Import Demand in Differentiated Products

Torsten Søchting Jaccard
Dartmouth College

March 24, 2023

Abstract

This paper links product country-of-origin data to US household purchase records and estimates a model of import demand with flexible consumer heterogeneity and attribute-based import substitution patterns. Counterfactual price shocks applied to all imported varieties reveal concentrated costs in urban US counties, while China-specific price shocks are strongly anti-rural. I provide evidence that differences in retail market composition drive these disparate outcomes by acting both in addition to, and independently of, household income. I provide evidence that import substitutability is greater than under typical Armington assumptions and justify the modeling choice of this paper by providing evidence that variety attributes – such as quality and markups – are highly separable between the location of production and the firm which designed that variety.

This paper forms a chapter of my thesis, and I am indebted to my supervisor - Daniel Trefler - and committee members Victor Aguirregabiria and Peter Morrow for their support and guidance on this project. I have benefitted from discussions with Loren Brandt, Kevin Lim, Pamela Medina-Quispe, and members of the globalization cluster at Dartmouth College. I thank Sebastian Stumpner for generously providing data, and the Rotman China Initiative for financial support. Adam Morrison provided excellent research assistance. The analysis in this paper is my own, calculated based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at the University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are my own and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein. All remaining errors are my own.

1 Introduction

One cannot quantify the causes and consequences of international trade without first positing a model of import demand. In practice, this requires that the researcher decide on two *a priori* specifications: whether, and how, to incorporate domestic consumer heterogeneity and what model of import substitution is most appropriate for the setting in question. These specifications are crucial to analyses as fundamental as quantifying the consumption gains from trade, evaluating the distributional consequences of trade shocks, recovering the cost structure of imported goods, and modeling the competitive effects of trade. And yet there is remarkably little empirical evidence available to guide researchers in either dimension.

Consider the extent to which households differ in their relative demand for imported goods. A substantial body of literature has documented that low-income households spend a greater share of their income on tradeable sectors – such as manufactures and agricultural goods – when compared to high-income households (Fajgelbaum and Khandelwal [2016]; Carroll and Hur [2020]). Yet this finding is entirely compatible with a world in which low-income households do not spend a single dollar of expenditure on imported goods. If low-income households *only* consume domestically produced varieties *within* manufactures and agricultural goods, then sector-level expenditure shares offer little information regarding consumption-based exposure to trade shocks[1]. This concern is compounded by the fact that while sector-level expenditure shares can be reasonably expected to correlate with income, it is not immediately clear that income is the dimension of heterogeneity most relevant for understanding the distribution of *within*-sector import expenditure. What is needed are data linking detailed household characteristics to expenditure on imported varieties, rather than expenditure on tradeable sectors[2].

Specifying a relevant model of import substitution is similarly fraught with difficulties. Almost all current approaches assume that the substitutability of varieties is heavily influenced by the country in which those varieties were produced. Yet this assumption is problematic given the primacy of multinational firms in observed trade flows (Bernard et al. [2012]; Bernard et al. [2018]). Consider an American multinational relocating production from a plant in Texas to a *maquiladora* across the Mexican border. The *a priori* assumption that imported and domestic varieties are poor substitutes provides little guidance as to whether the variety in question should be considered as "American" or "Mexican". In the absence of detailed attribute data for each imported and domestic variety bought and sold within the domestic market one cannot construct a more intuitive model of import substitution.

[1] Throughout this paper I will refer to varieties as unique barcodes within a given product category. I follow Faber and Fally [2021] and define a firm as a brand.

[2] A number of papers study plausibly exogenous trade shocks to infer expenditure exposure to trade shocks, such as: Cravino and Levchenko [2017]; Bai and Stumpner [2019]; Cravino and Levchenko [2018]; Amiti et al. [2020]; Jaravel and Sager [2019]; Hottman and Monarch [2021]; and Auer et al. [2021]. This paper instead estimates a demand model in order to recover predicted welfare costs of trade shock simulations, as in Faber and Fally [2021] and Borusyak and Jaravel [2021].

In this paper, I make headway in addressing both of these concerns by providing a dataset linking individual barcodes to their country of origin. These data constitute all text on the packaging of approximately 65,000 unique barcodes within the personal care/cosmetics category. Packaged consumer goods sold within the US are required to provide some statement equivalent to "Made in . . .", which allows for the recovery of each barcode's production location and I link these data to purchase records and socio-economic characteristics for over 42,000 US households. I leverage the detailed nature of this dataset to estimate within-category import demand models for each category present in the data, and apply counterfactual price shocks across import origins to recover both the distributional costs of such shocks across households as well as elasticities of demand associated with each origin country. The detailed nature of these data naturally come at the expense of scope: the categories studied in this paper constitute approximately $240 of expenditure per household-year. The goal of this paper is therefore not to provide aggregate welfare statements but rather to provide an evaluation of assumptions which are common in the trade literature regarding import demand and in doing so provide evidence for biases which may be unobserved, or difficult to evaluate, in more aggregated data settings.

Within narrow product categories, wealthier households purchase varieties from wealthier origin countries. Conditional on income and a battery of socio-economic characteristics, I find that urban households exhibit greater import expenditure shares than rural households but that this pattern is completely reversed for expenditure on varieties produced in China. I show that the urban/rural dimension of consumer heterogeneity is quantitatively important and provide cross-sectional evidence that the rural bias of China import expenditure is heavily influenced by the prevalence of dollar stores in rural US ZIP codes. I therefore incorporate both income and population density when placing households into types and estimate separate model parameters for each type, thereby allowing for an unrestricted distribution of demand elasticities across the income-density space.

In constructing a model of import substitution, I leverage both the detailed text data available for each barcode as well as intuition from the discrete choice demand literature. The dataset used in this paper provides all text on the packaging of each barcode, which allows for a uniquely granular approach to understanding product differentiation. However it is not immediately obvious how to make use of descriptive text within the context of a random coefficients logit model of demand. Instead, I make use of text similarity measures and unsupervised clustering algorithms to place varieties into market segments based on the similarity of descriptive text on their packaging. These clusters are then implemented as nests in a Nested Logit demand system, which I refer to as an attribute-cluster model[3].

I use the estimated model to run counterfactuals in which country-specific varieties experience a relative price shock. These counterfactual price shocks are partial equilibrium in that I maintain the set of varieties and retail market characteristics fixed. The goal of this exercise is therefore

[3]While the core intuition of this model is standard in the discrete choice literature (McFadden [1974]; Berry [1994]; Berry et al. [1995]; Cardell [1997]), the application of this intuition to questions of international trade is rare. Notable exceptions include Goldberg [1995]; Head and Mayer [2019]; and Coşar et al. [2018].

to recover *ex post* the distributional short-run costs of import price shocks as well as patterns of import substitution, rather than the long-run effects of a persistent change in trade technology or policy[4]. A relative price shock to all imported varieties has concentrated welfare costs in the urban cores of American cities, with these counties exhibiting welfare costs that are 16%-26% greater than rural and low-income counties. Similar patterns hold for a price shock to Canada, Mexico, and European produced varieties, with the relative difference in costs between wealthy-urban counties and poor-rural counties increasing to 52% in the case of a Europe price shock. A price shock to imports from China exhibits the exact opposite pattern, with poor-rural counties exhibiting relative costs that are effectively double those experienced by their wealthy and urban counterparts. The county-level correlation of costs between a Europe and China price shock is -0.80, suggesting that the distributional outcomes of import price shocks depend strongly on the origin country receiving this shock.

When compared to previous estimates of the distributional costs of import price shocks, I highlight the importance of two novel elements in this analysis. First, I show that omitting the urban/rural dimension of consumer heterogeneity and relying solely on income heterogeneity when defining import demand leads to an underestimate of the variance in costs across US counties by a third, when averaged across shocks. These findings suggest that studies asking whether or not a trade shock is "pro-poor" or "pro-wealthy" – as in Hottman and Monarch [2021], Borusyak and Jaravel [2021], and Auer et al. [2021] – likely underestimate the true variance in costs across households.

Second, I show that expenditure-based approximations of the welfare costs associated with import price shocks exhibit non-linear bias when compared to a model with both expenditure and elasticity heterogeneity. Specifically, I show that elasticities of demand are high for both low-income households and for households in the most urban ZIP codes of the US. These two household types also exhibit, respectively, the lowest and highest import expenditure shares. Relying solely on an expenditure-based approximation of welfare costs would overestimate the relative costs for both of these household types. This non-linear bias is economically relevant in that the rank order of county-level costs in response to an aggregate import price shock changes when comparing an expenditure-based approximation to the full model. These results build on those found in Hottman and Monarch [2021] and Auer et al. [2021] by considering elasticity and expenditure heterogeneity along the urban/rural divide.

The combination of an attribute-based model of import substitution and flexible consumer heterogeneity allows for an *ex post* study of how demand elasticities differ across origin countries. Cross-price elasticities are determined by the similarity in descriptive text on the packaging of varieties while own-price elasticities are recovered by calculating the average demand elasticity of a

[4]This approach is in line with Borusyak and Jaravel [2021] and Faber and Fally [2021]. While I incorporate a pricing model which allows for strategic within-firm pricing decisions, I do not model the process through which multinationals choose production locations for each individual variety. The goal of this exercise is therefore to recover the demand curves facing imported varieties, conditional on the attributes of that variety.

given variety's representative consumer. In both cases, the relationship between demand elasticities and origin country characteristics is left completely unspecified *a priori*.

I estimate Armington elasticities – the elasticity of domestic market share with respect to changes in import prices – that are 19% greater than an alternative specification in which I nest varieties based on their production origin, as is common in the trade literature (often referred to as an "Armington" model)[5]. Interestingly, the Armington elasticity recovered under the attribute-clustering modeling of this paper is almost identical to the elasticity recovered under a number of alternative nesting assumptions including firm-specific nests, price deciles, and quality deciles.

That alternative substitution models yield stronger substitutability between domestic and foreign varieties than an Armington model is not all that surprising given that the Armington assumption represents a limit case: imported varieties are defined tautologically to be weak substitutes with domestically produced varieties. This assumption is often difficult to test, as it is rare to study markets with sales and price data for both foreign and domestic varieties. The dataset used in this paper, however, provides exactly such an opportunity, and I make use of this feature to test directly whether observed attributes at the variety level are primarily determined by the country in which that variety was produced.

By combining observed and model-derived barcode-level attributes – such as the own-price elasticity, marginal cost, mark-up, per-unit profit, quality, and total variable profit – with a data-driven distinction between the "production" and "design" origin countries of each barcode, I test whether barcode-level attributes are separable between production and design origins[6]. Conditional on production location, I find that for every 10% increase in the GDP per capita of a variety's *design* origin, the own-price elasticity of this variety decreases by 0.89%, marginal cost increases by 2.47%, mark-ups increase by 1.96%, per-unit profits increase by 5.99%, quality increases by 17.54%, and total variable profits increase by 18.40%[7]. For context, these estimates suggest that a variety designed by an American firm but produced in Mexico would exhibit markups, quality, and variable profits which are 15%, 2.54 times, and 2.71 times greater than a variety designed *and* produced in Mexico. This paper therefore provides some of the first direct evidence for a phenomon I term "attribute separability": the capacity for multinationals to endow the varieties they produce with attributes which are separable from the location/country in which they produce those varieties.

Taken together, the import demand model in this paper suggests three key shortcomings in our understanding of within-category import demand elasticities. First, imported varieties are more dispersed in the attribute space than one would estimate under standard Armington-style assumptions, which increases the relative substitutability of imported and domestic varieties. Second,

[5] The following papers rely on this Armington structure in estimating consumption outcomes associated with trade shocks: Amiti et al. [2020]; Hottman and Monarch [2021]; Borusyak and Jaravel [2021]; and Auer et al. [2021]. Feenstra et al. [2018] provide an overview of the literature estimating Armington elasticities.

[6] By "design origin", I refer to the origin country of the firm/brand which produced a given barcode. As an example: all barcodes produced by l'Oreal have a design origin in France regardless of their production location.

[7] I define quality as a "demand-shifter". That is, demand conditional on price. This definition is identical to Khandelwal [2010].

revealed prefence estimates of variety-level demand elasticities exhibit a negative relationship between the origin country GDP per capita of an imported variety and the elasticity of that variety's demand curve in the US market. Third, regardless of production location, varieties designed in wealthier origin countries exhibit significantly more inelastic demand than varieties designed in lower income origin countries, which suggests that the negative relationship between origin country income and the "fundamental" elasticity of demand associated with that origin country is even stronger than the relationship one would observe due to the presence of multinationals.

The goal of this paper is to guide future research by studying a narrow context and in doing so empirically establish specific pitfalls or shortcomings that may arise in more aggregated data settings. To this end, I consider the contributions of this paper four-fold. First, the expenditure-based exposure of households to trade shocks depends as much on the urban/rural divide as income, with geography and income both acting independently and in concert to shape import expenditure. Second, relying on expenditure shares alone to estimate household exposure to trade shocks leads to potentially severe non-linear bias when compared to a model which allows for consumer heterogeneity in both expenditure and elasticities of demand. Depending on the shock in question, this bias may be large enough to alter the rank ordering of costs across households and counties. Third, import substitution models which assume *a priori* that the production location of a variety is the defining attribute of that variety likely underestimate import substitutability by as much as 20% when compared to attribute-based alternatives. And fourth, I provide novel evidence that observed variety attributes – such as demand elasticities, markups, and quality – are in fact separable between production location and design location, which further emphasizes the need for attribute-based models of import substitution as well as greater recognition of the role played by multinationals in shaping trade elasticities.

The discussion of attribute separability in this paper provides two important caveats to common assumptions made in the trade literature. First, I find that estimates of the "fundamental" quality of imported varieties versus observed quality in the presence of multinationals exhibit much greater variation across origin countries. Specifically, I find that the presence of multinationals leads to an overestimate of the quality of varieties imported from China and Mexico by, respectively, a factor of 3.3 and a factor of 3.1. These findings have implications for our understanding of factor competition across quality ladder "rungs" in the presence of multinational activity. Second, the notion of attribute separability calls into question the common definition applied by researchers in empirical trade as to what constitutes a "variety". In the absence of more detailed data, researchers are often forced to assume that each category-by-country diad constitutes a variety[8]. The evidence provided in this paper, however, suggests that there is little reason to believe that an imported variety from China, for example, should be thought of as a "Chinese" variety. This observation

[8] This approach has been used in the pioneering works of Feenstra [1994] and Broda and Weinstein [2006], among others. When more detailed data are available, this definition is sometimes expanded to a firm-sector-country, as in Redding and Weinstein [2018]. Notice that the core problem still remains: one generally does not know whether this new firm is a US multinational with off-shored production or a legitimately "new" variety.

suggests a wedge between empirical attempts to quantify the gains from variety and the theory positing such gains, as described in Armington [1969], Krugman [1980], and Romer [1994].

This paper is organized as follows: Section 2 provides an overview of the data and two stylized facts which inform the demand system specification outlined in Section 3, where I also provide the results of model estimation. Section 4 provides conparative static results regarding the distributional costs of import price shocks while Section 5 pivots to studying how import demand elasticities differ across origin countries. Section 6 concludes and Appendix A provides additional tables and figures.

2 Data and Stylized Facts

This section provides an overview of the two key datasets used in this paper as well as two stylized facts which motivate the demand model to follow.

2.1 Data

Household Scanner Data: This paper uses the NielsenIQ Homescanner dataset. These data consist of a rotating panel of ∼60,000 American households. Each household records the price, date, and store that they visited for each shopping trip and the barcode-specific purchases made within each trip. These data include a full suite of socio-demographic information for each household. I follow Faber and Fally [2021] and proxy for household income using total expenditure per household member.

Barcode Country of Origin: A particular issue for applying the NielsenIQ data to questions of international trade is that barcodes do not contain explicit information as to their country of origin. I therefore merge the NielsenIQ Homescanner data with barcode-specific country-of-origin information purchased from Label Insight, Inc. (LI), a firm that specializes in extracting and organizing information found on the labelling of consumer packaged goods[9].

Label Insight uses an AI to extract from packaging the ingredients, branding, and any other text information that may be included for thousands of barcodes sold across the majority of retail chains in the US. Since imported goods in the US are required to contain some statement equivalent to "Made in . . .", the Label Insight AI incidentally recovers a country of origin for each barcode they collect. Naturally, Label Insight can only cover a segment of total consumption and their coverage is best for personal care products and packaged food. I therefore purchased the origin country, ingredients list, brand, and barcode description (as read directly off package) for approximately

[9] I am aware of one other paper that has used data from Label Insight in a trade/international macroeconomics context: Hottman et al. [2019] use similar data to estimate an Armington elasticity using retail data for 10 consumer packaged goods categories. Auer et al. [2021] link barcodes to an import/domestic classification for 8,689 products and 3,000 household purchase records using Swiss data, and Borusyak and Jaravel [2021] map US scanner data to brand-specific import propensities.

65,000 barcodes spanning 15 product categories within personal care and cosmetics[10].

Data Summary: The final raw dataset used in this analysis contains purchase-level observations that combine a barcode with a household, store, and date. I restrict my sample to the years 2015 - 2017 and only include households that are present in each year of that period[11].

This final dataset contains 42,074 households purchasing barcodes from 15 distinct personal care and cosmetic product categories over three years[12]. These purchases span 23,806 barcodes (1,689 brands) and 20.6 million USD in total expenditure (3.3 million units purchased). Given that observations are at the purchase level, these data comprise 2.75 million observations. Table A.1 provides an overview of the data for each category, whereas Table A.2 provides data on sales by origin country. The final merged dataset contains purchases from 44 distinct origin countries, with an overall import share of 14.4% which amounts to 3.0 million USD of expenditure on imported varieties over three years. By merging the NielsenIQ categories to the BEA's Consumer Expenditure Survey I find that US households spend, on average, $240 USD per year on the categories studied in this paper. The merged NielsenIQ-LI dataset used in this paper covers approximately 68% of all predicted expenditure within these categories, or $163 of expenditure per household-year.

2.2 Stylized Facts

(1) ZIP population density is as important as income to understanding import expenditure heterogeneity. Almost all studies to date assume *a priori* that income is the most relevant dimension of heterogeneity for understanding the distributional effects of trade shocks (Faber and Fally [2021]; Borusyak and Jaravel [2021]; Auer et al. [2021]). In this section I show that while income is a relevant dimension of differential household import expenditure, ZIP population density has just as much predictive power and acts independently of income.

Figure 1 provides estimates of a simple linear probability model in which the dependent variable is equal to one if a purchase in the NielsenIQ records is of a barcode from origin country m, and zero otherwise. I then estimate four separate models with dependent variable indicators equal to one, respectively, if the purchase was of any imported barcode, a barcode from Canada/Mexico, a barcode from China, or a barcode from Europe[13]. Observations are at the household-barcode-date level and I estimate the probability of purchasing a barcode from origin m across ten deciles of

[10] The US Customs and Border Protection (CBP) require that the country-of-origin printed on the label corresponds to the last country in which the good underwent a "substantial transformation".

[11] The LI data provides a current snapshot of the marketplace for personal care products and given the significant year-over-year turnover in varieties offered for these products, I find that years prior to 2015 exhibit a significant drop in merged barcodes. Changes to US trade policy began in 2018 which directly affected the categories studied in this paper so I choose 2017 as the final year of my sample.

[12] These categories are: Bathing Accessories, Body Soap, Dental Care, Deodorant/Perfume, Eye Cosmetics, Face Cosmetics, Facial Care, Hair Care, Hair Styling, Hand Soap, Lip Cosmetics, Nail Cosmetics, Shaving Care, Skin Care, and Toothpaste.

[13] I define "European" countries broadly speaking: all countries in the EU as well as the United Kingdom, Norway, Switzerland, and Iceland.

Figure 1: Purchase Propensity by Income and Origin

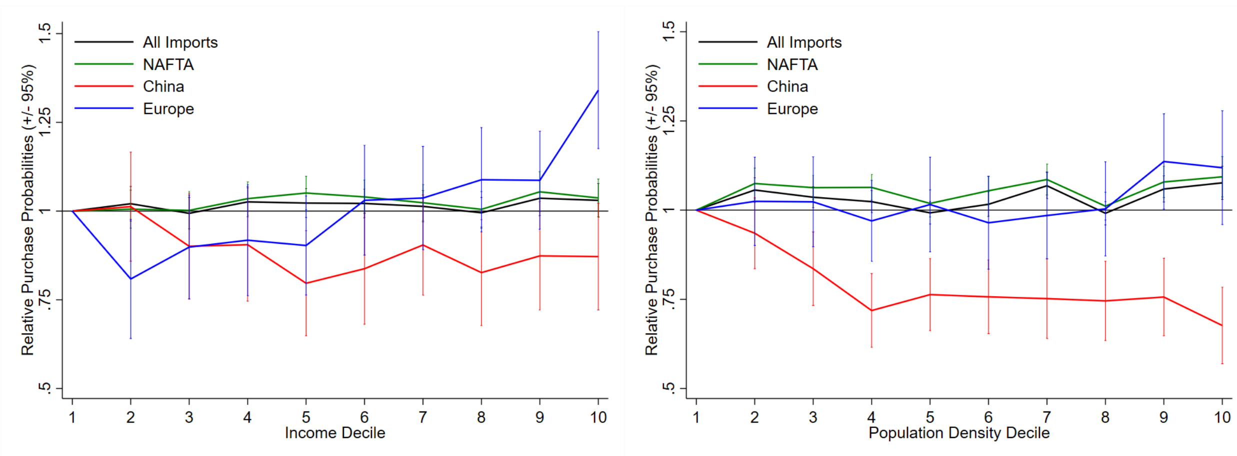


Figure 1 provides estimates of the relative probability of a given income decile (left) or population density decile (right) purchasing a barcode from a specific origin. All estimates represent output from a linear probability model with a dependent variable equal to one if a purchase is from origin m. Estimates reported as relative to the lowest income decile (left) and most rural density decile (right). All results include controls for education, race, age, presence of children, and married household heads. Region ($R = 5$), half-year, and product category fixed effects are also included. Obervations are weighted by total quantity of purchase, adjusted for weight. 95% confidence intervals are provided, and standard errors are clustered at the product-region-period level.

income and ten deciles of ZIP population density[14]. I additionally include controls for household education, race, composition, half-year fixed effects, and region fixed effects (I divide the US into five roughly equal regions). All models include product category fixed effects and so estimates should be interpreted as *within* product categories. The left panel provides the distribution of estimates across income deciles while the right panel provides the distribution of estimates across population density deciles (1 is most rural, 10 is most urban).

The relatively flat relationship between purchasing any imported variety and income matches results found in Borusyak and Jaravel [2021] and Auer et al. [2021], which is reassuring (Auer et al. [2021] find that for grocery purchases of Swiss households, the aggregate import share is effectively flat across the first four income quintiles but increases for the wealthiest households by 9% (Table 1, p.9)). I find that households in the wealthiest income decile are 34% more likely to purchase a European variety than households in the lowest income decile, but 13% *less* likely to purchase a Chinese variety. The positive gradient between European purchase probabilities and income, as well as the negative relationship between purchase probabilities and income for Chinese varieties, matches the general pattern of expenditure shares imputed by Borusyak and Jaravel [2021]. However the magnitude of differences is larger in my sample than what they find. While I discuss this discrepancy in greater detail in Section 4, and conclude that this is likely driven by

[14]Note that these estimates are conditional on each other and should be interpreted as the correlation between income and import purchase probabilities *conditional* on a household's ZIP population density (and *vice versa*).

the subset of product categories studied here, a key point of this paper is that neither Borusyak and Jaravel [2021] nor Auer et al. [2021] study the distribution of import expenditure across the urban/rural gradient, which is exactly the analysis described in the right panel of Figure 1.

The right panel of Figure 1 illustrates that conditional on income and a battery of other socio-economic characteristics, there is a strong negative correlation between the population density of a household's ZIP code and their probability of purchasing a variety produced in China. This pattern runs counter to the overall probability of purchasing any imported variety, which generally increases from rural to urban ZIP codes. The estimates in Figure 1 suggest that the urban/rural divide is an important dimension of heterogeneity to consider when mapping consumer heterogeneity in import expenditure shares. While a number of papers have looked at regional differences regarding expenditure exposure to trade shocks and generally found little in the way of consumer heterogeneity, the analysis here suggests that geography does play a key role in shaping import expenditure but across the urban/rural divide *within* regions, rather than necessarily *across* sub-national regions[15].

Consider three separate comparisons across households based on the estimates in Figure 1: the wealthiest to poorest income deciles, the most urban to most rural deciles, and a comparison between a household in both the most wealthy and most urban deciles versus a household in both the poorest and most rural deciles. For aggregate import purchase probabilities, these relative percentages are +3.0%, +7.6%, and +10.9%. For purchases of European-produced varieties, these relative percentages are +34.2%, +12.8%, and +51.3%, and for China-produced varieties these percentages are −12.8%, −32.0%, and −40.7%[16]. The standard deviation in purchase probabilities for both any imported variety and any Chinese variety is twice as large across density deciles than across income deciles, while the standard deviation of European variety probabilities is 2.5 times greater across income deciles than across density deciles.

These findings suggest that analyses asking whether origin-specific trade shocks are "pro-poor" or "pro-rich" may be missing a dimension of consumer heterogeneity that is at least as relevant as income to quantifying such outcomes: the urban/rural divide. Given the novelty of these findings, I provide descriptive evidence that changes in retail market characteristics across the urban/rural divide may provide a mechanism for understanding import expenditure heterogeneity in this dimension. Specifically, in Table A.3 in Appendix A I show that purchases at dollar stores are *seven* times more likely to be of a China-produced barcode than purchases made at a grocery store, conditional on all available household characteristics. Figure A.1 provides probability estimates of shopping at various retail formats across density deciles and illustrates that urban households are three times more likely to shop at a grocery store than rural households, but half as likely to shop at a dollar store. As a back-of-the-envelope calculation, these estimates suggest that based on retail market

[15]These papers include: Cravino and Levchenko [2018]; Bai and Stumpner [2019]; and Borusyak and Jaravel [2021]. Only Cravino and Levchenko [2018] find some relevance of region-level heterogeneity, with low-income households in Mexico facing differential exposure to import price shocks across regions.

[16]These estimates are effectively invariant to omitting either income or population density and estimating the unconditional distribution of purchase probabilities across either income or population density.

composition alone, urban households should be 35% less likely than rural households to purchase varieties produced in China. In fact this prediction almost exactly matches the true estimate in Figure 1 of a 32% differential.

These results suggest a more nuanced definition of household heterogeneity will be useful when specifying the demand model to follow, and I show in Section 4 that incorporating the urban/rural dimension of heterogeneity in the estimated model is quantitatively important to understanding the distributional consequences of import price shocks.

(2) 64% of import expenditure accrues to American-designed varieties, while 15% of expenditure on domestically produced varieties accrues to foreign firms. In order to specify a model of import demand, the researcher must take a stand on the model of substitution to implement *a priori*. It is exceedingly common in the trade and international macroeconomics literature to assume an Armington structure of substitution in which the production origin of a variety is the deciding factor in understanding the substitutability of that variety with all others. This assumption is hard to reconcile with the activity of multinationals: if an American firm off-shores production to Mexico, should their varieties be considered "American" or "Mexican"?

This question will form a central theme of this paper and the demand model to follow. As motivation, I make use of the detailed barcode-level data in this paper to define two separate origin countries for each barcode: a production origin and a design origin. The production origin will simply be where the barcode was produced, and matches the usual definition of an "origin" when studying trade from a customs data perspective. To arrive at a design origin for each barcode, I manually link the ∼1,700 firms/brands in my final dataset to a brand country of origin. This is either the headquarter location of that firm or – when these data are not available – the country in which that brand was founded.

Table 1 provides a decomposition of expenditure into these two definitions of variety origin. Each element of the table provides the percentage of expenditure accruing to that combination of production and design origin, with columns representing production origin and rows representing design origin. 85.5% of all expenditure accrues to domestically produced varieties and 14.5% accrues to foreign-produced varieties. This 14.5% constitutes the import share of expenditure one would infer using standard datasets. However by design location, these percentages change to 81.8% and 18.2%, respectively, which suggests that the penetration of foreign firms in the US market is over 25% greater than what one would estimate using customs data[17]. Lastly, I find that 9.3% of all expenditure accrues to domestically designed varieties produced off-shore, which constitutes 64% of all import expenditure[18].

The difference between production-based expenditure shares and design-based expenditure shares is even starker when calculated for individual origins. Based on production location, Canada,

[17] Blonigen and Soderbery [2010] find a similar result studying passenger vehicles.

[18] These data provide some of the first expenditure-based estimates at the intersection of multinationals and import consumption (Ramondo [2014]; Tintelnot [2017]; and Head and Mayer [2019]).

Table 1: Expenditure by Production and Design Origin

Design Origin	Production Origin Domestic	Production Origin Foreign	
Domestic	72.5	9.3	81.8
Foreign	13.0	5.2	18.2
	85.5	14.5	

Table 1 provides aggregate expenditure shares across all categories decomposed by the production and design origin of each barcode. Production origin refers to the location in which that barcode was produced, and design origin refers to the origin country of the brand/firm associated with that barcode.

Mexico and China account for 10.9% of all expenditure in my dataset. France, Germany, Italy and the UK account for only 1.2%. However when calculated using a design-based expenditure share, Canada, Mexico, and China account for only 0.2% of expenditure whereas the expenditure share for France, Germany, Italy, and the UK increases to 11.4%. The expenditure-weighted average origin country GDP per capita of varieties by production origin is $32 300 USD. By design origin, this value increases to $43 700 USD.

These results are striking in that they suggest a flaw in using customs data to study the variety gains from trade: increasing trade flows from China to the US do not necessarily imply an increase in the flow of Chinese-designed varieties to the US. Figure A.2 in Appendix A illustrates this fact directly by plotting the difference in the average design-origin GDP per capita of varieties exported from production origin m to the USA and production origin m's GDP per capita. There is effectively no relationship between these two quantities, implying that almost all varieties exported to the USA are designed in similarly wealthy countries regardless of production location[19].

Notice that these stylized facts suggest a tension between implementing a firm-based substitution model – as in Hottman et al. [2016] and Faber and Fally [2021] – and the standard production-based Armington model, as in Borusyak and Jaravel [2021] and Auer et al. [2021]. I find that 15.5% of all firm-category pairs sell varieties produced in more than one origin country and these same firms constitute 66.5% of aggregate expenditure. This poses a problem for specifying a model of substitution: should substitutability be based on production origin or design origin? I take the findings in this section as motivation to approach this problem from the perspective of attributes, and in doing so align this analysis more closely with standard techniques in the industrial organization literature. I turn to specifying this model now.

[19] Table A.2 provides a breakdown of expenditure shares across all countries by production and design.

3 Estimating a Model of Import Demand

Rather than relying on product-based demand systems - such as the CES or AIDS - I leverage the detailed data of this paper to estimate a demand system in the variety attribute space. I opt for the Nested Logit demand system for two reasons. First, in the absence of non-price continuous attributes the random coefficients logit model - as in Berry et al. [1995] - is as general as the Nested Logit (Grigolon and Verboven [2014]). The only non-price continuous attribute present in the data is the weight of a given variety. Second, the data presented in this paper contain descriptive text from the packaging of each barcode. These data provide a remarkably detailed approach to understanding variety differentiation but do not lend themselves easily to the Random Coefficients Logit model[20].

In addition, I only estimate the demand model introduced below for eight of the 15 product categories available in this paper: Hair Care, Body Soap, Skin Care, Deodorant, Toothpaste, Facial Care, Hair Styling and Face Cosmetics. Together, these eight categories represent 80% of all expenditure covered by the merged dataset. The categories not included mainly consist of cosmetics and often exhibit barcode-level own-price elasticities less than one. While it is intuitice that cosmetics would exhibit inelastic demand, these categories are not amenable to standard analyses of pricing and are therefore omitted from all discussion moving forward.

3.1 Nested Logit with Consumer Heterogeneity

This section provides the main estimating equations of this paper. Consider a barcode j of product category k belonging to brand $b(j)$. I define a market as a region-time (r, t) pair, where each time period consists of a two month period and each region refers to one of five non-overlapping geographic areas of the US[21]. For all equations to follow, I drop the subscript k as it is implied that each set of demand parameters is category-specific.

Assume that for each category k there exists an exhaustive set of market segments denoted by Γ_k with each individual segment indexed as $\gamma \in \Gamma_k$. Equation 1 then provides the demand curve – as derived in Cardell [1997] and Berry [1994] – for barcode j belonging to market segment γ.

$$\ln\left(\frac{s_{jrt}}{s_{0rt}}\right) = -\alpha \ln p_{jrt} + \boldsymbol{\delta_{b(j)}} + \sigma \ln(ns_{jrt\gamma}) + \zeta_{jrt} \tag{1}$$

The price of barcode j is calculated in dollars per unit weight, and I implement the constant-elasticity variation of the Nested Logit by taking the logarithm of price. s_{jrt} denotes the *within*-category quantity market share of variety j. When calculating quantity shares, I make use of the weight of each barcode in order to normalize quantities across barcodes. The market share of the

[20] Consider a matrix of bilateral similarity in text across all varieties, in which each element represents a similarity measure between varieties i and j. It is not clear how such a matrix could be transformed into a vector representing some continuous attribute for each variety that one might consider relevant for understanding substitution patterns.

[21] North-East, South-East, Mid-West, South-West, and Mountain/Pacific.

outside option is given by s_{0rt}, and I calculate this using the BLS Consumer Expenditure Survey (CEX). Specifically, I use the CEX to estimate average per-household monthly quantities consumed for each product category and define the aggregate market as this average quantity multiplied by the number of households present in a given market. I find a median values of 0.70 for s_{0rt} across all eight product categories, with a 25th percentile value of 0.67 and a 75th percentile value of 0.79. The within-nest market share of variety j in segment γ is given by $ns_{jrt\gamma}$, and $\delta_{b(j)}$ represent brand-level fixed effects which are allowed to vary freely across product categories.

Introducing consumer heterogeneity into Equation 1 is often accomplished by allowing parameters to vary linearly with some market characteristic, such as average household income. Given that household heterogeneity in this paper consists of two dimensions, I instead estimate a separate demand system for each household type h and define these types by clustering households based on their income and ZIP code population density. Each type h is allocated based on the k-medians clustering algorithm and I incorporate 15 different household types into this analysis. Table A.4 in Appendix A.1 provides an overview of descriptive characteristics for all 15 household types.

I therefore create a separate panel for each household type and product category at the barcode level within each bi-monthly period (t) and region of the US (r). Defining each (r, t) pair as a market, this implies 90 separate markets for each (h, k) combination. Equation 2 then provides the main estimating equation of this paper.

$$\ln\left(\frac{s_{jhrt}}{s_{0hrt}}\right) = -\alpha_h \ln p_{jrt} + \boldsymbol{\delta_{h,b(j)}} + \sigma_h \ln(ns_{jhrt\gamma}) + \zeta_{jhrt} \tag{2}$$

3.2 Pricing Model

I assume that firms set prices across all varieties within a given category that they produce in order to maximize brand-level profit. Define N_{kb} as the set of varieties within product category k produced by firm b. Aggregate profit Π_{kb} is then given by:

$$\Pi_{kb} = \sum_{j \in N_{kb}} q(p_j)[p_j - c_j] \tag{3}$$

where c_j represents a constant marginal cost that differs for each barcode j. Each firm sets the price vector across all varieties N_{kb} in order to maximize Equation 3, which leads to the following optimal pricing equation for each barcode j:

$$c_j = p_j + s_j \Big[\sum_{j' \in N_{kb}} \frac{\partial s_j}{\partial p_{j'}} \Big]^{-1} \tag{4}$$

Note that the price and market share of each variety is observed, while the vector of own- and cross-price elasticities in the Nested Logit demand system can be readily calculated. Given the consumer heterogeneity approach taken in this model, I assume that firms set a single price for the entire market while the elasticity vectors represent a sales-weighted average across all household types for each barcode. Firms therefore are aware of the average elasticity of demand of the representative consumer for each barcode and cannot price discriminate across household types, which is akin to the assumptions made in Faber and Fally [2021].

3.3 Attribute-Cluster Model of Import Substitution

The Nested Logit demand system outlined in Equation 1 requires that the researcher construct an *a priori* nesting hierarchy across varieties within a given product category. These nests allow for stronger and weaker substitutes within and across nests. As mentioned in Section 2, the distinction between design and production origin complicates this decision, as it suggests a discrepancy between nesting barcodes based on the firm they belong to versus the country those goods were produced in, as in a standard Armington framework.

This paper instead takes motivation from the industrial organization literature which emphasizes attribute similarity as a measure of substitutability across varieties. Given the detailed text data available, I make use of text similarity measures and unsupervised clustering to place varieties into nests of similarity with regards to each barcode's packaging. I remove all common words such as "the", or "and", before applying the "qgram" measure of text similarity to create a matrix of bilateral text similarity across all barcodes within a given product category[22]. I then apply the Partitioning Around Medoids (PAM) clustering algorithm to this similarity matrix in order to place barcodes into clusters/nests of text similarity[23]. I refer to the resulting nesting structure as an "attribute-clustering" model of substitution.

The use of text-based data to model product differentiation forms a key contribution of this paper[24]. As an example, when clustering deodorant varieties into nests of strong and weak substitutability, there are a number of nests differentiated by the presence of words such as "Men", "Women", "Sport", or "Antiperspirant" on the packaging. This approach also allows for the brand of each variety to play a role by increasing the probability that two varieties from the same brand end up in the same nest. However this brand effect can be overcome if all other text on the packaging is similar enough for any two barcodes belonging to different brands. Such a granular understanding of the attribute space would be difficult to incorporate into a random coefficients structure as it is not immediately obvious how one could map text similarity into a single cardinal measure or characteristic[25]. By studying attribute-based substitution in a dataset with foreign and

[22] Appendix A.2 provides more detail along with an example of how the "qgram" measure calculates text similarity.

[23] I discipline the number of clusters/nests using the Silhouette Method. See Appendix A.2 for more detail.

[24] Clustering has been used in the marketing literature to understand market segments since at least the 1980s (Punj and Stewart [1983]; Green et al. [1990]; Iacobucci et al. [2000]). Recently, Harding and Lovenheim [2017] use clustering to create aggregate product categories when estimating the AIDS model of demand.

[25] In concurrent and ongoing work Almagro and Manresa [2021] provide a data-driven approach to selecting the nests within a nested logit model. The approach implemented here is in response to a separate concern: when attributes consist of lengthy text data, it is not clear how to incorporate this information within a random coefficients

domestically produced goods competing alongside each other, this paper provides a unique lens through which to study questions of import substitutability without imposing a functional form on the relationship between these elasticities and origin country characteristics.

3.4 Identification

As is well-discussed in the demand estimation literature, the inclusion of price and within-nest market share in Equation 2 almost certainly raises concerns of endogeneity. I make use of the standard "BLP" style instruments as made popular by Berry et al. [1995] and Gandhi and Houde [2019]. These include the sum of attributes of competing products for any given barcode j as well as differentiation instruments capturing the degree of differentiation of any given product from its rivals. To this end, I include as instruments for price the count of barcodes within the same nest as barcode j that are produced by the same brand, the count of barcodes produced in the same origin country, the sum of text dissimilarity across barcodes within the same nest as j, the sum of text dissimilarity across barcodes not in the same nest as j, and the average text dissimilarity across all barcodes within and without the same nest of j.

The intuition for these instruments lies in the extent to which product differentiation allows for higher markups for any given barcode. The text differentiation instruments in particular are useful in that they capture the extent to which a given nest is differentiated from all other nests as well as how differentiated a given barcode within that nest is from all competing barcodes within the same nest. The identifying assumption in this case is that the attributes of competing products are fixed in the short run.

Lastly, I employ the count of all varieties within a given nest as an instrument for the within-nest market share. This is a common technique when estimating a nested logit model – as in Khandelwal [2010] – as the number of varieties within a nest provides a denominator-driven shifter of within-nest market share that is plausibly uncorrelated with demand shocks for any given barcode within that nest (the numerator). The identifying assumption, similar to the price instruments, is that the number of barcodes within a given nest is fixed in the short run.

3.5 Estimation Results

This section provides descriptive features of the 120 separate demand systems estimated in this paper. For the sake of exposition, this section omits detailed results for all regressions, however I provide the distribution of key parameter estimates in Figure A.3 in Appendix A.1. The first-stage K-P F-statistics across all 120 models have a mean value of 26, with less than ten of the 120 models failing the "rule-of-thumb" minimum value of 10. The price disutility estimates (-$\hat{\alpha}$) are all negative with only five of the 120 estimates indistinguishable from zero at 95% confidence levels. The mean estimate is -1.62 with a standard deviation of 0.77. Estimates of σ all lie within the required range

logit model of demand.

of $[0, 1]$, with eleven of the 120 estimates indistinguishable from zero at 95% confidence levels. The mean estimate is 0.19 with a standard deviation of 0.13.

Both α and σ combine to determine the own-price elasticity for any given variety[26]. I estimate a mean elasticity of 2.14 across all 120 demand models, and a population-weighted average elasticity across all categories of 2.20. Note that these estimates are lower than what is usually found when estimating elasticities using the Feenstra [1994] method within a nested CES setting, but they are in line with what is often found in studies of industrial organization (Nevo [2001]). In fact, these estimates line up almost identically with those found in Faber and Fally [2021], who also use the NielsenIQ data and report an average elasticity of 2.2. Table A.5 provides a detailed breakdown of median own-price elasticities across all household-category pairs.

Figure A.4 in Appendix A.1 projects household-type average own-price elasticities on income (left panel) and population density (right panel) rankings. Reassuringly, I find that wealthy households are more inelastic in their demand than low-income households, which echoes results in Hottman and Monarch [2021], Faber and Fally [2021], and Auer et al. [2021]. The results with respect to population density are more nuanced: I find that in general more urban households exhibit more inelastic demand than rural households, however this pattern strongly flips for the most urban household types. This leads to an "inverted check-mark" shape which will have important consequences for the counterfactual results to follow. Studies estimating elasticities along the urban/rural distribution are rare, but the results shown here are intuitive when viewed from a consideration set perspective: households in the most urban cores of US cities have considerably larger choice sets than rural and suburban households, and the high elasticities of demand for urban household types found in this paper likely reflect this fact[27].

4 Distributional Costs of Import Price Shocks

This section provides estimates of the distributional costs across US counties in response to origin-specific price shocks. Throughout this analysis I remain agnostic as to the cause of these shocks, and therefore bypass any discussion of pass-through into consumer prices. In all cases, I estimate uncompensated changes in utility following Small and Rosen [1981], and assume a relative price increase of 10ppt for all barcodes originating in the origin receiving a shock. Note that these exercises are all partial equilibrium in that I maintain the set of varieties and location production of these varieties fixed. In addition, I keep the retail market characteristics of each household fixed, which from the perspective of the model implies that the estimates of unobserved quality

[26] The own-price elasticity for any variety j in the constant elasticity Nested Logit demand system is given by: $\epsilon_{jj} = \frac{-\alpha}{1-\sigma}(1 - \sigma n s_j - (1-\sigma)s_j)$.

[27] An alternative interpretation is that the nominal income scale used by NielsenIQ does not fully account for increased cost-of-living, such as rent, in urban cores. This would imply that these urban household types are lower income, in real terms, than what is observed in the NielsenIQ data, thereby justifying the more elastic demand they exhibit. Notice that the flexible approach to estimating type-specific elasticities is particularly useful in this case, as I do not rely on any *a priori* functional form assumption relating observed nominal income to elasticities.

remain fixed across price shocks. In sum, this exercise should be viewed as a short-run analysis in order to recover household-level exposure to trade shocks, rathern than a long-run study of the full supply-side adjustment to such shocks.

4.1 Mapping Consumption Exposure to Production Origin Price Shocks

I study four price shocks applied to the production origin of barcodes in this section: (1) a shock to all imported varieties; (2) all varieties produced in Canada or Mexico; (3) varieties produced in Europe; and (4) varieties produced in China. Shocks (2) – (4) cover over 95% of all observed import expenditure. In each case, I calculate changes in welfare costs at the household-type-category level and calculate an expenditure-weighted average cost across all eight categories within a given household type. This analysis therefore abstracts from the aggregate cost to households across the entire consumption basket, but instead estimates relative costs within narrow product categories. Once household-type cost estimates are calculated for each shock, I use the NielsenIQ projection weights to calculate county-specific welfare costs for each shock[28].

Figure A.5 to Figure A.8 in Appendix A.1 provide maps of the relative costs across counties associated with the four price shocks mentioned above. In all cases, the lowest-cost county is normalized to one, and all costs are relative to this lowest-cost county. Consistent with the stylized facts presented in Section 2, I find that wealthy and urban counties in the US face greater costs from price shocks to all imported varieties, varieties produced in Canada and Mexico, and varieties produced in Europe, while this pattern is completely reversed for a shock to prices of varieties produced in China. The relative costs of a shock to all imported varieties is 16%-26% greater for wealthy and urban counties compared to low-income and rural counties, with the range of costs for a shock to Canada/Mexico imports exhibiting a similar magnitude (17%-25%). For varieties produced in Europe, the relative cost estimates increase to 23%-52% for urban/wealthy counties compared to rural/low-income counties. The reversal in costs caused by a China price shock is striking: the most rural and low-income counties exhibit welfare costs between 103%-123% greater than their urban and high-income counterparts.

While the direction of these estimates is consistent with evidence from Hottman and Monarch [2021], Borusyak and Jaravel [2021], and Auer et al. [2021], the magnitudes are generally much larger than what previous papers have found, particularly Borusyak and Jaravel [2021]. Borusyak and Jaravel [2021] impute expenditure shares on Chinese goods by linking the NielsenIQ brand-level expenditure at the household level to firm/brand level import data, and find that low-income households spend approximately six percent more on Chinese goods than wealthy households within the category of "Health and Household Products", which is similar to categories studied here.
I therefore construct a direct comparison with Borusyak and Jaravel [2021] (Table 1, p.16) by

[28] I limit this analysis to counties with at least five households present in the NielsenIQ data. For the maps to follow, I fill in the predicted cost for each county based on the state-level average if that county does not meet the five household minimum threshold.

Figure 2: Comparison of welfare Costs between China shock and Europe shock

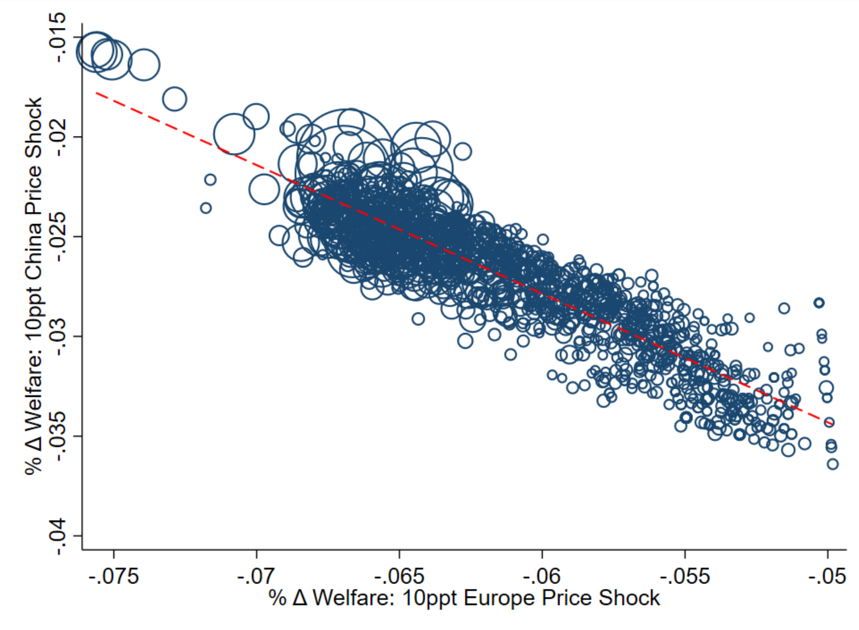


Figure 2 provides county-level costs for a 10ppt relative shock to China production origin prices (y-axis) and a 10ppt relative shock to Europe production origin prices (x-axis). Counties are weighted by population size and the red dotted line is a linear best-fit.

calculating observed expenditure shares for two income levels: households below and above a reported income of $60k USD. I find that observed relative China expenditure shares of the low-income group are 30.0% greater than the high-income group. I then impute these expenditure shares by multiplying income-brand expenditure shares by the within-brand China expenditure share, which is similar to the imputation strategy of Borusyak and Jaravel [2021], and I still find an expenditure gap of 27.0% between low-income and high-income households. It therefore seems likely that this differential is driven by the set of categories studied in this paper, which only forms a subset of "Health and Household Products". However even if the true expenditure shares are identical between this paper and Borusyak and Jaravel [2021], it is still possible to discuss the relative bias of welfare estimates when the urban/rural dimension of consumer heterogeneity is omitted and elasticities are assumed to be homogenous across household types, as in Borusyak and Jaravel [2021]. I discuss these biases in the following sections.

Given the novelty in detail of consumer heterogeneity studied in this paper, it is worth emphasizing the stark difference in outcomes across counties associated with price shocks to European and Chinese imported varieties. Figure 2 illustrates this difference explicitly by plotting county-level costs of a shock to China prices on the y-axis against county-level costs of a shock to European prices on the x-axis, along with a linear trend relating these two quantities. Each point marks a county, with the size of each marker representing the relative population of that county. While the general pattern of welfare costs simply reiterates the results described above, it is the strong negative correlation across counties that is a key feature of this analysis. The raw correlation is -0.80, suggesting that price shocks to different origin countries have remarkably different expenditure outcomes across counties within the US.

4.2 Comparison to Income Quintiles

The results found in this paper generally exhibit a larger variance in costs across households/counties when compared to those found in Borusyak and Jaravel [2021]. While it is beyond the scope of this paper to fully account for this discrepancy as I cannot comment on the matching procedure used in Borusyak and Jaravel [2021], this section instead focuses on what is omitted from our understanding of the distributional effects of import price shocks when we ignore the urban/rural component of consumer heterogeneity.

Table 2 provides expenditure-based estimates of welfare costs across counties based on the definition of consumer heterogeneity used in this paper – income-density clusters – and the more standard approach of income quintiles. The goal of this exercise is to strip away all demand elasticity heterogeneity and estimate the extent to which defining households based on income only – as is done in Hottman and Monarch [2021]; Borusyak and Jaravel [2021]; and Auer et al. [2021] – may lead to an underestimate of the true variation in exposure to trade shocks.

For each price shock, Table 2 provides the relative costs across counties at both the 90th/10th percentile and the 99th/1st percentiles. I find that across all four studied shocks, the income quintile approach to studying consumer heterogeneity in exposure to trade shocks significantly underestimates the variation across counties. Indeed the finding that income quintiles are effectively identical in their expenditure exposure to an aggregate import price shock mirrors the results found in Borusyak and Jaravel [2021] and Auer et al. [2021]. But this misses the significantly greater import expenditure that occurs in major urban centers, as these papers do not consider the role that the urban/rural divide plays in shaping import expenditure. At the 99th/1st percentile comparison, the income quintile definition of heterogeneity misses almost all of the 24% difference in costs across counties (Column (1)).

Similar results hold for a price shock to varieties produced in Mexico and Canada, whereas across the price shocks to European and Chinese varieties the income quintile approach generally captures about two thirds of the variation in costs across counties. This section therefore quantifies the biases associated with existing estimates of consumer exposure to trade shocks when dimensions of consumer heterogeneity other than income are omitted. Generalizing across shocks, I find that the omission of urban/rural heterogeneity leads to an underestimate of the variation in trade exposure across households by, on average, a third, and for aggregate import price shock, by almost 100%.

4.3 Elasticity Heterogeneity and the Bias of Expenditure Share Approximations

Auer et al. [2021] use similar data to this paper but in a Swiss context to estimate the consumption

Table 2: Comparing Income-Density Cluster Heterogeneity to Income Quintiles

	County Comparison	(1) All Imports	(2) NAFTA	(3) Europe	(4) China
Income-Density Clusters	90/10	1.06	1.06	1.16	1.35
	99/1	1.24	1.27	1.57	1.75
Income Quintiles	90/10	1.01	1.01	1.10	1.21
	99/1	1.02	1.03	1.22	1.52

Table 2 estimates expenditure share differences at the county-level for two different definitions of consumer heterogeneity. "Income-density clusters" refers to the clustering technique used in this paper, and "Income quintiles" refer to allocating households into income quintiles. The row "90/10" provides the ratio of relative expenditure-based costs of the county in the 90th percentile of costs to the county in the 10th percentile of costs. "99/1" provides the same but for the 99th and 1st percentile.

costs across Swiss households in response to a depreciation of the Swiss Franc. Similar to this paper, they find that expenditure on imported varieties is relatively flat across the income distribution, however they estimate more inelastic demand for high-income households and show that this leads to greater cost for high-income households in response to an aggregate import price shock. In this section I study whether similar mechanisms are at play in the US data by estimating county-level welfare costs under the full heterogeneous elasticity model and under a homogenous elasticity model in which all household types are assigned the average value of $\hat{\alpha}$ and $\hat{\sigma}$ within each category[29].

Figure 3 provides a county-level comparison between welfare costs of a 10ppt price shock to all imported goods under the assumption of homogenous elasticities (y-axis) and the full heterogeneous elasticity model used in this paper (x-axis). The dotted red line signifies a 45-degree line, along which county-level costs are identical under both models. Counties are once again weighted by their population.

The majority of counties illustrated in Figure 3 match the intuition discussed in Auer et al. [2021]: rural and low-income households with low aggregate import expenditure are also more elastic in their demand and therefore face even lower costs under the full model when compared to the homogenous elasticity model. Wealthier and larger counties exhibit both greater import expenditure and more inelastic demand, and this pattern emphasizes the results in Auer et al. [2021]: elasticity heterogeneity expands the cost gap between wealthy and poor counties/households in response to an aggregate import price shock. This paper builds on their finding however by incorporating a richer definition of household heterogeneity (income *and* urban/rural) as well as price shocks to different production origin countries, rather than all imports.

The bottom-left corner of Figure 3 illustrates a key finding of this paper: the highest import expenditure counties in the US are the urban cores of major metropolitan areas. In an expenditure-based approximation of welfare costs, these urban consumers would face the greatest costs of an

[29]Note that when estimating relative costs across household/counties, the homogenous elasticity model is akin to approximating welfare costs based on expenditure shares only, as in Borusyak and Jaravel [2021].

Figure 3: Comparison of welfare costs between homogenous and heterogenous elasticity models: 10ppt price shock to all imported goods

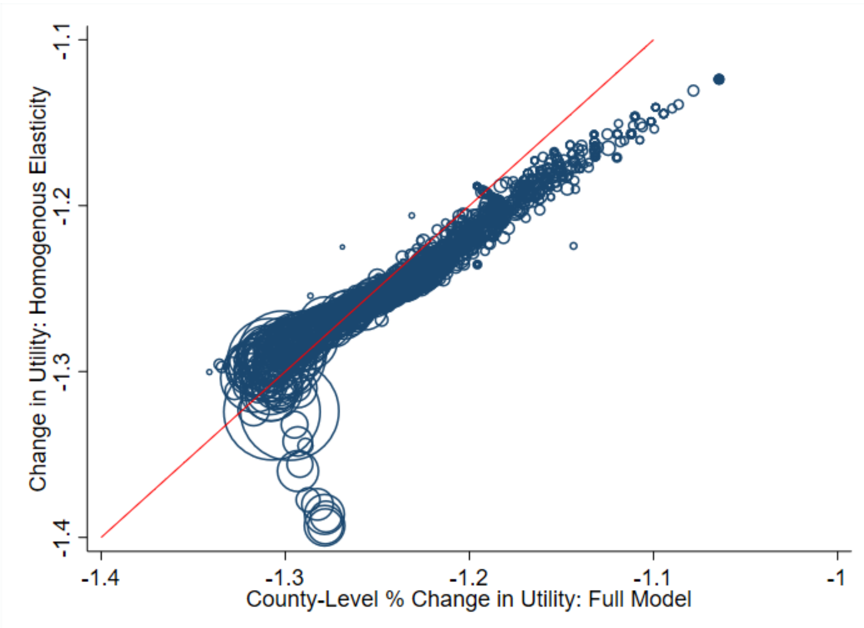


Figure 3 provides a county-level comparison of costs associated with a 10ppt price shock to all foreign produced goods under the homogenous elasticity model (y-axis) and the full model (x-axis). The dotted line represents $y = x$. Counties are weighted by population.

aggregate import price shock. However under the full model, notice that the curve bends back on itself, with these large metropolitan areas exhibiting substantial negative bias under the homogenous elasticity model. This is due to the "inverted check-mark" relationship between demand elasticities and population density discussed in Section 3.5: households living in the urban centers of major American cities exhibit surprisingly elastic demand and are therefore more likely to simply substitute away from a price shock to imported goods.

To my knowledge, this paper is the first to find a non-linear bias associated with welfare costs approximated from expenditure shares only. This non-linearity in bias is particularly important given that the bias is largest for both the highest and lowest import expenditure counties – those counties with either predominantly poor and rural households and those composed of the most urban households – and is large enough in magnitude to alter the ranking of costs across counties in response to an aggregate import price shock. Figure A.9, Figure A.10, and Figure A.11 provide similar scatterplots for a price shock to Canada/Mexico varieties, European varieties, and Chinese varieties, respectively. In all three cases, the relative ranking of costs across counties is generally preserved across models, however the non-linear bias is present across all shocks. The key result to take away from this study is that while incorporating elasticity heterogeneity may increase the cost differential between counties at the 90th and 10th percentile of costs, as in Auer et al. [2021], it likely decreases this differential when comparing counties in the 99th percentile to the 1st.

This section has highlighted the importance of incorporating both expenditure and elasticity heterogeneity across domestic households when estimating the distributional costs of import price shocks. Notice, however, that this flexible specification – along with the attribute-cluster model of substitution described in Section 3.3 – allows for an *ex post* study of how own-price and cross-price elasticities differ across origin countries, as these elasticities can be estimated via revealed preference. The remainder of this paper provides an analysis of these elasticities.

5 Import Elasticities and Attribute Separability

The demand model specified in this paper introduces two key departures from standard approaches to modeling import substitution. First, the nesting structure in this paper links attribute similarity across varieties to their substitutability. When taken to the country level, this implies that countries with more similar allocations of varieties across these attribute nests will be, in the aggregate, closer substitutes. Second, the flexible mapping between consumer types and demand parameters allows for an *ex post* study of how a variety's representative consumer shapes the own-price elasticity of demand for that variety in the aggregate. For any variety j one can denote the own-price elasticity for this variety as $e_{jj} = \sum_h w_j^h e_{jj}^h$, where w_j^h denotes the expenditure share (adjusted for population) of variety j deriving from households of type h. Notice that a variety with expenditure concentrated in low-elasticity households will exhibit more inelastic demand than a variety with expenditure concentrated in high-elasticity households. The same holds true for origin countries and their representative consumers.

This section provides an analysis of how both these modeling choices – attribute-based substitution and consumer heterogeneity – shape observed country-level own-price demand elasticities in the US market. I begin in Section 5.1 by illustrating how the attribute-based model leads to a greater estimate of import substitutability when compared to an Armington-based substitution model, which isolates the role played by the nesting structure in this model. I then illustrate the role played by consumer heterogeneity and provide revealed preference justification for the attribute-based model in Section 5.2 and Section 5.3. Finally, Section 5.4 brings these elements together to illustrate how import demand shapes trade elasticities across both import origins and firms.

5.1 Armington Elasticities under Attribute-Clustering

I start by comparing the attribute-based nesting structure used in this paper with other modeling approaches that have been used in the trade and international macroeconomics literature. The most prominent of such approaches is the production-based Armington model of import substitutability, in which the substitution nests represent the country in which a variety was produced[30]. While this assumption is often useful when using customs data, as it allows the researcher to assume that all

[30]In the trade literature studying expenditure effects of trade shocks, the following papers rely on this assumption: Amiti et al. [2020]; Hottman and Monarch [2021]; Borusyak and Jaravel [2021]; and Auer et al. [2021].

imported varieties (observed) exhibit identical substitution elasticities with all domestic varieties (unobserved), it is also a difficult model to validate for precisely the same reason. The dataset used in this paper, however, provides this opportunity.

Table 3 provides a comparison of the Armington elasticity (the elasticity of American-produced market share with respect to an aggregate price shock to all imported varieties) across all product categories under varying assumptions regarding the nesting structure of substitution. For each alternative nesting structure, I maintain the same set of parameter estimates $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\sigma}}$ and I remove consumer heterogeneity by assuming the same average parameter values for α and σ across types (I explore the role of consumer heterogeneity in the next section). Column (2) normalizes the Armington elasticities derived from the production-based Armington nesting structure, and Columns (3) – (7) provide Armington elasticity estimates relative to this normalization for, respectively: the text-based nesting structure used in this paper, a design-based Armington model, nests based on firms/brands, nests based on price deciles within a given category, and nests based on quality deciles within a given category[31]. The firm nesting structure mirrors the assumptions applied in Hottman et al. [2016] and Faber and Fally [2021] whereas the last two specifications – price and quality deciles – mirror the intuition from the "quality ladders" literature (Schott [2008]; Khandelwal [2010]; Hallak and Schott [2011]; Crozet et al. [2012]).

Table 3 therefore provides one of the first analyses as to how *a priori* assumptions regarding import substitutability lead to differing Armington elasticities. The first key result is that import substitutability is greater across all categories under any specification other than the production-based Armington model. This is intuitive, if rarely discussed explicitly: the production-based Armington model exhibits a limit case of import substitutability in which imports are assumed to be weak substitutes with domestic varieties by definition. What is striking, however, is the extent to which almost all of the attribute-based substitution models exhibit fairly consistent estimates of import substitutability. The text-based model used in this paper (Column (3)), a design-based Armington model (Column (4)), and the price/quality decile models ((6) and (7)) all exhibit average Armington elasticities which are approximately 20% greater than those found under the production-based Armington model.

While attributes are rarely mentioned explicitly in the trade literature, assumptions such as CES and Armington can easily be mapped into the language of attributes and attribute similarity. Consider a constant elasticity of substitution demand system with no nesting structure. The implicit assumption is that all varieties are identically differentiated in their attributes from all other varieties. The attribute-based interpretation of a production-based Armington model is that the production origin of a variety serves as the defining attribute of that variety. Attribute similarity – and therefore substitutability – can be viewed as purely a statement of identifying where a variety was produced. These assumptions are important in that they ultimately determine the scarcity of

[31] Note that for each variety j I proxy demand-shifter quality as $\zeta_j = \ln s_j + \hat{\alpha}_k \ln p_j$, with α_k taken as the average estimate of $\hat{\alpha}_{hk}$ across all household types within a given category.

Table 3: Armington Elasticity Estimates Under Various Specifications

	Nesting Attribute						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	w_k	Prod.	Text	Design	Firm	Price	Quality
Hair Care	0.23	1.00	1.27	1.26	1.14	1.28	1.28
Body Soap	0.17	1.00	1.11	1.10	1.05	1.11	1.11
Skin Care	0.14	1.00	1.34	1.23	1.10	1.34	1.34
Deodorant	0.14	1.00	1.10	1.09	1.06	1.13	1.13
Toothpaste	0.12	1.00	1.12	1.17	1.12	1.17	1.17
Facial Care	0.09	1.00	1.23	1.24	1.14	1.27	1.27
Hair Styling	0.07	1.00	1.19	1.13	1.09	1.20	1.20
Face Cosmetics	0.04	1.00	1.08	1.13	1.03	1.14	1.13
Average	1.00	1.00	1.19	1.18	1.10	1.21	1.21

Table 3 provides estimates of the Armington elasticity across all eight product categories and a number of nesting specifications. w_k provides the expenditure weight by category. For each specification I maintain the estimates of α and σ from the attribute-cluster model, but adjust the nesting structure. Column (2) provides nests based on production location and I normalize these elasticities to 1.00 for all categories. All other specifications are reported relative to Column (2), and are, from left to right: the text-cluster model of this paper, design origin, firm-specific nests, price decile nests, and demand-shifter "quality" decile nests.

attributes in the domestic economy and therefore the potential gains from trade in alleviating that scarcity. If a US consumer can only access whatever aspect of a variety that makes it "French" by consuming a variety produced in France then the production-based Armington model is appropriate, the consumption gains from trade are large, and the attributes associated with imported varieties are tautologically scarce in the domestic economy.

But recall the stylized fact from Section 2 discussing the production versus design origin of imported varieties. As we have seen, over 60% of all import expenditure in fact accrues to American firms with off-shored production. In this case the attributes associated with imported varieties are not necessarily scarce in the domestic economy, which is reflected in the estimates of Table 3: when one allows for a model in which attributes are not tautologically defined to differ across production locations, the substitutability of foreign and domestic varieties increases.

To place these estimates in some context, albeit in a crude way, consider the welfare formula derived by Arkolakis et al. [2012], in which the welfare cost of moving to autarky for the US would be $\hat{U} = 1 - \lambda^{-1/e}$, with λ representing the domestic share of expenditure (86% for the dataset used in this paper) and e the elasticity of import penetration with respect to changes in trade costs. Assuming that all supply elasticities incorporated in e are invariant across demand-side specifications, one can use the estimates in Table 3 to estimate the relative costs of moving to autarky under various nesting models. I find that the relative cost of moving to autarky would be 19.8% greater under the production-based Armington model when compared to the text-based attribute model used in this paper. Similarly, increasing the import penetration share by 10% would increase welfare by 20.0% more under the production-based Armington model when compared to the text-based model of this paper.

The results in this section raise a larger question: how can one test whether or not the production location of a variety is a relevant attribute in the eyes of consumers? I turn now to a study of the effects of consumer heterogeneity on own-price elasticities at the variety level, and in doing so provide direct evidence that the attributes of varieties are in fact "separable" between the firm which designed that variety and where that variety was produced.

5.2 Own-Price Elasticities and Consumer Heterogeneity

This section narrows the analysis of the previous section to understand how variety-specific own-price elasticities differ across origins. Recall that for each barcode, j, I calculate the expenditure-weighted average own-price elasticity of demand for this barcode as $e_{jj} = \sum_h w_j^h e_{jj}^h$. In this section I estimate projections of e_{jj} on design and production origin characteristics of j and recover correlations between variety production and design origins and own-price elasticities[32].

When projected on fixed effects for whether a variety was produced/designed abroad, I find that imported varieties exhibit own-price elasticities that are on average, 2.3% more elastic than domestically produced varieties, while varieties designed abroad exhibit own-price elasticities that are 1.4% more elastic than varieties designed in the USA. Both estimates are significant at levels of 99%. When I cross these two fixed effects, I find that domestically produced and designed varieties exhibit the most inelastic demand, with foreign produced *and* designed varieties exhibiting own-price elasticites that are 3.1% greater by comparison. While these estimates are statistically significant, they are not all that relevant from an economic perspective: if one assumes an elasticity of 2.20 for domestic varieties (the average in my estimated sample), these estimates suggest imported varieties would exhibit an average elasticity of 2.27.

However this exercise masks considerable heterogeneity within imported and foreign-designed varieties. In Table A.6 I provide estimates of a similar projection exercise in which I restrict my sample to only imported varieties and estimate correlations between the own-price elasticity of a barcode and the GDP per capita of both the production origin and design origin of that variety. In the most rigorous specification (Column (7)), I estimate these relationships conditional on each other, as well as conditional on the price of variety j and category-nest fixed effects. I estimate an elasticity between e_{jj} and the production origin GDP per capita of variety j of -0.016 which is significantly distinguishable from zero at a 10% confidence level. This result is intuitive in that it suggests that wealthier countries produce exports with more inelastic demand than low-income countries. But again, the magnitude of this elasticity is underwhelming, as it suggests a variety

[32] In each case I include category-specific fixed effects, so estimates should be interpreted as *within*-category.

produced in France would exhibit an own-price elasticity only 1.3% more *inelastic* than a variety produced in Mexico.

The elasticity between variety-specific own-price elasticities and *design* origin GDP per capita, however, reveals a key dimension of heterogeneity across imported varieties. Within production origin countries and within category-nest pairs I estimate this elasticity to be -0.089, which is over five *times* greater than the elasticity between own-price elasticities and the production origin GDP per capita of variety j. To put these estimates in context : conditional on production location, a French-designed variety exhibits an own-price elasticity of demand that is 9.5% lower than a Mexican-designed variety. Assuming an elasticity of 2.20 for the Mexican-designed variety, this implies an elasticity for the French-designed variety of 1.99.

This section provides results which suggest a two-fold importance of incorporating flexible demand elasticities within a model of import demand. First, when estimated via revealed preference, there is a negative relationship between the production origin country GDP per capita of an imported variety and the elasticity of the demand curve faced by that variety when sold in the US. This result is intuitive given the earlier findings of this paper: wealthy households exhibit both more inelastic demand and higher expenditure on imports from wealthier countries. The second finding, however, is that this relationship is much smaller than the relationship between a variety's *design* origin GDP per capita and the elasticity of demand faced by that variety in the US. This second result is crucial to understanding why standard Armington-style assumptions might fail to capture the nuances of import demand, and I turn to providing a more in-depth analysis of this result in the next section.

5.3 Attribute Separability at the Barcode Level

This section provides direct evidence for a phenomenon I term attribute separability: the capacity of multinational firms to endow the varieties they produce with attributes which are separable from the location of production. Specifically, I estimate projections of variety attribute x_j on production and origin characteristics of variety j for the following observed and model-derived attributes: marginal cost, mark-ups, per-unit profits, demand-shifter quality, and per-variety profits.

Table 4 provides output from this projection exercise. For each attribute, I provide estimates from two specifications. In the first specification, I include only production origin characteristics: the distance of variety j's production origin from the USA and the GDP per capita of variety j's production origin. In the second specification, I keep these two production origin characteristics and add two design origin characteristics: the GDP per capita of variety j's design origin and a dummy variable equal to one if variety j was designed by a US firm (and is therefore considered "off-shored"). Notice that when estimating gravity relationships one would tend to include all origin countries and account for zeros by specifying a PPML model. In this case, it is less clear how to account for zeros when estimating the relationship between, for example, markups and distance from the US. Instead, I simply include the varieties and countries which are present in the data and

Table 4: Projecting Barcode-level Attributes on Origin Variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Log Dependent Variable									
	Marginal Cost		Markup		Per-Unit Profit		Quality		Per-Barcode Profit	
Production Origin Variables (logs):										
Distance to US	0.458	0.448	-0.014	-0.019	0.424	0.405	-0.421	-0.480	-0.937	-1.001
	(0.037)	(0.037)	(0.013)	(0.013)	(0.035)	(0.035)	(0.063)	(0.062)	(0.059)	(0.058)
GDP per capita	0.685	0.647	-0.055	-0.076	0.620	0.544	0.508	0.265	0.164	-0.089
	(0.052)	(0.053)	(0.019)	(0.019)	(0.049)	(0.050)	(0.088)	(0.089)	(0.082)	(0.084)
Design Origin Variables (logs):										
GDP per capita		0.247		0.196		0.599		1.754		1.840
		(0.124)		(0.043)		(0.117)		(0.208)		(0.194)
USA Brand FE		-0.177		-0.089		-0.333		-1.115		-1.153
		(0.066)		(0.024)		(0.062)		(0.110)		(0.103)
Example: Production Origin = Mexico:										
Design: Mexico		1.00		1.00		1.00		1.00		1.00
Design: France		1.23		1.18		1.66		4.38		4.71
Design: USA		1.12		1.15		1.44		2.54		2.71
N	2993	2993	2993	2993	2993	2993	2993	2993	2993	2993
R^2	0.449	0.451	0.319	0.323	0.633	0.637	0.464	0.483	0.185	0.220
Category FE	x	x	x	x	x	x	x	x	x	x

Table 4 provides regression estimates for a projection of barcode-specific attributes on to production and design origin characteristics. The sample is limited to only those barcodes produced outside of the USA. All dependent variables are included in logarithms. The third panel provides a comparison of three varieties, all produced in Mexico, but with design origins of Mexico, France and the USA, respectively. These estimates are derived entirely from the estimates in Panels 1 and 2. Note that Mexico, France, and the USA have values of log GDP per capita (in thousands of USD), respectively, of: 2.88, 3.73, and 4.05. Standard errors are provided in parentheses.

estimate these projections using OLS. The estimates in Table 4 should therefore be interpreted as conditional on each variety j achieving enough revenue to profitably export to the US.

The production-based estimates are intuitive, in that I find a strong positive relationship between marginal cost and both production distance from the USA and production origin GDP per capita. For every 10% increase in production origin distance from the USA, marginal cost increases 4.58%, and for every 10% increase in production origin GDP per capita, marginal cost increases 6.85%. Similarly, I find a positive correlation between the production origin GDP per capita of a variety and that variety's quality and profits. For every 10% increase in the production origin GDP per capita, variety-level quality and profits increase 5.08% and 1.64%, respectively.

When including design-based origin characteristics, however, it becomes clear that the attributes in question are highly separable between firms and countries. Conditional on where a variety was produced, I find that for every 10% increase in design origin GDP per capita, marginal cost increases by 2.47%, markups by 1.96%, per-unit profits by 5.99%, quality by 17.54%, and per-variety profits by 18.40%. The third panel of Table 4 provides a direct comparison across three varieties using the estimates from the first two panels: I compare these five attributes for varieties produced in Mexico but designed in, respectively, Mexico, France, and the USA. I find that the French and

American varieties exhibit quality estimates that are 4.38 and 2.54 *times* greater than the quality of the Mexican designed variety, and that these French and American varieties accrue profits that are 4.71 and 2.71 *times* greater than the Mexican designed variety. At the variety level, the French and American varieties exhibit per-unit profits 66% and 44% greater than the Mexican variety, with the relative role of marginal cost and markups in explaining this discrepancy approximately 50/50. The capacity of multinational firms to charge higher mark-ups and exhibit greater quality is particularly striking given the relatively unrestricted approach to estimating these two attributes in this paper. Mark-ups are determined by the relative demand elasticity of a variety's representative consumer, while "quality" in this case simply refers to a demand-shifter which is derived from market share and price, both of which are observed directly in the raw purchase records.

There are two broad takeaways from the results in Table 4. First, incorporating consumer heterogeneity in elasticities as well as expenditure shares across import origins allows the researcher to recover intuitive patterns in the relationship between imported variety attributes and origin country income. Wealthier countries tend to export higher marginal cost, lower demand elasticity, and higher quality varieties than poorer countries. That this relationship can be uncovered through revealed preference, rather than an *a priori* specification, is useful as a benchmark for research moving forward. Second, the separability results shown here suggest that one what must distinguish between "fundamental" capabilities of production origin countries and the observed attributes associated with varieties which are produced in those locations but designed elsewhere. While this distinction is important for understanding the elasticities of trade, as I discuss in the following (and final) section of this paper, this distinction is important for a number of other research questions in international trade as well[33].

As an example of one such research area in which these findings of attribute separability are applicable, consider the "quality ladders" literature mentioned earlier. Generally speaking, this literature seeks to understand the relative quality of imported varieties and how this differs across origin countries. In doing so, one can attempt to understand how factors compete internationally either within or across rungs on the quality ladder. In Figure A.12 I make use of the estimates in Columns (7) and (8) of Table 4 to compare observed average quality estimates for each production origin country alongside a measure of the "fundamental" quality associated with that country, when the effects of firm composition and design are stripped away. The results of this exercise are striking: I find that quality estimates which fail to account for the role of multinational production over-estimate the average quality of Chinese exports by a factor of 3.3 and Mexican varieties by a factor of 3.1. Importantly, especially in reference to the quality ladders literature, the quality gap between the top five and bottom five countries increases by a factor of six when measured as "fundamental", rather than observed, quality.

[33] Alviarez [2019] makes a similar argument with respect to country-level productivity and comparative advantage. While their analysis focuses on the distribution of productivity across sectors within a country and the role that multinational activity plays in shaping that distribution, the conceptual similarity lies in defining a "fundamental" country-level productivity versus the observed productivity attributable to multinational activity.

Similarly, these findings call into question the pervasive assumption made in the trade literature that a "variety" can be empirically defined as a product-country pair, or even a product-firm-country pair, when estimating the welfare gains of trade using customs data. Empirically, this approach has been used by Feenstra [1994]; Broda and Weinstein [2006], and Redding and Weinstein [2018], which build on the theory of Armington [1969], Krugman [1980], and Romer [1994]. However if attributes are separable between multinational firms and production locations, and multinationals are dominant actors in international trade, then one must use caution when assuming that customs data alone allows for an understanding of the variety gains from trade.

The final section of this paper brings together the components of this model to study how demand elasticities differ across origin countries in the differentiated products studied in this paper, and provides a decomposition exercise to fully illustrate the role of each component.

5.4 Demand Elasticities by Origin Country

This section provides analysis for the two largest (by expenditure) product categories studied in this paper: hair care and body soap. Figure 4 plots own-price demand elasticities at the country level for hair care (left) and body soap (right) against each country's GDP per capita. In this case, the origin of each variety is the production origin, as one would find using customs data. Both panels compare the estimated elasticity for each country-product pair across four specifications: an Armington substitution model with no consumer heterogeneity ("Armington: Constant"), the attribute-based clustering model with no consumer heterogeneity ("Attribute: Constant"), the attribute-based model with consumer heterogeneity ("Attribute: Heterogeneity"), and the attribute-based model with elasticities defined as the "fundamental" demand elasticities associated with each production origin ("Attribute: Fundamental").

The "fundamental" model refers to a model in which I take the expenditure-weighted values of α_j and σ_j for each barcode and project these on category fixed effects, the production origin GDP per capita of variety j, and the design origin GDP per capita of variety j (as in Table 4). I then use the estimated relationships between α_j, σ_j, and origin characteristics to adjust both estimates by forcing the design origin GDP per capita to be equivalent to the production origin GDP per capita. This is the same exercise as described earlier for comparing estimated quality versus "fundamental" quality of exports, except in this case applied to demand parameters for each barcode.

These four alternative models highlight the importance of both the attribute-based substitution model of this paper as well as the barcode-level consumer heterogeneity which allows for *ex post* recovery of variety-level elasticities based on revealed preference. There are a number of points to discuss. First, the level shift upwards in demand elasticities across all origins when moving from the production-based Armington model to the attribute-based model simply mirrors the results presented earlier: in the attribute-space, imported varieties are less differentiated from domestically produced varieties than under Armington-style assumptions. Second, the addition of consumer heterogeneity alongside the attribute-based demand model rotates the relationship between origin

Figure 4: Country-Product Demand Elasticities under Various Specifications

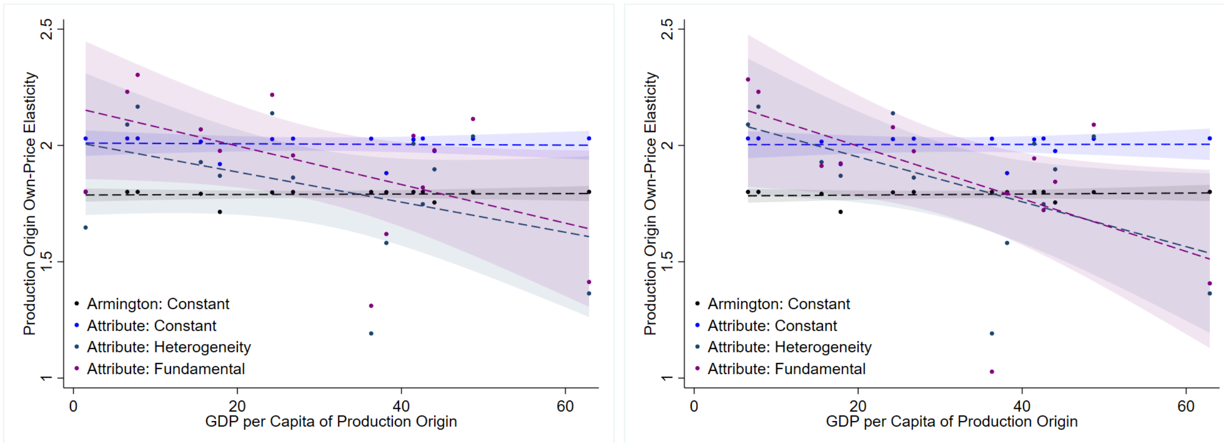


Figure 4 provides country-category own-price elasticity estimates for the two largest, be expenditure, categories: Hair Care (left) and Body Soap (right). These models are, respectively from top to bottom in the legend: an Armington model with no consumer heterogeneity, an attribute model with no consumer heterogeneity, an attribute model with consumer heterogeneity, and an attribute model with consumer heterogeneity applied to "fundamental" origin country demand. Linear fit lines as well as 95% confidence bands are provided.

demand elasticity and GDP per capita which confirms the results found earlier regarding consumer heterogeneity: wealthier origin countries tend to export varieties with sales concentrated in inelastic US consumers. Finally, adjusting demand elasticities to reflect the revealed demand elasticities that are "fundamental" to each production origin – in that they attempt to strip away any effects of multinational activity – further steepen the relationship between origin country income and demand elasticity.

By bringing together the key elements of this model, this section highlights the importance of incorporating consumer heterogeneity within our understanding of trade elasticities. While a significant body of literature has studied the effect of income differences across countries in their relative demand for *sectors*, and in doing so rationalize observed trade flows across countries at the sector level, these results suggest that even within narrow product categories and within a single destination country, consumer heterogeneity plays an important role in shaping trade elasticities[34].

6 Conclusion

Understanding the characteristics of import demand both across households and across import origins remains an imperative task for any study of the distributional costs of trade shocks and/or the cost structure of imported goods. This paper provides evidence that a number of assumptions common to this literature struggle to match the underlying variety-level data and quantifies the biases

[34] Both Adao et al. [2017] and Lind and Ramondo [2023] estimate correlated trade elasticities across origin countries based on correlated factor endowments, while Fieler [2011] and Caron et al. [2014] incorporate non-homothetic preferences to study sector-level demand and trade flows.

associated with these assumptions. This paper suggests a number of avenues for future research, including: quantifying the potentially causal role that retailers play in shaping household-level import expenditure, re-evaluating the consumption gains from trade in the presence of multinational production and attribute separability, and quantifying the extent to which multinational firms are constrained in their ability to separably endow their varieties with attributes across production locations and across sectors/products.

References

Rodrigo Adao, Arnaud Costinot, and Dave Donaldson. Nonparametric counterfactual predictions in neoclassical models of international trade. *American Economic Review*, 107(3):633–689, 2017.

Milena Almagro and Elena Manresa. Data-driven nests in discrete choice models. *No full draft available*, Slides, 2021.

Vanessa Alviarez. Multinational production and comparative advantage. *Journal of International Economics*, 119:1–54, 2019.

Mary Amiti, Mi Dai, Robert C Feenstra, and John Romalis. How did china's wto entry affect us prices? *Journal of International Economics*, 126:103339, 2020.

Costas Arkolakis, Arnaud Costinot, and Andrés Rodríguez-Clare. New trade models, same old gains? *American Economic Review*, 102(1):94–130, 2012.

Paul S Armington. A theory of demand for products distinguished by place of production. *Staff Papers-International Monetary Fund*, pages 159–178, 1969.

Raphael Auer, Ariel Burstein, Sarah Lein, and Jonathan Vogel. Unequal expenditure switching: Evidence from switzerland. 2021.

Liang Bai and Sebastian Stumpner. Estimating us consumer gains from chinese imports. *American Economic Review: Insights*, 1(2):209–24, 2019.

Andrew B Bernard, J Bradford Jensen, Stephen J Redding, and Peter K Schott. The empirics of firm heterogeneity and international trade. *Annu. Rev. Econ.*, 4(1):283–313, 2012.

Andrew B Bernard, J Bradford Jensen, Stephen J Redding, and Peter K Schott. Global firms. *Journal of Economic Literature*, 56(2):565–619, 2018.

Steven Berry, James Levinsohn, and Ariel Pakes. Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, pages 841–890, 1995.

Steven T Berry. Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, pages 242–262, 1994.

Bruce A Blonigen and Anson Soderbery. Measuring the benefits of foreign product variety with an accurate variety set. *Journal of International Economics*, 82(2):168–180, 2010.

Kirill Borusyak and Xavier Jaravel. The distributional effects of trade: Theory and evidence from the united states. Technical report, National Bureau of Economic Research, 2021.

Christian Broda and David E Weinstein. Globalization and the gains from variety. *The Quarterly journal of economics*, 121(2):541–585, 2006.

N Scott Cardell. Variance components structures for the extreme-value and logistic distributions with application to models of heterogeneity. *Econometric Theory*, pages 185–213, 1997.

Justin Caron, Thibault Fally, and James R Markusen. International trade puzzles: A solution linking production and preferences. *The Quarterly Journal of Economics*, 129(3):1501–1552, 2014.

Daniel R Carroll and Sewon Hur. On the heterogeneous welfare gains and losses from trade. *Journal of Monetary Economics*, 109:1–16, 2020.

A Kerem Coşar, Paul LE Grieco, Shengyu Li, and Felix Tintelnot. What drives home market advantage? *Journal of international economics*, 110:135–150, 2018.

Javier Cravino and Andrei A Levchenko. The distributional consequences of large devaluations. *American Economic Review*, 107(11):3477–3509, 2017.

Javier Cravino and Andrei A Levchenko. The geographic spread of a large devaluation. In *AEA Papers and Proceedings*, volume 108, pages 562–66, 2018.

Matthieu Crozet, Keith Head, and Thierry Mayer. Quality sorting and trade: Firm-level evidence for french wine. *The Review of Economic Studies*, 79(2):609–644, 2012.

Benjamin Faber and Thibault Fally. Firm heterogeneity in consumption baskets: Evidence from home and store scanner data. *Review of Economic Studies*, 2021.

Pablo D Fajgelbaum and Amit K Khandelwal. Measuring the unequal gains from trade. *The Quarterly Journal of Economics*, 131(3):1113–1180, 2016.

Robert C Feenstra. New product varieties and the measurement of international prices. *The American Economic Review*, pages 157–177, 1994.

Robert C Feenstra, Philip Luck, Maurice Obstfeld, and Katheryn N Russ. In search of the armington elasticity. *Review of Economics and Statistics*, 100(1):135–150, 2018.

Ana Cecilia Fieler. Nonhomotheticity and bilateral trade: Evidence and a quantitative explanation. *Econometrica*, 79(4):1069–1101, 2011.

Amit Gandhi and Jean-François Houde. Measuring substitution patterns in differentiated-products industries. *NBER Working paper*, (w26375), 2019.

Pinelopi Koujianou Goldberg. Product differentiation and oligopoly in international markets: The case of the us automobile industry. *Econometrica: Journal of the Econometric Society*, pages 891–951, 1995.

Paul E Green, Jonathan Kim, and Frank J Carmone. A preliminary study of optimal variable weighting in k-means clustering. *Journal of Classification*, 7(2):271–285, 1990.

Laura Grigolon and Frank Verboven. Nested logit or random coefficients logit? a comparison of alternative discrete choice models of product differentiation. *Review of Economics and Statistics*, 96(5):916–935, 2014.

Juan Carlos Hallak and Peter K Schott. Estimating cross-country differences in product quality. *The Quarterly journal of economics*, 126(1):417–474, 2011.

Matthew Harding and Michael Lovenheim. The effect of prices on nutrition: comparing the impact of product-and nutrient-specific taxes. *Journal of Health Economics*, 53:53–71, 2017.

Keith Head and Thierry Mayer. Brands in motion: How frictions shape multinational production. *American Economic Review*, 109(9):3073–3124, 2019.

Colin Hottman and Ryan Monarch. Estimating unequal gains across us consumers with supplier trade data. *Journal of International Economics*, 2021.

Colin Hottman, Surabhi Ghai, et al. Exchange rates, product variety, and substitution in us scanner data. In *2019 Meeting Papers*, number 1059. Society for Economic Dynamics, 2019.

Colin J Hottman, Stephen J Redding, and David E Weinstein. Quantifying the sources of firm heterogeneity. *The Quarterly Journal of Economics*, 131(3):1291–1364, 2016.

Dawn Iacobucci, Phipps Arabie, and Anand Bodapati. Recommendation agents on the internet. *Journal of Interactive Marketing*, 14(3):2–11, 2000.

Xavier Jaravel and Erick Sager. What are the price effects of trade? evidence from the us and implications for quantitative trade models. 2019.

Amit Khandelwal. The long and short (of) quality ladders. *The Review of Economic Studies*, 77 (4):1450–1476, 2010.

Paul Krugman. Scale economies, product differentiation, and the pattern of trade. *The American Economic Review*, 70(5):950–959, 1980.

Nelson Lind and Natalia Ramondo. Trade with correlation. *American Economic Review*, 113(2): 317–353, 2023.

Daniel McFadden. The measurement of urban travel demand. *Journal of public economics*, 3(4): 303–328, 1974.

Aviv Nevo. Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 69(2): 307–342, 2001.

Girish Punj and David W Stewart. Cluster analysis in marketing research: Review and suggestions for application. *Journal of marketing research*, 20(2):134–148, 1983.

Natalia Ramondo. A quantitative approach to multinational production. *Journal of International Economics*, 93(1):108–122, 2014.

Stephen Redding and David E Weinstein. Accounting for trade patterns. *Princeton University, mimeograph*, 22, 2018.

Paul Romer. New goods, old theory, and the welfare costs of trade restrictions. *Journal of development Economics*, 43(1):5–38, 1994.

Peter K Schott. The relative sophistication of chinese exports. *Economic policy*, 23(53):6–49, 2008.

Kenneth A Small and Harvey S Rosen. Applied welfare economics with discrete choice models. *Econometrica: Journal of the Econometric Society*, pages 105–130, 1981.

Felix Tintelnot. Global production with export platforms. *The Quarterly Journal of Economics*, 132(1):157–209, 2017.

A Appendix

A.1 Additional Tables and Figures

Table A.1: Descriptive Statistics by Product Category

Category	Expenditure *$M*	Quantity *units*	Brands *#*	Varieties *#*	Import Share *%*
Hair Care	3.75	577,628	209	4,492	7.0
Body Soap	2.70	461,470	195	2,718	8.5
Skin Care	2.27	256,086	252	2,820	22.3
Deodorant	2.25	403,441	207	1,702	20.1
Toothpaste	2.03	410,837	33	589	19.7
Dental Care	1.76	262,278	53	602	13.1
Facial Care	1.40	128,627	101	1,417	19.2
Hair Styling	1.20	194,534	110	1,296	15.1
Hand Soap	0.76	191,531	135	1,298	7.9
Face Cosmetics	0.66	71,896	54	2,110	16.6
Eye Cosmetics	0.57	75,452	34	1,172	15.7
Shaving Care	0.42	105,243	39	354	13.5
Lip Cosmetics	0.39	94,346	93	1,064	14.0
Nail Cosmetics	0.26	48,580	90	1,639	16.2
Bath Accessories	0.14	24,725	84	533	17.7

Table A.1 provides summary statistics for the 15 product categories studied in this paper, including aggregate expenditure by all households over three years, aggregate units purchased, the number of unique brands and varieties, and finally the expenditure import share. Back to Section 2.

Table A.2: Descriptive Statistics by Origin Country

	By Production			By Design	
Expenditure:	Aggregate %	Imports %		Aggregate %	Imports %
USA	85.55	-	USA	81.84	-
Canada	5.30	36.7	France	7.08	39.0
Mexico	4.84	33.5	Switzerland	3.56	19.6
China	0.80	5.5	Germany	2.30	12.7
Ireland	0.66	4.6	Australia	2.24	12.4
Germany	0.46	3.2	UK	1.80	9.9
France	0.45	3.1	Japan	0.36	2.0
Australia	0.24	1.7	Italy	0.20	1.1
Israel	0.22	1.5	Canada	0.18	1.0
UK	0.19	1.3	South Africa	0.11	0.6
Italy	0.19	1.3	Israel	0.08	0.4
Morocco	0.16	1.1	Belgium	0.06	0.3
Spain	0.16	1.1	Mexico	0.04	0.2
Thailand	0.14	1.0	Sweden	0.02	0.1
Taiwan	0.13	0.9	Netherlands	0.02	0.1
Philippines	0.08	0.6	Finland	0.02	0.1

Table A.2 provides the aggregate expenditure shares by origin country, with Columns (1)-(3) denoting shares by barcode production origin and Columns (4)-(6) denoting shares by design origin. Not shown (in descending order of %) for production origin: India, Belgium, Turkey, Brazil, Switzerland, Luxembourg, Poland, Japan, Finland, Greece, Czech Republic, Netherlands, Slovakia, Sweden, Togo, Jamaica, Hungary, Chile, New Zealand, Guatemala, Ghana, Bulgaria, Malaysia, Tunisia, Peru, Dominican Republic, Sri Lanka, Colombia. Not shown (in descending order of %) for design origin: Georgia, China, India, South Korea, St. Vincent and Grenadines, Philippines, Ireland, Somalia, Spain, Brazil, Jamaica, Colombia, UAE, Austria, Singapore, New Zealand, Serbia, Russia, Benin, Cuba, Denmark, Bulgaria, Turkey, Jordan, Ghana, Ukraine, Nigeria, Malaysia. Back to Section 2.

Table A.3: Relative Purchase Probabilities by Retail Format and Origin

	(1) Import	(2) NAFTA	(3) China	(4) Europe
Other	0.077	0.063	0.004	0.004
	(-)	(-)	(-)	(-)
Dollar Store	0.064	0.035	0.011	-0.001
	(0.008)	(0.007)	(0.001)	(0.000)
Discount Store	0.018	0.018	-0.001	-0.000
	(0.004)	(0.004)	(0.001)	(0.000)
Drug Store	0.017	0.018	-0.002	0.000
	(0.005)	(0.004)	(0.001)	(0.000)
Grocery Store	0.051	0.045	-0.002	0.005
	(0.005)	(0.005)	(0.001)	(0.001)
N	2.75M	2.75M	2.75M	2.75M
R^2	0.040	0.041	0.009	0.008
Controls:				
Income Decile	X	X	X	X
Pop. Density Decile	X	X	X	X
$\bar{\boldsymbol{x}}_i$	X	X	X	X
Category FE	X	X	X	X
Half-Year FE	X	X	X	X
Region FE	X	X	X	X

Table A.3 provides estimates of the relative probability of purchasing a variety from a given origin conditional on the format visited. All regressions are of a linear probability model with a binary outcome indicating whether a purchase was from origin m. All models include controls for household income decile, population density decile, education, race, age, presence of children, and married household heads. I also include region, half-year, and product category fixed effects. Estimates are relative to the "Other" retail format category, which is the reference category, and I provide the mean purchase probability for each origin in this row. Observations are weighted by the aggregate quantity, in weight, or each purchase. Standard errors are provided in parentheses and are clustered at the product-region-half-year level. Back to Section 2.

Table A.4: Household Type Descriptive Statistics

h	(1) Count #	(2) Weight	(3) **Income** 1000 USD	(4) **Density** Pop/Sq.mile	(5) %College	(6) %Black	(7) %Retired	(8) %Urban
1	2279	0.060	**28.4**	**614**	33.5	10.7	45.8	97.1
2	3902	0.090	**88.7**	**1049**	74.2	12.5	29.0	98.4
3	3425	0.074	**81.9**	**192**	65.5	7.5	28.4	86.7
4	2958	0.071	**30.7**	**33**	30.2	5.3	42.5	9.9
5	1602	0.058	**61.4**	**9821**	59.7	28.0	30.2	99.9
6	3258	0.068	**74.4**	**84**	55.7	6.0	31.9	45.0
7	2399	0.053	**77.4**	**32**	55.6	4.0	28.8	11.0
8	2159	0.050	**51.7**	**8**	38.5	3.8	37.1	2.7
9	3909	0.095	**81.2**	**2159**	67.9	13.7	29.6	99.9
10	2489	0.045	**42.5**	**268**	42.5	8.8	42.3	91.3
11	2901	0.075	**33.3**	**1967**	40.2	16.3	43.8	99.6
12	1238	0.050	**11.0**	**1652**	32.4	15.4	37.2	98.7
13	2573	0.073	**24.0**	**119**	29.5	6.5	41.2	58.3
14	2752	0.050	**51.4**	**981**	51.2	14.8	39.4	98.9
15	4193	0.089	**82.9**	**480**	68.9	11.5	30.0	96.2
Total	42,037	1.000	57.6	1276	51.6	11.1	35.2	75.0

Table A.4 provides descriptive statistics for the household clusters h used in this paper. Column (1) provides the number of households in each cluster, and Column (2) provides the weight assigned to typ h based on the NielsenIQ projection factors. Columns (3) and (4) provide average income and average ZIP population density for each type h, and are in bold as these are the only two dimensions used to construct the types h. Column (5) to Column (8) provide percentage of households within each type that, respectively, have a college degree, are Black, are retired, and live in an urban ZIP based on the USDA ERS RUCA classification system. Back to Section 3.1.

Table A.5: Median Elasticity Estimates

Income	28.4	88.7	81.9	30.7	61.4	74.4	77.4	51.7	81.3	42.5	33.3	11.0	24.0	51.4	82.9
Pop. Density	613	1049	192	33	9821	84	32	8	2159	268	1967	1652	119	981	480
h	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Category	Median Own-Price Elasticity: ϵ_{jj}														
Hair Care	-2.67	-0.91	-2.21	-1.90	-6.60	-2.26	-1.54	-4.26	-2.33	-2.17	-1.33	-1.02	-2.31	-1.39	-1.63
Body Soap	-3.71	-1.63	-2.96	-2.47	-2.77	-2.63	-1.39	-0.10	-0.10	-3.10	-0.13	-1.55	-4.68	-1.87	-2.23
Skin Care	-1.99	-0.67	-1.87	-4.43	-6.79	-1.79	-2.13	-5.03	-3.95	-2.90	-2.40	-0.88	-3.07	-1.28	-1.56
Deodorant	-2.44	-1.11	-1.90	-1.80	-2.47	-1.80	-1.48	-1.71	-1.24	-2.01	-1.09	-0.39	-2.89	-1.83	-1.35
Toothpaste	-2.21	-1.79	-2.14	-2.18	-4.18	-1.72	-1.70	-3.24	-2.17	-2.13	-1.47	-0.84	-2.23	-1.47	-1.72
Facial Care	-1.29	-0.86	-1.35	-1.45	-3.99	-1.18	-1.16	-2.56	-1.78	-1.16	-0.83	-0.70	-1.47	-0.85	-1.49
Hair Styling	-3.48	-1.69	-3.99	-3.15	-0.86	-4.86	-3.25	-5.30	-1.30	-3.75	-1.24	-2.39	-5.72	-3.04	-5.77
Face Cosmetics	-1.35	-1.31	-1.31	-1.39	-1.05	-1.17	-1.74	-1.55	-1.51	-0.77	-1.23	-1.98	-1.88	-1.42	-1.42

Table A.5 provides the median elasticity for each household type and the eight product categories included in the estimated models of this paper. Household type characteristics are given in the top two rows, with average income reported in thousands of USD, and population density reported as average population per square mile. Back to Section 3.5.

Table A.6: Projecting Origin Own-Price Elasticity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Dependent Variable: Barcode Log Own-Price Elasticity (e_{jj})						
$\ln GDPpc_j^p$	-0.025^{***}		-0.021^{**}		-0.023^{***}	-0.019^{**}	-0.016^{*}
	(0.008)		(0.008)		(0.008)	(0.009)	(0.009)
$\ln GDPpc_j^d$		-0.075^{***}		-0.083^{***}	-0.072^{***}	-0.091^{***}	-0.089^{***}
		(0.016)		(0.017)	(0.016)	(0.018)	(0.018)
$\ln p_j$							-0.007^{*}
							(0.004)
Constant	0.721^{***}	0.927^{***}	0.707^{***}	0.960^{***}	0.996^{***}	1.057^{***}	1.047^{***}
	(0.026)	(0.063)	(0.028)	(0.066)	(0.067)	(0.074)	(0.074)
N	2993	3034	2993	3024	2993	2970	2970
R^2	0.563	0.569	0.573	0.590	0.566	0.586	0.587
FE:							
Category	x	x	x	x	x	x	x
Nest						x	x
Prod. Origin				x			
Design Origin			x				

Table A.6 provides estimates from a linear regression at the barcode level. I restrict the sample to only foreign-produced barcodes, and estimate correlations between the estimated own-price elasticity of each barcode e_{jj} and the origin characteristics of that same barcode: the GDP per capita of the country in which j was produced (superscript p) and the GDP per capita of the country in which j was designed (superscript d). Columns (3) and (4) provide fixed effects to account for all variation explained by, respectively, design and production origins, whereas Columns (6) and (7) extend the category-specific fixed effects used throughout by including category-nest fixed effects. Column (7) includes the logarithm of the price of j as an additional control. Standard errors are provided in parentheses. Back to Section 5.2.

Figure A.1: Retail Format Trip Propensity by Population Density

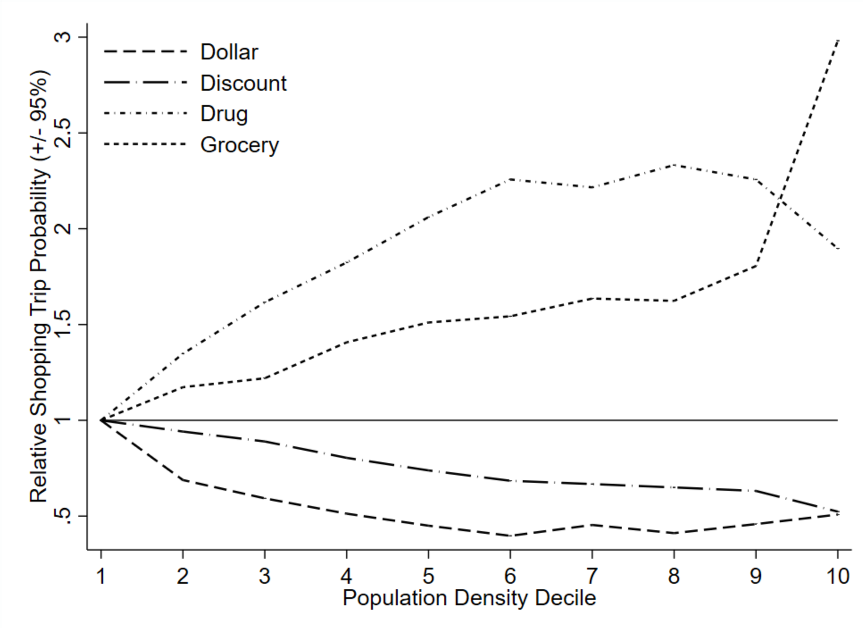
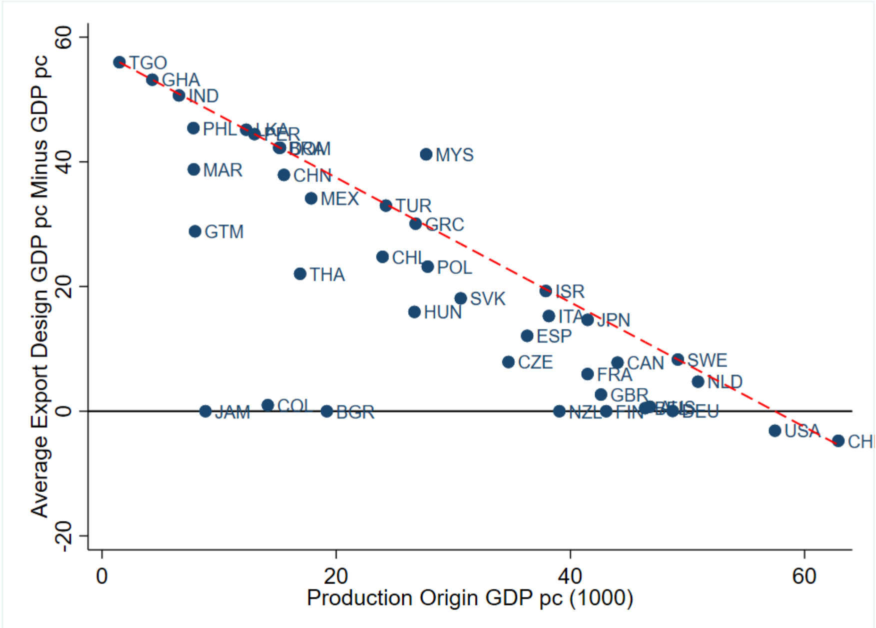


Figure A.1 provides estimates of four separate linear probabilit models estimates at the household-shopping-trip level. Each model corresponds to a retail format with the dependent variable a simple indicator if a given shopping trip was to the format in question. 95% confidence intervals are provided, and observations are weighted by the Nielsen projection factor weights for each household. All regressions include controls for household income decile, education, race, age, presence of children, married household heads, region ($R = 5$) and half-year time period. Back to Section 2.

Figure A.2: Production and and design based average GDP per capita of exports



[Figure A.2](#) provides expenditure-weighted average GDP per capita of design origin for all varieties exported from production origin m to the US. The x-axis provides the production origin GDP per capita of m, while the y-axis provides the differential between the average design-based GDP per capita of goods exported from m and m's GDP per capita. Countries on the solid horizontal line exhibit exports with an average design origin GDP per capita that is equal to the GDP per capita of that country, and all countries above this line exhibit exports to the US with design origin GDP per capita greater than m's GDP per capita. The red dash line denotes where a country would be if all of their exports to the US were designed by US firms. Back to Section 2.

Figure A.3: Distribution of estimates for $\hat{\alpha}$ (top-left), $\hat{\sigma}$ (top-right), first-stage K-P F-Statistics (bottom-left), and median elasticity (bottom-right)

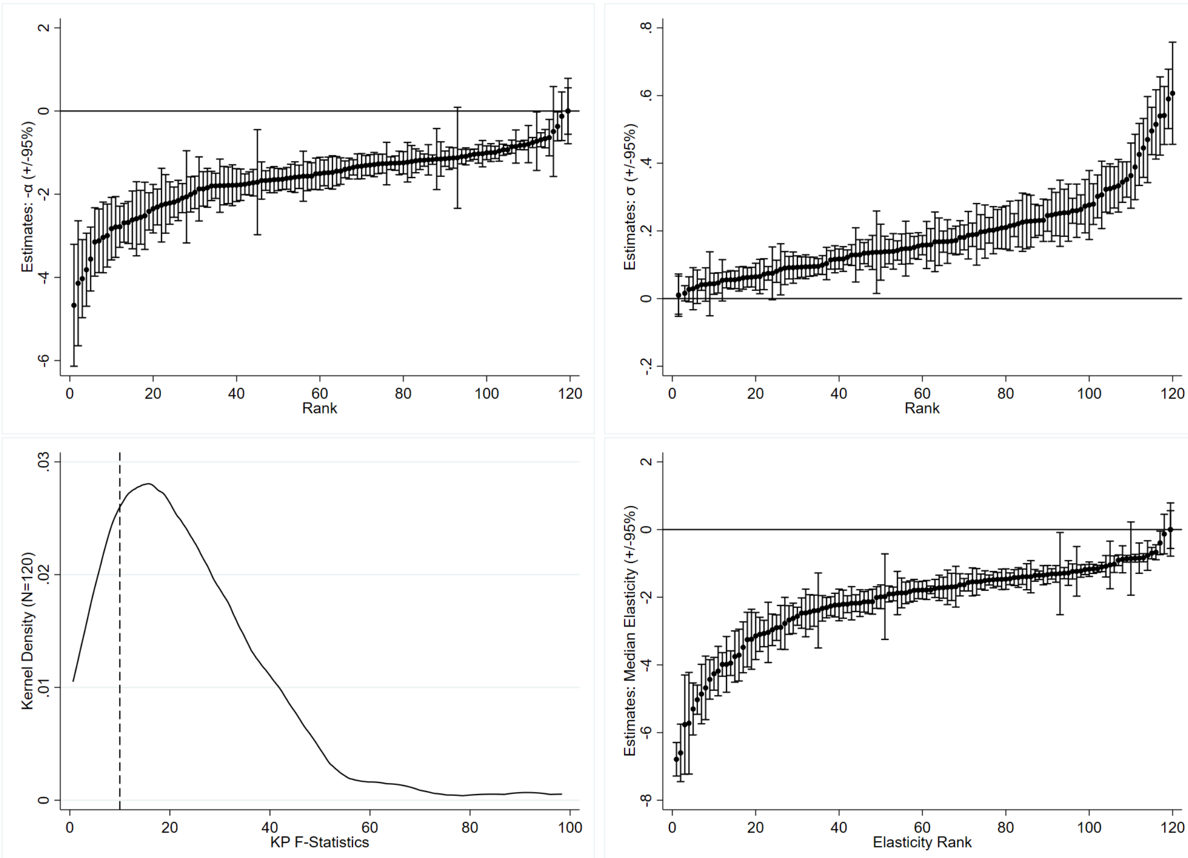
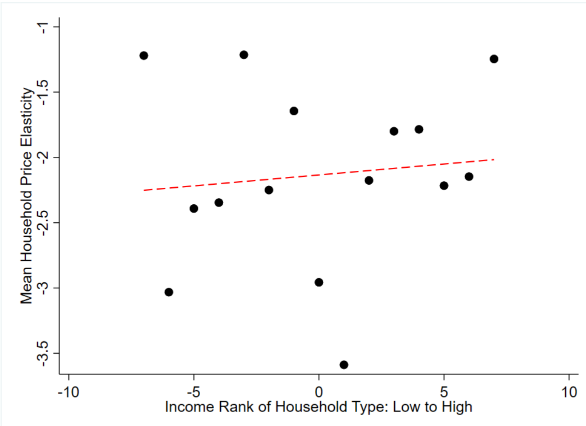


Figure A.3 provides distributions of the key parameters estimated within the 120 separate demand systems. From top-left clockwise, these distributions are for $\hat{\alpha}$, $\hat{\sigma}$, the median elasticity across all barcodes, and the distribution of first-stage K-P F-statistics. Back to Section 3.5.

Figure A.4: Average household type elasticity projected on income rank (left) and population density rank (right)



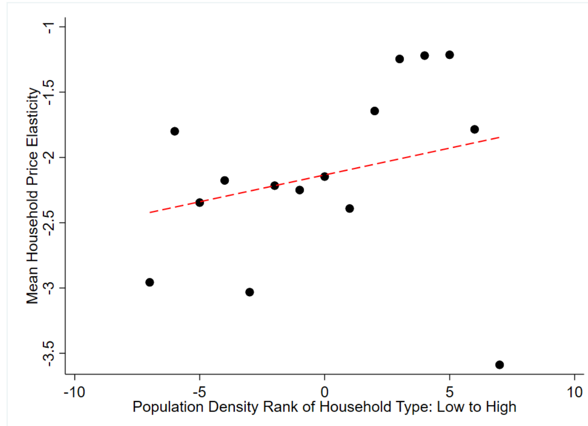


Figure A.4 average household-type own-price elasticities across all categories projected onto household income rank (left) and population density rank (right). Both rankings start at low income and low population density (-7) and increase to the highest income/population density types at $+7$. Back to Section 3.5.

Figure A.5: County-Level Welfare Costs of 10ppt Shock to all Imports by Production

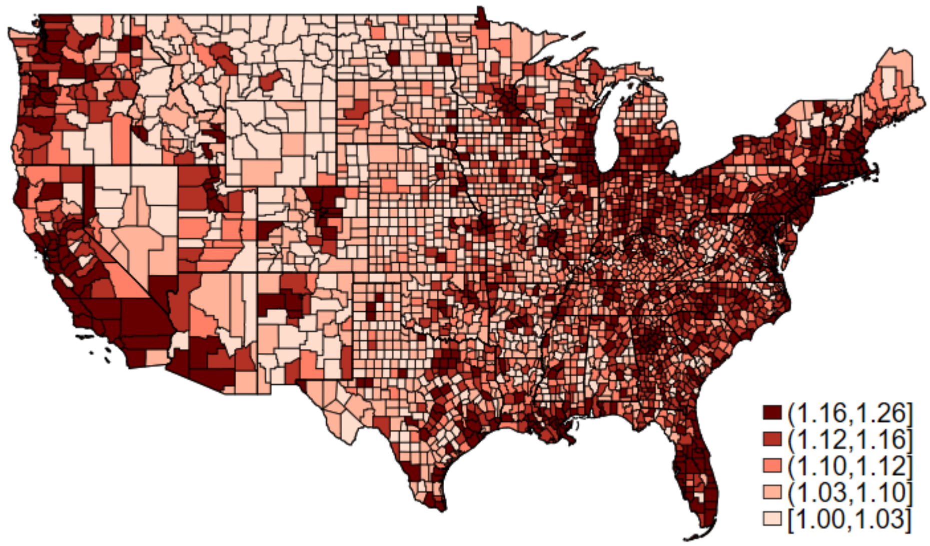


Figure A.5 provides estimates of the county-level relative welfare costs of a 10ppt increase in the relative price of all varieties produced abroad. Counties are categorized into quintiles of increasing relative costs. Back to Section 4.1.

Figure A.6: County-Level Welfare Costs of 10ppt Shock to CAN/MEX by Production

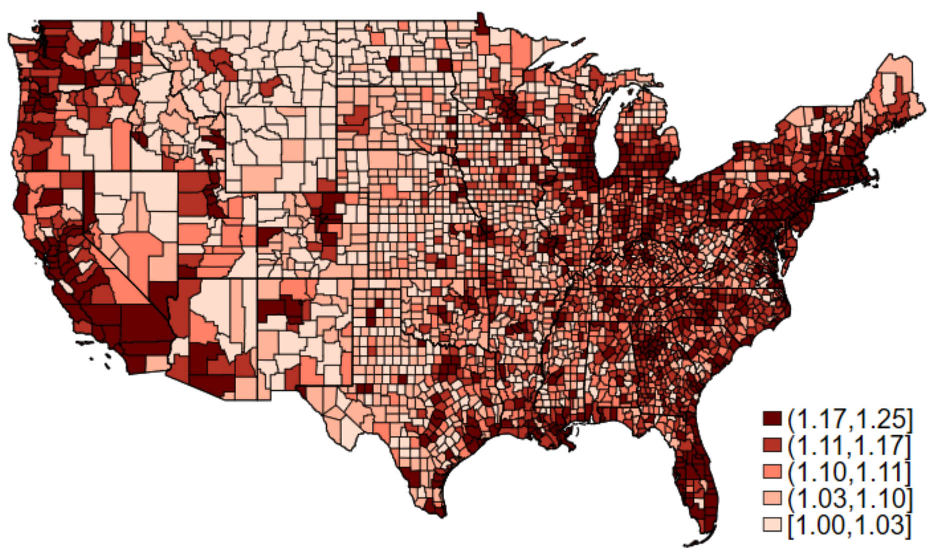


Figure A.6 provides estimates of the county-level relative welfare costs of a 10% tariff increase for all NAFTA varieties, excluding the USA. Counties are categorized into quintiles of increasing relative costs. Back to Section 4.1.

Figure A.7: County-Level Welfare Costs of 10ppt Shock to Europe by Production

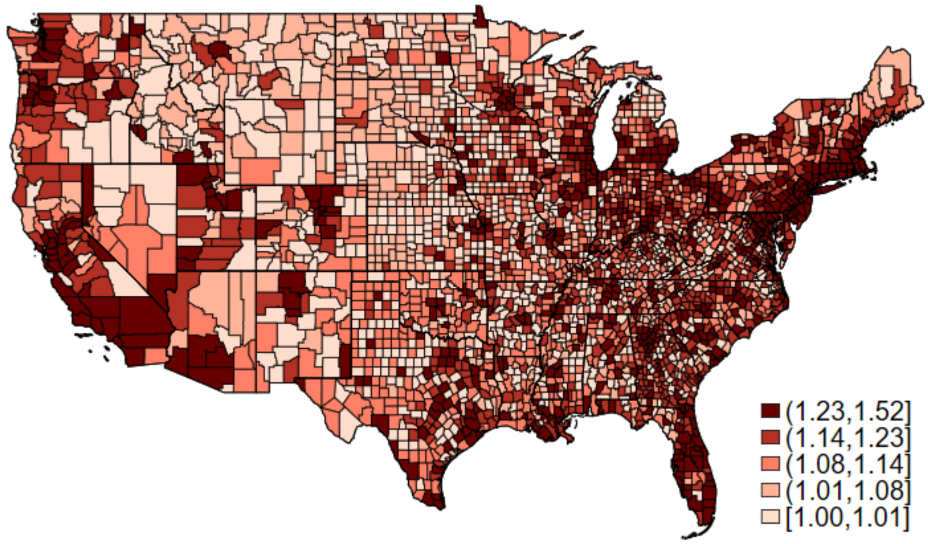


Figure A.7 provides estimates of the county-level relative welfare costs of a 10% tariff increase on all European varieties. Counties are categorized into quintiles of increasing relative costs. Back to Section 4.1.

Figure A.8: County-Level Welfare Costs of 10ppt Shock to China by Production

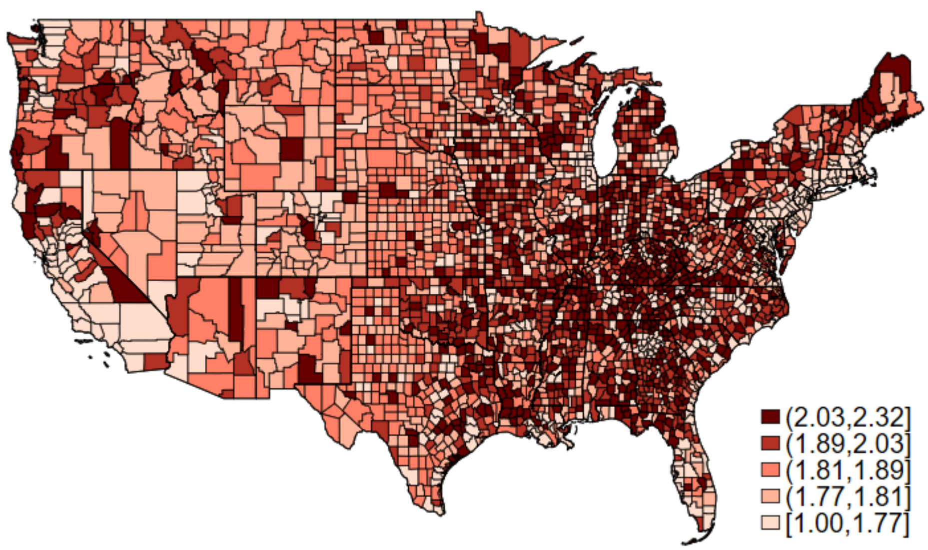


Figure A.8 provides estimates of the county-level relative welfare costs of a 10% tariff increase on all Chinese varieties. Counties are categorized into quintiles of increasing relative costs. Back to Section 4.1.

Figure A.9: Comparison of welfare costs between homogenous and heterogeneous elasticity models: 10ppt price shock to CAN/MEX goods

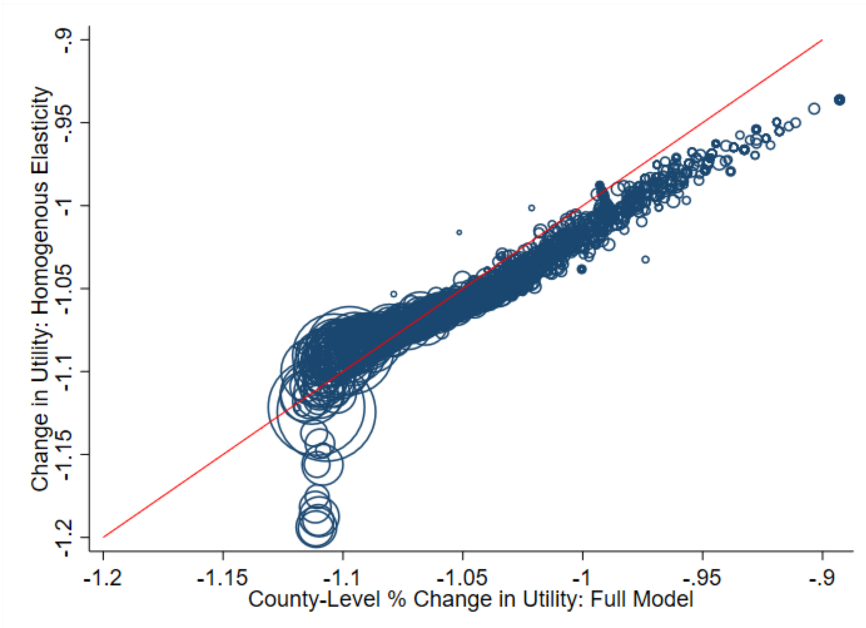


Figure A.9 provides a county-level comparison of costs associated with a 10ppt price shock to CAN/MEX goods under the homogenous elasticity model (y-axis) and the full model (x-axis). The dotted line represents $y = x$. Back to Section 4.2.

Figure A.10: Comparison of welfare costs between homogenous and heterogeneous elasticity models: 10ppt price shock to Europe goods

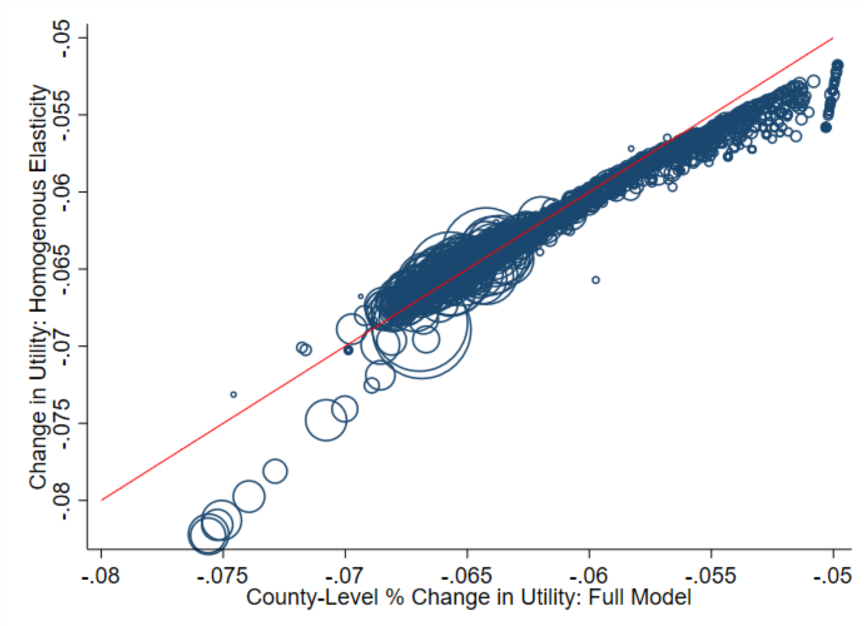


Figure A.10 provides a county-level comparison of costs associated with a 10ppt price shock to Europe goods under the homogenous elasticity model (y-axis) and the full model (x-axis). The dotted line represents $y = x$. Back to Section 4.2.

Figure A.11: Comparison of welfare costs between homogenous and heterogeneous elasticity models: 10ppt price shock to China goods

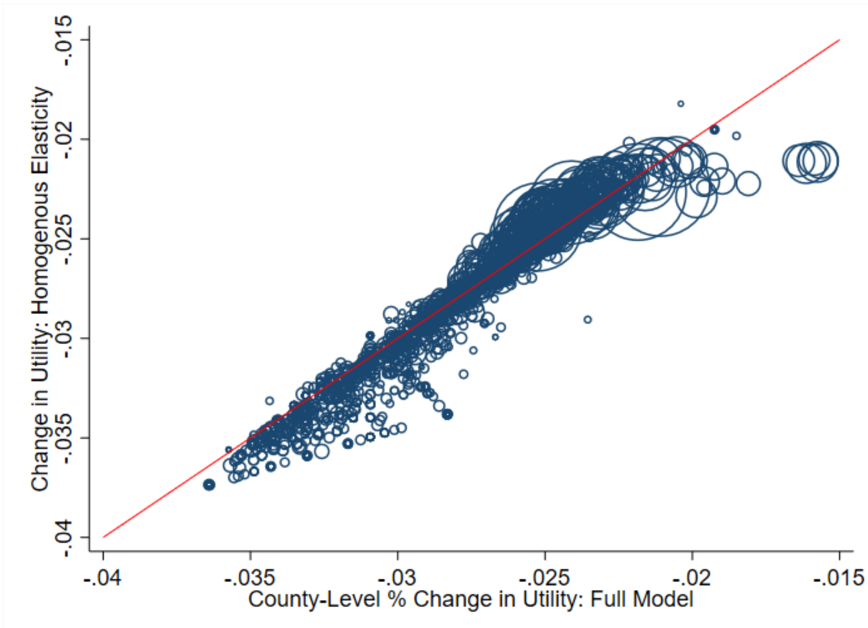


Figure A.11 provides a county-level comparison of costs associated with a 10ppt price shock to China goods under the homogenous elasticity model (y-axis) and the full model (x-axis). The dotted line represents $y = x$. Back to Section 4.2.

Figure A12: Production origin quality with and without multinationals

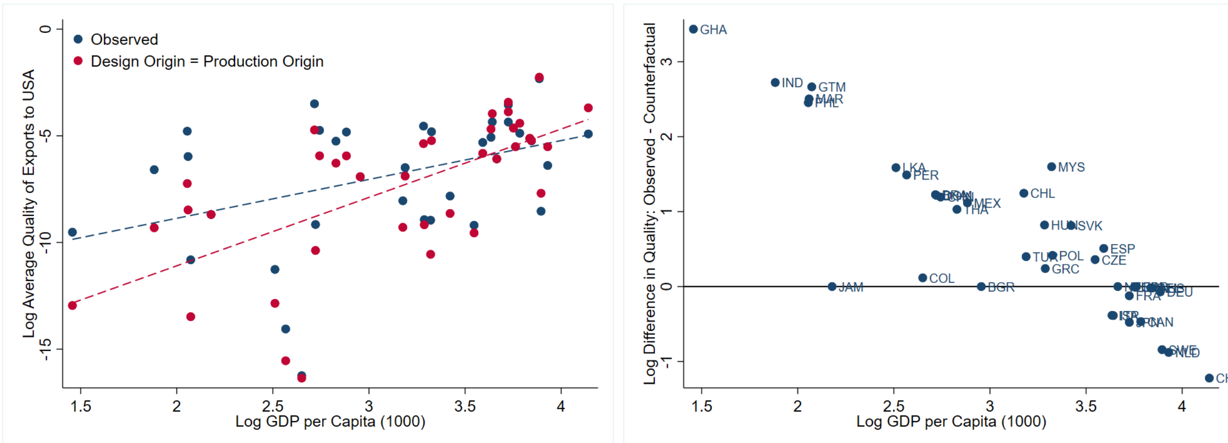


Figure A.12 provides estimates of the average demand-shifter of barcodes produced in a given production origin compared to a counterfactual in which the design origin of all varieties is forced to be identical to the production origin. The left panel then plots the observed quality measures (blue) versus the counterfactual measures (red), while the right panel simply plots this difference directly. Back to Section 5.3.

A.2 Text Analysis and Clustering Algorithm Details

This Appendix provides a description of the method used to turn label packaging into a measure of text similarity across all bilateral variety pairs within the same category. Label Insight utilizes an AI to read off all text from the packaging of the consumer packaged goods in their database. These text data then constitute strings of various length that provide the brand of each barcode as well as a description/branding. In order to clean the data, I remove all references to the brand of each variety that appear in the label packaging, before then adding the brand name back to the beginning of each description. The goal is to have the brand of each variety be considered in a similar way and only appear once in the text description.

I then apply the "qgram" method of turning two strings into a continuous measure of (dis)similarity. Consider two strings, A and B, with character counts $\mid A \mid$ and $\mid B \mid$. The qgram algorithm calculates, for each increasing value of q from one to $\min(\mid A \mid, \mid B \mid)$, the number of consecutive sequences of length q that are common to both A and B. This value is then divided by the absolute difference in possible sequences of length q in order to normalize this measure.

Consider the following three examples from the toothpaste category:

- **A**: "COLGATE ANTICAVITY ANTIGINGIVITIS TOOTHPASTE"
- **B**: "COLGATE TOTAL PROTECTION TOOTHPASTE: SPEARMINT"
- **C**: "SENSODYNE TOOTHPASTE FOR SENSITIVE TEETH: COOL MINT"

The distance measures for these three descriptions are the following: $\Sigma^t_{AB} = 26$, $\Sigma^t_{AC} = 35$, and $\Sigma^t_{BC} = 27$. As one might expect, **A** and **B** are deemed the most similar due to being of the same brand. The "mint" connection between **B** and **C**, however, makes them more similar than **A** and **C**, which share very little in common other than the word toothpaste. In order to understand how brand labels play a role, consider that the distance between **B** and **C** decreases to 15 when the brand of **C** is changed from "Sensodyne" to "Colgate". Similarly the distance between **A** and **B** decreases from 26 to 24 when the word "MINT" is added to the end of **A**'s description.

I know provide a description of the Partitioning Around Medoids (PAM) clustering algorithm as well as the Silhouette Method of selecting the "optimal" number of clusters to implement. The PAM algorithm takes as input a dataset of N observations with a distance measure between each observation i and j of Σ_{ij}, and a pre-determined number of clusters K. The PAM algorithm is similar to the K-Means/K-Medians algorithm in that the PAM algorithm seeks to define a centroid for each cluster and then minimize the distance between each observation and the centroid of that observation's assigned cluster. The PAM algorithm differs from K-Means in that the centroids in the PAM algorithm consist of data points - which are called "medoids" - whereas in the K-Means approach the centroid of each cluster is a point in the observation space, rather than an observation itself. In this case, given that I am feeding into the algorithm a dissimilarity measure rather than a set of real observations, the PAM algorithm provides a viable alternative to the K-Means algorithm, which requires real data points from which to calculate the centers of each cluster.

The algorithm begins by "greedily" selecting K of the N observations to act as medoids, and assigns all other observations to be in a cluster with their closest medoid. The algorithm then iteratively considers swapping a non-medoid observation with a medoid observation and considers the cost of doing so. In this case the cost refers to the distance between all observations and their respective medoids once the new clusters have been formed. This process continues iteratively until their are no additional swaps that would lower the overall cost of the clusters formed.

As mentioned in the text, this process requires both a pre-determined distance measure between all observations as well as pre-specified number of clusters. In order to select the number of clusters that leads to the best "fit" of the data, I employ the Silhouette Method. This method iterates over a range of cluster counts determined by the researcher and selects the number of clusters that maximize the "Silhouette Width".

Define the Silhouette Width for any given observation as $s(i)$. To calculate $s(i)$, the algorithm first calculates the average distance between an observation i and all observations belonging to the same cluster. Define this average distance as $a(i)$. The algorithm then calculates the additional average distances from observation i to the observations that form every other cluster. This leads to a separate measure of average distance for observation i to every other cluster. Define $b(i)$ as the average distance from i to the "nearest" cluster other than that which i has been assigned to[35]. The Silhouette Width $s(i)$ can then be calculated as:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

As shown, $s(i)$ will be closest to 1 when the difference $b(i) - a(i)$ is large, meaning that observation i is closer to the observations within it's cluster compared to observations in the "nearest neighbouring" cluster. The Silhouette Method then averages across all $s(i)$ for all observations $i \in N$ and calculates a total score that is bounded by $[0, 1]$, with a value of 1 signifying clusters that are perfectly distinct from one another. By iterating over possible cluster counts K, I use the Silhouette Method to then select the number of clusters that returns the largest Silhouette Score. Given that the algorithm is computationally demanding, I only search over cluster counts that are multiples of five. I also begin at $K = 20$ and increase the cluster number by two until $K = 40$, the Silhouette Method then selects which cluster count provides the closest fit to the data.

[35]Think of this as the average distance from i to i's nearest neighboring cluster.

Table A.8: Text Cluster Summary for Deodorant

Brand	Descriptive Text	Cluster
Jason	Calming Deodorant Stick: Lavender	1
Honest	Lavender Vanilla Deodorant	1
Schmidt's	Natural Deodorant: Lavender and Sage	1
Arm & Hammer	Antiperspirant Deodorant: Confidence	2
Secret	Antiperspirant Deodorant: Brazil Rainforest Mist	2
Axe	Antiperspirant: Black Chill	2
DKNY	Eau de Parfum Spray	3
Luxe	Hair and Body Perfume Mist: Aqua Moon	3
BOD	Body Spray: Blue Surf	3

Table A.8 provides an example of the market segments created by the text-similarity approach to placing varieties into nests. I provide three examples of the barcode descriptive text in each of the first three market segment/nests for deodorant barcodes.