Here prices p are market town prices. The left hand side of the equation measures welfare of someone with income y facing prices and availability in the market town, given by p, n and n_i. The right hand side indicates that when availability changes to n_i' income would have to change to $y\xi$ to keep welfare unchanged.[33] Note that y drops out (as a result of the homotheticity assumption) and that once ξ has been calculated the compensating variation (relative to income) follows as $\xi - 1$.

Similarly, we can estimate the price effect by solving

$$\widetilde{u}(p_{ij}, n, n_i) = \xi\widetilde{u}(p_{ij}', n, n_i) \tag{16}$$

for ξ. This gives the compensating variation for the case when availability is the same as in the market town but prices change, as a result of distance from the market town, from p to p'. We use the markup formula (see Equation 6) to calculate these p' prices.

The two calculations allow us to compare the magnitude of the two welfare effects: the price effect that is typically taken into account in empirical work and the variety or availability effect that is typically ignored. Figures 5 and 6 allow the comparison of these effects across space. Figure 5 displays the welfare loss due to the loss in variety across the time to travel to the nearest town, while Figure 6 displays the welfare loss due to the increases in prices across travel time, using equations 14 and 15 above to calculate these effects. The average welfare loss due to the fall in variety is on average 19% of incomes (rising to a maximum of 49%), while that due to the price effect is much lower, at an average of 1.75% (rising to a maximum to 8.3%), .The figures also drive home the fact that the losses increase with travel time to the nearest town, an effect sharper for the loss due to variety again with an average increase of 4% in the

forced to abandon the issues raised by inequality in incomes. This is perhaps less vital in the Ethiopian setting where inequality is low, given the equitable distribution of the key asset, land.

[33] Note that we change n_i but not n: we assume that while availability changes within groups all groups remain available.

welfare loss in variety for a 1% increase in travel time in contrast to the loss due to prices at 0.36% for a 1% increase in travel time.

6 Conclusion

We have described a model where consumers are affected by transport costs, not only through changes in prices, but also through changes in the set of available goods. In this model a reduction in transport costs increases consumer welfare in three ways. First, it increases the value of income from sale of crops. Secondly, it lowers the price of consumer goods. Finally, it increases the number of varieties available at a given distance form a market town. We have extended this framework to take account of heterogeneity in the size of markets and in the distribution of income. This model was tested using data collected in a purpose-designed survey of villages in Ethiopia, each served by a single larger market at the district level. We find that variety declines sharply with transport costs. It is also strongly affected by the size of the local market and the distribution of income.

While the new economic geography has led to a large number of empirical studies on the effect of international trade on variety, there is comparatively little empirical work on the effect of domestic trade on the number of varieties available to consumers and only so for developed countries. This is likely to be a much more important issue in developing countries where transport costs are very much higher. This is one of the few papers that addresses the loss in variety in consumption with remoteness: consumer choice fades away with distance in developing countries. The magnitude of the effects we found in Ethiopia suggests that when comparing incomes across space within countries it is not sufficient to correct for price differentials, keeping the basket of goods fixed. That ignores the availability effect which according to our results dramatically affects the options open to consumers in distant locations. Distance reduces welfare not only by inducing a deterioration of the terms of trade of a village trading with a market town but also reduces the number of goods that will be available to consumers in the village. The size of the welfare effects we find depend both on domestic trade costs within countries and the spending on consumer goods - for instance, the Nigerian case (see Atkin and Donaldson (2015)) where both intra-trade costs and consumer spending are far higher offers a sharp contrast to Ethiopia. We describe a spatial equilibrium where variety falls across space despite households demonstrating a similar taste for variety in goods. The framework assumes that households are fixed in space: given the restrictions on land ownership and the difficulties of migration in Ethiopia, this seems consonant with the context.

What do these results tell us? First, they demonstrate that there are significant welfare losses to low variety in manufactures and that intra-trade costs (both mark-ups and transport costs) matter in affecting this margin. Second, the average costs to consumers of about 19% of consumer expenditures on manufactures suggests that the costs of losing items across space is not small, despite the fact that Ethiopian consumers are poor and their share of spending on manufactures is low. These costs will increase with rising incomes unless domestic trade costs fall as well. It should also be emphasised that the items lost are usually part of a basic set of necessary consumption items, even in a poor setting, comprising basic clothing, footwear, hygiene, kitchen and housewares and even educational items.

We have shown how variety in manufactures fades with distance and associated transport costs. Reductions in transport costs increase consumer welfare through higher incomes, lower prices - and increased variety. This also has implications for welfare measurement: poverty is underestimated since people in remote places have little to choose – but equally, when changes such as infrastructure investments raise availability, rates of decline in poverty will be underestimated too.

References

C. Adam, D. Bevan and D. Gollin, 2016, ‘Rural-Urban Linkages, Transport Costs and Public Investment: The Case of Tanzania’, forthcoming in *World Development.*

D. Atkin and D. Donaldson, 2015, ‘Who’s Getting Globalized? The Size and Implications of Intranational Trade Costs’, NBER Working Paper No. 21439, National Bureau of Economic Research.

T. Arethun and B. P. Bhatta, 2009, ‘Contribution of Rural Roads to Access and Participation in Markets: Theory and Results from Northern Ethiopia’, *Journal of Transportation Technologies*, 2(1), pp. 65-174.

A. Banerjee, E. Duflo, and N. Qian, 2012, ‘On the Road: Access to Transportation Infrastructure and Economic Growth in China’, NBER Working Paper 17897, National Bureau of Economic Research.

Benassy, J-P, 1996. ”Taste for variety and optimum production patterns in monopolistic competition,” Economics Letters, Elsevier, vol. 52(1), pages 41-47, July.

A. B. Bernard, J. B. Jensen, S. J. Redding, and P. K. Schott, 2007, ‘Firms in International Trade,’ *Journal of Economic Perspectives*, 21(3), pp. 105-130.

C. Broda and D. E. Weinstein, 2006, ‘Globalization and the Gains from Variety’, *Quarterly Journal of Economics*, 121(2), pp. 541-585.

A. Burstein and G. Gopinath, 2013, ‘International Prices and Exchange Rates’, forthcoming in the *Handbook of International Economics.*

P. Buys, U. Deichmann, and D. Wheeler, 2010, ‘Road Network Upgrading and Overland Trade Expansion in sub-Saharan Africa,’ *Journal of African Economies*, 19(3), pp. 399–432.

L. Casaburi, R. Glennerster and T. Suri, 2013, ‘Rural Roads and Intermediated Trade: Regression Discontinuity Evidence from Sierra Leone’, *IGC Working Paper*, London School of Economics.

T. Chaney, 2008, ‘Distorted Gravity: The Intensive and Extensive Margins of International Trade’, *American Economic Review*, 98(4), pp.1707-1721.

V. Couture, B. Faber, Y. Gu, and L. Liu, 2018, 'E-Commerce Integration and Economic Development: Evidence from China', *NBER Working Paper* 24384.

D. Donaldson, 2016, 'Railroads of the Raj: Estimating the Impact of Transportation Infrastructure', forthcoming in the *American Economic Review.*

S. Dercon, D. O. Gilligan, J. Hoddinott and T. Woldehanna, 2009, 'The Impact of Agricultural Extension and Roads on Poverty and Consumption Growth in Fifteen Ethiopian Villages', *American Journal of Agricultural Economics,* 91(4), pp. 1007-1021.

A. K. Dixit and J. E. Stiglitz, 1977, 'Monopolistic Competition and Optimum Product Diversity', *American Economic Review*, 67(3), pp. 297-308.

J. Eaton, S. Kortum, and F. Kramarz, 2011, 'An Anatomy of International Trade: Evidence from French firms', *Econometrica*, 79(5), pp. 1453-1498.

M. S. Emran and F. Shilpi, 2012, 'The Extent of the Market and Stages of Agricultural Specialization', *Canadian Journal of Economics*, 45(3), pp. 1125-1153.

The Ethiopian Road Authority, Manual for Formulation and Implementation Framework, various years, at: http://www.era.gov.et/.

P. L. Fackler and B.R. Goodwin, 2001, "Spatial Price Analysis", Chapter 17 in: Handbook of Agricultural Economics, Volume 1, Part B, Pages 971-1024

J. Fan, L. Tang, W. Zhu and B. Zou, 2016, "The Alibaba Effect: Spatial Consumption Inequality and the Welfare Gains from e-Commerce". Available at: http://dx.doi.org/10.2139/ssrn.2707293

E. Glaeser, J. Kolko and A. Saiz, 2001, 'Consumer City', *Journal of Economic Geography*, 1(1), pp. 27-50.

M. Gonzalez-Navarro and C. Quintana-Domeque, 2016, 'Paving Streets for the Poor: Experimental Analysis of Infrastructure Effects', forthcoming in the *Review of Economics and Statistics.*

D. Gollin and R. Rogerson, 2010, 'Agriculture, Roads, and Economic Development in Uganda', NBER Working Papers 15863, National Bureau of Economic Research.

D. Gollin, M. Kirchberger and D. Lagakos, 2017, "In Search of a Spatial Equilibrium in the Developing World", NBER Working Paper 23916.

J. Handbury, 2012, 'Are Poor Cities Cheap for Everyone? Non-Homotheticity and the Cost of Living Across U.S. Cities', mimeo.

J. Handbury and D. E. Weinstein, 2014, 'Goods Prices and Availability in Cities', *Review of Economic Studies,* 82(1), pp. 258-296.

R. Hillberry and D. Hummels, 2008, 'Trade Responses to Geographic Frictions: a Decomposition Using Micro-Data', *European Economic Review,* 52(3), pp. 527-550.

D. Hummels, and P. J. Klenow 2005, 'The Variety and Quality of a Nation's Exports', *American Economic Review*, 95(3), 704–723.

D. Hummels, 2007, 'Transportation Costs and International Trade in the Second Era of Globalization', *Journal of Economic Perspectives*, 21(3), pp. 131-154.

H. G. Jacoby, 2000, 'Access to Markets and the Benefits of Rural Roads', *Economic Journal*, 110(465), pp. 713-737.

H. G. Jacoby and B. Minten, 2009, 'On Measuring the Benefits of Lower Transport Costs', *Journal of Development Economics*, 89(1), pp. 28-38.

S. R. Khandker, S. Bakht and G. B. Koolwal, 2009, 'The Poverty Impact of Rural Roads: Evidence from Bangladesh', *Economic Development and Cultural Change,* 57(4), pp. 685-722.

W.A. Lewis, 1954, 'Economic Development with Unlimited Supplies of Labour', *The Manchester School,* 22, pp. 139–91.

N. Li, 2011, 'An Engel Curve for Variety,' mimeo.

P. Krugman, 1991, 'Increasing Returns and Economic Geography,' *Journal of Political Economy*, 99(3), pp. 483-499.

P. R. Krugman, 1995, *Development, Geography, and Economic Theory,* Cambridge, MA.: MIT Press.

B. Minten and S. Kyle, 1999, 'The Effect of Distance and Road Quality on Food Prices, Marketing Margins, and Traders' Wages: Evidence from Zaire', *Journal of Development Economics*, 60(2), pp. 467-495.

B. Minten, B. Koru and D. Stifel, 2013, 'The Last Mile(s) in Modern Input Distribution: Pricing, Profitability, and Adoption,' *Agricultural Economics*, 44(6), pp. 629-646.

K. M. Murphy, A. Shleifer and R. Vishny, 1989, 'Income Distribution, Market Size, and Industrialization', *Quarterly Journal of Economics*, 104(3), pp. 537-564.

A. Shiferaw, M. Söderbom, E. Siba and G. Alemu, 2012, 'Road Networks and Enterprise Performance in Ethiopia: Evidence from the Road Sector Development Program', Working Paper, International Growth Centre, London, U.K.

S. J. Redding and D. M. Sturm, 2008, 'The Costs of Remoteness: Evidence from German Division and Reunification', *American Economic Review*, 98(5), pp. 1766-1797.

S. Rozelle, J. Huang and V. Benziger, 2003, "Continuity and Change in China's Rural Periodic Markets", *The China Journal*, 49 pp. 89-115

W. Skinner, 1964 "Marketing and Social Structure in Rural China", *Journal of Asian Studies* No. 24, pp. 3-42; 195-228; 363-399

D. Stifel and B. Minten, 2008, 'Isolation and Agricultural Productivity,' *Agricultural Economics*, 39(1), pp. 1-15.

A. Storeygard, 2016, 'Farther on down the Road: Transport Costs, Trade and Urban Growth', *Review of Economic Studies,* 83(3), pp. 1263-1295.

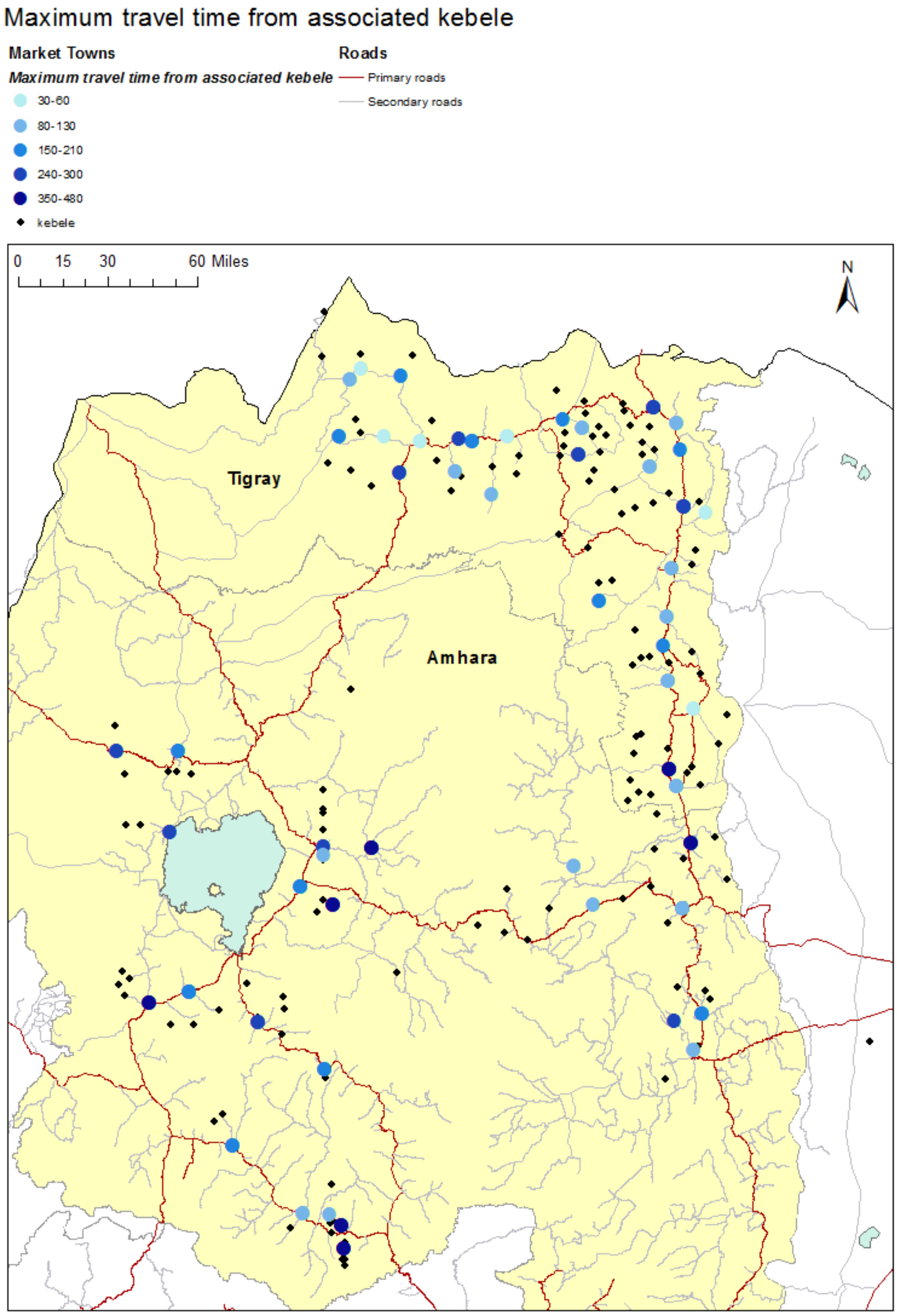
E. Van Leemput, 2016, 'A Passage to India: Quantifying Internal and External Barriers to Trade ', Working Paper.

Table 1 Why do traders trade in particular items?

Reasons for not trading in extra items	%
License regulations	38
Capital constraints	33
Low demand	20
Transport costs	5
Other miscellaneous	4

Notes: Data are from traders interviewed in 150 villages. They were asked how many items they carry and why they do not carry any other items, even within the same category such as processed foods or clothing.

Figure 1 Map of sampled villages and associated market towns for two regions



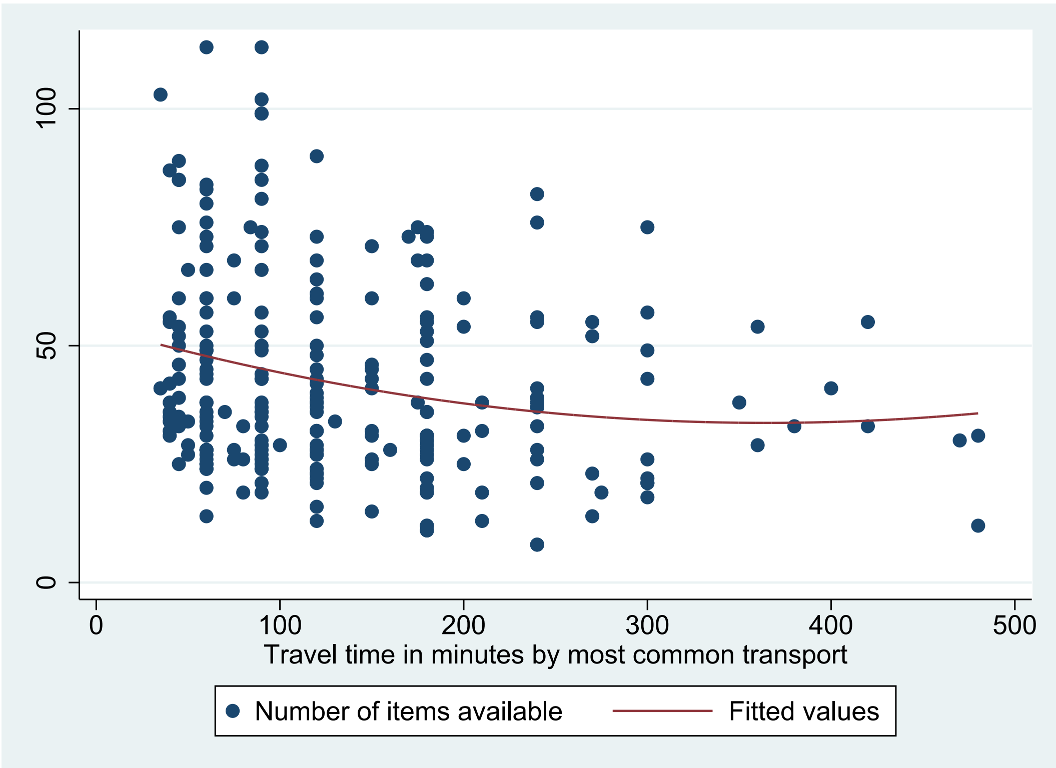
Notes: The map represents locations of market towns and the associated villages by distance in two of the four surveyed regions, the Amhara and Tigray regions in Ethiopia.

Table 2 The sampling frame for villages by connections to town

Number of villages	Mean	SD	Max
In district	10.1	6.7	42
Directly connected to one town	7.4	5.7	30
Directly connected to more than one town	2.6	3.2	17
% connected to one town	73	28.7	100

Notes:These summary statistics are based on the first round survey conducted to list all the market towns and villages linked by one road

Figure 2 Fading choice: Number of items available by travel time to nearest town



Notes:The figure shows a scatter plot and a quadratic fit of the total count of items available in each village relative to the time taken to travel (in minutes) to the nearest market town.

Table 3 Road placement: Logit estimates of type of village road and correlates

	Type of Road (Base = Track)
	Tarmac or Stone Road
Coefficients	
Share of poor households	-0.621
	(1.069)
Median income	0.061
	(2.010)
Household density	0.004
	(0.004)
Market size (ln)	0.518
	(0.382)
Constant	-0.661
	(1.587)
District Fixed Effects	Yes
Observations	134
Pseudo R^2	0.107

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table displays logit regression of the odds of a good road of tarmac or stone relative to the base of a track. District fixed effects determine road type in 151 villages out of 292 sampled villages, while for the remainder, basic village-level characteristics are uncorrelated with road type.

Table 4 Household Characteristics by Income Group

	Poor	Rich
% saying yes to whether?		
Difficulty obtaining consumer goods in last year	55	83
Would produce more for market if travel easier	57	86
Would travel more frequently if travel easier	66	87
Main reasons to travel to town (%)		
Buy inputs and sell produce	10	10
Buy consumer goods	15	15
Both	75	75
Main items reported unavailable locally (item group) (%)		
Clothing and linens	48	44
Processed foods	30	34
Leather goods, footwear (plastic)	12	14
Household goods, kitchenware	8	8
Household characteristics		
Mean household size	5.7	7.9
Annual median income (birr)	5500	40,000
Male household head %	80	94
Number of households	906	908

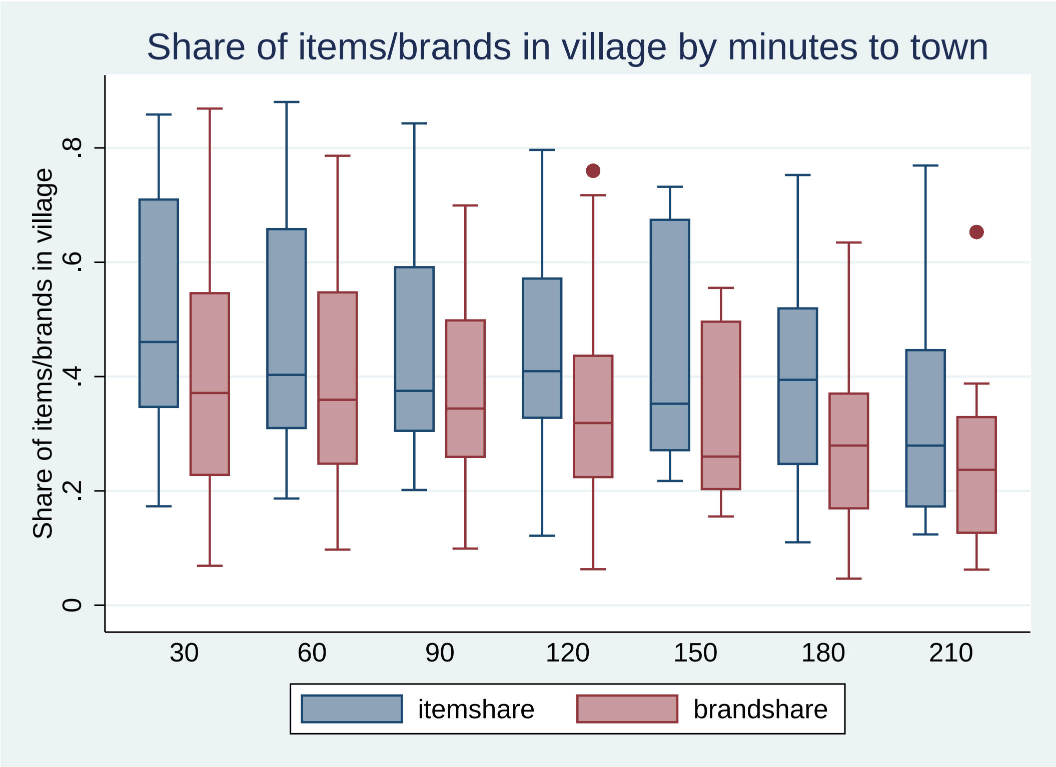
Notes: The data here come from the survey of six households in each village, chosen so that three were deemed to be representative of poor households in the villages and the other three deemed relatively rich. This was in consultation with village-level officials.

Table 5 Summary of Village Characteristics

Variable	Mean	Standard Deviation
Items available in village	45.00	22.00
Items available in nearest town	97.00	20.30
Brands of items available in village	73.3	39.92
Brands of items available in town	206.64	74.26
Distance to town in kms	22.94	12.43
Travel time in minutes by most common transport	117.24	93.37
Whether Tarmac/Stone road	.69	-
Number of villages using local market	6.84	5.34
Household density	65.32	68.47
Share of poor households	.13	.21
Reliable electricity	.26	-
Reliable cellphone	.40	-

Note:We do not report standard deviations for binary variables.

Figure 3 Share of items and brands in village relative to town by travel time



Notes: The figure shows the number of items (and brands of items where known) relative to the total number of items and the total number of brands of items available in the nearest market town by minutes of travel to the town.

Table 6 Generalised negative binomial estimates by items and brands in village

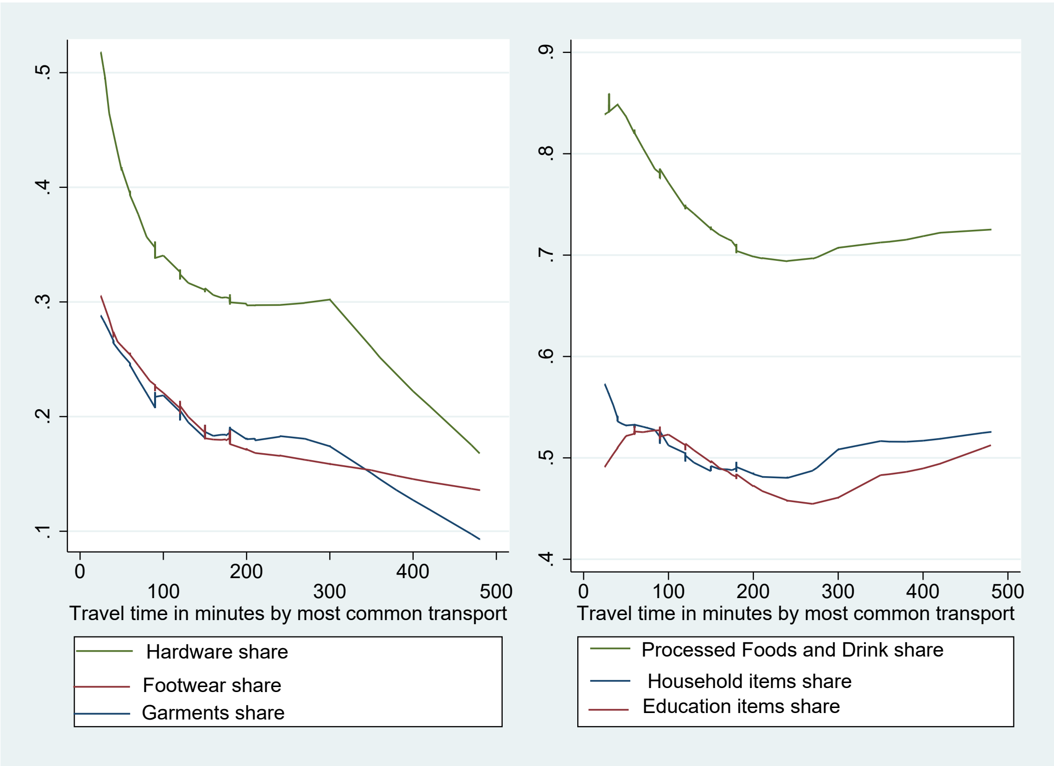
	Items available			Brands available		
	(1)	(2)	(3)	(4)	(5)	(6)
ln(Travel time to Town)	-0.092***	-0.092***	-0.069***	-0.117***	-0.119***	-0.093***
	(0.0239)	(0.024)	(0.026)	(0.027)	(0.029)	(0.028)
Household Density	-0.0001	0.0000	-0.0001	-0.0000	0.0001	0.0000
	(0.0003)	(0.0003)	(0.0003)	(0.0004)	(0.0004)	(0.0004)
ln(Market Size)	0.255***	0.255***	0.231***	0.249***	0.247***	0.223***
	(0.042)	(0.041)	(0.042)	(0.044)	(0.0448)	(0.047)
Share of Poor Households	-0.172	-0.205	-0.164	-0.142	-0.353	-0.325
	(0.164)	(0.192)	(0.182)	(0.209)	(0.261)	(0.248)
Income Spread		0.233***	0.218***		0.193**	0.165**
		(0.067)	(0.066)		(0.089)	(0.0840)
Income Spread*Share of Poor		0.112	0.0251		0.357	0.285
		(0.192)	(0.182)		(0.314)	(0.277)
Reliable electricity			0.146***			0.155**
			(0.057)			(0.062)
Reliable cellphone			0.0226			0.078
			(0.048)			(0.052)
Constant	-0.743***	-0.864***	-0.940***	-0.861***	-0.945***	-1.095***
	(0.162)	(0.167)	(0.172)	(0.299)	(0.304)	(0.296)
Variables for overdispersion						
Region=3	0.717*	0.726*	0.895**	0.484	0.487	0.651**
	(0.418)	(0.423)	(0.404)	(0.317)	(0.317)	(0.306)
Region=4	-0.421	-0.380	-0.358	-0.976***	-0.932***	-0.918***
	(0.462)	(0.451)	(0.463)	(0.340)	(0.328)	(0.315)
Region=6	-0.270	-0.386	-0.550	-0.436	-0.478	-0.605
	(0.475)	(0.515)	(0.571)	(0.413)	(0.435)	(0.471)
Constant	-3.115***	-3.160***	-3.257***	-2.383***	-2.407***	-2.49***
	(0.326)	(0.332)	(0.319)	(0.245)	(0.243)	(0.231)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	292	285	285	292	285	285
Pseudo R^2	0.121	0.125	0.129	0.113	0.115	0.119

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Market size is measured as the number of nearby hamlets (including the main village sampled) that use the periodic market and fixed shops in the village. The share of poor households is obtained from the number of households in the village who are part of the PSNP programme targeted at poor villages, while income spread measures the difference in incomes between the richest and poorest households in the sample, with the mean standardised to zero, a mean-preserving spread.

Figure 4 Availability across space by item group



Notes: The graph displays lowess graphs of shares of items available in the village relative to the associated market town, disaggregated by item group to examine heterogeneity in the fall in variety over travel time. Heavier, bulkier items such as processed food and drink and hardware display a sharper fall in availability by travel time.

Table 7 Prices between source town (P) and destination (P_j)

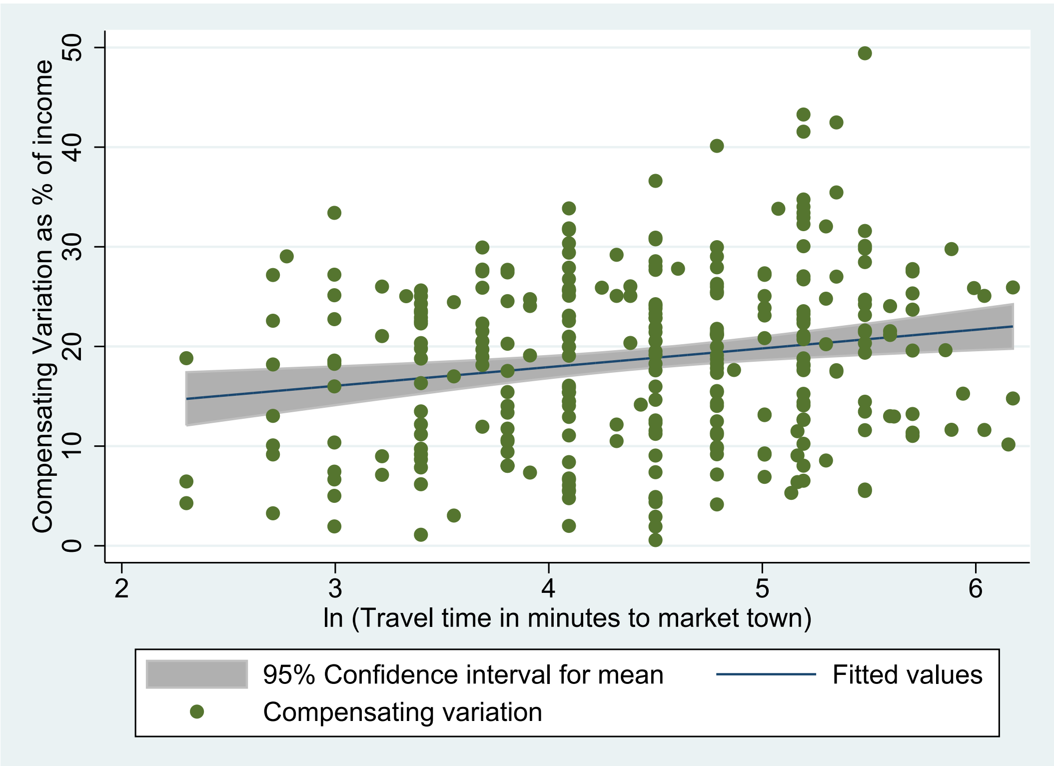
$P_j = \frac{P}{\theta} + \frac{\beta S_j}{\theta}$	Variety survey (2016) Prices in village	Price Survey (2016) Prices in village	National Price Survey (2010-16) Prices in 118 market towns
Source price (town)	1.146***	1.103***	1.144***
Coeff=$\frac{1}{\theta}$	(0.0213)	(0.009)	(0.007)
Travel time (minutes)	0.007***	0.005***	0.003***
	(0.002)	(0.0006)	(0.0007)
Year-Month fixed effect			Yes
Item fixed effect			Yes
Observations	1370	1971	32,212
Adjusted R^2	0.978	0.991	0.956

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

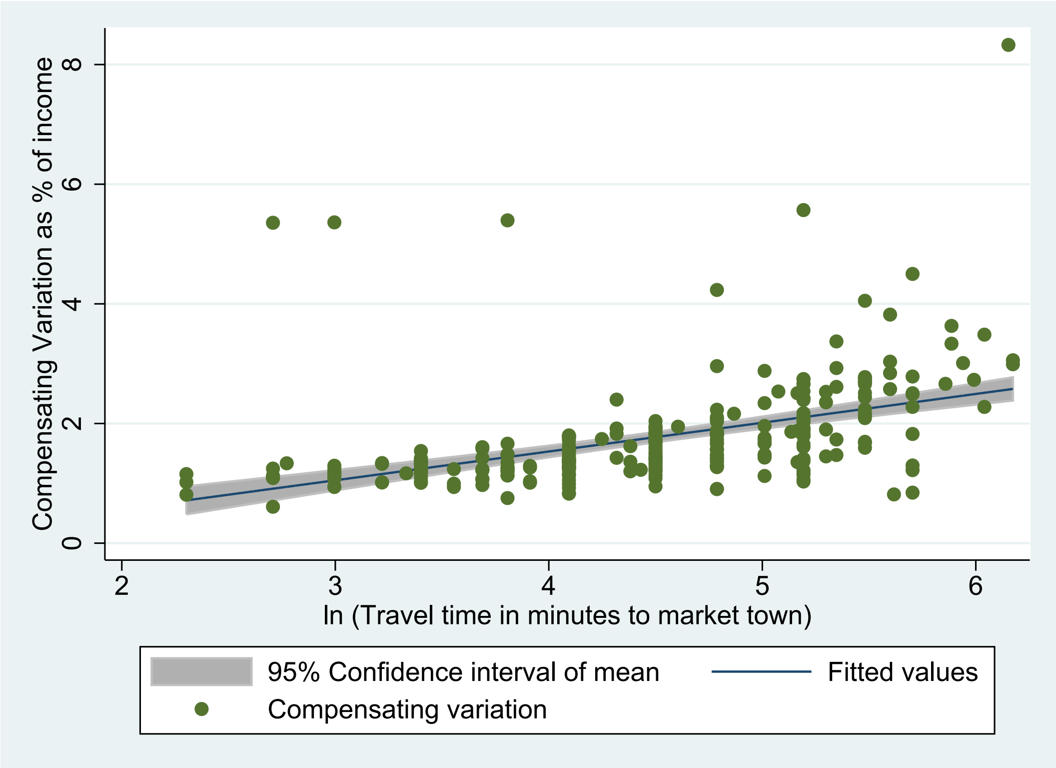
Notes: Column (2) uses the prices of branded items in village and associated towns in identical weight or volume units from the availability survey with data from 187 villages and 77 market towns. Column (3) uses similar prices on branded items from a follow-up survey on prices alone in 115 villages and 43 market towns as a consistency check. Column (4) uses data (on 13 items whose source town is established) from the National Price Survey conducted by the Central Statistical Authority in 118 market towns from 2010-2016. We use these years since the road quality (and thus travel time) between source town and market town can be assumed to be unchanged. Note that this last set of prices reflects national differences rather than within-district variation in as in Columns 2 and 3.

Figure 5 Welfare Costs of Remoteness : Compensating variation due to fall in variety



Notes: The graph above describes the loss in welfare due to the decline in variety with remoteness from town, valued using prices in town. (See equation 15)

Figure 6 Welfare Costs of Remoteness : Compensating variation due to higher prices in village



Notes: The graph above describes the loss in welfare due to higher prices in the village compared to market town for brands available in both town and village. (See equation 16)