INTERNAL GEOGRAPHY AND THE EFFECT OF IMPORTS ON NON-HOMOTHETIC CONSUMPTION PATTERNS WITHIN THE U.S.

Volodymyr Lugovskyy*
Alexandre Skiba†

February 2, 2018

Abstract

The rich households are known to purchase more expensive varieties of a given durable good than the poorer ones. In this paper, we show that there is a significant variation in these rich-to-poor price differences across U.S. states and that it is related to internal geography and access to imports across states. Our identification of exogenous variation in imports relies on differences in the internal distance of imports across states and products. We find that (1) the differences in purchase prices of durables between rich and poor households are more pronounced in states that receive more imports and (2) those differences are driven by the effect of imports on the poor, not the rich, households. Both findings are consistent with the predictions of a novel model that combines non-homothetic preferences with specific transportation costs.

---

[*] Department of Economics, Indiana University, Wylie Hall Rm 301, 100 S. Woodlawn, Bloomington, IN 47405-7104; e-mail: `vlugovsk@indiana.edu`

[†] Department of Economics and Finance, University of Wyoming, Dept 3985, 1000 E University Ave, Laramie, WY 82071, e-mail: `askiba@uwyo.edu`

1 Introduction

In the U.S., richer households buy more expensive varieties of durable goods than the poorer ones.[1] What is less known, is that these price differences vary significantly across U.S. states. In this paper, we highlight the extent of these previously unexplored differences and attempt to explain them using the data on consumer prices, state-level imports, and internal distance. Specifically, our findings suggest that better access to imports increases the extent of the difference in consumer prices of products purchased by the rich versus poor households. Understanding this relationship is important because the effects of international trade on the U.S. economy are heterogeneous across occupations, industries, and geographic locations. These heterogeneities play an increasingly important role in shaping the U.S. economy and politics.

In this paper, we analyze consumer prices from Consumer Expenditure (CE) Surveys between 2009 and 2012. CE Surveys are collected by the Census Bureau for the Bureau of Labor Statistics—the only comprehensive and detailed dataset on the households expenditures on durable goods (Attanasio & Pistaferri, 2016). We first confirm that, ranked by the expenditures on nondurable goods,[2] the top quintile (i.e., rich) households pay significantly higher prices for televisions, bicycles, refrigerators, and other durables than the bottom quintile (i.e., poor) households. What is unexpected, however, is by how much the relative price paid by rich (henceforth, RP-rich) for a given good varies across states. For example, in 2012, the difference in the average prices of televisions purchased by the richest and poorest quintile was only 50% in Florida while over 200% in Illinois.

Furthermore, we find the spatial differences in RP-rich are driven primarily by the variation in prices paid by the poor across states and that the internal distance

---

[1] See, for example, Bils & Klenow (2001).

[2] Ranking based on expenditures is in line with the literature, since expenditures on consumption are considered to be a better measure of well-being than income (see, for example, Goldberg & Pavcnik, 2007), since it allows to account for the inter-temporal shift of resources take into account changes in prices, and to alleviate the reporting problems. Importantly, in our dataset, expenditures on nondurables and income are highly correlated which allows us to label households with high expenditures on nondurable as rich and with low expenditures on nondurables as poor.

of trade is negatively correlated with the RP-rich. Subsequently we develop a theory to understand these patterns, and then propose an estimation approach that verifies that these patterns are borne out by the data under a more careful econometric investigation.

In order to guide our interpretation of the stylized facts and to motivate our empirical exercise, we construct a model in which consumers can adjust their quality choices based on their income and the magnitude of trade friction. The preferences are defined over one nondurable and many durable goods. Domestic supply of durables is insufficient, so that imports of each durable good is strictly positive. Each durable good is divided into two types, Low and High, whereas the utility curvature parameter of High is greater than that of Low, as in Fieler (2011). As shown by Fieler (2011), the difference in the curvature parameters, makes demand for High more income elastic than Low. We show that, for any quantity above one, High and Low types can also be interpreted as high-quality and low-quality varieties of a given durable, since, ceteris paribus, households derive greater utility from each unit of High than from each unit of Low.

On the supply side, we assume that higher-quality goods require higher marginal cost of production and that transportation costs are specific and equal in magnitude for both types. In equilibrium, we derive the conditions under which (i) transport costs increase demand for quality and—what sets us apart from the previous research—(ii) the demand for quality of poorer households is more sensitive to transportation costs that that of richer households.

To measure the systematic variation in the RP-rich across states, we estimate product and state specific quality Engel curves. Quality Engel curves, introduced by Bils & Klenow (2001), describe the relationship between the price of a purchased durable product and the household expenditures on nondurable goods.[3] Estimation of the quality Engel curves offers a more systematic picture of the RP-rich because estimation uses information on the entire income distribution and allows controlling for the demographic characteristics of the households.

---

[3] The latter is used to trace out differences in household total expenditure and therefore income as is commonly done in the literature

The novelty of our estimation is that we let the relationship between the household income and the purchase price of a given product depend on the access to imports of that product in the state of the household's residence. The product-level data on state imports comes from the US Census imports by the state of intended destination available at six-digit codes of the Harmonized System (HS) classification.

A key element of our estimation is identification of the exogenous variation in import access. In order to resolve likely endogeneity between imports and consumption, we use measures of internal distance of imports that reflect differences in internal geography. To construct those measures we use information on import routes from the Commodity Flow Survey (CFS) available through the Freight Analysis Framework (FAF). to calculate internal and external distance of the major import routes.[4] Our constructed measures reveal significant variation in the internal distance of imports across states and, more importantly, products. This is precisely the variation that allows us to trace out the exogenous variation in per capita imports that is determined by internal geography but not by other state and product specific characteristics (e.g., cost of living, heterogeneity in local product demand shocks, etc)

We find that the differences in the access to imports have a significant impact on the relationship between incomes and purchase prices. In the states with more imports, the relationship between the prices and incomes is stronger. That is, imports increase the slopes of the quality Engel curves. Interestingly, in the states with the lowest per capita imports, the Engel curves are nearly flat. The estimates also reveal that the slopes of the quality Engel curves vary across states and products due to the geographic differences in the purchase prices of the poor households.

This paper makes several distinct contributions. First, our findings most directly contribute to the literature investigating differential effects of trade on consumers with various incomes (Porto, 2006; Nicita, 2009; Broda & Romalis, 2009; Faber, 2014; Fajgelbaum & Khandelwal, 2016). We contribute by adding the geographic dimension and focusing on how the differences between the rich and the poor vary across states within the US. By virtue of the focus on differential trade effects,

---

[4] The identifying assumption is that the set of major routes is affected by geography but not by small differences in demand for the origin-specific varieties.

our findings tangentially relate to the broader literatures that focus on the within country heterogeneity of trade effects across industries (Autor *et al.* , 2013) , skills (Dix-Carneiro & Kovak, 2015), and income (Faber, 2014; Fajgelbaum & Khandelwal, 2016); as well as he emerging literature on the effects of income distribution on trade in vertically differentiated goods (Choi *et al.* , 2009; Fieler, 2011; Fajgelbaum *et al.* , 2011).

Second, the analysis of the geographic variation leads us to the novel conclusion that the geography matters primarily for the poor. This empirical finding is easy to rationalize by a theory that allows for differences in import access in a model with non-homothetic preferences. This combination of non-homothetic preferences with a model of imports has a more general implication because we find that the way non-homothetic preferences manifest themselves in consumption is affected by supply considerations, in particular the differential access to imports.

Third, our treatment of internal geography also contributes to the literature on the effects of internal geography on trade (Atkin & Donaldson, 2015; Coçar & Fajgelbaum, 2016) by introducing specific transportation costs to capture the effect of internal distance. This is interesting because even though prices of the identical varieties of imported goods do not vary much across the US regions, theoretically, even low internal trade frictions can substantially affect the average price of the narrowly defined good consisting of the high and low-priced varieties due to the Alchian-Allen effect.

2 Stylized patterns of purchase prices

This section describes patterns of raw consumer expenditure data from the interview section of the CE for 2009-2012.[5]

Our focus on the durables has several benefits: the durables are a larger part of consumer expenditures and the share of imports in consumption is higher for

---

[5] See the Data Appendix for the details on the CE data and on the cross-walk between product codes.

durables than for the nondurables.[6]

We begin our investigation of patterns by focusing on a specific CE product: televisions.[7] Figure 1 presents a simple plot of average purchase prices of televisions for the consumers in the richest quintile against corresponding averages for the poorest quintile.[8] As expected, in most states households in the richest quintile on average buy more expensive televisions than households from the poorest quintile.

What is less expected is how substantially this difference varies across states, from under 30% in New Jersey, Nevada, Massachusetts to over 300% in Missouri, South Carolina, and Kentucky. The pattern of the scatterplot in Figure 1 also reveals that most of the variation in the differences between the rich and the poor is driven by the variation of the average prices paid by the poorest quintile across states and much less by the variation in prices for the richest quintile. It appears that the choices of the poor are more affected by the factors related to the internal geography.

To illustrate that these patterns generalize to other goods, we express the log-price of every expenditure relative to the product mean and then calculate the median of those de-meaned prices for the fifth income quintile versus the first quintile. The results are shown in Figure 2. Every dot represents the median product de-meaned log price for the 5th and 1st income quintile for a given state. The pattern in Figure 2 is very similar to the illustrative case of television in Figure 1 suggesting that the patterns from the example of televisions carry over into the plot for all goods. Formally the patterns are stated as follows.

Pattern 1. *Across US states, households in the poorest quintile tend to purchase*

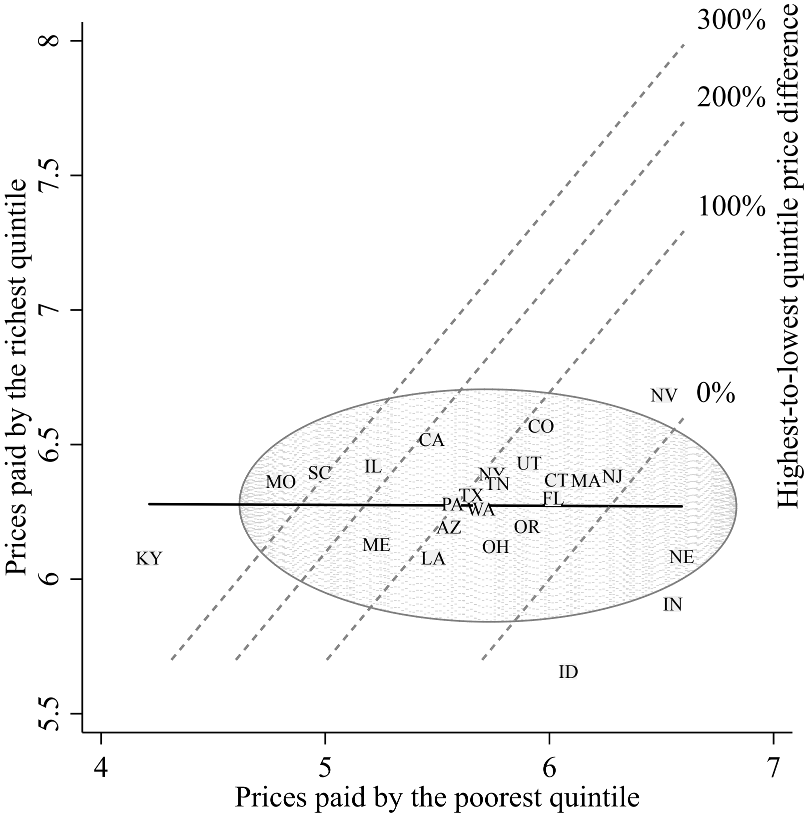
---

[6] In the growing literature utilizing scan data, the main focus is on studying inter- and international variation in prices of the exactly same varieties. (Broda & Weinstein, 2006; Atkin & Donaldson, 2015) Our focus instead, is on how purchasing prices of durable goods by poor and rich households vary across U.S. states, whereas households may adjust their variety choices (within the same product type) due to internal geography.

[7] Televisions provide a good illustrative case (1) it is a well defined product in the CE as opposed to for example "clocks and accessories"; (2) there is a substantial amount of variation in features and prices; (3) almost all of TV's consumed in the US are imported; (4) they are well represented in CE through multiple purchases throughout the states

[8] We construct national quintile cutoffs based on the household expenditures on nondurables to match our empirics and existing literature. The pictures look quantitatively similar if we use the national income percentile rank.

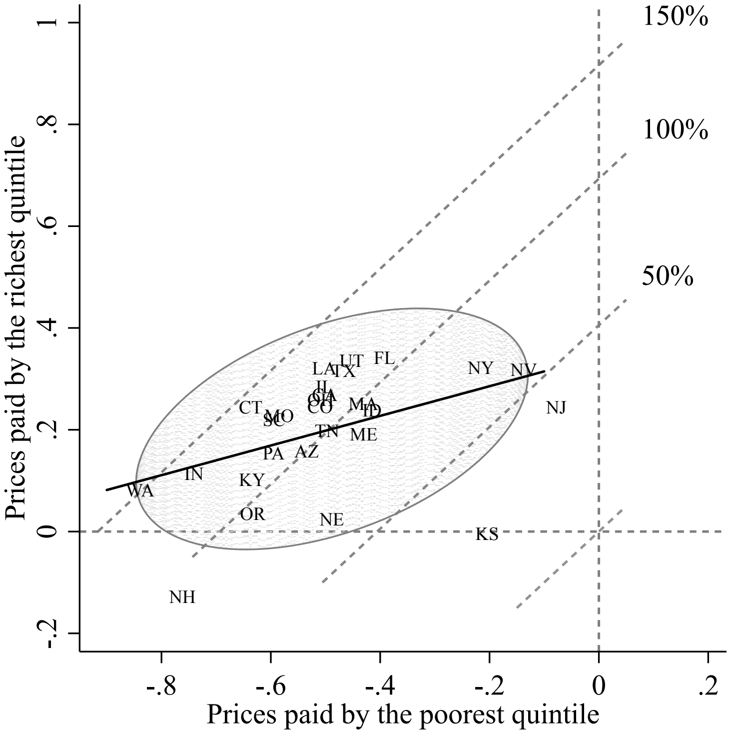
Figure 1: Log prices of televisions paid by rich and poor households across states



Notes: Every data point represents a unique state-product-income group combination

cheaper varieties of the same durable goods than the households in the richest quintile.

Figure 2: Median deviations of the prices from the product means by state and national income quintiles

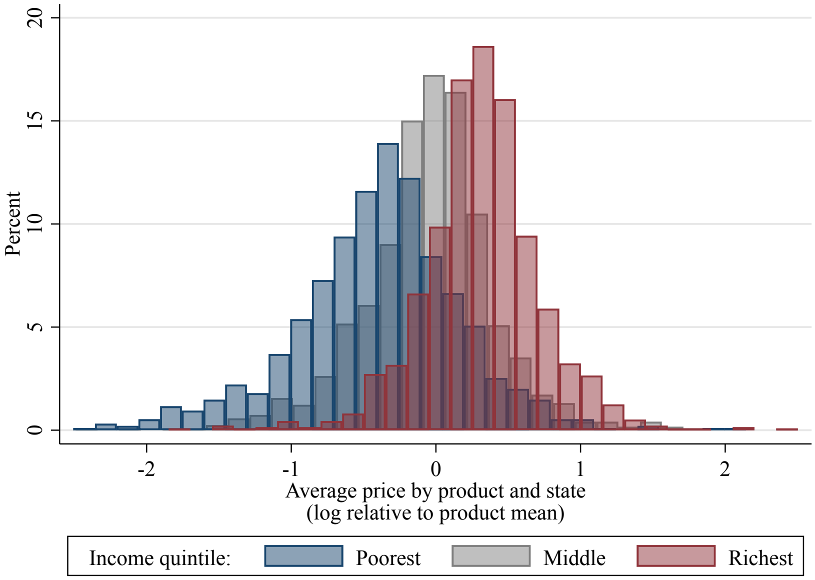


Pattern 2. *There is substantial spatial variation in the differences in prices paid by the rich and poor households.*

Pattern 3. *There is more variation in the prices paid by the poor across states than in the prices paid by the rich.*

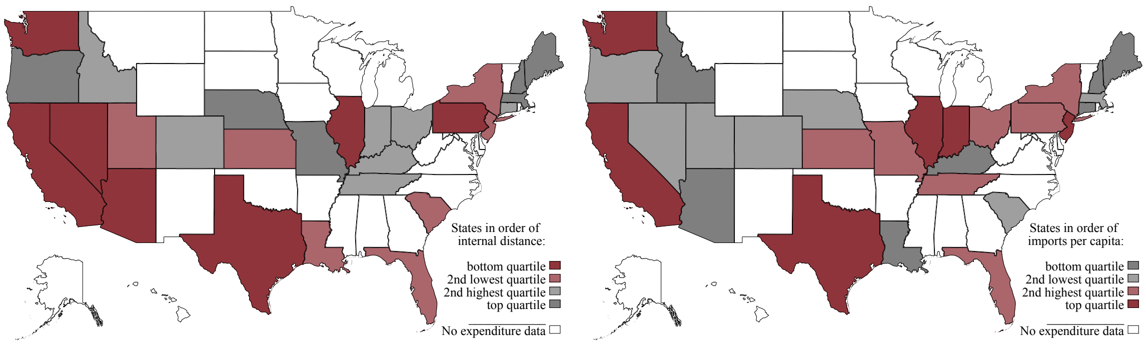
Another way to see that the households in the lower income quintiles buy cheaper varieties and have higher spatial dispersion (Patterns 1 and 2) is by plotting distributions of the state average prices of all de-meaned log prices for the top, middle, and bottom quintiles as shown in Figure 3. The de-meaned prices (in logarithms) paid by the households in the bottom quintile exhibit a higher dispersion than those paid by the top quintile households.

Figure 3: Poor households have a wider distribution of prices.



Notes: Prices are product de-meaned. Every data point represents a unique state-product-income group combination

Figure 4: Internal distance and import per capita



(a) Internal distance

(b) Per capita import volume

Notes: Internal distance of imports is calculated based on FAF routes. It is the simple average of the main routes. The routes are the combinations of exporting origin or country, state of entry, and state of destination. A route is considered to be one of the main ones if its cumulative share in the total product imports of the destination state is at least 50%. SCTG 2 digit trade is merged to the CE product codes and averaged for those good within state.

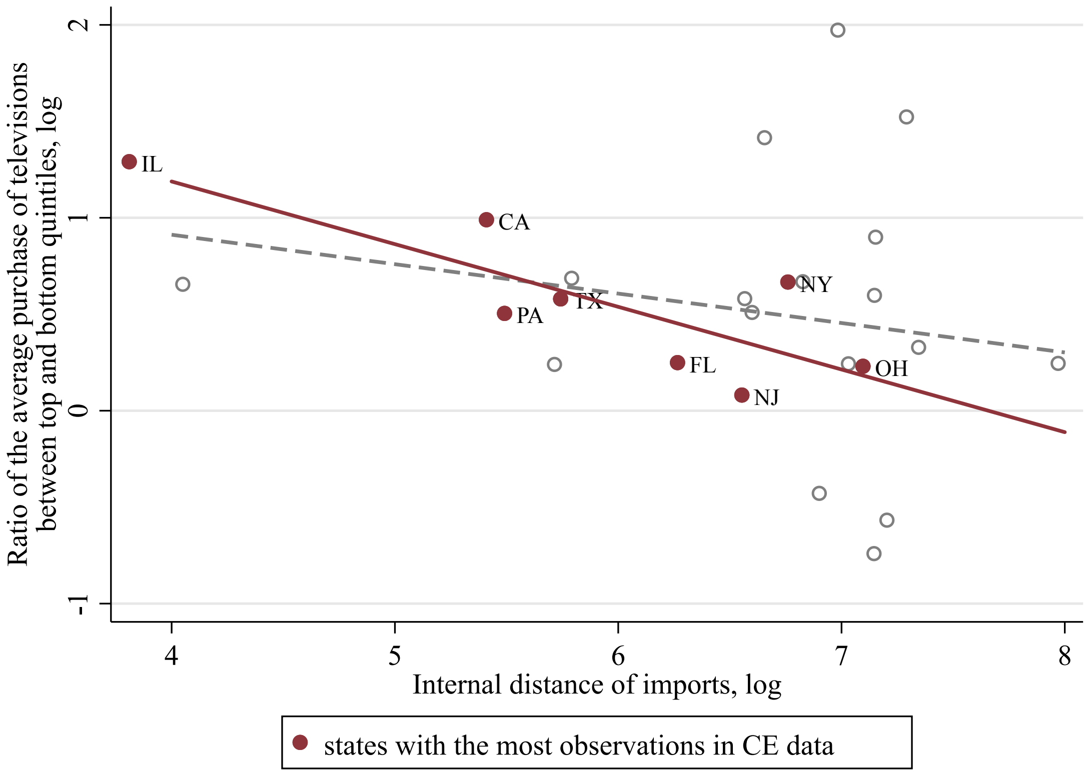
Considering that most (if not all) televisions sold in he US are imported, we next explore to what extent the patterns in Figure 1 can be related to the differences in access to imports. This relationship is the primary focus of our theory and emprics. In order to reflect internal differences in import access, we use internal distance of imports.[9] This is consistent with our empirical exercise, where we use internal geography to trace out exogenous variation in imports per capita. To illustrate the validity of this approach, the maps in Figure 4 illustrate the rank of states by imports of televisions per capita and by internal distance of television imports. The figure illustrates that states with lower internal distance tend to have greater per-capita imports of televisions. The correlation is not perfect, since endogenous differences in demand across states are also likely to play a role.

Pattern 4. *There is a negative correlation between rich-to-poor price differences and internal distance of imports.*

---

[9] The details of how we construct this measure are outlined at the end of this section.

Figure 5: Internal distance and rich-to-poor relative price of TVs.



Notes: The solid line is the line of best fit for the solid data points, while the dashed line represents best fit for the hollow points.

In order to document the pattern of spatial correlation between rich-to-poor price differences and internal distance of imports, we incorporate the information from Figure 1 by calculating the log ratio of the top to bottom quintile prices of televisions, and relating it to the internal distance of imports. There is a clear negative relationship, which is extremely pronounced when we focus on the states with the most observations on TV purchases in the CE data. Furthermore, the fitted (solid) line crosses zero on the vertical axis within the range of observed distances. This suggests that the difference in prices between rich and poor may be virtually negligible in locations with high internal distance.

A similar exercise can be performed for all goods. To this end we estimate relation between the log of ratio of prices paid by the richest and poorest quintile in a given state. Note that this transformation while simple would eliminate both state and product price means because they are the same for both the top and the bottom income quintiles. The estimated relationship is

$$\ln\widehat{\left(\frac{\text{Price rich}}{\text{Price poor}}\right)}_{\text{state,product}} = 1.053 - 0.047 \ln\left(\text{Internal distance}\right)_{\text{state,product}}.$$

To put the estimated coefficient in perspective, consider predicted rich-to-poor price ratios for the state-product combination with the 1st and 99th percentile of internal distance among all state-product pairs. For the 1st percentile of distance, internal distance predicts 98% difference and for the 99th percentile the model predicts 143% difference in prices paid by the rich and the poor. This variation in the measure of internal distance varies not only across the states but also across the products within a given state. For example Illinois has log of internal distance between 3.44 and 6.64, New York between 3.34 and 7.03, while for the remote Nebraska the log of distance varies between 7.05 and 7.514.

3 Theoretical Framework

In this section, our goal is explore how consumers with different income levels adjust their quality choices in response to trade frictions. To this end, we employ a model with specific trade costs and preferences defined over one non-durable and multiple quality-differentiated durable goods.

3.1 The Environment

There are two countries, Foreign and Home, $g = 1, 2, 3, ..., G$ durable goods and a numeraire nondurable good indexed by 0. Each durable goods is divided into two types, Low and High, and their consumption levels by household h are labeled as $\underline{x}_{hg}$ and $\overline{x}_{hg}$, respectively. We set the model from Home's perspective. Household h

in Home chooses the quantities of the numeraire and durable goods to maximize the following utility function:

$$U_h = x_{h0}^{\frac{\sigma_0 - 1}{\sigma_0}} + \sum_{g=1}^{G} \left(\underline{\lambda}_g \underline{x}_{hg}^{\frac{\underline{\sigma}_g - 1}{\underline{\sigma}_g}} + \overline{\lambda}_g \overline{x}_{hg}^{\frac{\overline{\sigma}_g - 1}{\overline{\sigma}_g}} \right), \tag{1}$$

where $\underline{\lambda}_g, \overline{\lambda}_g > 0$ are the multiplicative preference parameters, and $\sigma_0, \underline{\sigma}_g, \overline{\sigma}_g > 1$ are the preference parameters determining the curvature of utility for the goods. Note that, for any $\overline{x}_g = \underline{x}_g \geq 1$, if

$$\text{Assumption 1:} \qquad \overline{\lambda}_g \geq \underline{\lambda}_g \qquad \text{and} \qquad \overline{\sigma}_g > \underline{\sigma}_g, \tag{2}$$

both the utility obtained from consumption of $\overline{x}_g$ is always greater than the utility obtained from $\underline{x}_g$. Thus, under Assumption 1, we can interpret Low and High types, as the low-quality and high-quality varieties of good g.[10]

The budget constraint of household h is given by:

$$x_{h0} + \sum_{g=1}^{G} \left(\underline{x}_{hg} \underline{p}_g + \overline{x}_{hg} \overline{p}_g \right) = I_h, \tag{3}$$

where $\underline{p}_g$ and $\overline{p}_g$ denote the prices of the low- and high-quality types of good g, while I_h is h's income.

3.2 Production and Trade Costs

In both countries, all goods are produced by perfectly competitive firms. Labor is the only production factor. For simplicity, we normalize wages in both countries to be equal to one in terms of the numeraire.[11] The production technologies of durables

---

[10] Using a multiplicative parameter to preference parameter, λ, as a quality shifter is rather standard in the trade literature. Linking quality to the curvature parameter σ is rather novel. In a different context, σ is linked to the Quality Engel curves by Bils & Klenow (2001).

[11] Formally, we assume that the production technology of the numeraire requires one unit of labor in all countries, the numeraire can be traded costlessly across countries, and the numeraire sector is sufficiently large so that, in equilibrium, it is produced in both countries.

are characterized by the constant marginal costs of production, which may differ across countries and goods. We assume that Home has an insufficient capacity to satisfy the local demand for durable goods. Thus, even if for some goods Home has a lower marginal cost than Foreign, Home still imports these goods and Home's prices are determined by the prices of imports.

In Foreign, the marginal costs are $\underline{c}_g$ and $\overline{c}_g$, for the low- and high-quality varieties, respectively, where

$$\text{Assumption 2:} \qquad \overline{c}_g > \underline{c}_g \quad \forall g. \tag{4}$$

For imported (from Foreign to Home) goods there is also specific (i.e., per-unit) transportation cost τ_g. Thus, for imported goods, Home's consumers face the following delivered price:

$$\overline{p}_g = \overline{c}_g + \tau_g \qquad \underline{p}_g = \underline{c}_g + \tau_g, \tag{5}$$

for high-quality and low-quality varieties, respectively.

3.3 Equilibrium Prices and Quantities

In equilibrium, the ratio of the marginal utilities (which can be derived form equation (1)) is equal to the ratio of corresponding prices, which, for durable goods, are expressed by equation (5). This allows us to express the quantities of durables in terms of the numeraire as:

$$\underline{x}_{hg} = \left[\frac{\frac{\underline{\sigma}_g-1}{\underline{\sigma}_g}\underline{\lambda}_g}{\frac{\sigma_0-1}{\sigma_0}(\underline{c}_g+\tau_g)}\right]^{\underline{\sigma}_g} x_{h0}^{\frac{\underline{\sigma}_g}{\sigma_0}} \qquad \overline{x}_{hg} = \left[\frac{\frac{\overline{\sigma}_g-1}{\overline{\sigma}_g}\overline{\lambda}_g}{\frac{\sigma_0-1}{\sigma_0}(\overline{c}_g+\tau_g)}\right]^{\overline{\sigma}_g} x_{h0}^{\frac{\overline{\sigma}_g}{\sigma_0}} \tag{6}$$

Budget constraint (3) can then be re-written as:

$$x_{h0} + \sum_{G=1}^{G}\left\{\left[\frac{\frac{\underline{\sigma}_g-1}{\underline{\sigma}_g}\underline{\lambda}_g}{\frac{\sigma_0-1}{\sigma_0}(\underline{c}_g+\tau_g)}\right]^{\underline{\sigma}_g} x_{h0}^{\frac{\underline{\sigma}_g}{\sigma_0}}(\overline{c}_g+\tau_g) + \left[\frac{\frac{\overline{\sigma}_g-1}{\overline{\sigma}_g}\overline{\lambda}_g}{\frac{\sigma_0-1}{\sigma_0}(\overline{c}_g+\tau_g)}\right]^{\overline{\sigma}_g} x_{h0}^{\frac{\overline{\sigma}_g}{\sigma_0}}(\underline{c}_g+\tau_g)\right\} = I_h \tag{7}$$

The implicit solution to the utility maximization problem is given by equations (5), (6), and (7).

For our empirical exercise, we are interested in the average price paid by household h for a given durable good g, which is a weighted-average price of the high-quality and low-quality varieties of g : $P_{hg} \equiv \frac{\underline{x}_{hg}\underline{p}_g + \overline{x}_{hg}\overline{p}_{hg}}{\underline{x}_{hg} + \overline{x}_{hg}}$.

3.4 Predictions

As is evident from equation (5), transport cost τ has a positive linear effect on the prices of both low- and high-quality varieties. At the same time, it has negative, and asymmetric effect on the equilibrium quantities of the the low- and high-quality varieties. That is why, the direction and the magnitude of the effect of τ on the weighted-average price P_{hg} is more complex: it depends on how τ affects the share of the high-quality variety in the total mix, which is defined as $\overline{s}_{hg} \equiv \frac{\overline{x}_{hg}}{\underline{x}_{hg} + \overline{x}_{hg}}$.

To illustrate this concept consider an increase in transport cost by the amount of $\Delta\tau$, while holding the expenditures on the numerarie good, x_{h0}, constant.[12] If the effect of this increase on the share $\overline{s}_{hg}$ is positive, the share of the more expensive variety (High) increases, and thus the average price increases by an amount greater than $\Delta\tau$ (Alchian-Allen effect). If, on the other hand, the effect is negative, the share of High will decrease and the average price will either decrease or increase by an amount lesser than $\Delta\tau$ (reverse Alchian-Allen effect). Formally, we express it as

$$\epsilon_{\overline{s},hg} \equiv \left.\frac{\partial \overline{s}_{hg}}{\partial \tau_g}\right|_{x_{h0}=Const} > 0 \iff \epsilon_{P,hg} \equiv \left.\frac{\partial P_{hg}}{\partial \tau_g}\right|_{x_{h0}=Const} > 1.$$

Next, we provide the conditions for the Alchian-Allen effect to hold.

Proposition 1. The Alchian-Allen effect.
Holding the expenditures of household h on the numeraire good constant, h's average consumer price of good g, P_{gh}, increases in the specific transport cost τ_g by more than an increase in τ_g if and only if the ratio of price over the income elasticity parameter

---

[12] We choose to explore the expenditures on the numerarie constant for a better connection between the model and empirical part, in which we compare the consumer prices of households with the same level of expenditure on the numerarie.

σ_g is greater for the higher-priced variety:

$$\epsilon_{P,hg} > 1 \iff \frac{\overline{c}_g + \tau_g}{\overline{\sigma}_g} > \frac{\underline{c}_g + \tau_g}{\underline{\sigma}_g}.$$

Proposition 1 is in line with the previous theoretical literature on Alchian-Allen effect in that the AA effect exists when goods differ in prices and when the substitution with respect to the outside good does not substantially differ between the two qualities (Borcherding & Silberberg, 1978). Intuitively, if $\underline{\sigma}_g = \overline{\sigma}_g$ specific transportation costs increase the relative demand for the higher priced good. Allowing σ_g's to vary introduces differences in income effects and requires an additional restriction that the resulting income effect does not offset the decrease in the relative delivered price of the more expensive product. The next proposition is novel: it shows how the magnitude of the AA effect varies depending on the households expenditures on the numeraire good.

Proposition 2. The magnitude of the Alchian-Allen effect.
Conditional on its existence, the magnitude of the AA effect decreases in the expenditure on the numeraire good if and only if the equilibrium consumption of higher-priced variety exceeds that of the lower-priced variety:

$$\textbf{If} \quad \epsilon_{P,hg} > 1 \quad \textbf{then} \quad \frac{\partial \epsilon_{P,hg}}{\partial x_{h0}} < 0 \iff \overline{x}_{hg} > \underline{x}_{hg}.$$

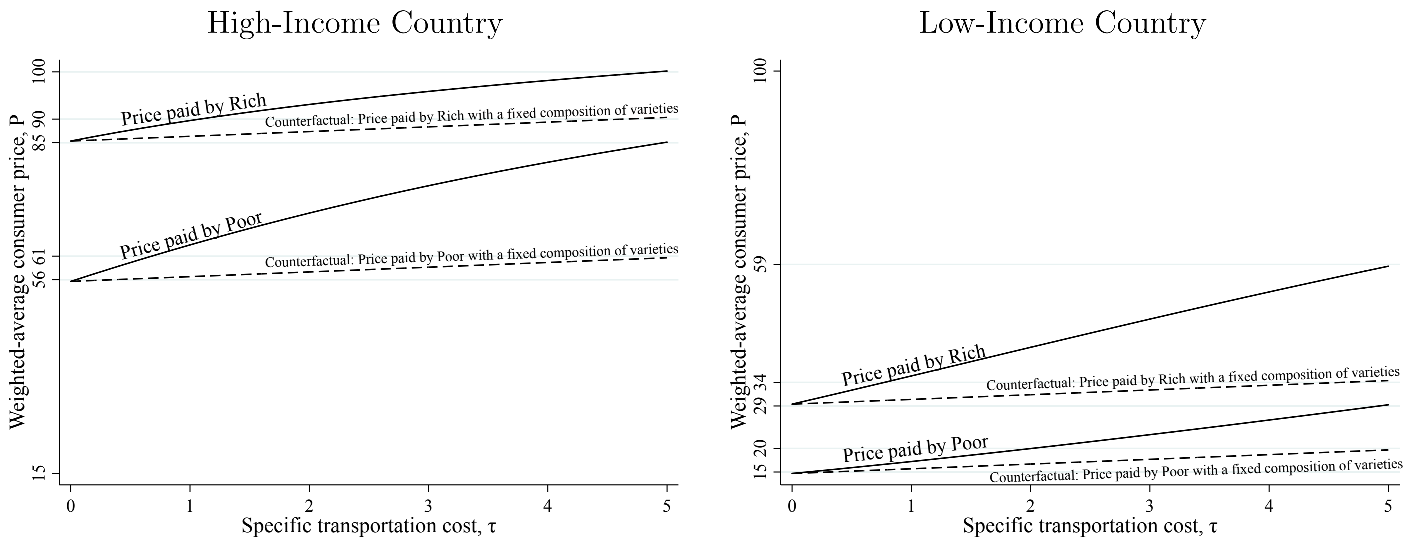
This will be always the case for households with sufficiently high expenditures on the numeraire good.

The intuition of this result is closely related to the income elasticities of the lower and higher-quality varieties. Recall that, by construction, the lower-quality variety has a lower income elasticity parameter than the higher-quality variety ($\underline{\sigma}_g < \overline{\sigma}_g$). Thus, households below a certain income threshold consume more units of the lower-quality (necessity) variety than of the higher-quality (luxury) variety. For these households, an increase in τ causes a relatively strong income and relatively weak substitution effect, and thus weak Alchian-Allen effect. With higher income and expenditures on the numeraire, the substitution effect becomes relatively stronger compared to the income effect, but only up until certain point. Namely, when the households already consume more of the higher-quality than of the lower quality variety, additional increase in income makes the substitution effect relatively weaker, because the share of the higher-quality varieties is already high and is bounded by one from above. For example, when the share of the higher quality varieties is very close to one, the magnitude of the AA is negligible and the average price increases by the same amount as τ.

The intuition of Proposition 2 is further illustrated with Figure 6, which considers two countries: High-Income and Low-Income. In the High-Income country, both Rich and Poor households have sufficiently high equilibrium expenditures on the numeraire so that they consumer more of higher-quality than of the lower-quality variety. As a result, the magnitude of the AA is stronger for the poorer households and the weighted-average consumer prices of Rich and Poor households converge as the transport cost increases. In the Low-Income country, both Rich and Poor households have relatively low equilibrium expenditures on the numeraire and thus consume more of lower-quality than of the higher-quality variety. Consequently, the AA is stronger for the Rich household and the weighted-average prices of Rich and Poor diverge as transport cost increases.

Figure 6 also helps us illustrate that even small differences in transportation costs can generate large swings in average prices. In fact the model is capable of generating a range of average price responses to transportation costs. By comparing the dashed and solid lines, we can see that even a modest increase in transportation costs is capable of substantially moving the average price.

Figure 6: Simulated average consumer prices paid by Rich and Poor households.



High-Income Country

Low-Income Country

Notes: The figure shows how the consumer price P varies in transport cost τ across different types of households.
The common (across all graphs and lines) parameters are: $\overline{c}_g = 100$, $\underline{c}_g = 10$, $\sigma_0 = 4$, $\underline{\sigma}_g = 4$, $\overline{\sigma}_g = 8$, $\underline{\lambda}_g = 1$, $\overline{\lambda}_g = 10$.
Lines differ by the household's expenditure on the numeraire: $x_{0,Rich} = 15,000$, $x_{0,Poor} = 3,000$ for High-Income Country; $x_{0,Rich} = 800$, $x_{0,Poor} = 160$ for Low-Income Country.

The dashed lines have a slope of one: assuming away the AA effect, P increases by exactly the same amount as τ.

Next, we extend our analysis of the AA effect to a more general case with many varieties of a given durable good. This extension allows for a closer link with our empirical exercise, since our data are likely to contain multiple varieties with varying prices and quality levels. Specifically, we now assume that that each good g consists of an arbitrary number of variety types, N_g, indexed by $n = 1, 2, 3, ..., N_g$, where

$$\sigma_1 \leq \sigma_2 \leq \sigma_3 \leq ... \leq \sigma_{N_g}, \quad c_1 \leq c_2 \leq c_3 \leq ... \leq c_{N_g}, \text{ and } \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq ... \leq \lambda_{N_g},$$

with a strict inequality in these parameters holding for at least two varieties:

$$\sigma_1 < \sigma_{N_g}, \quad c_1 < c_{N_g}, \text{ and } \lambda_1 < \lambda_{N_g}.$$

Next we show how our previous results extend to a multi-variety model.

Proposition 3. The AA effect with multiple varieties.
In a multi-variety model, an increase in the specific transportation cost τ_g results in the greater than just transport cost increase in price (the Alchian-Allen effect) if the ratio of price over the income elasticity parameter σ_g is at least as great for each higher-indexed variety as it is for each lower-indexed variety with a strict inequality holding for at least two varieties:

$$\epsilon_{p,hg} > 1 \quad \textbf{if} \quad \frac{c_{gN_g} + \tau_g}{\sigma_{gN_g}} \geq ... \geq \frac{c_{g2} + \tau_g}{\sigma_{g2}} \geq \frac{c_{g1} + \tau_g}{\sigma_{g1}} \quad \textbf{and} \quad \frac{c_{gN_g} + \tau_g}{\sigma_{gN_g}} > \frac{c_{g1} + \tau_g}{\sigma_{g1}}.$$

In general, the condition and its intuition are very similar to those of the two-variety setting. The main difference is that this is a sufficient, but not a necessary condition.

Proposition 4. The magnitude of the AA effect with multiple varieties.
Conditional on its existence in the model with many varieties of good g, the magnitude of the AA effect decreases in the expenditure on the numeraire good if the equilibrium consumption of each higher-indexed variety is at least as high as it is of each lower-

indexed variety with a strict inequality holding for at least two varieties:

$$\frac{\partial \epsilon_{P,hg}}{\partial x_{h0}} < 0 \quad \textbf{if} \quad x_{hg1} \leq x_{hg2} \leq ... \leq x_{hgN_g} \quad \textbf{and } x_{hg1} < x_{hgN_g}.$$

$$\frac{\partial \epsilon_{P,hg}}{\partial x_{h0}} > 0 \quad \textbf{if} \quad x_{hg1} \geq x_{hg2} \geq ... \geq x_{hgN_g} \quad \textbf{and } x_{hg1} > x_{hgN_g}.$$

4 Empirics

The goal of this section is to document differences in the non-homothetic consumption patterns across states and investigate wether those differences can be attributed to the import-reducing effects of internal geography. To this end, we provide econometric evidence on the relationship between household-level purchase prices across income levels and local per capita imports due to variation in the geography of import routes by state and product. Intuitively speaking, we estimate to what extent the difference in the price of, say, televisions consumed by rich versus poor households in Illinois is systematically related to the imports of bicycles to Illinois instrumented by the internal distance of television imports to Illinois.

4.1 Empirical specification

Motivated by our theoretical predictions, in this section, we estimate the relationship between the expected (average) price of the goods purchased by a given household depending on the households expenditure on the nondurables.

Similarly to Bils & Klenow (2001) we trace out differences in consumption across income distribution by estimating slopes of the quality Engel curves. Their theoretical setup allows for direct interpretation of the slopes of the quality Engel curves in terms of the model parameters. Our model, on the other hand, is specifically geared to analyze the Alchian-Allen effect under non-homothetic preferences. As a consequence, we trade off the structural interpretation in favor of the reduced form estimation to focus on the heterogeneity in price-income relationship across US states.

Baseline specification for the quality Engel curve is given below. The indexing of variables reflects the structure of the data set. In our data every household h belongs to the importing state s, g indexes consumption goods, and t indexes the expenditure year.

$$\ln p_{ight} = \beta_1 \ln Y_{iht} + \beta_2 X + \gamma_{gt} + \gamma_{it} + \epsilon_{ight}, \tag{8}$$

where p - price; Y - income (expenditure on nondurables); X - controls; γ - fixed effects. Introducing a measure of trade both in levels and as an interaction with the nondurable expenditures results in our main specification.

$$\ln p_{ight} = \delta_1 \ln Y_{iht} + \delta_2 \left(\ln Y_{iht} \times T_{igt}\right) + \delta_3 T_{igt} + \delta_4 X + \gamma_{gt} + \gamma_{it} + \epsilon_{ight}, \tag{9}$$

where T - trade measure (per capita import volume, extensive margin, intensive margin).

4.2 Identification

Imports and consumption are clearly endogenous. Since our dependent variable is the purchase price, not the total volume, the concern for endogeneity arises from a possible connection of the distribution of volumes across origins that can be related to the error term in the price regression. This can occur, for example, due to unobserved local import demand factors that are specific to an origin. For instance, Ohio may have unusually high demand for German cars, which are more expensive than cars from other exporters. This preference idiosyncracy can affect the calculation of trade-weighted internal distance if German cars arrive through a port in Georgia. Our instruments are the measures of internal and external distance. For the exclusion restriction to be valid our measures of distance need to be uncorrelated with the error term in the price regression.

Geographical distance is routinely used in trade literature to measure the exogenous component of the international trade costs between the location of production and location of consumption.[13] Constructing a similarly exogenous measure of internal distance is not equally straightforward. The same argument for exogeneity does not apply because the decomposition into external and internal distance is not independent of the trade volume.

It is however equally misleading to completely ignore the exogenous of available feasibile routes. To do so, we rely on the insights from the transportation literature on route selection. Two important lessons emerge from the route selection literature. First, due to large fixed costs associated with establishing a transportation route there is a certain degree of lumpiness to the route formation, this is why there are *routes* instead of direct connection between every possible combination of locations. Second, the literature on route choice separates the decisions into (1) generation of the route choice set and (2) choice of routes. Based on those insights our instrumentation strategy is based on construction of the simple average distance of the main routes of imports of a given product to a given state. We are effectively assuming that if, for example, households in Colorado had the same characteristics as the households in Pennsylvania, Colorado's top state of entry for East Asian imports of furniture would still be California. By calculating a simple average we do not allow the trade volumes to affect our measure of distance. This is clearly not to say that the exact route that goods travel is not endogenous on trade. We are just assuming that unobservable error components are that could shift both trade and prices are not sufficiently large to affect the set of available routes. We will assume that the set of available feasible routes is exogenous even if the choice of routes from that set may be endogenous.

4.2.1 Measuring internal distance of imports

Pattern 5. *Internal distances of imports vary significantly both for a given good across U.S. states and for a given U.S. state across products. This correlation is caused by*

---

[13] According to Frankel and Romer (1999): "it is difficult to think of reasons that a country's geographic characteristics could have important effects on its income except through their impact on trade."

- *differences in transportation modes and ports of entry across goods;*
- *anything else?*

We use geographic distance to capture exogenous trade-reducing effect of cross-state difference in geography through distance-related trade costs.[14] Identifying exogenous variation in imports that is not related to consumption is central for building a compelling identification strategy for the effects of imports on the slopes of the quality Engel curves.[15] If imports traveled directly point-to-point from the place of their production to the place of their consumption, the relative location within the U.S. would be characterized by its geographic distance to the sources of imports. The shipping is however done indirectly. Few major gateways like ports and airports handle a large share of imports. As a result, different states receive their imports from different regions through different ports. Consider Figure 1 that shows shares of state in destination state's value of imports. Colored columns indicate that some states are significant gateways for imports to other state. California and New Jersey host some important water gateways. Kentucky and Tennessee are home to air shipping hubs. Texas and Michigan are points of entry for imports from Mexico and Canada by rail and truck.

---

[14] Distance-related trade costs are the empirical counterpart of the theoretical transportation costs.

[15] Note that in order to reflect differences in internal geographic location of states we need to use both measures of international (i.e. external) as well as intranational (i.e. internal) distance of imports. Colorado is further from China than California but it is also far from California, where most Chinese imports arrive to the U.S. by sea.

Table 1: Direct and indirect imports across U.S. states: Entry state's share of destination state's total imports.

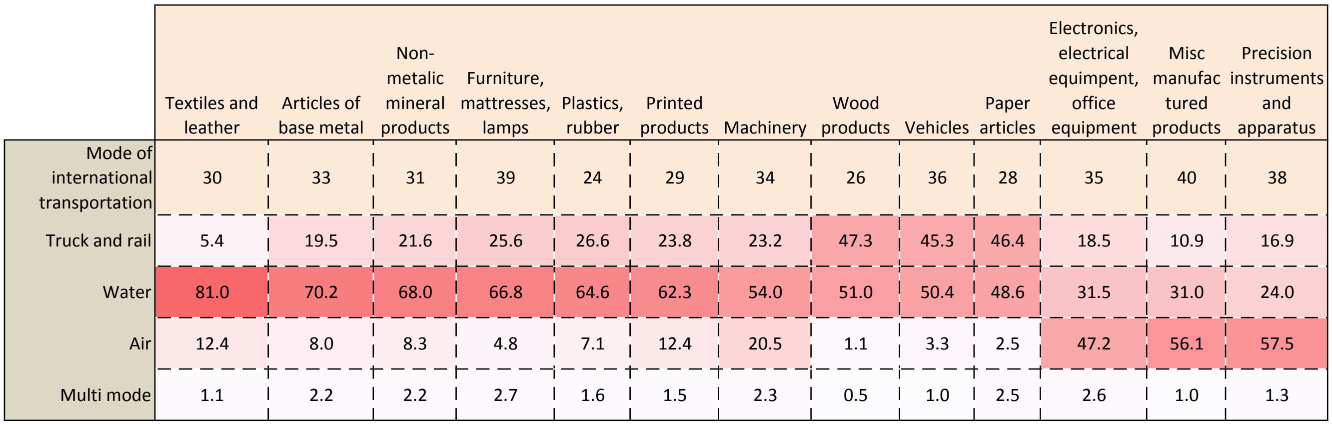
	State of entry																																													
Destination state	AK	AL	AR	AZ	CA	CO	CT	DC	DE	FL	GA	HI	ID	IL	IN	KS	KY	LA	MA	MD	ME	MI	MN	MO	MS	MT	NC	ND	NE	NH	NJ	NM	NV	NY	OH	OK	OR	PA	RI	SC	TN	TX	UT	VA	VT	WA	WI
AK	57.4				3.6						0.6						4.9					2.5						0.8						1.1						0.9	9.9	1.2				14.7	
AL		21.0			21.5					1.8	28.2			0.6			0.6	1.8			0.6	5.9	0.5		2.2									1.6						3.1	0.8	7.6				0.6	
AR		0.7			22.4					0.7	2.1			27.9			0.6	1.8				5.5						0.7			0.5			3.6	0.6					4.7	1.1	24.4				0.6	
AZ	7.0			27.1	42.2					1.9							2.2					3.3									0.6			1.3							4.8	5.9				1.5	
CA	1.7				70.3					1.4	1.0			1.9			0.8					3.8									1.9			3.4	2.3		1.1				3.4	4.7				0.7	
CO	3.3			1.2	25.5	10.6								1.5			4.3			0.6		4.0				3.6		1.2			10.1			1.9	2.7			0.7			7.8	15.7				2.3	
CT		0.5		1.0	6.6		2.1			0.6	1.4			0.9			1.4		1.2												26.6			37.9	1.0			0.6			2.5	9.1		0.5	1.1	2.8	
DC					3.1			21.4		3.4				4.2			2.3			10.5		0.9									2.6			28.7	0.8			1.2		5.3	6.6			5.1	0.6	2.0	
DE	2.8				3.4			0.8						0.8			1.0			3.0		5.5									14.0			21.9	0.7			30.8			4.0	6.8			0.9	1.3	
FL		0.7		0.5	5.4					65.7	4.4			1.0			1.9					0.8									1.5			3.6	1.9					0.7	2.6	6.6					
GA					9.9					2.3	64.3			0.7			0.6			0.6		1.8									2.0			1.9						5.1	1.1	6.1				1.4	
HI					7.3							84.8					0.7					0.9												0.7							2.2	0.8				1.5	
IA					13.4						0.6			29.2			1.3	0.7		1.2		12.2	0.9					4.1			5.5			0.9	0.5						0.7	24.7		1.1		1.5	
ID	24.2			1.2	29.7								1.0				1.4					1.8				0.9		0.9			0.6			3.8	0.8		1.0				21.4	1.0				9.4	
IL	0.5				10.6									61.1			0.6			0.7		5.8	0.5					0.9			1.8			0.9	0.8						1.3	8.6				2.7	
IN				0.5	19.7					0.9	0.6			33.3	2.6		1.9					13.8	0.5					0.8			4.6			1.1	1.4						0.9	11.6		0.9		2.4	
KS					40.6					1.0	0.9			6.3		1.7	2.7					8.2		5.1				1.7			6.8			1.1	1.5						1.7	14.1		1.9		3.0	
KY					24.6						1.7			17.4			16.2			0.7		6.6			1.3						3.8			1.0	6.7					0.5		10.0		4.8	0.5	1.3	
LA		1.0			7.0					0.8	0.7						1.2	45.1				6.4	0.8			1.2		0.9						2.8				0.6			6.4	22.5				0.9	
MA	1.3			2.5	16.5					5.8	1.0			0.9			4.2		26.1		0.7										7.5			16.4	1.0			1.4	0.9		3.8	4.1			2.2	1.7	
MD	2.3				1.8			2.1		1.1							0.9			63.8		0.6									2.7			6.7	0.5			2.0			3.3	7.0		2.9			
ME	4.1				2.3					1.0				0.8			0.6		10.3		37.5	0.5									9.8			10.4						0.8	7.8	2.6			10.1		
MI				4.2	2.4									4.0			0.6			0.8		45.4									1.0			4.4							0.7	33.3		0.5		0.6	
MN	1.0				30.9					0.5	2.5			4.1			2.0			0.5		3.4	5.8					3.6			5.3			0.8	1.0			0.8			6.0	8.7		5.3		16.4	
MO				0.5	40.1					1.2	1.0			4.3			0.8					11.2		3.1				1.4			6.9			2.7							1.2	20.8		0.6		1.8	
MS		1.3			43.9						5.5			4.9			0.9	3.4				4.2	3.3		5.8			1.0						1.0	0.8					6.0	3.7	12.1				0.7	
MT	0.6				2.5									1.3			1.6					13.4	0.6			39.7		6.2			2.9			1.9							2.5	2.2				22.2	
NC				1.5	8.9					9.1	13.7			1.8			1.0			1.9		0.9					12.0				2.3	0.8		7.0	1.9					11.1	2.2	11.7		9.1		1.4	
ND					6.4									2.0						1.1		7.1	0.6			10.5		39.0			18.2			0.8							2.0	6.7				4.2	
NE					29.4					0.7				12.2			1.0			2.9		10.7	2.0					3.7	1.0		9.5			1.6	0.7					1.5	0.8	8.3				12.2	
NH	0.8			1.1	11.9						0.6			1.1			2.4		24.0		1.0									2.9	11.4			10.7	1.5			1.5			4.0	13.7			8.9	0.6	
NJ	0.6				13.5					0.7	3.1			1.7			0.9		0.8	4.0											42.8			19.4	1.2		0.8	2.1	0.6	0.6	1.7	3.4				0.6	
NM	0.7			1.4	46.3												1.1					3.1				1.4		0.6			10.9	6.6		1.1	1.6						2.9	19.9				1.0	
NV	0.5			0.5	82.2												0.6					3.4									1.1		1.0	1.8							2.1	2.1				1.9	
NY					10.8					1.4	1.4			0.7			0.9														19.2			56.4	0.9					0.8	1.9	2.1				0.6	
OH					26.0						0.6			7.9			5.6	0.6		1.2		10.5						0.6			5.2			4.4	13.4					0.8	1.0	13.2		1.3		4.7	
OK					9.7									7.5			0.9	1.4				5.8				2.6		2.4			0.8			1.4	0.6	1.6					2.0	60.1				0.8	
OR					11.9						1.4			3.2			1.4			0.7		1.8									1.0			6.3			48.7				1.5	4.4				14.8	
PA					7.2			0.5		0.9				2.7			1.1			4.0		0.6						0.5			22.2			33.2	1.5			14.2			1.5	6.1				0.8	
RI				4.2	5.0					2.4	2.5						0.8		7.1												3.3			2.7	0.8			0.9	65.1		1.3	1.9			0.6		
SC					5.7					2.8	15.4			0.9			1.0	1.0			1.1	1.3					1.2				1.4			3.1						50.0	0.5	11.1		1.8			
SD	2.7				18.8									5.9			1.1					12.4	2.6			1.6		20.3			8.1			1.8	1.2						3.0	12.0		0.6		6.9	
TN					39.4					0.6	10.7			18.1			0.6	0.6				3.2									0.6			1.0	1.0					5.1	5.3	10.7				0.8	
TX					7.2					2.0							0.6					2.5				0.8						6.1		1.1	0.8						1.5	73.7					
UT				5.0	20.0						0.8						15.1					2.8				0.8					2.4	9.6		3.0							1.0	10.6	24.6			1.0	
VA					3.8			6.1			1.6			0.8			0.9			2.9		0.6					1.4	0.6			1.4			11.2	0.7			0.7		0.6	1.3	2.7		61.0			
VT					0.9												0.8		3.3												8.6			5.3				1.1			1.3	1.2			75.1		
WA					4.1									0.6			0.8					2.1												1.2	0.6		2.9				1.6	1.9				81.4	
WI	0.5			1.9	20.2					1.9	2.6			32.6			1.6					6.7	3.0					1.4			4.2			1.6							1.9	12.1		0.8		1.9	3.1
WV					31.5			5.3		1.5				2.6			4.3			1.4		3.8						1.0			2.5			11.2	1.4			5.0		0.7	2.5	11.1		4.5		7.9	
WY					22.7	7.9			0.5					1.2			0.7			0.6		7.2				37.0		2.5			2.3			3.4							1.8	6.8				3.7	

Notes: The shares are constructed using 2012 values from FAF.

There is also evidence that the internal cost of transportation differ from the international ones. For one thing, rail and truck are much more dominant mode of delivering good domestically than internationally. So, distinguishing between the international and intranational (sometimes called behind-the-border) distance of imports is crucial for characterizing effects of geographic position on trade.

We use the term “internal distance of imports” to denote the physical distance traveled by the imported goods from the border to the location of consumption. In order to calculate internal distance we use two primary sources of data. First, we use the data on U.S. imports by state of destination at 6-digit HS codes. These data lack the details on the location of entry into the U.S. Second, we use the FAF data that builds on CFS to construct flows of imports from the foreign origin to an entry location within the U.S. and then between those locations in the U.S. and the destination state. The entry locations within the U.S. are associated with state names, to which we will refer as the entry states. These data identify the origin of imports by six aggregated origins such as Europe or Southwest and Central Asia, and Mexico and Canada which are identified separately. The goods are classified according to the Standard Classification of Transported Goods used by the Commodity Flow Survey and designed to be hierarchically similar to the HS system. The data available on the FAF website is at 2-digit SCTG level. This disaggregation picks up some difference in the mode of transportation across different commodities. A two-digit product is for example ”Electronic and Other Electrical Equipment and Components, and Office Equipment” (code 35) that includes computers and other electric equipment and ”Plastics and Rubber” (code 24) that include tires. This level of disaggregation is crucial because the difference in the modes of transportation (air versus water) affects the route and therefore the distance (internal and external) that the products need to travel. Figure 2 shows shares for each mode of international transportation for various goods.

Table 2: Shares of international transportation modes in imports, by industry.



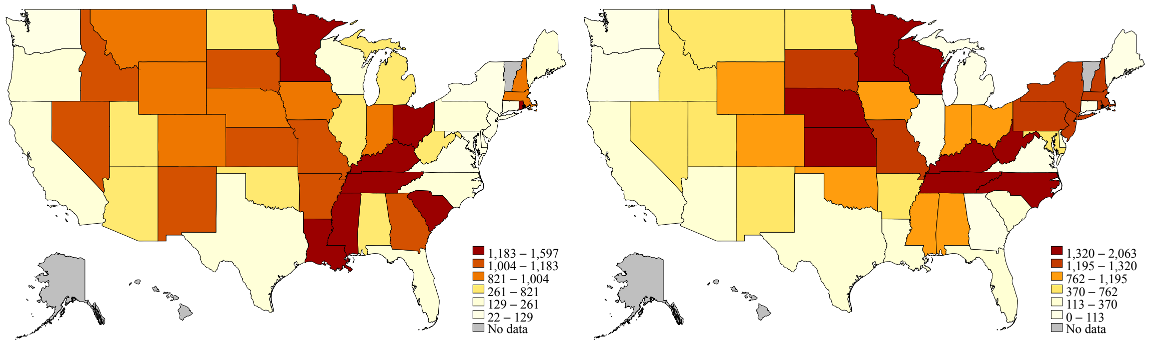
	Textiles and leather	Articles of base metal	Non-metalic mineral products	Furniture, mattresses, lamps	Plastics, rubber	Printed products	Machinery	Wood products	Vehicles	Paper articles	Electronics, electrical equimpent, office equipment	Misc manufac tured products	Precision instruments and apparatus
Mode of international transportation	30	33	31	39	24	29	34	26	36	28	35	40	38
Truck and rail	5.4	19.5	21.6	25.6	26.6	23.8	23.2	47.3	45.3	46.4	18.5	10.9	16.9
Water	81.0	70.2	68.0	66.8	64.6	62.3	54.0	51.0	50.4	48.6	31.5	31.0	24.0
Air	12.4	8.0	8.3	4.8	7.1	12.4	20.5	1.1	3.3	2.5	47.2	56.1	57.5
Multi mode	1.1	2.2	2.2	2.7	1.6	1.5	2.3	0.5	1.0	2.5	2.6	1.0	1.3

For each destination state and SCTG 2-digit product we define an import route as a combination of the exporting country and the state of entry. So, for electronics (code 35) imports to Indiana, a flow of trade from East Asia through Illinois is a route. Every route has an internal and an external component. First, using FAF data for every state imports of each SCTG product keep routes that amount to at least 50% of the destination state's imports of that commodity. The same entry state can be a part of more than one route from different foreign origins. Second, take the simple average of the internal distances of those routes. We calculate internal distance two different ways. First, we use population-pair-weighted greater circle distance between all population centers larger than 50,000 in each state. Second, we use a tonnage weighted distance obtained by dividing ton-miles by the tons available for 2012 through the FAF.

Figure 7 shows differences in internal distances across the U.S. states for four selected expenditure products. The figure illustrates that (1) there are substantial differences in the internal distance of the major import routes across states and (2) there are differences in the distribution of internal distance of trade across products. The variation of the internal distance of imports across states and products can be attributed to two important sources. First, there is an important hub and spoke system for entry of imports into the US This is more so for imports that enter by water. California and New York's share as an entry state for other states is fairly large. Second, this pattern differs across models of transportation. Third, the goods differ in the international mode of transportation. Forth, there is variation in the routing across sources of imports. Most notably imports from Canada and Mexico are more likely to enter the US by truck or rail.

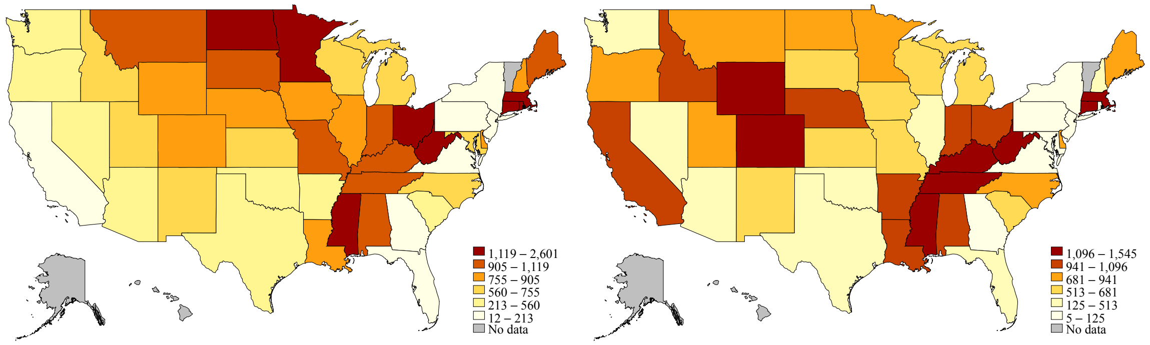
Establishing meaningful variation in the internal distance simultaneously across states and products is crucial for our identification strategy. This is because it allows us, at least mechanically, to use state-year fixed effects along with product-year fixed effects. Those fixed effects are very valuable because there is a variety of factors that are state specific that we might not be able to measure with available variables. Those would include real estate prices, cost of living, product-specific degree of import competition, state-specific preference shocks or business cycles.

Figure 7: Internal distance of imports for selected products



Furniture

Textiles



Consumer electronics

Cars

Notes: Internal distance of imports is calculated based on FAF routes. It is the simple average of the main routes. The routes are the combinations of exporting origin or country, state of entry, and state of destination. A route is considered to be one of the main ones if its cumulative share in the total product imports of the destination state is at least 50%. SCTG 2 digit trade is merged to the CE product codes and averaged for those good within state.

Figure 8 plots product-level internal distances against state averages. The measure of internal distance exhibits significant amount of variation around state means. In fact, state means explain 52% of the variation in the the internal distance leaving the rest unexplained. The standard deviation of the internal distance is 1.08, while the standard deviation of the mean-differenced internal distance is 0.75. There is however a possibility that deviation from the state means is product-specific. So, that imports of some goods have systematically higher internal distances than that one of the others. This could happen for example some goods tend to be imported by air and therefore closer to the final destination. In order to gauge how much variation can be attributed to the differences across products, we regress internal distance on the full set of state and product dummies. The R-squared increases very little to 53% suggesting that the internal distance is not product-specific. Highlighted points for Televisions in Figure 8 illustrate this point. The difference between internal distance and its state averages varies both in sign and magnitude across states.

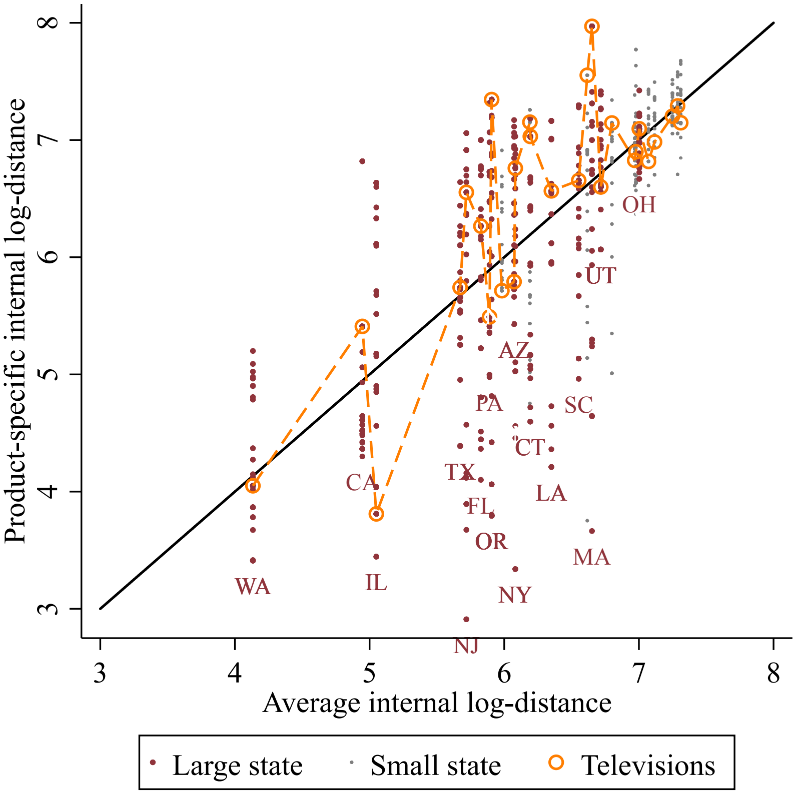
The measure of the external component of the entire route is measured similarly based on the distance of the major routes. For the external distance of the largest routes, we first combine the FAF data with the U.S. imports by state of destination. This allows us to eliminate countries in every region that do not export to a given state. So, if Mongolia does not export electronics to Alabama, we will not consider distance from Mongolia to the entry states of the main import routes of electronics to Alabama. Next we take a simple average of the external distance.from the state to the countries. We also construct a similar measure of the direct distance of imports.

There are two ways to calculate distance between states (including distance for the same-state shipments) based on the FAF data. First, we can use the data on the routes combined with our own calculations of the greater circle distances. Second, we can use the total measures of ton-miles and tons of commodity flows to construct the ton-weighted measure of state-to-state distances. The relationship between the two measures is illustrated in Figure 9. For distances between states, not surprisingly, most ton-weighted measures are close but somewhat longer than the greater-circle measures. For the import shipments within state[16] the difference is larger even

---

[16] A within state movement of imports would be for example imports arriving in a port in Cali-

Figure 8: Variation of product-specific internal distance across states.



Notes: Every data point represents a unique state-product-income group combination

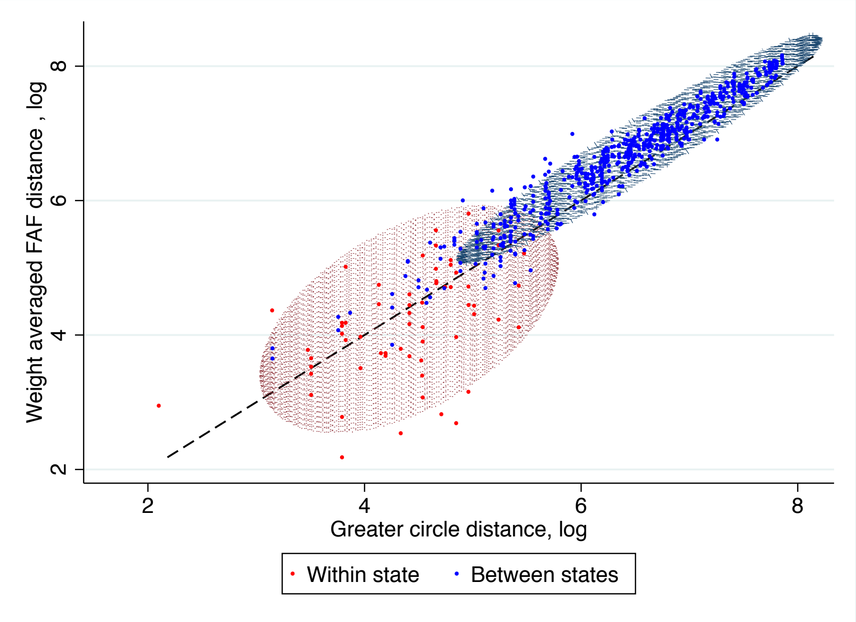
though the variation does not appear to be systematic.

Our measure of internal distance is related to but different from the concept of remoteness. The term remoteness is widely used in the gravity literature to account for differences in the price indices due to particular constant elasticity demand structure of the demand in gravity model. The specific meaning of remoteness depends on the model and the purpose for measuring it. A typical measure of remoteness involves calculation of the average inverse of exponential distance weighted by the GDP of all trading partners. Our consumption model differs from the traditional

---

fornia and then transported to the final consumer also in California

Figure 9: Two measure of internal distance



Notes: The confidence ellipses are constructed for all data but the dots represent a random 10% sub-sample to preent over-plotting. Freight Analysis Framework (FAF) distance is calculated by dividing total ton-miles by total tons.

gravity setup and therefore our measure of internal distance of imports is different.

There is still a possibility that the areas that receive more trade, even conditional on internal distance, have higher cost of living. For instance, the coastal areas are more likely to have ports, but they are also attractive places to live, which drives the prices or real estate up. This is an issue that is not easily controlled by the state-year fixed effects because higher cost of living may affect rich household differently from the poor biasing the estimates of the quality slopes. In order to account for such possibility and make our estimates conditional on the cost of living differences, all specifications include interactions between a measure of cost of living and expenditure on the non-durables. We use two measures of costs of living: cost of living index produced by the Council for Community and Economic Research[17] and the Living Wage Calculator created by by Dr. Amy K. Glasmeier at the MIT[18].

A possible alternative to this specification would involve interacting expenditure on nondurables directly with the exogenous measures of distance. Such a reduced form approach is problematic on several grounds. But considering it, highlights several features of our instrumentation approach. First, there are many reasons why internal distance may affect consumption through location of economic activity. We are interested only in the effect of the internal distance that works through the trade costs and imports. Second, the goods are likely to differ in the sensitivity of their imports to internal distance. For example, bulky goods may be affected by internal transportation more than light goods. Goods that are transported by air might be not particularly sensitive to internal distance. By using imports instrumented by internal distance, we automatically allow for the differential sensitivity of imports to internal distance. So, the trade volumes instrumented by distance is particularly well suited for our purpose of tracing out differences in import access across goods and states.

---

[17] For more information see `http://coli.org/`

[18] For more information see `http://livingwage.mit.edu/`

4.3 Results

Results of the estimation are presented in Tables 4 and **??**. The two tables differ only in the cost of living variable used in the estimation. The interaction between trade and income is universally positive indicating that income has stronger positive effect on the prices of purchases if a the state receives higher imports of that good. The instruments pass the customary Cragg-Donald test for weak instruments. The positive interaction between imports and income suggests that the household income tends to have a stronger effect on the purchase prices in the locations with higher imports per capita.

The sign of the interaction term alone is not sufficient to understand whether the increase in the slope of the quality Engel curve comes about from the changes of purchase prices by the low income households, high income households, or both. Figure 10 illustrates the predicted quality Engel curves over the range of observed incomes for several values of imports per capita. The lines intersect close to the top of the income distribution. Such location of the pivot point suggests that the quality Engel curves become steeper because the low income households shift their choices toward the lower priced varieties of a given product. Product quality Engel curves appear to be sloping up only in locations with high local per capita imports of those goods.

As one of the alternative specifications we decompose imports into intensive and extensive margins. By splitting imports into the extensive and intensive margins we are able to detect that the effect of imports works through the intensive margin. This is not surprising simply because at the level of CE Survey products and for the states present in the Survey, if a product is imported into the US it tends to be imported into all states. Lacking meaningful variation in extensive margin, the intensive margin is the only source of variation in trade.

Table 3: Local imports and slopes of Engel curves (living wage).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Weighted	No	No	No	Yes	Yes	No	No	Yes	Yes
N. of states	27	27	27	27	27	16	16	16	16
Nondurable exp., iht	6.256	6.801	6.592	6.174	5.862	7.254	7.312	7.184	7.297
	(1.946)**	(2.637)*	(2.084)**	(2.435)*	(1.917)**	(2.732)*	(2.101)**	(2.766)*	(2.074)**
Imports/cap, igt	-1.081	-0.971		-0.731		-0.994		-1.006	
	(0.298)**	(0.354)**		(0.349)*		(0.350)**		(0.355)**	
(Nondur.exp)×(Imports/cap)	0.109	0.107	0.103	0.080	0.073	0.111	0.110	0.113	0.110
	(0.034)**	(0.039)**	(0.038)**	(0.039)*	(0.036)*	(0.037)**	(0.036)**	(0.038)**	(0.036)**
Living wage, i	5.648								
	(1.860)**								
(Nondur.exp)×(Living wage)	-0.683	-0.729	-0.705	-0.631	-0.592	-0.777	-0.782	-0.773	-0.780
	(0.209)**	(0.286)*	(0.229)**	(0.263)*	(0.211)**	(0.292)**	(0.225)**	(0.295)*	(0.222)**
State pop, it	0.140								
	(0.029)**								
Median h.h. income, it	0.203								
	(0.139)								
Family size, iht	-0.207	-0.213	-0.206	-0.215	-0.209	-0.207	-0.199	-0.204	-0.197
	(0.021)**	(0.022)**	(0.012)**	(0.022)**	(0.012)**	(0.022)**	(0.012)**	(0.022)**	(0.013)**
Age of ref. person, iht	0.047	0.055	0.056	0.043	0.045	0.065	0.066	0.044	0.046
	(0.027)+	(0.031)+	(0.019)**	(0.033)	(0.020)*	(0.030)*	(0.019)**	(0.032)	(0.019)*
Housing exp., iht	0.051	0.059	0.061	0.050	0.053	0.065	0.066	0.058	0.059
	(0.009)**	(0.011)**	(0.007)**	(0.011)**	(0.007)**	(0.012)**	(0.007)**	(0.012)**	(0.007)**
C.D.F-stat	294.9	146.5	455.0	147.7	501.4	133.7	484.1	128.6	467.8
N.obs.	67,297	67,297	67,289	67,297	67,289	67,060	66,964	67,060	66,964
FE	gt	gt it	git	gt it	git	gt it	git	gt it	git

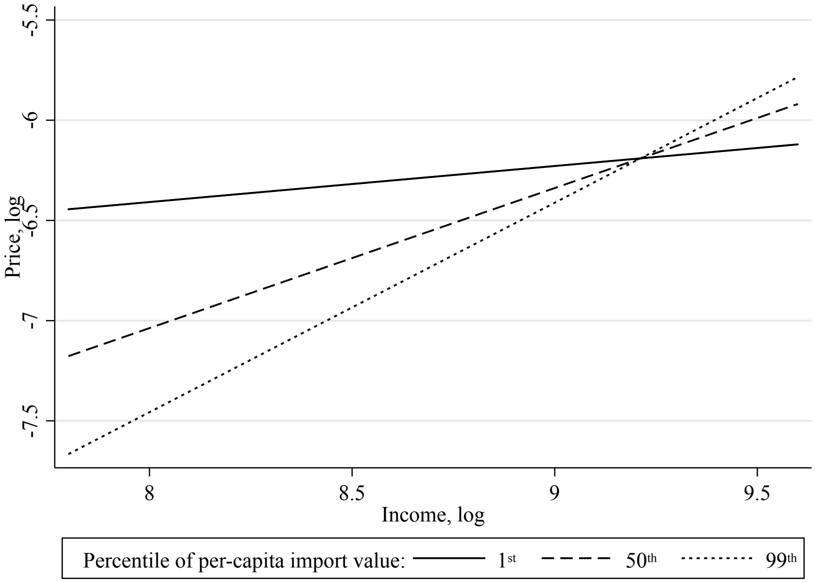
Notes. + $p < 0.1$; * $p < 0.05$; ** $p < 0.01$. Standard errors in all specifications are clustered at the level of included fixed effects. Every specification includes dummies for the population size of the primary sampling unit. All variables are in logarithms.

Table 4: Local imports and slopes of Engel curves (cost-of-living index).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Weighted	No	No	No	Yes	Yes	No	No	Yes	Yes
States	27	27	27	27	27	16	16	16	16
Nondurable exp., iht	0.945	1.057	1.020	1.225	1.202	1.115	1.081	1.084	1.078
	(0.457)*	(0.435)*	(0.486)*	(0.475)*	(0.449)**	(0.423)*	(0.512)*	(0.461)*	(0.517)*
Imports/cap, igt	-1.236	-1.118		-0.864		-1.113		-1.136	
	(0.337)**	(0.399)**		(0.382)*		(0.400)**		(0.410)**	
(Nondur.exp)×(Imports/cap)	0.127	0.123	0.119	0.095	0.086	0.124	0.122	0.127	0.124
	(0.038)**	(0.044)**	(0.042)**	(0.043)*	(0.039)*	(0.043)**	(0.040)**	(0.045)**	(0.040)**
Cost of living, i	4.151								
	(1.249)**								
(Nondur.exp)×(cost of living)	-0.503	-0.512	-0.491	-0.451	-0.418	-0.527	-0.516	-0.531	-0.519
	(0.142)**	(0.173)**	(0.153)**	(0.155)**	(0.141)**	(0.181)**	(0.150)**	(0.188)**	(0.150)**
State pop, it	0.143								
	(0.027)**								
Median h.h. income, it	0.228								
	(0.119)+								
Family size, iht	-0.208	-0.213	-0.206	-0.215	-0.209	-0.207	-0.199	-0.204	-0.197
	(0.022)**	(0.022)**	(0.012)**	(0.022)**	(0.013)**	(0.022)**	(0.012)**	(0.022)**	(0.013)**
Age of ref. person, iht	0.046	0.053	0.055	0.042	0.045	0.064	0.065	0.043	0.045
	(0.027)+	(0.031)+	(0.019)**	(0.033)	(0.020)*	(0.030)*	(0.019)**	(0.031)	(0.019)*
Housing exp., iht	0.051	0.058	0.060	0.050	0.053	0.064	0.066	0.058	0.059
	(0.009)**	(0.012)**	(0.007)**	(0.011)**	(0.007)**	(0.012)**	(0.007)**	(0.012)**	(0.007)**
C.D.F-stat	262.6	146.1	402.0	147.2	444.3	133.4	427.8	128.3	408.9
N.obs.	67,297	67,297	67,289	67,297	67,289	67,060	66,964	67,060	66,964
FE	gt	gt it	git	gt it	git	gt it	git	gt it	git

Notes. + $p < 0.1$; * $p < 0.05$; ** $p < 0.01$. Standard errors in all specifications are clustered at the level of included fixed effects. Every specification includes dummies for the population size of the primary sampling unit. All variables are in logarithms.

Figure 10: Local imports and slopes of Engel curves



Notes. The range of income is between the $5th$ and the $95th$ percentile of expenditures on nondurables.

4.4 Caveats and discussion

There are two reasons stemming from data limitations that may weaken our results: local import substitution and imperfect between location of imports and location of consumption.

4.4.1 Substitution with domestic production

We do not have data on domestic production by state. The concern here is that domestic production in the states that face higher barriers to importation would develop domestic production to substitute for lacking imports. In extreme, this substitution would exactly offset differences in imports making them irrelevant. Such a limiting pattern would work against us detecting any effect of imports. In order to see if this is a problem we related total expenditures from CE data with the total imports.

$$\ln(\widehat{imports})_{igt} = \underset{(0.030)^{***}}{0.83} \ln(expenditure)_{igt} + \hat{\gamma}_{gt}$$

The elasticity is significant and sufficiently close to one allowing us to exclude the possibility of such a drastic import substitution pattern. Note, this pattern of substitution would worry us only if the relation between expenditure and imports was not significant. The absolute size of import penetration is not important if it is similar or constant across states.

4.4.2 Matching the location of consumption with the import destination

Our identification is based on a match of the location of the import arrivals and location of consumption. It is not necessarily the case that the importation documents contain destination state of consumption. A product may be shipped to Oklahoma but consumed in Texas. Such trans-shipments would introduce a measurement error into our measure of trade and potentially attenuate our results. As it was in the case with the domestic production, the specification with state-product-year fixed effects provide a robustness check against this possibility.

4.4.3 Representativeness of the CE survey

Since our analysis is aimed at recovering population parameters we include all households in our analysis. It should be noted here that CEX sample is regionally representative for four of the aggregated Census regions. It is not representative at the state level. For the purposes of our analysis it is important that the sample is not stratified by state but it is also unbiased. In our case we just need to connect a consumption shifter (internal distance or trade) with the location of consumption. We are not constructing state-level estimate, we are uncovering a population parameter that depends on spatial variation in consumption location. For this reason, for us it is only necessary that the sample be unbiased, and it is. According to the documentation provided with the CE survey public use interview files: "the CE sample was not designed to produce precise estimates for individual states. Although state-level estimates that are unbiased in a repeated sampling sense can be calculated for various statistical measures, such as means and aggregates, their estimates will generally be subject to large variances."

5 Conclusion

Our results are consistent with the theory we propose. There is however a possibility that a model with differences in the number of available varieties may help explain differences in slopes of the quality Engel curves by affecting the set of available varieties. While theoretically possible, such possibility would be difficult to detect with our data because at the level of CE products there is not a lot of variation in the set of available varieties across states. Not surprisingly, our experiment with extensive margin in the quality Engel curve regressions was not fruitful. Related to the previous point, the measures of trade and distance may subsume important differences in the quality of imported varieties. For example, imports of high quality custom bicycles may be less responsive to distance as they might be arriving from Europe instead of East Asia and by air instead of vessels. As it is in the case with extensive margin, features of our data preclude us of successfully differentiating routes by the source country depending on the products. This is because the FAF data does not offer detail on the country of origin.

One useful implication of our results is that the differences in import access can affect how rich and poor respond to increases in income through fiscal policy. This is because a household's ability and willingness to consume will be affected by the access to imports.

Data appendix

5.1 Consumer Expenditure Survey

Consumer expenditure survey is a nationally representative survey of purchases by households. More information is available from `https://www.bls.gov/cex/`. We use the interview files that record purchase of durables. For each household we know the month of purchase, the amount paid, state where the household lives, total expenditures on nondurables by that household.

5.1.1 Identification of state

To identify location of consumption we use state codes (STATE variable) form the CE interview files. CE covers 46 states. CE does not cover New Mexico, North Dakota, Iowa, Wyoming. Rhode Island, Oklahoma, and Virginia are listed among available states in User's Documentation files but there is no data on durables consumption in our sample. however not all observations have a reliable measure of state code. The state code is not available for Arkansas, Mississippi, Montana, North Carolina, South Dakota (these states are not included). In other cases the state may have been recoded for another state (either some of this state's observations have been recoded, or it contains some re-codes from other states). These states are also excluded from the estimation. These states include: Alabama, Delaware, Georgia, Maryland, Minnesota, Wisconsin. We also exclude Alaska and Hawaii because of their remoteness from the other U.S. states, which makes it difficult to characterize their internal and external distance in the same way as we can for other states. We also eliminate three smallest states by total expenditure on nondurables captured in the sample: West Virginia, Delaware, and District of Columbia. The largest of them accounts for just over 10th of a percent of total expenditures on nondurables. This leaves us with the following 27 states: Arizona, California, Colorado, Connecticut, Florida, Idaho, Illinois, Indiana, Kansas, Kentucky, Louisiana, Maine, Massachusetts, Missouri, Nebraska, Nevada, New Hampshire, New Jersey, New York, Ohio, Oregon, Pennsylvania, South Carolina, Tennessee, Texas, Utah, Washington.

5.1.2 Defining nondurable expenditure

In order to calculate expenditures on nondurables we use total expenditures on food, alcoholic beverage, tobacco, apparel, footwear, reading materials, and gasoline.

Every household is interviewed four times in five consecutive quarters (e.g. December, March, June, September). Some households do not participate in all interviews. In order to construct the instruments the measure of the past expenditures in the case of missing interviews we used two earliest if available to instrument for the most recent.

5.1.3 Defining housing expenditure

We measure household expenditures on housing by adding two expenditure aggregates available from the CE survey: EOWNDWLP - owned home outlays last quarter including mortgage principal and interest, property taxes, maintenance, insurance, and other expenses. (it is obtained by summing up two other aggregates OWNDWEPQ and EMRTPNOP) and RENDWEPQ - rented dwelling last quarter.

5.2 Product code cross-walks

Concordance between the Harmonized System codes and the Standardized Classification of Transported Goods used in FAF comes from Appendix E of Southworth, Frank, et. al., The Freight Analysis Framework Version 3 (FAF3) A Description of the FAF3 Regional Database And How It Is Constructed, prepared for the Federal Highway Administration, 2011. URL:`http://faf.ornl.gov/faf3/Data/FAF3ODDoc611.pdf` accessed August 7, 2014. Bureau of Transportation Statisics has a statement about the SCTG codes and their compatibility with other product classifications `https://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/commodity_flow_survey/html/classification.html`.

Concordance between the harmonized system and CE product codes was constructed manually using verbal product descriptions from USITC `https://hts.usitc.gov/` and U.S. Census Bureau `https://uscensus.prod.3ceonline.com/`.

Appendix: Proofs

Proof of Proposition 1:

$$\overline{s}_{gr} = \frac{1}{1 + \left[\frac{\frac{\overline{\sigma}_g-1}{\overline{\sigma}_g}\overline{\lambda}_g}{\frac{\sigma_0-1}{\sigma_0}(\overline{c}_g+\tau)}\right]^{-\overline{\sigma}_g}\left[\frac{\frac{\underline{\sigma}_g-1}{\underline{\sigma}_g}\underline{\lambda}_g}{\frac{\sigma_0-1}{\sigma_0}(\underline{c}_g+\tau)}\right]^{\underline{\sigma}_g} x_0^{\frac{\underline{\sigma}_g-\overline{\sigma}_g}{\sigma_0}}} \tag{10}$$

$$\frac{\partial \overline{s}_{gr}}{\partial \tau} = \frac{\left[\frac{\frac{\overline{\sigma}_g-1}{\overline{\sigma}_g}\overline{\lambda}_g}{\frac{\sigma_0-1}{\sigma_0}(\overline{c}_g+\tau)}\right]^{-\overline{\sigma}_g}\left[\frac{\frac{\underline{\sigma}_g-1}{\underline{\sigma}_g}\underline{\lambda}_g}{\frac{\sigma_0-1}{\sigma_0}(\underline{c}_g+\tau)}\right]^{\underline{\sigma}_g} x_0^{\frac{\underline{\sigma}_g-\overline{\sigma}_g}{\sigma_0}}\left[\frac{\underline{\sigma}_g(\overline{c}_g+\tau)-\underline{\sigma}_g(\underline{c}_g+\tau)}{(\overline{c}_g+\tau)(\underline{c}_g+\tau)}\right]}{\left\{1 + \left[\frac{\frac{\overline{\sigma}_g-1}{\overline{\sigma}_g}\overline{\lambda}_g}{\frac{\sigma_0-1}{\sigma_0}(\overline{c}_g+\tau)}\right]^{-\overline{\sigma}_g}\left[\frac{\frac{\underline{\sigma}_g-1}{\underline{\sigma}_g}\underline{\lambda}_g}{\frac{\sigma_0-1}{\sigma_0}(\underline{c}_g+\tau)}\right]^{\underline{\sigma}_g} x_0^{\frac{\underline{\sigma}_g-\overline{\sigma}_g}{\sigma_0}}\right\}^2}. \tag{11}$$

The equation above is positive if and only if

$$\frac{\underline{\sigma}_g(\overline{c}_g + \tau) > \underline{\sigma}_g(\underline{c}_g + \tau)}{(\overline{c}_g + \tau)(\underline{c}_g + \tau)}$$

□

Proof of Proposition 2. The elasticity of the Alchian-Allen effect is then given by:

$$\frac{\partial \overline{s}_{gr}}{\partial \tau}\frac{\tau}{\overline{s}_{gr}} = \frac{\left[\frac{\underline{\sigma}_g(\overline{c}_g+\tau)-\underline{\sigma}_g(\underline{c}_g+\tau)}{(\overline{c}_g+\tau)(\underline{c}_g+\tau)}\right]}{\left[\frac{\frac{\overline{\sigma}_g-1}{\overline{\sigma}_g}\overline{\lambda}_g}{\frac{\sigma_0-1}{\sigma_0}(\overline{c}_g+\tau)}\right]^{\overline{\sigma}_g}\left[\frac{\frac{\underline{\sigma}_g-1}{\underline{\sigma}_g}\underline{\lambda}_g}{\frac{\sigma_0-1}{\sigma_0}(\underline{c}_g+\tau)}\right]^{-\underline{\sigma}_g} x_0^{\frac{\overline{\sigma}_g-\underline{\sigma}_g}{\sigma_0}} + 1}, \tag{12}$$

from which it is easy to see that conditional on the numerator being positive, the elasticity decreases in x_0. □

References

Atkin, David, & Donaldson, Dave. 2015. Who?s Getting Globalized? The Size and Implications of Intra-national Trade Costs. *Manuscript*.

Attanasio, Orazio P., & Pistaferri, Luigi. 2016. Consumption Inequality. *Journal of Economic Perspectives*, **30**(2), 3–28.

Autor, David H., Dorn, David, & Hanson, Gordon H. 2013. The China Syndrome: Local Labor Market Effects of Import Competition in the United States. *American Economic Review*, **103**(6), 2121–68.

Bils, Mark, & Klenow, Peter J. 2001. Quantifying Quality Growth. *The American Economic Review*, **91**(4), 1006–1030.

Borcherding, Thomas E, & Silberberg, Eugene. 1978. Shipping the Good Apples Out: The Alchian and Allen Theorem Reconsidered. *Journal of Political Economy*, **86**(1), 131–138.

Broda, Christian, & Romalis, John. 2009. The Welfare Implications of Rising Price Dispersion. *Manuscript*.

Broda, Christian, & Weinstein, David E. 2006. Globalization and the Gains from Variety. *The Quarterly Journal of Economics*, **121**(2), 541–585.

Choi, Yo Chul, Hummels, David, & Xiang, Chong. 2009. Explaining Import Quality: the Role of the Income Distribution. *Journal of International Economics*, **77**, 265–276.

Coçar, A. Kerem, & Fajgelbaum, Pablo D. 2016. Internal Geography, International Trade, and Regional Specialization. *American Economic Journal: Microeconomics*, **8**(1), 24–56.

Dix-Carneiro, Rafael, & Kovak, Brian K. 2015. Trade Liberalization and the Skill Premium: A Local Labor Markets Approach. *American Economic Review*, **105**(5), 551–557.

Faber, Benjamin. 2014. Trade Liberalization, the Price of Quality, and Inequality: Evidence from Mexican Store Prices. *Mimeo.*

Fajgelbaum, Pablo, Grossman, Gene M., & Helpman, Elhanan. 2011. Income Distribution, Product Quality, and International Trade. *Journal of Political Economy*, **119**(4), 721 – 765.

Fajgelbaum, Pablo D., & Khandelwal, Amit K. 2016. Measuring the Unequal Gains from Trade. *The Quarterly Journal of Economics*, **131**(3), 1113–1180.

Fieler, Ana Ceclia. 2011. Nonhomotheticity and Bilateral Trade: Evidence and a Quantitative Explanation. *Econometrica*, **79**(4), 1069–1101.

Goldberg, Pinelopi Koujianou, & Pavcnik, Nina. 2007. Distributional Effects of Globalization in Developing Countries. *Journal of Economic Literature*, **45**(1), 39–82.

Hummels, David, & Skiba, Alexandre. 2004. Shipping the Good Apples Out? An Empirical Confirmation of the Alchian-Allen Conjecture. *Journal of Political Economy*, **112**, 1384–1402.

Nicita, Alessandro. 2009. The price effect of tariff liberalization: Measuring the impact on household welfare. *Journal of Development Economics*, **89**(1), 19–27.

Porto, Guido G. 2006. Using survey data to assess the distributional effects of trade policy. *Journal of International Economics*, **70**(1), 140–160.