

McCully, Brett A.; Jaccard, Torsten; Albert, Christoph

Working Paper

Immigrants, Imports, and Welfare: Evidence from Household Purchase Data

CESifo Working Paper, No. 12278

Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: McCully, Brett A.; Jaccard, Torsten; Albert, Christoph (2025) : Immigrants, Imports, and Welfare: Evidence from Household Purchase Data, CESifo Working Paper, No. 12278, Munich Society for the Promotion of Economic Research - CESifo GmbH, Munich

This Version is available at:

<https://hdl.handle.net/10419/334634>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

CES ifo

12278
2025

November 2025

Working Papers

Immigrants, Imports, and Welfare: Evidence from Household Purchase Data

Brett A. McCully, Torsten Jaccard, Christoph Albert

CES ifo

Imprint:

CESifo Working Papers

ISSN 2364-1428 (digital)

Publisher and distributor: Munich Society for the Promotion
of Economic Research - CESifo GmbH

Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740

Email office@cesifo.de
<https://www.cesifo.org>

Editor: Clemens Fuest

An electronic version of the paper may be downloaded free of charge

- from the CESifo website: www.ifo.de/en/cesifo/publications/cesifo-working-papers
- from the SSRN website: www.ssrn.com/index.cfm/en/cesifo/
- from the RePEc website: <https://ideas.repec.org/s/ces/ceswps.html>

Immigrants, Imports, and Welfare: Evidence from Household Purchase Data*

Brett A. McCully[†]  Torsten Jaccard[‡]  Christoph Albert[§]

November 19, 2025

Abstract

Who buys imports? By augmenting U.S. grocery purchase data to include origin countries of both products and households, we provide the first evidence that immigrants exhibit substantially stronger preferences for imported consumer goods than natives. We develop and estimate a quantitative trade model to show that immigrants also reduce trade costs and expand the effective market size for foreign goods, thereby increasing local import supply for all households. Overall, however, immigrants generate considerably more local import expenditure via their own purchases than via spillovers to natives, with profound implications for the distributional costs of a negative trade shock, such as an import tariff: the average within-county difference in welfare costs between immigrants and natives is over six times the across-county standard deviation in native household costs.

Keywords: Gains from trade, heterogeneous preferences, spillover effects.

JEL Categories: F22, J31, J61, R11.

*We are grateful to Barthélemy Bonadio, Daniel Chiquiar, Keith Head, Wookun Kim, Munseob Lee, Andrei Levchenko, Hong Ma, Peter Morrow, Riccardo Turati, and Jonathan Vogel for helpful comments, as well as participants at various conferences and seminars. We thank Ilaria Malisan and Andrea Montalbano for research assistance. The analysis in this paper is our 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 our own and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein. All errors are our own.

[†]Collegio Carlo Alberto; brett.mccully@carloalberto.org

[‡]Vancouver School of Economics; tjaccard@mail.ubc.ca

[§]University of Turin and Collegio Carlo Alberto; christoph.albert@carloalberto.org

1 Introduction

Who buys imports? This question is fundamental to understanding the distributional effects of trade shocks, such as tariff increases, which incur greater costs for households spending more on imported goods. The limited existing evidence focuses almost exclusively on differential exposure to international trade by consumer income (Fajgelbaum and Khandelwal 2016; Cravino and Levchenko 2017; Borusyak and Jaravel 2021; Galle et al. 2023; Auer et al. 2024).

This paper is the first to explore how the consumer gains from trade vary with nativity. Immigrants’ preferences and habits may differ from those of natives (Fernández and Fogli 2009; Atkin 2016; Miho et al. 2023), and to the extent that these differential preferences affect immigrants’ demand for foreign goods, they will be differently impacted by trade shocks. Moreover, immigrants may also affect natives’ gains from trade by increasing local import availability (Gould 1994, Head and Ries 1998, Parsons and Vézina 2018), thereby shaping the distribution of the gains from trade across natives.

Quantifying the separate effects of immigrants on import demand and supply requires detailed data that simultaneously capture households’ nativity, geographic location and the origin countries of their purchases. We construct, to our knowledge, the first dataset to contain all three elements by augmenting U.S. grocery scanner data at the household level to include the origin country of both households and products.¹

We obtain three key results. First, immigrants have a stronger preference for imports than natives, leading them to accrue a disproportionate share of the gains from trade. Second, immigrants increase import supply by reducing the fixed costs of importation and increasing market size, thereby enhancing natives’ gains from trade. Third, the majority of immigrant-induced local import expenditure is attributable to purchases made by immigrants themselves, rather than via spillovers to native import expenditure.

Regarding import demand, our data allow us to provide the first direct test of whether immigrants differ from natives in their import expenditure and by how much. Applying

¹We are unaware of any alternative data that provide information on both household nativity and import shares. For example, the Consumer Expenditure Survey, which may be combined with other datasets to measure import expenditure shares on cars as in Borusyak and Jaravel (2021), does not ask about respondents’ nativity.

the general welfare formula of [Arkolakis et al. \(2012\)](#), immigrants accrue consumer gains from trade within the grocery sector that are 42% greater on average than those of native households. Furthermore, we show that within-county variation across households in import expenditure explains the vast majority of this differential, highlighting the role of preferences rather than sorting of immigrants into locations with high import supply. Our estimates suggest that, on average, immigrants spend 25% more on all grocery imports and 127% more on imports specifically from their own origin country when compared to an observationally identical native household living in the same U.S. county. Nativity is substantially more important than income in explaining variation in the gains from trade. The within-county difference in import expenditure between a household in the highest income quintile and an observationally identical household in the lowest income quintile is only +8.5%, a third of the within-quintile gap between an immigrant and an observationally identical native household.

Regarding import supply, we develop and estimate a heterogeneous firms model of international trade à la [Melitz \(2003\)](#) extended to allow immigrants to affect local import costs. We leverage unique features of our dataset, including barcode-level price and variety counts by origin, in order to estimate the various channels through which immigrants affect local import supply. The leave-out push-pull instrumental variables developed by [Burchardi et al. \(2019\)](#) provide exogenous variation in the local population of immigrants by origin country. Our estimates imply that immigrants reduce the fixed costs of trade but have a negligible effect on variable trade costs.²

We combine the estimated import demand and supply effects of immigrants within our full model and run three counterfactual exercises. In our first counterfactual, we consider what would happen to imports and native welfare in the absence of preference and trade cost effects of immigrants while holding the local population size fixed. National grocery import expenditure would fall by almost 9%, akin to doubling prevailing tariffs applied to grocery goods. Immigrants’ stronger preferences for imports explain 60% of the immigrant-import

²That immigrants reduce fixed trade costs, but not variable trade costs, aligns with the previously untested assumption of [Peri and Requena-Silvente \(2010\)](#). We confirm this result within the NielsenIQ store-level scanner dataset: immigrants have no effect on import prices across stores, consistent with uniform pricing within large retail chains ([DellaVigna and Gentzkow 2019](#)). While we observe a positive effect of immigrants on import sales across stores, the immigrant effects documented in our household purchase records occur primarily through shopping trips to smaller retailers not present in the NielsenIQ retailer panel.

elasticity, with the remaining 40% attributable to changes in import supply common to all households. Approximately half of the import supply effect occurs via reduced fixed costs of trade, while the remainder is due to a novel channel we recover through our structural model: immigrants' heightened import preferences encourage foreign firm entry into local markets which in turn increases import variety for all households. The welfare gains for natives due to immigrant-induced trade—equivalent to about 0.1% of grocery expenditure, or \$800 million—are 70% less than one would infer from aggregate regional trade flow data and standard welfare formulas, such as [Arkolakis et al. \(2012\)](#), as immigrants disproportionately increase local import expenditure via their own purchases.³

In a second counterfactual exercise, we show that the presence of immigrants yields significant consumption welfare gains for natives via an additional market size effect. If immigrants and their expenditure were to entirely disappear, native welfare from grocery consumption decreases by almost 8%—equivalent to over \$59 billion in lost welfare nationally—primarily due to the loss of product variety resulting from smaller markets.⁴ These losses are highly concentrated among high-income and urban households.

Our final counterfactual simulates the heterogeneous impact of an increase in import tariffs leveraging our full model of import demand and supply. Specifically, we simulate a variable cost shock to all imported grocery goods, and recover the distributional welfare costs across households and counties. The average within-county difference in costs between immigrants natives is over six times larger than the across-county standard deviation in costs for native households. Variation in costs across geographies and income groups explain a fraction of the variation in costs between immigrants and natives.

This paper provides the first direct evidence that immigrants, as consumers, derive substantially larger gains from trade than native households in the U.S. By encouraging foreign product entry into local grocery markets, immigrants also generate import consumption benefits for natives. Still, reductions in the immigrant population and higher costs of trade both incur relatively lower welfare losses for natives, especially those with lower education.

³Our estimated welfare effect is sizable. For comparison, [Brinatti and Morales \(2025\)](#) find that a 20% increase in immigrants raises native welfare by 0.1%, while [Caliendo et al. \(2019\)](#) and [Hsieh and Ossa \(2016\)](#) estimate the welfare impact of the China shock to be between 0.03% and 0.2%.

⁴While this scenario is unrealistic, it provides a useful benchmark (comparable to the standard gains-from-trade exercise) by which to value immigration in terms of native expenditure.

Related literature. By estimating consumer heterogeneity in exposure to trade shocks we contribute to a growing literature on the impact of trade across consumers (Fajgelbaum and Khandelwal 2016; Cravino and Levchenko 2017; Bai and Stumpner 2019; Hottman and Monarch 2020; Borusyak and Jaravel 2021; Faber and Fally 2022; Jaccard 2023; Auer et al. 2024). This paper is the first to document that immigrants are substantially more exposed to trade shocks than non-immigrants. When assessing international trade exposure, nativity matters more than income, which has been the focus of prior work.

We also add to the literature on immigrants’ trade impact by leveraging our unique data to unpack the mechanisms driving this relationship, which allows us to compute the welfare gains across heterogeneous households from immigrant-induced trade creation. In contrast, existing studies almost exclusively use aggregated data on region-to-region trade flows, such that the separate effects of immigrants on import demand and supply cannot be observed directly (Gould 1994; Head and Ries 1998; Combes et al. 2005; Peri and Requena-Silvente 2010; Parsons and Vézina 2018; Steingress 2018; Bonadio forthcoming).⁵

Thus, our study contributes to the ongoing public discourse on the benefits and costs of immigration. A vast literature has focused on the way in which immigrants affect the labor market outcomes of native workers (e.g., Card 2001, Borjas 2003, Ottaviano and Peri 2012, Dustmann et al. 2017, Monras 2020, Burstein et al. 2020). We quantify a novel margin by which immigrants benefit natives: increasing local imported product variety.⁶ Furthermore, while studies on the effects of immigration on the labor market carefully consider distributional effects (e.g., Dustmann et al. 2013 and Llull 2018), the consumption-side distributional effects have thus far been ignored.

This paper proceeds as follows. We describe our data and present stylized facts in Section 2. In Section 3, we characterize heterogeneity in preferences between immigrants and natives and therefore in the gains from trade. Section 4 microfoundations and estimates how immigrants affect import supply. Section 5 describes our counterfactuals using the full quantitative model. Section 6 concludes.

⁵The demand effects documented here may be stronger for non-durable consumer goods, as opposed to the intermediate and capital goods included in many prior studies.

⁶Iranzo and Peri (2009), Di Giovanni et al. (2015), and Aubry et al. (2016) study the aggregate variety effects of immigration but with a focus on immigrants expanding production in high-productivity locations.

2 Data and Stylized Facts

2.1 Expenditure on Consumer Packaged Goods

We use two datasets that link household characteristics—including country of birth—to grocery import expenditures: the NielsenIQ household panel scanner dataset and barcode country-of-origin data from Label Insight Inc.

NielsenIQ Household Panel Scanner Data: These data consist of a panel covering approximately 90,000 U.S. households and all grocery purchases at the barcode level. Variables include detailed household demographic information and county of residence as well as barcode-level expenditure, price, date, and store for each purchase.⁷

In 2008, NielsenIQ distributed the “Tell Me More About You” survey, which included questions about respondents’ birth place.⁸ Thus in subsequent years we observe the country of origin for the subset of households present in the data since 2008. If household members have mixed nativity and only one member was born abroad, we assign the household to the country of the immigrant member. In the rare case that a household has multiple foreign-born members from different origins, we assign the household to the larger country of origin as measured by the total immigrant population in the U.S.

Barcode Country of Origin: We merge the NielsenIQ data with barcode-specific country-of-origin information purchased from Label Insight Inc., a firm that specializes in scraping and organizing information found on the labels of consumer packaged goods.⁹ Label Insight uses a computer vision algorithm to extract text from the packaging for thousands of barcodes sold across major retail chains in the U.S. Since imported goods in the U.S. are required to declare on their labels the country in which the good underwent the most “substantial transformation,” the algorithm incidentally recovers the origin country for each collected barcode. Label Insight only covers a subset of all barcodes and their coverage is best for food and beverages, alcohol, personal care products, and cosmetics.

We leverage data on the origin country for over 600,000 barcodes in these grocery product

⁷In robustness exercises in Section 4.5, we also make use of the NielsenIQ store-scanner data pooled over the year 2016. This dataset provides weekly barcode-level price and units sold by over 35,000 U.S. stores.

⁸See Bronnenberg et al. (2012) for more details regarding this survey.

⁹See Jaccard (2023) for a more detailed discussion of this dataset.

categories. Given the universality of barcodes, these data can be directly merged with the purchase records from NielsenIQ. Figure C.1 documents the distribution of production origin countries in the merged scanner data. As expected, Mexico and Canada constitute just over half of all import expenditure, with Thailand, China, and Italy rounding out the top five. Overall, our sample contains 74 origins with positive imports. The average household’s import expenditure share is approximately 9%.¹⁰

The sample of NielsenIQ households interviewed in 2008, and therefore with nativity information, attrits over time. On the other hand, the Label Insight data on products’ country of origin has been improving in recent years. To maximize data quality in both datasets, we restrict our analysis to the years 2014-2016, which we pool to a single cross-section at the household level. Our final sample consists of 19,822 households, of which 1,237 are immigrants.¹¹

Household-Level Coverage of Import Expenditure: Our final merged dataset exhibits \$764 billion of expenditure and is at the household-import origin level of aggregation. When compared to estimates from the BLS Consumer Expenditure Survey, the grocery categories studied in this paper account for approximately a third of all expenditure on tradables, with this share increasing to almost half if one excludes passenger vehicles and energy products. Jaravel (2019) estimates a within-grocery coverage of NielsenIQ expenditure data of approximately 80%, while the Label Insight data covers approximately 60% of expenditure within the NielsenIQ sample.

2.2 Immigration Data

We measure the contemporary immigrant population by origin with the pooled 2006–2010 sample of the American Community Survey (ACS). We obtain exogenous variation in these

¹⁰Throughout this paper we make use of the projection factor weights provided by NielsenIQ when presenting aggregated statistics. These weights are a population projection based on the representativeness of each household, and sum to 120 million households, which roughly matches the aggregate total for the U.S.

¹¹Immigrants and natives have a virtually identical rate of survival to our final dataset (39% versus 40%), and this difference is not significant at the 90% confidence level. For households with purchase records in 2014-2016, 23% can be linked to the 2008 nativity survey. This rate is virtually unchanged (24%) for the top half of households by income. We find that sample representativeness and differential attrition does not significantly drive our baseline results, as discussed in Appendix A.3.

contemporary immigrant populations via interactions of historic immigrant inflows, a strategy described in Section 4.3. Historic inflows are measured from the decennial Censuses from 1880-1930 and 1970-2000.¹² We provide additional details on data construction in Appendix A.1.

2.3 Stylized Facts

The combined dataset described above constitutes the first direct measurement of import expenditure by country of birth. We leverage this novel feature to demonstrate three stylized facts which characterize import consumption heterogeneity by nativity.

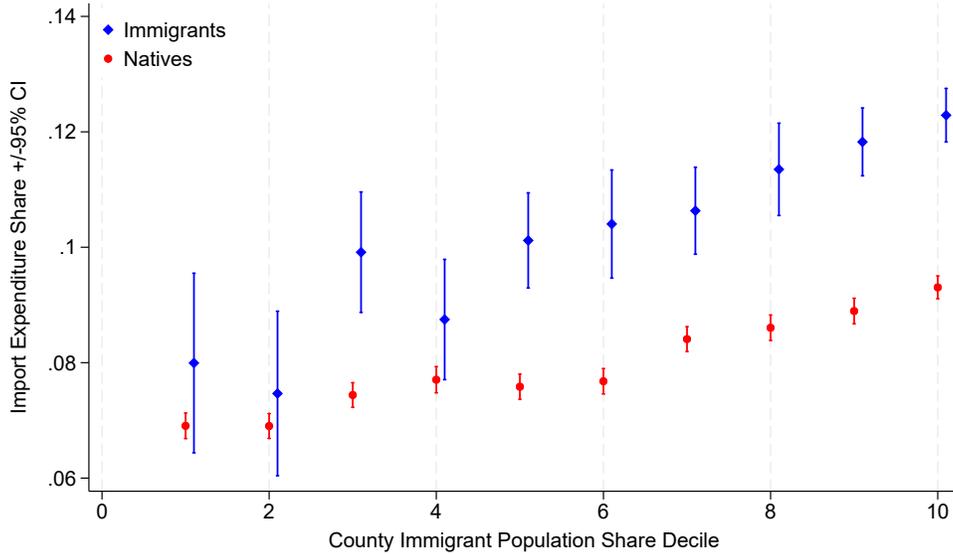
Fact 1: Household-level import expenditures are increasing in county immigrant population shares. Figure 1 plots the average import expenditure share across county deciles based on the immigrant population share. Both native and immigrant households exhibit a strikingly positive relationship between the presence of immigrants and the propensity to purchase imported goods. Relative to the lowest decile, native households living in the most immigrant-intensive decile of counties exhibit import expenditure shares which are 35% higher on average. For immigrant households, this differential increases to +50%. The figure represents the first direct evidence of a positive correlation between household-level import expenditure and local immigrant population shares.

Fact 2: Average import expenditure shares are 38% greater for immigrant households compared to non-immigrant households. Figure 1 also shows that immigrants exhibit stronger import demand than native households even within the same county decile. We also quantify this difference in mean import expenditure by regressing the household-level import expenditure on a dummy for whether a household is an immigrant household and find an unconditional mean difference in import expenditure between immigrants and natives of +3.1 percentage points. When compared to the average import expenditure share of non-immigrant households, this estimate represents a 38% differential.¹³

¹²The 1940, 1950 and 1960 censuses did not ask about the year of the respondent's immigration.

¹³See Figure C.2 for a histogram of import expenditure shares for native and immigrant households.

Figure 1: Immigrants, Natives, and Import Expenditure



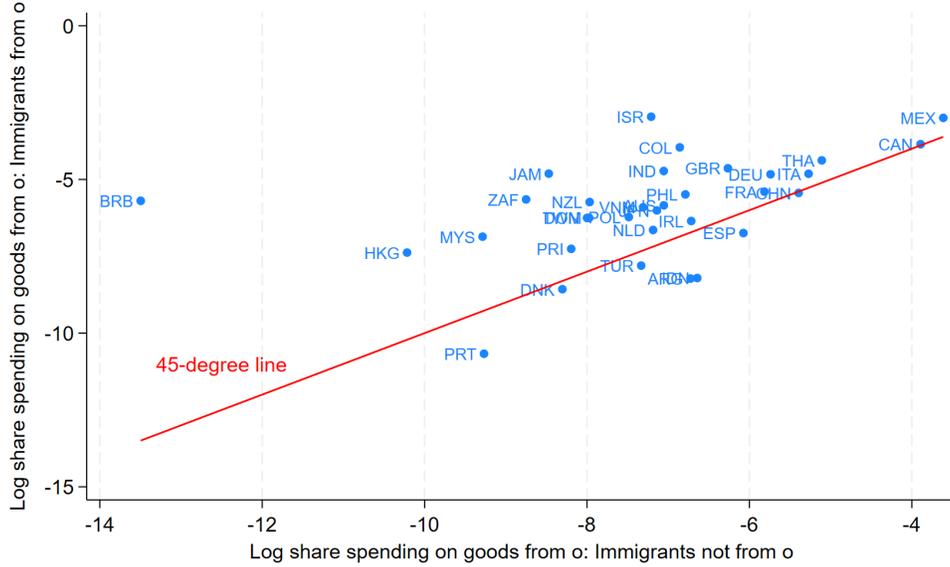
Notes: The figure plots average import expenditure shares by county immigrant population share deciles and nativity. Each point on the graph is the corresponding nativity-by-county decile fixed effect coefficient, with 95% confidence intervals plotted. These fixed effects are recovered from a linear regression at the household level in which household import expenditure shares are the dependent variable. Counties are placed into deciles based on the immigrant population share of that county, and households are grouped into two categories: immigrants (blue diamonds) or natives (red circles). 95% confidence intervals are provided, and all observations are weighted by the NielsenIQ projection factors.

Fact 3: Immigrants spend over twice as much as on goods from their origin country as immigrants not from that origin. For each origin country o , we calculate the share of expenditures on goods from o by both immigrants from o and immigrants not from o . Figure 2 depicts this relationship. The 45-degree line in red plots where immigrants not from o and immigrants from o would exhibit identical expenditure on imports from o .

We find that most origins lie above the 45-degree line, suggesting that immigrants exhibit stronger demand for imports from their specific country of origin. For the 33 countries in our sample with nonzero expenditures by both immigrants from that origin country and those not from that origin, the median relative expenditure share on goods from origin o by immigrants from o is 1.9 times greater than the expenditure on goods from o by immigrant households not from o .¹⁴ To our knowledge, this paper is the first to provide direct evidence

¹⁴Note that this estimate represents the weighted median relative expenditure across origins. The mean estimate is 81.6, but this is driven by outliers. When weighted by origin-specific aggregate expenditure shares, the mean difference is 2.8. Thus the median estimate of 1.9 represents a conservative figure.

Figure 2: Immigrants Tend to Spend more on Goods from their Origin



Notes: The figure shows the relationship between spending on goods imported from immigrants’ own country (the y-axis) and spending on goods from that country by other immigrants (x-axis). The red line is the 45-degree line, which indicates no preference by immigrants for goods imported from their origin country relative to immigrants from other origins. Data come from the NielsenIQ Household Panel 2014-2016 and Label Insight. NielsenIQ projection factor weights are used to construct expenditure shares.

that the preference persistence documented in [Logan and Rhode \(2010\)](#) and [Atkin \(2016\)](#) exists for U.S. immigrants with respect to goods imported from their origin country.

3 Immigrants and Import Demand

How are the consumer gains from trade distributed by nativity? Little is known about the distribution of import expenditure across households, and the few studies that do exist focus on non-homothetic demand and therefore variation across income levels ([Fajgelbaum and Khandelwal 2016](#); [Cravino and Levchenko 2017](#); [Borusyak and Jaravel 2021](#); [Auer et al. 2024](#)). In this section, we quantify the role of nativity and household country of origin in shaping variation in import expenditure across households.

3.1 Import Expenditure and the Consumer Gains from Trade

The three stylized facts illustrated in Section 2.3 suggest immigrants spend more on imports than natives within the grocery sector. To a first-order approximation, this variation in import expenditure leaves immigrants more exposed to trade shocks than natives, regardless of the underlying demand system. Consider household $h \in H$ characterized by income X_h and expenditure $X_h(\omega)$ per variety $\omega \in \Omega$, such that $X_h = \sum_{\omega \in \Omega} X_h(\omega)$. All households face the same schedule of prices \mathbf{p} . For some infinitesimally small change in prices $d \ln \mathbf{p}$ and total income $d \ln X_h$, the equivalent variation is defined as the change in X_h that generates an equivalent change in indirect utility, holding prices fixed. We refer to the equivalent variation of h as $d \ln W_h$. As discussed in Fajgelbaum and Khandelwal (2016), this equivalent variation can be expressed as:

$$d \ln W_h = \sum_{\omega} [-d \ln p(\omega)] s_h(\omega) + d \ln X_h \quad (1)$$

where $s_h(\omega) = X_h(\omega)/X_h$ denotes the household-specific expenditure share of variety ω . The first term on the right-hand side captures the expenditure channel, through which changing prices affect welfare, and is the exclusive focus of this paper. Households allocating a higher share of expenditure to imported varieties are more vulnerable to import-specific price shocks, such as tariffs. The second term of equation (1) captures the income effect of trade—the subject of a vast literature on the labor market impacts of trade openness—which we hold fixed throughout our analysis.

In order to quantify differential welfare impacts of trade shocks by nativity, we make two simplifying assumptions: (1) all households face the same trade elasticity, θ , which we calibrate to $\theta = 5$ (Head and Mayer 2014; Costinot and Rodríguez-Clare 2014), and (2) the import supply micro-foundation adheres to the assumptions described in Arkolakis et al. (2012).¹⁵ We define the grocery consumption gains from trade for household h as the proportional change in welfare due to an external shock to foreign production and trade costs, denoted by $\hat{W}_h = W'_h/W_h$. The gains from trade in groceries are therefore a function

¹⁵That is, the import supply model generates a structural gravity equation (Head and Mayer 2014).

of the observed household-level expenditure share on U.S. goods $\pi_{US,h}$:

$$\hat{U}_h = (\hat{W}_h)^{1-\mu_h} = (\pi_{US,h})^{-\frac{1}{\theta}(1-\mu_h)},$$

where $1 - \mu_h$ denotes the expenditure share of household h on groceries. In practice, we bin households by income quintiles $y(h)$ when constructing μ_h , such that $\mu_h = \mu_{y(h)}$. We compute $\mu_{y(h)}$ using the BLS Consumer Expenditure Survey (CEX), and find that grocery expenditure shares are monotonically decreasing in income.¹⁶ Having calibrated θ and $\mu_{y(h)}$, and directly observing $\pi_{US,h}$ in our expenditure data, we recover the cross-sectional distribution in consumer gains from grocery trade.

Table 1 provides estimates of both conditional and unconditional differences in the average gains from trade in groceries by nativity, both within groceries (\hat{W}_h) and in the aggregate when fixing non-grocery gains from trade to zero (\hat{U}_h). In Columns 1 to 3, the dependent variable is $(\hat{W}_h - 1) \times 100$ such that estimates represent the percentage point gaps across households. Column 1 provides estimates of the unconditional gap between immigrants and natives, while column 2 provides a within-county comparison, thus controlling for geographic sorting of households. When only one household is present in a given county, that household drops out of the sample when adding county fixed effects. Column 3 includes a vector of household controls and provides estimates associated with income quintiles.¹⁷

Within groceries, immigrants exhibit unconditional gains from trade that are, on average, 42% greater than the average non-immigrant (2.45% versus 1.72%, column 1). When controlling for county of residence, and therefore local import supply, immigrants still exhibit gains that are 36% greater than their within-county non-immigrant neighbors. Column 3 highlights the extent to which nativity drives import expenditure heterogeneity substantially more than income. Households in the highest income quintile exhibit within-grocery gains from trade that are 0.15 percentage points greater than their within-county neighbors

¹⁶We sum CEX annual expenditure over food at home, housekeeping supplies, personal care products, and alcohol. We obtain the following grocery expenditure shares by income quintile, in order of lowest to highest income: 0.134, 0.119, 0.110, 0.104, and 0.087.

¹⁷The household controls are dummy variables indicating h 's race, ethnicity, household size bins, bins for the number and age of children in the household, highest education among household members, marital status and age. When constructing household income, we use the grocery expenditure per capita approach to placing households in income quintiles as in [Faber and Fally \(2022\)](#).

Table 1: Consumer Gains from Trade by Nativity

	Dep. var.: $\% \Delta W_h$			Dep. var.: $\% \Delta U_h$		
	(1)	(2)	(3)	(4)	(5)	(6)
Immigrant from anywhere	0.725*** (0.033)	0.635*** (0.034)	0.562*** (0.036)	0.076*** (0.004)	0.065*** (0.004)	0.062*** (0.004)
2nd poorest quintile			0.030 (0.028)			-0.024*** (0.003)
Median quintile			0.082*** (0.030)			-0.034*** (0.003)
2nd richest quintile			0.133*** (0.032)			-0.039*** (0.004)
Richest quintile			0.146*** (0.034)			-0.069*** (0.004)
Constant	1.722*** (0.009)	1.739*** (0.009)	2.556*** (0.543)	0.190*** (0.001)	0.192*** (0.001)	0.337*** (0.061)
N	19,822	19,223	19,223	19,822	19,223	19,223
County FE		✓	✓		✓	✓
HH Controls			✓			✓
Sample Weights	✓	✓	✓	✓	✓	✓
Estimator	OLS	OLS	OLS	OLS	OLS	OLS

Notes: The table presents regression results at the household-country level. Observations weighted using NielsenIQ household weights. The reference income quintile is the poorest. Standard errors clustered two-ways at the household and origin-by-destination levels. Household controls are dummy variables indicating h 's race, ethnicity, household size bins, bins for the number and age of children in the household, highest education among household members, marital status and age. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

in the bottom income quintile. Yet conditional on income, immigrant households exhibit gains from trade that are 0.56 percentage points greater than an equivalent native household. Reassuringly, the positive relationship between income and within-grocery import expenditure documented here matches existing research. This paper, however, is the first to observe import expenditure by nativity conditional on income, and the first to highlight the considerable heterogeneity in consumption exposure to trade shocks by household nativity.¹⁸

¹⁸[Borusyak and Jaravel \(2021\)](#) document a grocery import expenditure share of 11.5% for U.S. households earning \$150,000 a year versus 10.3% for U.S. households earning \$10,000 a year (Figure 2). Within the welfare calculation considered here, their findings imply high-income households exhibit gains from trade that are 13% greater than low-income households. For Swiss households, estimates in Table 1 of [Auer et al. \(2024\)](#) suggest a relative gain from trade to high-income households of 14% versus low-income households. Column 3 of Table 1 implies a relative gain to high-income households of 9% versus low-income households.

Columns 4 to 6 provide the same analysis and a similar set of results, but scaled by household expenditure on groceries. Given that groceries exhibit a negative income elasticity, high-income households benefit less from the grocery-specific gains from trade than low-income households. Still, immigrant effects persist: the unconditional average gains from trade for immigrants are 40% greater than for native households.

While these estimates focus on the consumer gains from trade, the same import expenditure differential documented here also leaves immigrants more exposed to increases in trade costs, such as tariffs. As a first-order approximation, the fact that the unconditional immigrant consumer gains from trade are 42% greater than those of natives also suggests that immigrants face welfare losses from tariffs of an equivalent magnitude relative to natives. While researchers tend to focus on whether the cost of tariffs are borne by poor or wealthy consumers, or the extent to which intra-national geographies may differ in their exposure to tariffs, the analysis provided here suggests that even within narrow income bins and within U.S. counties, immigrants bear the cost of tariff increases to a greater degree than any observed variation across income groups.

3.2 Preferences and Import Demand

Next, we formalize the import expenditure heterogeneity documented above by introducing and estimating a model of import demand. In order to maintain consistency with the welfare formula developed by [Arkolakis et al. \(2012\)](#), and to simplify the incorporation of a general equilibrium import supply model in [Section 4](#), we assume households exhibit a two-tiered demand system. In the upper tier, households have Cobb-Douglas preferences over differentiated grocery goods, our focus in this paper, and a non-grocery good, q_0 .¹⁹

Households aggregate grocery variety consumption with constant elasticity of substitution (CES) preferences. There is a continuum of differentiated varieties $\Omega_{o,c(h)}$ associated with each origin country $o \in \mathcal{O}$ and h 's county of residence $c(h)$. We allow for household heterogeneity in income Y_h , origin-specific preferences denoted by z_{oh} , and substitution

¹⁹As discussed previously, we calibrate the Cobb-Douglas grocery expenditure share parameter $1 - \mu_{y(h)}$ using income quintile data from the CEX.

elasticity $\sigma_h > 1$. Preferences are therefore represented by the following utility function:

$$U_h = q_0^{\mu_{y(h)}} \left[\sum_{o \in \mathcal{O}} z_{oh}^{1/\sigma_h} \int_{\omega \in \Omega_{o,c(h)}} q_{oh}(\omega)^{\frac{\sigma_h-1}{\sigma_h}} d\omega \right]^{\frac{\sigma_h}{\sigma_h-1}(1-\mu_{y(h)})}. \quad (2)$$

Equilibrium expenditure on the differentiated grocery sector is $X_h = (1 - \mu_{y(h)})Y_h$. Demand at the origin-household level is then:

$$X_{oh} = z_{oh} p_{o,c(h)}^{1-\sigma_h} X_h P_h^{\sigma_h-1},$$

where $p_{o,c(h)}$ denotes some origin-county-level price index and P_h the household-level price index. That is, we assume all households within the same county face the same schedule of prices by import origin, such that within-county variation in origin-specific import expenditure is attributable to variation in the demand shifter z_{oh} .²⁰ In the steps to follow, it will be useful to consider variation in the relative expenditure of households compared to expenditure on domestically produced U.S. goods, $X_{US,h}$. For any variable x , we therefore define $\tilde{x}_{oh} \equiv x_{oh}/x_{US,h}$.²¹ Furthermore, we normalize the perceived quality of imports such that $z_{US,h} = 1$ for all households, which yields the following expression for relative expenditure at the household-origin level:

$$\tilde{X}_{oh} = z_{oh} \left(\frac{p_{o,c(h)}}{p_{US,c(h)}} \right)^{1-\sigma_h}. \quad (3)$$

We next proceed to estimate the two exogenous preference terms in (3): h 's evaluation of the quality of goods from o , z_{oh} , and the elasticity of substitution across varieties σ_h .

3.3 Estimating Heterogeneity in Perceived Import Quality

How do households differ in their perceived quality of imported goods? In this section, we make use of our remarkably detailed household purchase data, including household nativity, to empirically recover the relationship between exogenous household characteristics and per-

²⁰We check whether immigrants and natives pay different prices for the same barcode within counties. We find no evidence for differential prices faced by these sub-populations. Results available upon request.

²¹Head and Mayer (2014) refer to this normalization as a ‘‘ratio method.’’

ceived import quality, z_{oh} . We therefore assume a functional form for the preference vector z_{oh} that relates observed socioeconomic household characteristics to import demand:

$$z_{oh} = \exp\left(\beta_z I_{o,c(h)} + \delta J_h + \zeta_1 \mathbf{1}[o(h) \neq US] + \zeta_2 \mathbf{1}[o(h) = o]\right) \times \eta_{oh}^z. \quad (4)$$

J_h is a vector of observed household characteristics and $o(h)$ maps each household h to its country of birth. Motivated by our stylized facts, ζ_1 captures the strength of immigrants' taste for goods from all foreign countries, and ζ_2 captures the strength of immigrants' home-biased preferences à la [Atkin \(2016\)](#) and [Logan and Rhode \(2010\)](#). I_{oc} denotes the population share in county c of immigrants from origin o , and we allow for the possibility that the local immigrant population share from o endogenously affects preferences for imports from o via the parameter β_z .²² The goal of this section is to estimate differences in exogenous preferences between immigrants and natives, such that this endogenous preference term will be absorbed by fixed effects at the origin-county level. We return to estimating β_z in Section 4. η_{oh}^z denotes an idiosyncratic preference term at the household-origin level.

In our baseline specification we assume that the price sensitivity of households σ_h varies at the origin-county level, i.e., $\sigma_h = \sigma_{c(h)}$. We then gather together price terms and the elasticity of substitution in a origin-county fixed effect, $\psi_{o,c(h)} = \left(\frac{p_{o,c(h)}}{p_{US,c(h)}}\right)^{1-\sigma_{c(h)}}$ from equation (3). We thus obtain the following estimable equation:

$$\tilde{X}_{oh} = \exp\left(\psi_{o,c(h)} + \delta J_h + \zeta_1 \mathbf{1}[o(h) \neq US] + \zeta_2 \mathbf{1}[o(h) = o]\right) \times \eta_{oh}^z, \quad (5)$$

where we have gathered all terms related to prices, entry, price sensitivity, and county-specific average preferences into the origin-county fixed effects $\psi_{o,c(h)}$. We also consider a less restrictive specification in which prices and the price sensitivity of households vary at the county-by-income-quintile level. In this less restrictive specification, fixed effects are at the origin-county-income quintile-level: $\psi_{o,c(h),y(h)}$.

Estimates of ζ_1 and ζ_2 quantify how much immigrants differ from native households within the same county, or county-income-quintile, in their propensity to purchase imported

²²In this way, we do not treat preferences as a primitive, but instead allow one's preferences to be at least partially determined by one's cultural and social context ([Bowles 1998](#); [Atkin et al. 2021](#)).

goods from all origins (ζ_1) and imports specifically from their own origin country (ζ_2). Our estimation procedure projects observed import expenditure onto household characteristics, with county-level and county-origin-level factors—such as prices, variety availability and local average preferences—controlled for via fixed effects. Our identifying assumption is that there are no unobserved factors shaping variation in prices and product availability within county-origin pairs which are correlated with observed household characteristics. A violation of this assumption would be if immigrant households cluster spatially within counties in ethnic enclaves which feature ethnic grocery stores, and therefore face a different product and price mix. In our results below, we show that estimates of ζ_1 and ζ_2 are virtually unchanged when including ZIP Code-origin fixed effects, supporting our identifying assumption.

To estimate equation (5) we make use of the pseudo-Poisson maximum likelihood (PPML) estimator to deal with the large number of zeros observed in our outcome variable (Santos Silva and Tenreyro 2006). We show in Appendix A.2 that our headline estimates are robust to alternative estimators, although it should be noted that these alternative estimators are not necessarily consistent with the theory outlined here. Table 2 provides estimates of ζ_1 and ζ_2 for three separate specifications of equation (5). Columns 1 and 2 include origin-county fixed effects without household controls J_h , columns 3 and 4 add in J_h , and columns 5 and 6 control for origin-county-income-quintile fixed effects.²³

Across specifications, our estimates of ζ_1 and ζ_2 are remarkably stable. When estimated alongside ζ_2 , ζ_1 fluctuates between 0.22 and 0.26 and is highly significant. ζ_2 fluctuates between 0.58 and 0.61, and is also highly significant. To place these estimates in context, consider column 4. For two households in the same county and identical in their income, education, age, and family structure, the immigrant household allocates a 25% higher share of expenditure on all import origins than the native household. For imports specifically from this immigrant household’s origin country, their expenditure is 127% greater.²⁴

Appendix Table C.1 reports estimates of ζ_1 and ζ_2 when implementing origin-by-ZIP Code fixed effects, such that the identifying variation is generated by import expenditure heterogeneity across households within the same ZIP Code. These estimates are effectively

²³Note that when comparing households within income quintiles, we lose approximately a third of our sample size due to separation (Correia et al. 2019), hence the reduction in N in columns 5 and 6.

²⁴That is, $e^{\hat{\zeta}_1} - 1 = 0.25$ and $e^{\hat{\zeta}_1 + \hat{\zeta}_2} - 1 = 1.27$.

Table 2: Estimates of Perceived Import Quality

	Dep. var.: Rel. expenditure share on goods from o (\tilde{X}_{oh})					
	(1)	(2)	(3)	(4)	(5)	(6)
Immigrant from anywhere (ζ_1)	0.31*** (0.027)	0.26*** (0.027)	0.27*** (0.028)	0.22*** (0.028)	0.28*** (0.030)	0.23*** (0.030)
Immigrant from o (ζ_2)		0.58*** (0.064)		0.60*** (0.065)		0.61*** (0.077)
N	911,385	911,385	911,385	911,385	619,741	619,741
Origin-County FE	✓	✓	✓	✓		
HH Controls			✓	✓	✓	✓
Origin-County-Income FE					✓	✓
Sample Weights	✓	✓	✓	✓	✓	✓

Notes: The table presents results from estimating equation (5) using PPML at the household-country level. Observations are weighted using NielsenIQ household weights. Standard errors are clustered two-ways at the household and origin-by-destination levels. Household controls are dummy variables indicating: race, ethnicity, household size bins, bins for the number and age of children in the household, highest education among household members, household income quintiles, marital status and age groups. The estimation sample size is less than our total sample, and falls across columns, due to fixed effects causing separation in the sense of [Correia et al. \(2019\)](#). *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

unchanged from our baseline, other than a marginal reduction in $\hat{\zeta}_2$ associated with our most stringent fixed effects at the origin-by-5 digit ZIP Code level ($\hat{\zeta}_2 = 0.41$). We maintain counties as our baseline geographic unit to remain consistent with the level of variation used to estimate the local import supply effect of immigrants in Section 4.

Appendix Table C.2 provides import preference estimates associated with the entire vector of household characteristics included in J_h . While the positive, but relatively small, relationship between income and import expenditure reflects the estimates in Table 1, we also find a strong relationship between education and import expenditure. Households with at least a postgraduate degree exhibit import expenditure that is 27% greater than an otherwise identical household. An immigrant household with a postgraduate degree therefore exhibits import expenditure on all origin countries that is 58% greater than a native household with no college education living in the same county.

We unpack immigrants' general preference for imports, ζ_1 , by examining the interaction with other variables—such as geographic proximity between immigrant origin and the U.S.—in Appendix A.4, with the results shown in Appendix Table A.5. We find strong evidence

for Hispanic and Asian households generally preferring goods from Latin America and Asia, respectively, over imports from elsewhere, while there is little evidence for other bilateral geographic, linguistic, or historic political proximity measures affecting immigrants’ import demand. In Appendix Table C.3, we show that the strength of immigrants’ home bias, ζ_2 , is largely uniform across education (column 1) and income (column 2), with the exception of a marginally significant drop of the home bias for immigrants in the top income quintile. In column 3, we interact the home-bias indicator with an index of cultural distance between the origin country and the U.S. (Spolaore and Wacziarg 2016).²⁵ Home bias becomes stronger as cultural distance from the U.S. increases. Finally, given that immigrants might culturally assimilate during their stay, we allow the home bias to vary by time in the U.S. in column 4 and obtain a small, marginally significant negative effect of an additional year since immigration.²⁶

As a final robustness exercise, we consider the possibility that immigrants and natives simply consume different products, given the profusion of specialty goods and specialty grocery stores (Neiman and Vavra 2023). If purchases are highly segregated by nativity, our baseline preference estimation may not be compatible with CES preferences, which generate non-zero expenditure on all available products. We find only modest evidence of such purchase segregation in our data. Among the barcodes in our sample, just 1.5% are purchased by immigrants alone; 48.1% by natives alone; and the remaining 50.4% by both. These barcodes are a very small part of each household’s consumption basket when weighted by expenditure: just 0.5% of immigrants’ grocery expenditures are on the set of barcodes they exclusively purchase. We further re-estimate equation (5), restricting the dependent variable to be computed only over the set of barcodes that we observe in both immigrant and native purchase records. The results, reported in Appendix Table C.4, are little different from the baseline: we recover $\hat{\zeta}_1 = 0.24$ and $\hat{\zeta}_2 = 0.54$.

We therefore find compelling evidence that households exhibit heterogeneous grocery import quality evaluations, with nativity playing a significant role in shaping this heterogeneity.

²⁵We use the “total” index based on the World Value Survey; see section 6.3.3 in Spolaore and Wacziarg (2016), and normalize it to have a mean of zero and a standard deviation of one.

²⁶We proxy arrival year in the U.S. with the year in which the immigrant moved to their state of residence as of 2008.

As highlighted in equation (3), however, variation in σ_h could also plausibly generate import expenditure variation if imports systematically differ in price when compared to domestically produced goods. While we are able to flexibly control for variation in $\sigma_{c(h)}$ when estimating the components of z_{oh} , we are unable to recover σ_h directly. In the following section, we empirically test for heterogeneity in σ_h by nativity and are unable to rule out that immigrants and natives exhibit the same value of σ_h .

3.4 Estimating Elasticities of Demand

This section provides estimates of σ both for all households as a pooled sample and for sub-samples based on nativity. We ultimately do not find strong evidence that immigrants' elasticity of substitution differs significantly from natives.

The data requirements associated with estimating elasticities of demand lead us to a few data modifications. First, we restrict our sample to the set of barcodes purchased by at least one household in every time period. This sample restriction alleviates bias associated with zeros, as a positive shock to prices could simply lead to product exit, thereby biasing estimates of σ towards zero. Second, we aggregate geographies to the level of a Census Bureau Statistical Area (CBSA), in order to account for the sparse number of immigrant households in our sample.²⁷ Finally, we aggregate purchase records to the barcode-CBSA-bimonthly period, such that we have 18 time periods t over three years (2014-2016).

For a given barcode ω , total sales within CBSA c are $X_c(\omega) = \sum_{h \in \Lambda_c} X_h(\omega)$, where Λ_c denotes the set of households residing in c . We derive our estimating equation by summing household-barcode-specific expenditure given in equation (3) across households. We then take the logarithm of this sum and take first differences over time within barcode-CBSA pairs via the Δ_t operator, which generates our estimating equation:

$$\Delta_t \ln X_{ct}(\omega) = (1 - \sigma)\Delta_t \ln p_{ct}(\omega) + \alpha_{ct} + \epsilon_{ct}(\omega) \quad (6)$$

²⁷While we use counties as the level of analysis throughout the remainder of this paper, generating significant expenditure variation is crucial to estimating demand elasticities. With only 1,237 immigrant households in our entire sample, we aggregate to larger geography definitions in order to generate such variation.

We make use of standard “Hausman-Nevo” instruments, which are the average price of barcode ω in all other markets, $\Delta_t \ln p_{ct}^{IV}(\omega) = \Delta_t \ln[\sum_{c' \neq c} X_{c't}(\omega)/Q_{c't}(\omega)]$, as an instrument for $\Delta_t \ln p_{ct}(\omega)$ (Hausman 1996; Nevo 2001). We maintain the same prices for all subpopulations, which we calculate using all purchase records in the NielsenIQ HomeScanner data, and by taking a sales-weighted average of expenditure divided by units sold.

Appendix Table A.6 reports the resulting estimates, and we recover $\hat{\sigma} = 3.27$ in our preferred specification. Including origin-CBSA fixed effects, such that our estimating equation is consistent with an “Armington” demand structure of barcodes nested within origin countries, yields a statistically indistinguishable estimate of $\hat{\sigma} = 3.30$ (column 4).

Comparing estimates across household types is made difficult by the disparity in sample size associated with our two main household types: native and immigrants. As shown in columns (5), (7), and (8) of Table A.6, $\hat{\sigma}$ decreases in magnitude mechanically with the number of observations.²⁸ To gauge the importance of sample size for the smaller immigrant coefficient, we take the number of immigrant households and randomly sample an equivalent number of native households with replacement twenty times, estimating a separate σ for each native subgroup. As shown in Figure A.1, the distribution of native estimates overlaps considerably with the confidence interval of our immigrant estimate. We thus conclude that there is insufficient evidence to assume σ varies by nativity, and proceed by calibrating $\sigma = 3.30$ for both immigrants and natives.²⁹

Throughout this section, we have absorbed supply-side factors in fixed effects and compared import demand across local heterogeneous households. We turn now to developing and estimating a model of local import supply with the potential for endogenous immigrant effects on trade costs and market size.

²⁸Given the relatively small number of immigrant households in our sample (1,237), the set of barcodes endogenously selected into the estimating dataset for immigrants exhibits, on average, aggregate market shares that are 14 times larger than barcodes in the native sample. The number of barcodes is also much smaller: 439 for immigrants and 7,240 for natives. If high market share barcodes exhibit more inelastic demand, estimates of σ for immigrants will be biased towards zero due to this smaller sample size.

²⁹We consider an alternative specification in which heterogeneity occurs via the “Armington” elasticity of substitution across product origin countries (Armington 1969). Using exchange rates as instrumental variables for country-CBSA-period prices, we cannot rule out homogeneity in the Armington elasticity by nativity: 3.78 (SE=0.39) for natives and 3.91 (SE=0.85) for immigrants. See Appendix A.5 for details. We run our immigration counterfactuals in Section 5 assuming $\sigma = 3.9$ and find broadly similar results (results available upon request).

4 Immigrants and Import Supply

In this section we examine how immigrants affect local import supply. Specifically, we microfound a model of import supply in which immigrants may induce greater local import supply by reducing local trade costs and increasing market size. Combining the supply-side with the demand system introduced in Section 3, we estimate the key parameters of the model. In Section 5, we use the full estimated model to quantify the effect of immigrants on native import expenditure and the differential effects of a tariff increase.

4.1 A Heterogeneous Firms Model of Import Supply

This section develops a heterogeneous firms model in which immigrants may affect (i) import prices via variable trade costs, (ii) variety availability via fixed trade costs and market size, and (iii) native preferences for imported goods. We opt for the Melitz (2003) model because it features fixed costs and therefore market size effects, a key channel through which immigrants may affect local variety availability above and beyond any direct effect of immigrants on trade costs (Iranzo and Peri 2009; Di Giovanni et al. 2015; Aubry et al. 2016).

Preferences: Household preferences are as in equation (2): utility is determined by a CES aggregator across grocery varieties. Recall that country-specific import preferences z_{oh} consist of an exogenous component associated with parameters δ , ζ_1 , and ζ_2 , as well as an endogenous component $\beta_z I_{oc}$, which characterizes how preferences may diffuse from immigrants to their neighbors.

Firms: Each country $o \in \mathcal{O}$ has exogenous size Y_o and marginal cost of production w_o . Trade is characterized by origin-by-county iceberg trade costs τ_{oc} and fixed costs of exporting f_{oc} . Each firm draws some productivity φ from a Pareto distribution with shape parameter $\theta > \sigma - 1$. The set of potential entrant firms in each origin is proportional to the size of that origin Y_o . The cost of providing q units to destination county c by a firm in origin o with productivity φ is $C_{oc}(q) = \frac{w_o \tau_{oc}}{\varphi} q + f_{oc}$. The county is the most granular geographic unit at which firms make pricing and entry decisions, with each county treated as an independent market.

We allow both types of trade costs to vary according to a vector of distance measures d_{oc} , the local immigrant population share I_{oc} , and an unobserved component:

$$\tau_{oc} = \exp\left[-\frac{1}{\theta}(\rho^\tau d_{oc} + \beta^\tau I_{oc})\right] \times \eta_{oc}^\tau, \quad (7)$$

$$f_{oc} = \exp\left[-\left(\frac{\sigma-1}{1+\theta-\sigma}\right)(\rho^f d_{oc} + \beta^f I_{oc})\right] \times \eta_{oc}^f, \quad (8)$$

where η_{oc}^τ and η_{oc}^f represent idiosyncratic deviations in trade costs across county-origin pairs with a mean of one. β^τ captures the strength of the the variable cost channel of immigrants and β^f the fixed cost channel of immigrants. We normalize domestic trade costs $\tau_{US,c} = f_{US,c} = 1$.³⁰ Firms price according to monopolistic competition and thus set constant mark-ups. The optimal price for any variety ω from origin o in county c is

$$p_{\omega(o),c} = \frac{\sigma}{\sigma-1} \frac{w_o \tau_{oc}}{\varphi(\omega)}. \quad (9)$$

Gravity: Household-level expenditure on goods from origin o can be expressed as:³¹

$$X_{oh} = e^{\alpha_o} X_h P_h^{\sigma-1} \tau_{o,c(h)}^{-\theta} \left(\frac{f_{o,c(h)}}{S_{c(h)} z_{o,c(h)}} \right)^{-\left(\frac{\theta}{\sigma-1}-1\right)} z_{oh}, \quad (10)$$

where $S_{c(h)}$ is real grocery sector expenditure in county c .³² Average county-level preferences $z_{o,c(h)}$ are an expenditure-weighted average of the preference shifter z_{oh} across all households in county c , X_h is total grocery expenditure by h , and P_h the price index faced by h . The real size of origin o relative to the U.S. is defined as $\exp(\alpha_o) \equiv \lambda Y_o w_o^{-\theta}$, where λ is a collection of parameters defined in Appendix B.1.

Increasing returns to scale due to fixed costs feature in equation (10) via the term raised to the exponent $\theta/(\sigma-1)-1$, and can be divided into two components. $S_{c(h)}$ measures total real expenditure by all households on all goods, thereby capturing the typical notion of market size. $z_{o,c(h)}$ scales total expenditure by origin and captures the extent to which average preferences in a given county are potentially biased towards goods from a specific

³⁰The normalization terms $\frac{1}{\theta}$ and $\frac{\sigma-1}{1+\theta-\sigma}$ simplify subsequent expressions but are not necessary.

³¹We relegate the full derivation of model equations to Appendix B.1.

³²Formally: $S_{c(h)} = \sum_{h' \in \Lambda_{c(h)}} X_{h'} P_{h'}^{\sigma-1}$, where $\Lambda_{c(h)}$ is the set of households residing in $c(h)$.

origin due to the observable characteristics of households residing in county c .³³

We consider a gravity equation of expenditure on imports relative to expenditure on U.S. produced goods for a given household. Household-by-origin expenditure (relative to expenditures on domestic produced goods) is

$$\tilde{X}_{oh} = e^{\alpha_o} (\tau_{o,c(h)})^{-\theta} \left(\frac{f_{o,c(h)}}{z_{o,c(h)}} \right)^{-\left(\frac{\theta}{\sigma-1}-1\right)} z_{oh}. \quad (11)$$

This relative formulation differences out factors specific to the household and county which may increase import expenditures, such as household income X_h and the strength of the local labor market (as captured by total expenditure $S_{c(h)}$).³⁴ Plugging in our functional form assumptions—equations (4), (7), and (8)—we obtain a household-level gravity equation

$$\tilde{X}_{oh} = \exp \left(\alpha_o + \rho d_{o,c(h)} + \beta I_{o,c(h)} + \ln \bar{z}_{o,c(h)}^{\frac{\theta}{\sigma-1}-1} + \ln \bar{z}_{oh} \right) \times \eta_{o,c(h)} \times \eta_{oh}^z \quad (12)$$

with the following definitions:

$$\begin{aligned} \rho &= \rho^\tau + \rho^f, \\ \beta &= \beta^\tau + \beta^f + \left(\frac{\theta}{\sigma-1} \right) \beta^z, \\ \eta_{o,c(h)} &= \eta_{o,c(h)}^\tau \times \eta_{o,c(h)}^f, \\ z_{oc} &= e^{\beta^z I_{oc}} \bar{z}_{oc} \quad \text{and} \quad z_{oh} = e^{\beta^z I_{oc}} \bar{z}_{oh}. \end{aligned}$$

A key difference between equation (12) and equation (5) from Section 3 is the identifying variation. In equation (12), we leverage variation *across* county-origin pairs, which is the level at which the immigrant population share varies. In equation (5) we leverage variation *within* county-origin pairs across households, which is the level at which household preferences vary.

³³The preference shifter $z_{o,c(h)}$ represents the heterogeneous preference extension beyond the standard Melitz (2003) framework. As preferences shift towards goods from origin o , more firms are able to cover the fixed costs of supplying county c , which further increases the market penetration of imports from o to county c .

³⁴As we noted with the comparable relative estimation from Section 3, this is an example of the ‘ratio-type’ estimation strategy, as termed by Head and Mayer (2014).

4.2 Estimation strategy for identifying the import supply channels

We provide a three-step estimation strategy to recover the key parameters of the model. We start by estimating the total spillover effect of immigrants on import purchases, β . We then estimate the effect of immigrants on prices to recover β^τ . Finally, we estimate the extensive margin effect of immigrants to separately recover β^f and β^z .

Step 1: Estimating β . The spillover effect of immigrants on natives, net of market size effects, is captured by β in equation (12). To make this equation estimable, we must impute the unobserved county-average preference term $\bar{z}_{o,c(h)}$. We do so using the individual preferences \bar{z}_{oh} estimated in Section 3. Because our sample of NielsenIQ HomeScanner households may not be representative at the county-level, we use the ACS when computing $\bar{z}_{o,c(h)}$. We denote imputed preference terms with an apostrophe, e.g., \bar{z}'_{oh} . As discussed in Section 3, we calibrate $\theta = 5$ and we use our estimate of $\sigma = 3.3$.

With the preference terms in hand, we then difference out the imputed preference term $\mathcal{Z}_{oh} \equiv \bar{z}'_{oh}(\bar{z}'_{o,c(h)})^{\frac{\theta}{\sigma-1}-1}$ from the relative expenditure share \tilde{X}_{oh} . The deflated expenditure term, $\tilde{X}_{oh}/\mathcal{Z}_{oh}$, captures variation in import expenditure across households that is solely attributable to variation in trade costs. That is, $\tilde{X}_{oh}/\mathcal{Z}_{oh}$ represents import expenditure by household h conditional on the implied import demand associated with the observed characteristics of h and the observed characteristics of all other households within the same county. The residual variation in \tilde{X}_{oh} net of \mathcal{Z}_{oh} therefore captures the effect of immigrants on local import expenditure that is common to all households and unexplained by observed household characteristics.

Since the remaining right-hand-side variables in (12) only vary at the county-origin level, we aggregate to the county-origin level. To do so, we take the county-origin averages of the household-origin-level ratios, such that our dependent variable is $\tilde{X}_{oc} \equiv \sum_{h \in \Lambda_c} \kappa_h \frac{\tilde{X}_{oh}}{\mathcal{Z}_{oh}}$, with κ_h representing household sampling weights.³⁵ We thus obtain the following estimating

³⁵Specifically, κ_h are the NielsenIQ-provided household weights divided by the sum of weights across all households in c . Note that, since the independent variables only vary at the country-origin level, point estimates and standard errors are identical to those obtained from running the regression at the household-level when appropriately clustering at the county level. We show the equivalence also theoretically in Appendix B.2.

equation to identify β :

$$\tilde{X}_{oc} = \exp(\alpha_o + \rho d_{o,c} + \beta I_{o,c}) \times \eta_{oc}, \quad (13)$$

Estimating β alone, however, does not facilitate computing the welfare effect of immigrants on natives because immigrants may change the preferences of natives. We next leverage model restrictions and data characteristics in order to decompose β into its separate components: β^f , β^τ and β^z .

Step 2: Estimating β^τ . Immigrants may benefit natives by reducing the price of imported goods. We test this hypothesis using the highly detailed price data available in our household purchase records. Because the county-origin price only depends on variable trade costs net of county and variety fixed effects, we can leverage equation (9) to isolate β^τ . To do so we aggregate our data to the barcode-by-county level, incorporate the functional form of τ_{oc} from equation (7), and estimate the resulting price equation:

$$\ln p_{\omega(o),c} = \psi_c + \psi_{\omega(o)} - \frac{\beta^\tau}{\theta} I_{oc} - \frac{\rho^\tau}{\theta} d_{oc} - \frac{1}{\theta} \ln \eta_{oc}^\tau, \quad (14)$$

where ψ_c and $\psi_{\omega(o)}$ represent county- and barcode-level fixed effects.³⁶

Using data at the barcode level, equation (14) allows us to identify the effect of immigrants on the price of imported varieties rather than the effect of immigrants on the average price of imported goods. This is an important benefit of our dataset, as our model predicts that an average price measure would conflate the effect of immigrants on within-variety prices and the effect of immigrants on variety entry and exit.³⁷

³⁶Since barcodes are unique to origin countries, $\psi_{\omega(o)}$ absorbs variation in production costs w_o across origins.

³⁷When constructing $p_{\omega(o),c}$ we pool all purchases of barcode ω in county c and calculate a sales-weighted average price. One concern may be that retail firms price nationally (DellaVigna and Gentzkow 2019), leaving us with little variation in price across counties. However, we are leveraging household purchase data, not store-level data. This means that the average price across retail establishments within county varies for consumers. We graphically depict the average price variation across counties relative to the average variation in immigrant population shares in Appendix Figure C.4. Consistent with the above logic, we find substantial variation in average prices (net of barcode fixed effects) across counties. We further address the issue of national pricing using StoreScanner data in Section 4.5.

Step 3: Estimating β^f and β^z . We follow [Chaney \(2008\)](#) and derive expressions for both the extensive margin elasticity of imports with respect to immigrant population share and the total expenditure elasticity of imports with respect to immigrant population share. When $\beta^\tau \approx 0$, which we show empirically to hold in [Section 4.4](#), this derivation yields two equations and two unknowns: β^f and β^z . We thus estimate the extensive margin effect of immigrants on trade directly by replacing \tilde{X}_{oh} in [equation \(13\)](#) with \tilde{N}_{oh} : the count of barcodes from origin o in household h 's consumption basket divided by the count of barcodes from the U.S. in household h 's consumption basket. While the full derivation is provided in [Appendix B.3](#), our functional form assumptions, [equations \(4\)](#) and [\(8\)](#), yield the following two expressions regarding the import expenditure elasticity and extensive margin import elasticity, respectively:

$$\frac{\partial \ln \tilde{X}_{oh}}{\partial I_{oc}} = \beta^f + \left(\frac{\theta}{\sigma - 1} \right) \beta^z,$$

$$\frac{\partial \ln \tilde{N}_{oh}}{\partial I_{oc}} = \beta^f + \left(\frac{\theta}{\sigma - 1} - 1 \right) \beta^z.$$

Intuition follows directly from the results derived in [Chaney \(2008\)](#): firms enter a new market if and only if they can cover their fixed costs. Immigrants thus facilitate firm entry by (i) reducing fixed costs, or (ii) increasing the intensive margin of sales-per-variety in their market. Yet expenditure per barcode only increases if preferences become more favorable to the firm in question, as a reduction in fixed costs has no impact on within-firm sales, conditional on exporting. This intuition is reflected in the equations above: as the total trade elasticity and the extensive margin trade elasticity converge, $\beta^z \rightarrow 0$ and $\beta^f \rightarrow \beta$.

4.3 Identification and Instrumental Variables

Estimates of β , β^f , β^z , and β^τ recovered by applying OLS to, respectively, [equations \(13\)](#) and [\(14\)](#), are likely biased due to endogenous sorting of immigrants into locations with high import availability. For example, low transportation costs between New York and Italy may independently expand both the local import expenditure on Italian goods and the population of Italian immigrants. Immigrants may also be attracted to locations in which there are specialty retailers selling goods from the immigrants' origin who simultaneously

hire immigrants disproportionately from that origin.

To deal with all such origin-by-county specific confounders, we adopt the instrumental variable approach of [Burchardi et al. \(2019\)](#).³⁸ The instrument interacts the arrival into the U.S. of immigrants from origin country o (the push) with the attractiveness of destination d to all immigrants (the pull) during a given historical decade D . We follow [Burchardi et al. \(2019\)](#) and leave out both the continent of origin country o when computing the pull component and leave out the Census region of county c when constructing the push component. Formally, the instrument is defined as

$$IV_{o,c}^D = L_{o,-r(c)}^D \times \frac{L_{-\mathcal{C}(o),c}^D}{L_{-\mathcal{C}(o)}^D}, \quad (15)$$

where $r(c)$ is the Census region of county c , and $\mathcal{C}(o)$ the set of countries on o 's continent. $L_{o,-r(c)}^D$ is the number of immigrants from o settling in the U.S. outside the Census region of county c in decade D and $L_{-\mathcal{C}(o),c}^D/L_{-\mathcal{C}(o)}^D$ is the fraction of immigrants arriving to the U.S. in decade D from outside the continent of o and settling in county c .

The identification assumption is that any confounding factors which make a given county more attractive to both immigrants and importing firms from a given country do not simultaneously affect the interaction of (i) the settlement of immigrants from other continents in past decades with (ii) the total number of immigrants arriving in past decades from the same country but settling in a different Census regions.

We use equation (15) to predict immigrant inflows into the U.S. for all decades between 1880 to 2000, with Appendix Table A.8 providing the first-stage estimates. When estimating equation (14), we use 2SLS to incorporate these instruments. When estimating equation (13), however, we encounter many zeros in household expenditure and variety shares. We again use PPML and account for the non-linearity of PPML by taking a control function approach ([Petrin and Train 2010](#); [Morten and Oliveira 2024](#)). In particular, we add the residuals from the first-stage instrumental variable regressions as controls in our main specifications.³⁹ We

³⁸We provide only a brief description of the instrumental variable strategy here, as our approach follows closely that of [Burchardi et al. \(2019\)](#). We refer the interested reader to Appendix A.6 for more details, as well as to the growing literature making use of this same identification strategy: [Terry et al. \(2024\)](#), [Choi et al. \(2024\)](#), [McCully \(2024\)](#), and [Bonadio \(forthcoming\)](#).

³⁹[Atalay et al. \(2019\)](#) demonstrate in a Monte Carlo simulation that the control function approach gen-

bootstrap standard errors to account for the two stages of our estimation approach.

4.4 Estimation Results

We start by estimating the total effect of immigrants on imports using equation (13), in which expenditure is deflated by household- and county-level preferences.⁴⁰ Columns 1 and 2 of Table 3 provide estimates of β with and without the use of the instrument from Burchardi et al. (2019).⁴¹ In our preferred control function specification, we estimate $\hat{\beta} = 1.01$. Consistent with immigrants' endogenously choosing to living in places with a greater supply of imported goods, the coefficient drops, albeit modestly, between columns 1 and 2.⁴²

With the estimate for β in hand, we now turn to decomposing β into its constituent channels. We estimate the effect of immigrants on local prices using equation (14), with the estimates appearing in Table 4. In columns 1 and 2, we use variation across all barcodes regardless of how regularly we observe them across counties. To address concerns about products sold in only a handful of counties driving our results, we restrict the sample to barcodes which we observe in at least 10 counties in the NielsenIQ data in columns 3 and 4. In columns 2 and 4 we instrument for the bilateral immigrant-population share using the leave-out push-pull instrumental variables defined in equation (15).

The OLS estimates are statistically significantly negative, suggesting more immigrants from a given origin correspond to lower prices for goods imported from that origin. However, the estimated coefficients shrink and become statistically insignificant when adjusting for endogenous immigrant location choice in columns 2 and 4. This suggests that immigrants

erates estimates when using PPML that are quite close to the true data generating process. They further show that the PPML control function estimates are very similar to those produced by the related GMM estimation strategy developed by Wooldridge (1997) and Windmeijer (2000).

⁴⁰We plot the normalized distribution of imputed household- and county-level preferences, \bar{z}'_{oh} and $\bar{z}'_{o,c(h)}$, in Appendix Figure C.3.

⁴¹Burchardi et al. (2019) find no effect of immigrants on trade using state-level trade data and state fixed effects. When we aggregate our household data to the state level and control for state fixed effects, our headline estimates remain positive and significant. We provide further details regarding these differing results in Appendix A.7.

⁴²In Appendix Table C.5, we estimate equation (13) without adjusting the dependent variable for the preference term \mathcal{Z}_{oh} in columns 1 and 2. We also estimate equation (13) using the full sample (including those for which we do not observe nativity) in columns 3 and 4. In both cases, the estimated coefficient is around three times larger than those in Table 3, underlining the importance of adjusting for household heterogeneity in import preferences when identifying the spillover effect β .

Table 3: Estimates of Gravity Equation

	Adjusted import expenditure		Adjusted import variety count	
	(1)	(2)	(3)	(4)
Immigrants/Pop. 2010	1.07*** (0.20)	1.01** (0.49)	0.86*** (0.12)	0.82* (0.48)
First-stage residuals		0.068 (0.54)		0.039 (0.52)
N	167,860	167,860	167,860	167,860
Country FE	✓	✓	✓	✓
Distance & latitude difference	✓	✓	✓	✓
1st-stage F-statistic		103.4		103.4

Notes: The table presents the results of estimating equation (13) using PPML. In columns 1 and 2, the dependent variable is the preference-adjusted origin-specific import expenditure relative to expenditure on US goods, averaged at the county level: \tilde{X}_{oc} . In columns 3 and 4, expenditure is replaced by count of different varieties (barcodes) purchased per household, also averaged at the county level: \tilde{N}_{oc} . The first-stage residual term is taken from a first-stage regression of all the instruments on the immigrant-population share. Observations are weighted using NielsenIQ household weights. Standard errors are computed from 1,000 bootstrapped repetitions. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 4: Estimates of Variable Cost Parameter using Variation in Prices

	Dependent variable: Log Average Barcode Price			
	(1)	(2)	(3)	(4)
Immigrants/Pop. 2010	-0.046*** (0.015)	0.0017 (0.037)	-0.047*** (0.015)	-0.0061 (0.041)
N	2,294,880	2,294,880	2,231,690	2,231,690
Barcode FE	✓	✓	✓	✓
County FE	✓	✓	✓	✓
Distance & latitude difference	✓	✓	✓	✓
1st-stage F-statistic		18.2		17.3
Sample	All	All	>10 Counties	>10 Counties
Num. unique UPCs	270,345	270,345	174,230	174,230

Notes: The table presents two-stage least square regression results at the barcode-county level of estimating equation (14). Standard errors clustered at the barcode and origin country level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

tend to sort into counties in which the price of goods from their origin country is lower.

The point estimate in column 2 implies that a one percentage point increase in the share of the local population from o decreases prices by 0.0017 percent, implying $\hat{\beta}^\tau = (-\theta) \times -0.0017 = 0.0085$. Due to its small magnitude and statistical insignificance, we

conclude that $\hat{\beta}^\tau \approx 0$. This represents the first direct test of how immigrants affect variable trade costs, and the null result is consistent with the previously untested assumption made by [Peri and Requena-Silvente \(2010\)](#).

We next estimate equation (13) but replace the relative expenditure term \tilde{X}_{oh} with the relative variety count share \tilde{N}_{oh} in order to recover the extensive margin import elasticity of immigrants on import expenditure. Columns 3 and 4 of Table 3 provide estimates of the extensive margin effect of immigrants on import expenditure. Our preferred coefficient estimate in column 4 implies that a one percentage point increase in immigrants from a given origin raises the share of varieties purchased from that origin by 0.82 percent.

Given that the estimate in column 4 of Table 3 is close in magnitude to that in column 2, we conclude that the primary spillover channel through which immigrants affect non-immigrant households is the fixed-cost channel. A formal test of the hypothesis that the two coefficients are equal yields a p-value of 0.5, such that we cannot reject $\beta^f = \beta$ and $\beta^z = 0$. Solving instead for the parameters based on the exact point estimates from columns 2 and 4, we recover $\beta^f = 0.60$ and $\beta^z = 0.19$. In our counterfactual analysis in Section 5, we use $\beta_f = \beta = 1.01$ as our baseline parameter specification value but also repeat the counterfactual exercises under the assumption $\beta^f = 0.60$.

4.5 Robustness Checks

StoreScanner. A drawback of the HomeScanner data is that one only observes a small sample of households for the typical county. A potential solution is to turn to NielsenIQ’s StoreScanner dataset, which tracks barcode-level sales across over 35,000 stores in the U.S. We argue, however, that the composition of consumer spending in StoreScanner stores differs meaningfully from HomeScanner purchases, to such an extent as to reduce the usefulness of StoreScanner data for the questions studied here.

Our analysis of the StoreScanner data is based on the 2016 extract of the data merged with the Label Insight data. StoreScanner is based on a subset of retail companies that opt in to sending NielsenIQ data on their sales, which tend to be larger chains, whereas HomeScanner data captures all purchases which we can link to barcode characteristics via Label Insight. This is relevant as households in our HomeScanner sample make 68% of their

grocery expenditures in non-StoreScanner stores. Of particular relevance for our extensive margin and price regressions, HomeScanner purchase records contain vastly more unique barcodes. Among the approximately 275,000 barcodes that we observe in either HomeScanner or StoreScanner datasets, nearly 90% appear exclusively in the HomeScanner dataset. Put differently, in HomeScanner we observe nine times more unique barcodes than we observe in StoreScanner.

In Appendix Table C.6, we show that our baseline results only hold for HomeScanner purchases made at non-StoreScanner stores. Panel A shows the results when restricting to purchases made in StoreScanner stores and Panel B for purchases made outside of StoreScanner stores. While we find statistically insignificant effects of immigrants on import purchases made at StoreScanner stores, the estimated effects are strongly positive for non-StoreScanner store purchases.⁴³ Hence, the positive effect of immigrants on import purchases from our baseline results are driven by shopping at non-StoreScanner stores.

We next analyze the effect of immigrants on import expenditure in the StoreScanner data. We estimate the following gravity model:

$$\tilde{X}_{os} = \exp \left(\alpha_o^{SS} + \alpha_{p(s)}^{SS} + \rho^{SS} d_{o,c(s)} + \beta^{SS} I_{o,c(s)} \right) \times \eta_{o,c(s)}^{SS} \quad (16)$$

where s denotes the store, $c(s)$ the county in which s is located, and $p(s)$ the parent company who operates s . The parent company fixed effect controls for decisions made at the nationwide retailer company level, such as product sourcing or uniform pricing (DellaVigna and Gentzkow 2019).

Two caveats should be kept in mind when evaluating these results. First, as noted above, StoreScanner captures distinct shopping behavior relative to HomeScanner data. Second, because we cannot observe the characteristics of the shoppers at a given store, we can no longer adjust for the preferences of consumers shopping at each store. The resulting coefficient estimates therefore may combine both supply and demand forces.

We report the estimation results of equation (16) in Appendix Table C.7. Consistent with our baseline HomeScanner analysis, immigrants significantly raise import expenditure

⁴³Note that the number of observations varies slightly between the two panels as a few households shop either exclusively at StoreScanner stores or exclusively at non-StoreScanner stores.

(column 2). A one percentage point increase in the share of the population from country o raises expenditures on goods from o relative to domestic goods by 5%. We also estimate a statistically significantly positive effect of immigrants on the extensive margin of imported variety counts (column 4). The magnitude, however, is much smaller than the coefficient for import expenditure in contrast to our baseline HomeScanner findings.

Why might our extensive margin results differ between HomeScanner and StoreScanner? The fixed-cost reduction channel may operate primarily in non-chain stores, which are generally excluded from StoreScanner, perhaps via direct ownership by immigrants who have social and professional connections with their home country. Moreover, chains typically make stocking decisions at the headquarters for the entire national retailer (Jaccard 2023), and thus connections to immigrants local to a store are less likely to affect varieties offered there.

We provide suggestive evidence for this channel in Appendix Figure C.5, which plots β^{SS} from equation (16) when the dependent variable is the ratio of the number of barcodes from o to the number from the U.S. Each point on the graph represents a separate regression, with the sample expanding from using retail chains present in only a single U.S. state for the furthest left estimate to using all retail chains in the data for the furthest right estimate. Among smaller retailer chains, present in only a few states, the results concord with our HomeScanner findings: more immigrants from a given origin increase the number of barcodes sold by stores from that origin. As bigger, national chains are added to the sample, however, the coefficient converges to zero.

Finally, we examine how prices are affected by the local immigrant population in the StoreScanner data. We estimate the following specification:

$$\ln p_{\omega s} = \psi_{c(s)} + \psi_{\omega, p(s)} + \beta^{p, SS} I_{o, c(s)} + \rho^{p, SS} d_{o, c(s)} + \ln \eta_{\omega s}^{SS}, \quad (17)$$

where $p_{\omega s}$ is the expenditure-weighted average price charged by store s throughout 2016. The barcode-parent company fixed effect $\psi_{\omega, p(s)}$ adjusts for potential uniform national pricing that many retailers engage in (DellaVigna and Gentzkow 2019). We cluster standard errors two ways at the barcode-parent company level and at the county-origin level. We weight each observation by the amount of expenditure on the given barcode at the given store.

Appendix Table C.8 reports the results. Similar to our baseline HomeScanner analysis, we find no statistically significant effect of immigrants on prices when using the push-pull instrumental variables (columns 2 and 4).

Excluding immigrants. In our baseline estimation, we adjust for local and household demand factors by dividing the dependent variable by \mathcal{Z}_{oh} . If there are, however, remaining unadjusted differences between immigrants and natives, or if our functional form assumption is sufficiently erroneous, we may still be picking up immigrants' differential demand for imports. We redo our baseline expenditure analysis excluding immigrant households entirely in Appendix Table C.9. The results are very close to those reported in Table 3, suggesting that the \mathcal{Z}_{oh} term largely deals with differences in preferences across households and counties.

5 Immigrants, Imports, and Welfare

This section uses the general equilibrium model introduced above to characterize both the effect of immigrants on import expenditure in the U.S. and the distributional costs of an increase in variable trade costs, such as an import tariff.

To generate counterfactual results which are representative of the U.S. as a whole, we leverage the pooled 2013-2017 ACS sample. In particular, we use the results from estimating equation (5) with the NielsenIQ data to map household socioeconomic characteristics to predicted relative import expenditure shares. We use crosswalks provided by [Burchardi et al. \(2019\)](#) to generate county-specific immigrant population shares based on the Public Use Microdata Area of residence. To compute dollar-equivalent utility values, we assume that each household spends \$7,500 on grocery and personal care products, which is close to the value computed from the Consumer Expenditure Survey.

5.1 Immigrants and Imports

What are the import and welfare effect of immigrants via grocery consumption? We leverage the model introduced in Section 4.1 and the parameters estimated in both Section 3 and Section 4.2 to quantify the effect of immigrants. We do so via two counterfactual exercises.

In the first, we remove all channels through which immigrants affect import expenditure, effectively turning immigrants into otherwise identical native households. In the second counterfactual, we remove immigrant households entirely, thereby including the direct expenditure effect of immigrants on market outcomes.⁴⁴

We summarize the results in Table 5. The results from the first counterfactual scenario—turning immigrants into natives—appear in the first row. Summing across households, we find that aggregate U.S. expenditures on imports of grocery and personal care items fall by 8.8%.⁴⁵ We also find that removing all immigrant spillover effects yields an average welfare loss of 0.1%, amounting to a welfare-equivalent fall of around \$8 per household-year.⁴⁶ With just over 100 million native-born households in the U.S., this amounts to an dollar-equivalent welfare loss of nearly \$800 million.

To understand which mechanisms drive the counterfactual effects, we shut down each channel one at a time. The results appear in the middle section of Table 5. The key finding is that immigrants generate 50 percent more import expenditure via their preferences than by reducing trade costs. Removing the effect of immigrants on fixed trade costs reduces import expenditure by 1.5%, implying that this channel contributes 17% to the total effect. The indirect effect of immigrants’ preferences through the market size term $z_{o,c(h)}$ is a fall of import expenditure by 1.9%.

The preferences of immigrants have the strongest effect on trade volumes among all three channels: removing the demand channel by setting immigrants’ import preference parameters, ζ_1 and ζ_2 , to zero causes a decline of 5.3%, about 60% of the total effect of all immigrant channels.

In the analysis above, we assumed $\beta^f = \beta = 1.01$. In Appendix Table C.10, we present the counterfactual effects under the assumption $\beta_f = 0.6$, the parameter we obtain based on the exact point estimates in Table 3. As one would expect, the fall in the parameter

⁴⁴Note that our counterfactuals only allow for partial-equilibrium adjustment in the consumption space. The labor market effects of immigrants are outside the scope of our framework. For details on how we compute counterfactual changes to trade flows and utility, see Appendix B.4.

⁴⁵It is worth noting that our immigrants-imports elasticity lies in between the range of estimates surveyed by Felbermayr et al. (2015), 0.12–0.15, and the null result reported in Burchardi et al. (2019).

⁴⁶The estimated welfare effect is sizable. For comparison, Brinatti and Morales (2025) find that a 20% increase in immigrants raises native welfare by 0.1%, while Caliendo et al. (2019) and Hsieh and Ossa (2016) estimate the welfare impact of the China shock to be between 0.03% and 0.2%.

Table 5: Counterfactual Results Summary

Counterfactual exercise: Removing ...	(1) Change (%) grocery import expenditure	(2) Change (%) grocery welfare natives	(3) Change (\$) welfare per native HH
... all immigrant channels	-8.8	-0.104	-7.8
... trade cost channel	-1.5	-0.047	-3.5
... market size channel	-1.9	-0.058	-4.4
... demand channel	-5.3	–	–
... all immig. channels + expenditure	-26.9	-7.9	-593

Notes: This table shows the change in outcomes under various counterfactual scenarios. The first counterfactual, with results shown in the first row, removes all channels through which immigrants affect grocery import expenditures—i.e. we set $\zeta_1 = \zeta_2 = \beta_f = 0$ and recalculate \bar{z}_{oc} —but keeps total household expenditure constant. The next three rows show the results when only the following parameters change one at a time: β_f (fixed trade cost); \bar{z}_{oc} (preference-driven market size); and ζ_1 and ζ_2 (demand). In the last row, we remove all immigrant channels and the grocery expenditures made by immigrants (counterfactual 2), equivalent to removing immigrants from the U.S. population. For column 3, we assume that each household spends \$7,500 on grocery and personal care products, which is close the value computed from the Consumer Expenditure Survey.

from 1.01 to 0.6 implies a proportional decrease in the counterfactual effect of removing the trade cost channel on import expenditure from -1.5 to -0.9. Accordingly, the welfare impact associated with this channel is reduced.

The change in import expenditure associated with removing immigrant channels is generally larger than the associated change in welfare. In particular, the effect of immigrants on fixed trade costs and market size are the only channels through which immigrants increase native welfare via increased import expenditure. These two channels capture around 40% of the total trade volume effect of immigrants on local import expenditure. A naive application of the ACR welfare formula to their aggregate trade-creating effect would therefore overestimate the welfare gains to natives by 70%.⁴⁷

The last row of Table 5 shows the second main counterfactual, in which we remove all channels of the first counterfactual as well as immigrants’ grocery expenditures. In the model,

⁴⁷Assuming an initial import expenditure share of 9%: $[(1 - 0.09(1 - 0.088))/(1 - 0.09)]^{(-1/\theta)} - 1 = -0.00173$. The naive welfare estimate is therefore -0.173%, as opposed to our estimated -0.1%.

this corresponds to a reduction in $S_{c(h)}$, the real market size of county c , as opposed to only changing the preference-driven market size z_{oc} as in the previous counterfactual. Accounting for this additional expenditure market size effect leads to a decline in import expenditure by almost 27%. The average change in grocery consumption welfare for natives in this scenario is -7.9% or a welfare-equivalent fall of \$593 per household-year.⁴⁸ In the aggregate, this amounts to a dollar-equivalent welfare loss of over \$59 billion. Thus, removing immigrants' expenditures leads to a sizable welfare effect on natives.⁴⁹ As the scale of demand decreases, fewer varieties—both domestic and foreign—are available for native consumers.⁵⁰

The average effects shown in Table 5 mask considerable geographic heterogeneity. Specifically, Appendix Figure C.6a maps the county-level change in import expenditure associated with removing all immigrant effects, but maintaining immigrant expenditure. For 75% of U.S. counties, immigrants increase local import expenditure by less than 5%. For urban areas such as Los Angeles (CA), Dallas (TX), and Queen's (NY), however, immigrants increase local import expenditure by, respectively, 31%, 23%, and 22%. As suggested by Table 5, the effect of immigrants on native import expenditure is considerably less, with almost 95% of all U.S. counties exhibiting immigrant-induced changes in import expenditure by natives of less than 5%.

The clear urban bias of immigrant-induced import expenditure suggests heterogeneity in native exposure due to sorting of natives by income into urban and rural counties. We confirm this in Figure C.6b, in which we plot the relative benefits to natives of immigrant-induced trade by income deciles. When removing immigrant effects, but maintaining immigrant expenditure, natives in the highest income decile exhibit costs that are a third greater relative to the lowest income decile. Only a quarter of this gap occurs within counties due

⁴⁸In previous versions of this paper, when we were still calibrating $\sigma = 5$, we presented lower estimated welfare effects of removing immigrants. The change in welfare is driven almost entirely by our use of the $\sigma = 3.3$ now estimated in our data in Section 3.4.

⁴⁹For comparison, Piyapromdee (2021) estimates that a counterfactual 25% increase in the immigrant stock would increase native welfare by 1.3% when considering both labor and housing market effects. Albert and Monras (2022) compute a 1.6% welfare increase for natives resulting from immigrant consumption patterns. We view our computed welfare effects to be complementary to these previous figures, given our focus on the novel margin of grocery consumption.

⁵⁰We also run our immigration counterfactuals assuming $\sigma = 3.9$, the point estimate for immigrants from the Armington setup discussed in Section 3. The top line welfare impact of removing all immigrant channels falls to -0.065 and the bottom line welfare impact of removing all immigrants entirely falls to 3.7 percent.

to non-homothetic preferences for imported goods. When we also remove immigrant expenditure, the relative costs for households in the highest income decile increases to over 60%, highlighting the strong assortative matching between high-income natives and immigrants.

While an expansive literature studies the heterogeneous labor market effects of immigrants on native income groups (e.g., [Dustmann et al. 2013](#)), we are the first to study distributional effects on the consumption side, enabled by our highly detailed household-level data. The estimates presented here shed light on the remarkable variation in the consumer gains from immigrant populations across income groups within the U.S. Of particular note is the striking pattern of high-income urban native households benefiting substantially—even within the same county—relative to lower-income natives.

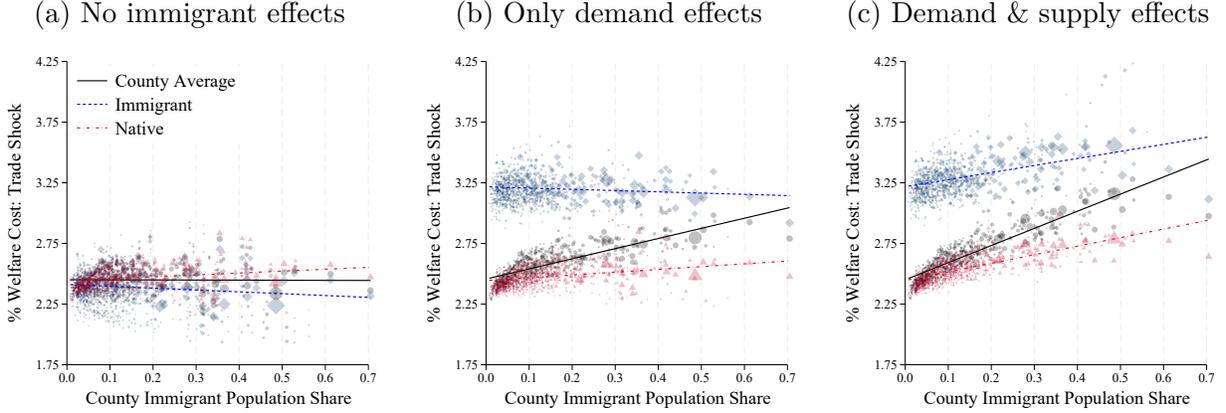
5.2 The Distributional Effects of a Trade Cost Shock

In this section, we consider the relative importance of the demand heterogeneity outlined in [Section 3](#) and the import supply effect of immigrants estimated in [Section 4](#) for understanding the distributional consequences of a trade cost shock, such as an import tariff. We focus our analysis on the policy-relevant counterfactual scenario of an across-the-board 25% increase in variable trade costs.

[Figure 3](#) plots county welfare losses associated with this shock against county immigrant population shares. All losses are in percentage terms, with black circles denoting county averages while blue diamonds and red triangles denote losses specific to immigrants and to natives, respectively. Panel (a) depicts losses in the model with no immigrant effects on import supply and demand. As expected, without any immigrant effects there is now no relationship between county average welfare losses and import supply either in the aggregate or within native and immigrant sub-populations. Furthermore, the level of welfare loss is quite close between immigrants and natives. In this restricted version of the model, immigrants actually exhibit slightly lower losses due to a trade cost shock than natives, as immigrants earn a lower average income.

Panel (b) allows for heterogeneous import preferences between immigrants and natives—as documented in [Section 3](#)—while restricting immigrant supply effects to zero. The level of welfare loss from a trade cost increase is now sharply higher for immigrants than for natives

Figure 3: Welfare Costs of a 25% Variable Trade Cost Shock



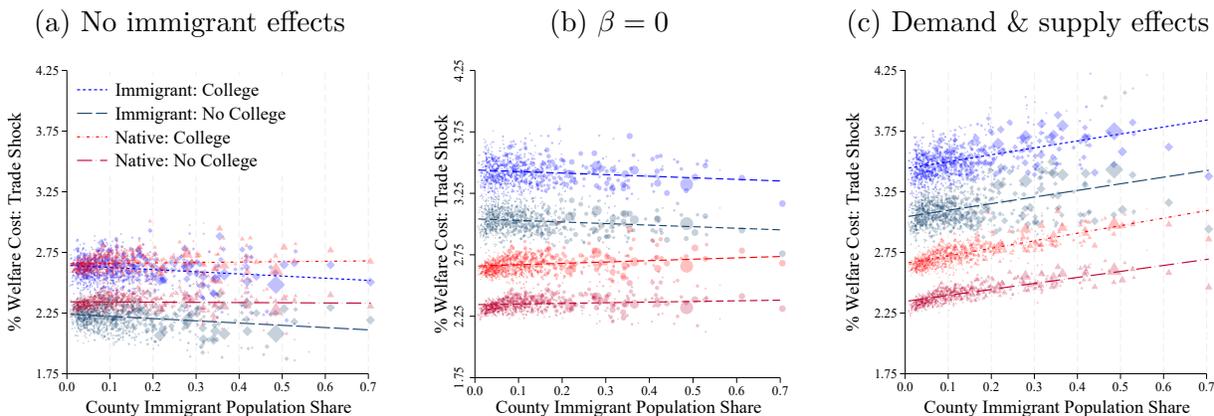
Notes: The figures plot the average county welfare cost within grocery expenditure of a 25% variable cost shock to all imported goods against the county immigrant population share. All markers weighted by total county population. Black circles denote county average costs. Blue diamonds and red triangles provide costs for immigrants and natives, respectively. From left to right, in (a) we plot costs under the model with no immigrant demand or supply effects where $\beta = \zeta_1 = \zeta_2 = 0$, in (b) the model with only immigrant demand effects, $\beta = 0$ but ζ_1 and ζ_2 as estimated, and in (c) the full model with both demand and supply effects of immigrants. Linear best-fit lines are provided in all cases.

and the county average welfare loss increases with the immigrant population share as the composition of local preferences becomes more skewed towards imports.

Finally, panel (c) allows for all immigrant effects, such that immigrants generate both increased import demand and supply in their county of residence. While the level difference in welfare losses between immigrants and natives remains large, these losses are increasing for both groups in the local immigrant population share, consistent with the supply-side spillovers we document. Still, the within-county gap in welfare costs across immigrants and natives is over six times larger than the across-county standard deviation in costs for natives, thereby dominating all variation associated with immigrant spillovers, geography or observable native household characteristics. Immigrants therefore generate considerable variation in the exposure of U.S. counties to trade shocks, but this effect is largely attributable to the composition of households by nativity, rather than a true increase in the variation in exposure to trade shocks of natives themselves.

Figure 4 further decomposes trade shock exposure across counties by both nativity and education, such that we recover average costs for four population types: immigrants with a college degree (blue), immigrants without a college education (light blue), natives with a

Figure 4: Welfare Costs of a 25% Variable Trade Cost Shock by Nativity and Education



Notes: The figures plot the average county welfare cost within grocery expenditure of a 25% variable cost shock to all imported goods against the county immigrant population share. All markers weighted by total county population. Bright blue diamonds denote immigrants with a college degree, teal blue diamonds denote immigrants without a college degree. For natives, bright red denotes natives with a college degree, and dark red natives without a college degree. Linear best-fit lines are provided in all cases.

college degree (red) and natives without a college education (light red). Again, we consider three versions of the model: (a) no immigrants effects, (b) import preference heterogeneity by nativity, and (c) the full model with immigrant effects on demand and supply.

Two additional results stand out from this education-nativity decomposition when compared to the findings shown in Figure 3. First, while existing estimates of heterogeneous exposure to trade cost shocks tend to focus on exposure by income or education, we find that the average within-county gap in welfare losses between immigrants and natives without a college degree is nearly nine times larger than the across-county variation in costs for natives without a college degree. Second, the largest welfare loss gap within county is between college educated immigrants and non-college natives: welfare losses are 46% greater for the former.

6 Conclusion

This paper introduces and quantifies a novel dimension of consumer heterogeneity with respect to the gains from trade: household nativity. By linking U.S. scanner data to both barcode and household origin country data, we provide evidence that immigrants exhibit

grocery consumer gains from trade that are 42% greater than those of native households. 60% of the aggregate immigrant effect on import expenditure is attributable to immigrants' differential expenditure on imports, with the remainder driven by immigrants easing supply-side constraints on imports.

This study makes two core contributions. First, we show that the gains from trade are unequally distributed by nativity, a novel margin of heterogeneity. Within the grocery sector, nativity is substantially more important than income, the standard margin the literature has examined, in generating unequal gains from trade through the expenditure channel. Moreover, immigrants raise natives' gains from trade by reducing trade costs.

Our second key contribution is to the debate about the costs and benefits of immigration. While a vast literature focuses on the labor market impact of immigrants, we are the first to rigorously quantify the consumption-side benefits of immigrants. Our quantitative estimates can be combined with existing labor- and housing-market estimates of the welfare effect of immigrants to compute a more comprehensive net welfare effect of immigration.

Our results are consistent with recent political trends across socioeconomic groups in the U.S. Immigrant, high-income and highly educated households pay an outsized share of the cost of tariffs, and high-income and urban natives pay an outsized share of the cost of immigrant removals. These same households tend to oppose tariffs and anti-immigrant policies (O'Rourke and Sinnott 2006; Hanson et al. 2007; Tabory and Smeltz 2017; Stantcheva 2022). Further research is needed to understand the relationship between consumption preferences and voting behavior.

A challenge for future research will be to apply and extend our framework to other dimensions of international trade and investment. A large share of international trade deals with intermediate inputs, for example, which immigrants may facilitate (Ariu 2022). Extending the analysis to trade in durable final goods, while not feasible with available U.S. data, may be another avenue for future research.

References

Albert, Christoph and Joan Monras, "Immigration and Spatial Equilibrium: The Role of Expenditures in the Country of Origin," *American Economic Review*, November 2022, 112 (11), 3763–3802.

- Ariu, Andrea**, “Foreign workers, product quality, and trade: Evidence from a natural experiment,” *Journal of International Economics*, 2022, 139, 103686.
- Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare**, “New trade models, same old gains?,” *American Economic Review*, 2012, 102 (1), 94–130.
- Armington, Paul S**, “A Theory of Demand for Products Distinguished by Place of Production,” *Staff Papers-International Monetary Fund*, 1969, pp. 159–178.
- Atalay, Enghin, Ali Hortaçsu, Mary Jialin Li, and Chad Syverson**, “How wide is the firm border?,” *Quarterly Journal of Economics*, 2019, 134 (4), 1845–1882.
- Atkin, David**, “The caloric costs of culture: Evidence from Indian migrants,” *American Economic Review*, 2016, 106 (4), 1144–1181.
- , **Eve Colson-Sihra, and Moses Shayo**, “How do we choose our identity? A revealed preference approach using food consumption,” *Journal of Political Economy*, 2021, 129 (4), 1193–1251.
- Aubry, Amandine, Michał Burzyński, and Frédéric Docquier**, “The welfare impact of global migration in OECD countries,” *Journal of International Economics*, 2016, 101, 1–21.
- Auer, Raphael, Ariel Burstein, Sarah Lein, and Jonathan Vogel**, “Unequal expenditure switching: Evidence from Switzerland,” *Review of Economic Studies*, 2024, 91 (5), 2572–2603.
- Bai, Liang and Sebastian Stumpner**, “Estimating US consumer gains from Chinese imports,” *American Economic Review: Insights*, 2019, 1 (2), 209–24.
- Bonadio, Barthélémy**, “Migrants, trade and market access,” *Review of Economics and Statistics*, forthcoming.
- Borjas, George J**, “The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market,” *Quarterly Journal of Economics*, 2003, 118 (4), 1335–1374.
- Borussyak, Kirill and Xavier Jaravel**, “The distributional effects of trade: Theory and evidence from the United States,” Technical Report, National Bureau of Economic Research 2021.
- Bowles, Samuel**, “Endogenous preferences: The cultural consequences of markets and other economic institutions,” *Journal of Economic Literature*, 1998, 36 (1), 75–111.
- Brinatti, Agostina and Nicolas Morales**, “Firm heterogeneity and the impact of immigration: Evidence from German establishments,” 2025.
- Bronnenberg, Bart J, Jean-Pierre H Dubé, and Matthew Gentzkow**, “The evolution of brand preferences: Evidence from consumer migration,” *American Economic Review*, 2012, 102 (6), 2472–2508.
- Burchardi, Konrad B, Thomas Chaney, and Tarek A Hassan**, “Migrants, ancestors, and foreign investments,” *Review of Economic Studies*, 2019, 86 (4), 1448–1486.
- Burstein, Ariel, Gordon Hanson, Lin Tian, and Jonathan Vogel**, “Tradability and the Labor-Market Impact of Immigration: Theory and Evidence from the United States,” *Econometrica*, 2020, 88 (3), 1071–1112.
- Caliendo, Lorenzo, Maximiliano Dvorkin, and Fernando Parro**, “Trade and labor market dynamics: General equilibrium analysis of the china trade shock,” *Econometrica*, 2019, 87 (3), 741–835.
- Card, David**, “Immigrant inflows, native outflows, and the local labor market impacts of

- higher immigration,” *Journal of Labor Economics*, 2001, 19 (1), 22–64.
- Chaney, Thomas**, “Distorted gravity: the intensive and extensive margins of international trade,” *American Economic Review*, 2008, 98 (4), 1707–21.
- Choi, Jaerim, Jay Hyun, and Ziho Park**, “Bound by ancestors: Immigration, credit frictions, and global supply chain formation,” *Journal of International Economics*, 2024, 147, 103855.
- Combes, Pierre-Philippe, Miren Lafourcade, and Thierry Mayer**, “The trade-creating effects of business and social networks: Evidence from France,” *Journal of International Economics*, 2005, 66 (1), 1–29.
- Correia, Sergio, Paulo Guimarães, and Thomas Zylkin**, “Verifying the existence of maximum likelihood estimates for generalized linear models,” *arXiv preprint arXiv:1903.01633*, 2019.
- Costinot, Arnaud and Andrés Rodríguez-Clare**, “Trade theory with numbers: Quantifying the consequences of globalization,” in “Handbook of international economics,” Vol. 4, Elsevier, 2014, pp. 197–261.
- Cravino, Javier and Andrei A Levchenko**, “The distributional consequences of large devaluations,” *American Economic Review*, 2017, 107 (11), 3477–3509.
- DellaVigna, Stefano and Matthew Gentzkow**, “Uniform pricing in us retail chains,” *Quarterly Journal of Economics*, 2019, 134 (4), 2011–2084.
- Di Giovanni, Julian, Andrei A Levchenko, and Francesc Ortega**, “A global view of cross-border migration,” *Journal of the European Economic Association*, 2015, 13 (1), 168–202.
- Dustmann, Christian, Tommaso Frattini, and Ian P Preston**, “The effect of immigration along the distribution of wages,” *Review of Economic Studies*, 2013, 80 (1), 145–173.
- , **Uta Schönberg, and Jan Stuhler**, “Labor supply shocks, native wages, and the adjustment of local employment,” *Quarterly Journal of Economics*, 2017, 132 (1), 435–483.
- Faber, Benjamin and Thibault Fally**, “Firm heterogeneity in consumption baskets: Evidence from home and store scanner data,” *Review of Economic Studies*, 2022, 89 (3), 1420–1459.
- Fajgelbaum, Pablo D and Amit K Khandelwal**, “Measuring the unequal gains from trade,” *Quarterly Journal of Economics*, 2016, 131 (3), 1113–1180.
- Feenstra, Robert C**, “New product varieties and the measurement of international prices,” *The American Economic Review*, 1994, pp. 157–177.
- , **Philip Luck, Maurice Obstfeld, and Katheryn N Russ**, “In search of the Armington elasticity,” *Review of Economics and Statistics*, 2018, 100 (1), 135–150.
- Felbermayr, Gabriel, Volker Grossmann, and Wilhelm Kohler**, “Migration, international trade, and capital formation: Cause or effect?,” in “Handbook of the economics of international migration” 2015, pp. 913–1025.
- Fernández, Raquel and Alessandra Fogli**, “Culture: An empirical investigation of beliefs, work, and fertility,” *American Economic Journal: Macroeconomics*, 2009, 1 (1), 146–177.
- Galle, Simon, Andrés Rodríguez-Clare, and Moises Yi**, “Slicing the pie: Quantifying the aggregate and distributional effects of trade,” *Review of Economic Studies*, 2023, 90

- (1), 331–375.
- Ghai, Surabhi and Colin Hottman**, “Estimating Armington Elasticities using Scanner Data,” 2025.
- Gould, David M.**, “Immigrant Links to the Home Country: Empirical Implications for U.S. Bilateral Trade Flows,” *Review of Economics and Statistics*, 1994, 76 (2), 302–316.
- Hanson, Gordon H, Kenneth Scheve, and Matthew J Slaughter**, “Public finance and individual preferences over globalization strategies,” *Economics & Politics*, 2007, 19 (1), 1–33.
- Hausman, Jerry A.**, “Valuation of new goods under perfect and imperfect competition,” in “The economics of new goods,” University of Chicago Press, 1996, pp. 207–248.
- Head, Keith and John Ries**, “Immigration and trade creation: econometric evidence from Canada,” *Canadian Journal of Economics*, 1998, pp. 47–62.
- **and Thierry Mayer**, “Gravity equations: Workhorse, toolkit, and cookbook,” in “Handbook of international economics,” Vol. 4, Elsevier, 2014, pp. 131–195.
- Hottman, Colin J and Ryan Monarch**, “A matter of taste: Estimating import price inflation across US income groups,” *Journal of International Economics*, 2020, 127, 103382.
- Hsieh, Chang-Tai and Ralph Ossa**, “A global view of productivity growth in China,” *Journal of international Economics*, 2016, 102, 209–224.
- Iranzo, Susana and Giovanni Peri**, “Migration and trade: Theory with an application to the Eastern–Western European integration,” *Journal of International Economics*, 2009, 79 (1), 1–19.
- Jaccard, Torsten**, “Retailers and Import Penetration: Evidence from US Dollar Stores,” 2023.
- Jaravel, Xavier**, “The unequal gains from product innovations: Evidence from the us retail sector,” *Quarterly Journal of Economics*, 2019, 134 (2), 715–783.
- Llull, Joan**, “Immigration, wages, and education: A labour market equilibrium structural model,” *Review of Economic Studies*, 2018, 85 (3), 1852–1896.
- Logan, Trevon D and Paul W Rhode**, “Moveable feasts: A new approach to endogenizing tastes,” *manuscript (The Ohio State University)*, 2010.
- McCully, Brett A.**, “Immigrants, legal status, and illegal trade,” *Journal of International Economics*, 2024, 152, 104016.
- Melitz, Marc J.**, “The impact of trade on intra-industry reallocations and aggregate industry productivity,” *Econometrica*, 2003, 71 (6), 1695–1725.
- Miho, Antonela, Alexandra Jarotschkin, and Ekaterina Zhuravskaya**, “Diffusion of gender norms: Evidence from Stalin’s ethnic deportations,” *Journal of the European Economic Association*, 2023, 22 (2), 475–527.
- Monras, Joan**, “Immigration and wage dynamics: Evidence from the Mexican Peso crisis,” *Journal of Political Economy*, 2020, 128 (8), 3017–3089.
- Morten, Melanie and Jaqueline Oliveira**, “The effects of roads on trade and migration: Evidence from a planned capital city,” *American Economic Journal: Applied Economics*, 2024, 16 (2), 389–421.
- Neiman, Brent and Joseph Vavra**, “The rise of niche consumption,” *American Economic Journal: Macroeconomics*, 2023, 15 (3), 224–264.
- Nevo, Aviv**, “Measuring market power in the ready-to-eat cereal industry,” *Econometrica*, 2001, 69 (2), 307–342.

- O'Rourke, Kevin H and Richard Sinnott**, "The determinants of individual attitudes towards immigration," *European Journal of Political Economy*, 2006, 22 (4), 838–861.
- Ottaviano, Gianmarco IP and Giovanni Peri**, "Rethinking the effect of immigration on wages," *Journal of the European Economic Association*, 2012, 10 (1), 152–197.
- Parsons, Christopher and Pierre-Louis Vézina**, "Migrant networks and trade: The Vietnamese boat people as a natural experiment," *Economic Journal*, 2018, 128 (612), F210–F234.
- Peri, Giovanni and Francisco Requena-Silvente**, "The trade creation effect of immigrants: Evidence from the remarkable case of Spain," *Canadian Journal of Economics*, 2010, 43 (4), 1433–1459.
- Petrin, Amil and Kenneth Train**, "A control function approach to endogeneity in consumer choice models," *Journal of Marketing Research*, 2010, 47 (1), 3–13.
- Piyapromdee, Suphanit**, "The impact of immigration on wages, internal migration, and welfare," *Review of Economic Studies*, 2021, 88 (1), 406–453.
- Santos Silva, Joao MC and Silvana Tenreyro**, "The log of gravity," *Review of Economics and Statistics*, 2006, 88 (4), 641–658.
- Spolaore, Enrico and Romain Wacziarg**, "Ancestry, Language and Culture," in Victor Ginsburgh and Shlomo Weber, eds., *The Palgrave Handbook of Economics and Language*, London: Palgrave Macmillan, 2016, chapter 6, pp. 174–211.
- Stantcheva, Stefanie**, "Understanding of trade," Technical Report, National Bureau of Economic Research 2022.
- Steingress, Walter**, "The causal impact of migration on US trade: Evidence from political refugees," *Canadian Journal of Economics*, 2018, 51 (4), 1312–1338.
- Tabory, Sam and Dina Smeltz**, "The Urban-Suburban-Rural «Divide» in American Views on Foreign Policy," Technical Report, The Chicago Council on Global Affairs 2017.
- Terry, Stephen J, Thomas Chaney, Konrad B Burchardi, Lisa Tarquinio, and Tarek A Hassan**, "Immigration, Innovation, and Growth," 2024.
- Windmeijer, Frank**, "Moment conditions for fixed effects count data models with endogenous regressors," *Economics Letters*, 2000, 68 (1), 21–24.
- Wooldridge, Jeffrey M**, "Multiplicative panel data models without the strict exogeneity assumption," *Econometric Theory*, 1997, 13 (5), 667–678.

A Empirical Appendix

A.1 Data Construction

We aggregate decennial census waves across individuals aged 16 and above to the county-by-origin level, applying the Census’ individual sample weights. Immigrants are defined as those born outside the U.S. To compute decadal immigrant inflows from origin o into destination county c between two census years $t-10$ and t , denoted L_{oc}^t , we count only those respondents in t who immigrated to the U.S. between $t-10$ and t . Following [Burchardi et al. \(2019\)](#), in the first sample year the measure L_{oc}^{1880} includes all those that are either first-generation immigrants from o or second-generation immigrants whose parents were born in o .

Destination regions c are defined as 1990 counties and we use the transition matrices provided by [Burchardi et al. \(2019\)](#) to maintain consistent boundaries over time despite the Census changing geographic boundaries across waves. The U.S. geography of reference is called “Historic counties” until 1940; then county groups in 1970/1980; and finally public-use microdata areas (PUMAs) from 1980 to the present.

The latest available transition matrix provided by [Burchardi et al. \(2019\)](#) is for the year 2010, in which PUMAs are based on 2000 boundaries. Thus, for the 2013-2017 ACS sample, in which PUMAs are based on 2010 boundaries, we use the crosswalk provided by the Missouri Census Data Center to transition PUMAs to 2000 boundaries before applying the corresponding transition to 1990 counties.

A.2 Robustness of Perceived Import Quality Estimation

While the pseudo-Poisson maximum likelihood (PPML) estimator in order to account for the zeros in \tilde{X}_{oh} maintains theory-consistency in the estimating equation, one may be concerned that the relative weight PPML attributes to zeros could potentially bias our estimates ([Breinlich et al. 2024](#); [Tyazhelnikov and Zhou 2024](#)). We therefore provide two robustness checks and replicate our estimates in Table 2. Table A.1 applies OLS to estimate equation (5) with relative expenditure by h on goods from o measured in levels, and Table A.2 applies OLS to estimate equation (5) as a linear probability model, such that the dependent variable

is a binary indicator for $\tilde{X}_{oh} > 0$.⁵¹

In both cases, the estimates of ζ_1 and ζ_2 mirror those found when estimating our empirical model with PPML. In levels, we find that immigrants exhibit import expenditure shares that are 0.03 percentage points higher than within-county natives for all origins and 0.83 percentage points higher for their specific origin country. Both estimates are significant at 99% confidence levels, and suggest an import expenditure ratio for immigrants that is 25% greater than for native households.⁵² As shown in Table A.2, the probability of purchasing a nonzero amount of groceries from a given import origin follows a similar pattern, with all estimates positive and significant at 99% confidence levels.

A.3 Sample Representativeness

As discussed by Feenstra et al. (2023), the NielsenIQ household sample may not be perfectly representative of the U.S. population in terms of income and price sensitivity. We further find that the NielsenIQ data is not representative of the distribution of immigrant origin countries. Mexican-born immigrants, for example, make up 15% of immigrant households in NielsenIQ but 30% in the ACS.

NielsenIQ’s shortcomings in representativeness across immigrant origins may be driven by a combination of two factors. First, cross-sectionally NielsenIQ HomeScanner may miss some households if, for example, the survey module is not available in the language an immigrant household speaks. Secondly, there may be differential attrition across immigrant origins between 2008, when the “Tell Me More About You” was distributed, and the sample period of 2014-2016.⁵³

To gauge the importance of NielsenIQ’s uneven representation across immigrant origins in driving our results, we adjust the survey weights so that the weighted aggregate population shares of natives and immigrants of each origin reflect those measured in the pooled 2013-

⁵¹Table A.1 provides estimates of ζ_1 and ζ_2 associated with applying OLS to the following estimating equation:

$$\tilde{X}_{oh} = \Psi_{o,c(h)} + \delta J_h + \zeta_1 \mathbf{1}[o(h) \neq US] + \zeta_2 \mathbf{1}[o(h) = o] + \epsilon_{oh}.$$

⁵²The average per-origin import expenditure share is 0.12 percent, with an average probability of purchasing nonzero groceries from a given origin of 0.17.

⁵³The NielsenIQ survey is voluntary so households may drop out at any time.

2017 ACS sample. Using these adjusted weights, we replicate our baseline estimates of import demand preferences from Table 2 in Table A.3 and our baseline gravity equations from Table 3 in Table A.4. Both the estimates of immigrants’ import demand preferences and the estimates of their spillover effects are very similar to the baseline estimates. If anything, the spillover effects are slightly larger when using the adjusted weights.

As an additional check on potentially selective attrition between 2008 and our sample period of 2014-2016, we estimate equation (13) but using the full 2014-2016 NielsenIQ sample (which includes every household whether or not we observe its nativity) and therefore dropping the nativity variables from the estimation. The dependent variable is thus $\tilde{X}_{oc} \equiv \sum_{h \in \Lambda'_c} \kappa_h \tilde{X}_{oh}$, that is, the weighted average of origin-county relative expenditure without deflating by Z_{oh} , since we no longer observe the full vector of household characteristics required to predict Z_{oh} . We then compare the results with the same specification using our smaller main sample in order to investigate how much attrition matters for the estimates. Table C.5 shows that the IV estimate increases modestly from 2.84 with the nativity sample (column 2) to 3.09 with the full sample (column 4). However, these two estimates are statistically indistinguishable from each other. We conclude that differential attrition is unlikely to drive our estimate of the immigrant spillover effect.

A.4 Ethnicity and Alternative Measures of Origin Country Connectedness

In our baseline specification for estimating import demand heterogeneity, equation (5), we allow households to have specific preferences for (i) all imports and (ii) imports specifically from their origin country. Households may additionally exhibit specific preferences towards goods from countries close—geographically or culturally—to their origin country or, in the case of second-generation immigrants, the origin country of their parents. We test the importance of such specific household preferences in Table A.5 using various proxies.

We start by assessing whether first- and second-generation immigrant exhibit differential demand for goods from their *continent* of origin. The NielsenIQ data, however, do not ask about panelists’ parents’ countries of origin, so we cannot directly observe second-generation

immigrants. Instead, we use indicators for Hispanic and Asian ethnicity as a proxy. In the Current Population Survey, 77% of Asian and 55% of Hispanic non-immigrants have at least one immigrant parent, suggesting ethnic indicators are reasonable, albeit imperfect proxies for second-generation immigrants.

In column 1 of Table A.5, we show estimates of dummies indicating Hispanic or Asian ethnicity (which are both also included in our baseline specification). We further interact the ethnicity variables with dummies indicating imports originating in Latin America (defined as the set of origins we observe in the Americas excluding Canada) or Asia, respectively. We continue to control for whether the household was born abroad. Households of Hispanic origin spend somewhat less than natives on imports in general, although the effect is imprecisely estimated, but 15% more on imports from Latin America. Similarly, households of Asian origin spend 51% more on imports from Asian countries.

In column 2, we test whether immigrant households generally exhibit a specific preference for goods from their continent of origin (in addition to their country of origin). We do not find evidence of such a preference. Thus, differential preferences extending to the whole continent seem to be unique to ethnically Asian and Hispanic households.

In column 3, we add a set of dummy variables that capture cultural similarity between an immigrant's origin and the import origin: sharing the same language and past or current colonial relationships, all obtained from Conte et al. (2022). We find only a marginally statistically significant effect of a past colonial relationship on import demand.

A.5 Armington Elasticity Heterogeneity

As an alternative approach to incorporating demand elasticity heterogeneity within a framework consistent with the ACR welfare formula, we consider estimates of the Armington elasticity associated with substitution across origin countries.

Specifically, we assume a CES demand structure across baskets of grocery varieties, which are differentiated according to their origin country, with potential heterogeneity by household

type z . Formally, we consider preferences for household type z as:

$$U(z) = \left[\sum_{o \in \mathcal{O}} (\phi_o(z) q_o(z))^{\frac{\eta(z)}{\eta(z)-1}} \right]^{\frac{\eta(z)-1}{\eta(z)}},$$

where $o \in \mathcal{O}$ denotes origin countries. $\eta(z) > 1$ denotes the Armington elasticity for household type z , and $\phi_o(z)$ the quality, or demand shifter, of origin country o for households of type z . Expenditure by household type at the origin country level is therefore:

$$x_o(z) = \left(\frac{\phi_o(z)}{P_o} \right)^{\eta(z)-1} X(z) P(z)^{\eta(z)-1},$$

where $X(z)$ is total expenditure by households of type z . P_o denotes some price index over goods from origin o , and $P(z)$ denotes the aggregate price index for households of type z , which is given by the following:

$$P(z) = \left[\sum_o \left(\frac{\phi_o(z)}{P_o} \right)^{\eta(z)-1} \right]^{\frac{1}{1-\eta(z)}}.$$

Throughout, we assume that firms cannot price discriminate by nativity within a market region, such that the same price P_o applies to all household types. Let the elasticity of substitution of varieties within origin be σ .

We aggregate across barcodes within a given country-market-period, thereby alleviating some of the small-sample issue associated with our immigrant sample, and generate the following estimating equation:

$$\Delta_t \ln X_{oct} = (1 - \eta(z)) \Delta_t \ln P_{oct} + \alpha_{ctz} + \epsilon_{oct} \quad (\text{A.1})$$

where, as before, c denotes a city and t a bi-monthly period. As in Section 3.4, we calculate prices using a simple sales-weighted average across all purchase records associated with all NielsenIQ households. We apply this same price to all estimating equations associated with population subsamples (immigrants, natives, or pooled). We make use of exchange rates as an instrumental variable for origin-specific price changes $\Delta_t \ln P_{oct}$ and estimate equation (A.1) using 2SLS.

From a measurement perspective, this approach has much to recommend. First, by aggregating to the origin country level, we mitigate some of the concern associated with a lack of observations for the immigrant sample at the barcode level. While the expenditure associated with only 1,237 households is spread thinly across barcodes, expenditure across origin countries is less thinly distributed due to aggregation. Second, exchange rates offer a plausible and readily implementable instrumental variable for origin-specific average prices. Third, origin countries likely do not partake, as a whole, in strategic pricing. Thus, the fact that the immigrant sample will again be biased towards high market share origin countries does not necessarily imply that the immigrant sample will exhibit lower elasticities of demand, on average, than the non-immigrant sample.

Table A.7 provides estimates across two different estimation procedures. Columns 1-4 provide estimates associated with applying 2SLS directly to the data associated with each sub-population. For the entire NielsenIQ sample, we recover $\eta = 4.07$ (SE=0.53). [Ghai and Hottman \(2025\)](#) use similar scanner data in the U.S. to recover a median Armington elasticity of 3.0 across fourteen scanner data product categories, while [Feenstra et al. \(2018\)](#) estimate a range of Armington elasticities of between 3.1-4.1 for grocery food categories in the U.S. Thus our baseline estimate is at the higher end of estimates found elsewhere. Note that in order for the Armington model to satisfy the underlying preference specification, it must be that $\eta < \sigma$, and origin countries are seen as less substitutable than the varieties within each origin. [Ghai and Hottman \(2025\)](#) estimate a median within-origin price elasticity of 7.2 using a version of the estimator described in [Feenstra \(1994\)](#), thereby suggesting the Armington assumption is validated for a majority of their product categories. [Feenstra et al. \(2018\)](#), on the other hand, estimate a range of 3.2-4.1 for the within-origin elasticity, and therefore cannot reject that $\sigma = \eta$ for 70% of their product categories.

However it should be noted that the “Hausman-Nevo” instruments implemented for our estimation of σ tend to lead to demand estimates that are smaller in magnitude than those recovered using the method developed in [Feenstra \(1994\)](#). For example, the elasticity of 7.2 documented in [Ghai and Hottman \(2025\)](#) is considerably larger than estimates found in much of the industrial organization literature, and in particular those found in [Faber and Fally \(2022\)](#), who document a median price elasticity in U.S. scanner data of approximately

2 using instruments which are similar to those used in this paper. In any case, we are more concerned with understanding the extent to which immigrants and natives exhibit differing elasticities of demand, as opposed to necessarily recovering elasticities which are consistent with Armington preferences. Still, our estimates do not suggest that an Armington demand structure is empirically validated.

Columns 3 and 4 provide estimates of η for natives and immigrants, respectively. The small sample size associated with immigrants clearly leads to a reduction in the first-stage strength of our instrument, thereby biasing estimates of η upwards in magnitude. The first-stage F-statistic in this case is only 2.11 for immigrants, versus 51.84 when estimated for the entire NielsenIQ sample. Thus, the Armington elasticity recovered for natives is 5.61 and the Armington elasticity for immigrants is estimated as 10.59 (SE=6.88).

Columns 5-8 offer an alternative approach, in which we first estimate the first-stage residuals using data for the entire NielsenIQ sample, and then add these residuals to each second-stage estimating equation, which we estimate via OLS. That is, we make use of information regarding all purchase information in order to generate the first-stage relationship between exchange rates and scanner prices, and use this estimated relationship to de-bias second-stage estimates of η for varying sub-groups. In this case, we remove bias associated with a weak first-stage and generate estimates of η which are clearly indistinguishable for immigrants and natives. Our immigrant estimate is 3.91 (SE=0.85), while our native estimate is 3.78 (SE=0.39).

When compared to our headline estimate of $\sigma = 3.30$, which we apply to both immigrants and natives, it is clear that we cannot reject, at least for immigrant households, that our within-origin price elasticity is equal to the Armington elasticity, thus validating the use of CES preferences over all varieties, as opposed to the Armington preferences described here.⁵⁴

⁵⁴Note that in Table A.6, we include origin-market fixed effects in our preferred specification, thereby aligning our estimating equation with that which would be derived from an Armington nested-CES preference structure.

A.6 Instrumental Variables: Details and First-Stage Estimates

This section provides a more detailed discussion regarding our implementation of the leave-out push-pull instrumental variables introduced by [Burchardi et al. \(2019\)](#). The same instrument has been used by a recent crop of papers studying the effects of immigration, such as [Terry et al. \(2024\)](#), [Bonadio \(forthcoming\)](#), [McCully \(2024\)](#), and [Choi et al. \(2024\)](#).

The immigration leave-out push-pull instrument interacts the arrival to the U.S. of immigrants from origin country o (push) with the attractiveness of different destinations to immigrants (pull) measured by the fraction of all immigrants to the U.S. who choose to settle in county c . A simple version of the instrument is defined as

$$IV_{o,c}^D = L_o^D \times \frac{L_c^D}{L^D},$$

where L_o^D is the number of immigrants from origin o coming to the U.S. in decade D , and L_c^D/L^D is the fraction of immigrants to the U.S. who choose to settle in county c in that decade.

Two key threats to the exogeneity of the instrument remain: a scale component and a spatial correlation component. The scale component is the threat that a single origin o constitutes a large share of the instrument's components for a given county c . A simple solution would be to leave out the bilateral immigration $L_{o,c}^D$ flows when constructing the instrument for the country-county pair $\{o, c\}$.

However, there might also be spatial correlation in confounding variables. For example, both Belgian and French immigrants and goods may go to Chicago for the same reason: many flight connections out of Paris, which is very accessible to potential Belgian immigrants by train. Leaving out Belgium-to-Chicago immigration flows when computing the instrument predicting these same immigration flows is therefore not sufficient, because now the French immigration flows to Chicago (used to predict Belgium-to-Chicago flows) are also contaminated with the confounding flight connections. To avoid such endogeneity, we again follow [Burchardi et al. \(2019\)](#) and leave out both the set of countries which share a continent with origin country o , $\mathcal{C}(o)$, and the Census region of county c , $r(c)$, to construct the instrumental variable that we defined in equation (15).

A violation of the identification assumption may occur if, say, immigrants and goods from France tend to flow to Chicago and immigrants and goods from South Korea flow to Miami in the same decade and for the same reason: a large number of flight connections. This violation is only quantitatively meaningful if the French are a large fraction of immigrants settling in Chicago, and if South Korean immigrants are a large fraction of the immigrants settling in Miami.

We use equation (15) to predict immigrant inflows into the U.S. decades spanning 1880 to 2000. [Burchardi et al. \(2019\)](#) extensively explore the validity of this instrumental variable and conduct numerous robustness checks for the instrument in the U.S. setting and find that it holds up quite well. Following [Burchardi et al. \(2019\)](#), we include five principal component terms which capture the variation of interactions of the instruments within county-origin pairs and across decades.⁵⁵

While the push-pull instrument may bear a passing resemblance to a standard shift-share instrument, we note two key differences. First, shift-share instruments are typically summed over a dimension (e.g., across origins), whereas the push-pull is not summed and thus retains two dimensions of variation. Second, the ‘share’ component of the push-pull is not lagged, unlike in the canonical shift-share style instrument, such as the ethnic enclave instrument proposed by [Card \(2001\)](#).

We show the first-stage results of the leave-out push-pull instruments using our HomeScanner data at the county-origin level in [Table A.8](#). We find that the push-pull instrument strongly predicts the contemporary bilateral immigrant population.

We estimate the first-stage three ways. In column 1, the specification is at the county-by-origin level, consistent with equation (13). Here we cluster standard errors at the county-level and weight by the sum of NielsenIQ household projection factors within the county. F-statistics differ from those appearing in the main text because we bootstrap them in the main text. The first-stage F-statistic is about 110, well above conventional thresholds of instrument relevance. Coefficients are typically positive and statistically significant, with the exception of 1900.

⁵⁵We compute 1,013 higher-order interaction terms, defined as $L_{o,-r(c)}^{D'} \times \dots \times L_{o,-r(c)}^D L_{-c(o),c}^D / L_{-c(o)}^D$ for each $D' < D \leq 2000$. We then compute five principal components which capture the variation contained within those 1,013 terms.

Columns 2 and 3 show estimates from data at the county-by-barcode level, as in equation (14). We again cluster standard errors at the county level. The F-statistics are near 20, and most coefficients are positive and statistically significant, with the exception of the earlier decades. Note that we are predicting immigrant populations (and not ancestry populations, as in [Burchardi et al. 2019](#)), and cohorts of immigrant groups likely change their location choices over time.

A.7 Comparison with the Results of Burchardi et al. (2019)

[Burchardi et al. \(2019\)](#) estimate a null effect of immigrants on trade.⁵⁶ We consider three possible explanations for our diverging results, and conclude that the choice of estimation strategy explains the difference between our results and those of [Burchardi et al. \(2019\)](#).⁵⁷ We then explain why we prefer our PPML estimator over the Heckman selection model used by [Burchardi et al. \(2019\)](#).

First, the primary explanatory variable of [Burchardi et al. \(2019\)](#) is the logarithm of one plus the count of individuals with a given ancestry measured in thousands instead of the foreign-born population share that we use. We first replace our previous explanatory variable, *Immigrants/Pop. 2010*, with *Ancestry/Pop. 2010*. Next, we take the functional form used in [Burchardi et al. \(2019\)](#), *Log Ancestry 2010*. Table A.9 shows that we still obtain positive coefficients with these alternative measures using our data and estimation strategy, though the coefficient on *Ancestry/Pop. 2010* is not statistically significant.

Second, [Burchardi et al. \(2019\)](#) use state-level data, whereas we leverage household- and county-level data, and [Burchardi et al. \(2019\)](#) use a two-step Heckman estimation strategy to account for selection into bilateral trading, whereas we apply the pseudo-Poisson maximum likelihood (PPML) estimator to account for zeros. We test whether the level of data aggregation or the estimation procedure resolve our diverging results. In Table A.10 we mimic the specification in [Burchardi et al. \(2019\)](#) more closely by aggregating our data to the state-origin level. We run regressions first using our PPML approach (columns 1 and 2).

⁵⁶The focus of [Burchardi et al. \(2019\)](#), however, was on how immigrants shape FDI.

⁵⁷Something we cannot test is the difference in product coverage. While our data covers groceries and related final goods, the data of [Burchardi et al. \(2019\)](#) covers all traded goods.

As explanatory variables we employ both *Log Ancestry 2010* and our preferred measure *Immigrants/Pop. 2010*. As in our baseline household- and county-level results, we continue to find a significantly positive effect of immigrants and ancestors on import volumes. Turning to columns 3 and 4, we apply the Heckman correction strategy of [Burchardi et al. \(2019\)](#). Here we obtain negative and insignificant coefficients. As a result, we conclude that the choice of estimation approach is important for the contrasting results between our study and [Burchardi et al. \(2019\)](#).

We argue that Poisson pseudo-maximum likelihood (PPML) estimation is more appropriate in our setting. [Santos Silva and Tenreyro \(2006\)](#) demonstrate that PPML performs well across a variety of settings, accommodating heteroskedasticity and measurement error even when the proportion of zeros is very high, as in our data. [Santos Silva and Winkelmann \(2024\)](#) shows that PPML performs well even when the conditional expectation function is misspecified and [Fally \(2015\)](#) shows that PPML is the only estimator which satisfies the adding-up constraints of structural gravity. As pointed out by [Santos Silva and Tenreyro \(2015\)](#), the Heckman approach makes strong assumptions on the distribution of the error terms and on the homoskedasticity of those errors. Furthermore, using the same vector of variables in both first and second stage of the Heckman approach leads to identification by functional form ([Puhani 2000](#); [Lewbel 2019](#)).

Table A.1: Estimator Robustness: ζ_1 and ζ_2 Estimated in Levels

	Dep. var.: Rel. expenditure share on goods from o (\tilde{X}_{oh})					
	(1)	(2)	(3)	(4)	(5)	(6)
Immigrant from anywhere	.00045*** (.000049)	.00036*** (.000042)	.00041*** (.00005)	.00031*** (.000043)	.00041*** (.000053)	.00032*** (.000046)
Immigrant from o		.0083*** (.0018)		.0083*** (.0018)		.0085*** (.002)
N	1480171	1480171	1480171	1480171	1306382	1306382
Origin-County FE	✓	✓	✓	✓		
HH Controls			✓	✓	✓	✓
Origin-County-Income FE					✓	✓
Sample Weights	✓	✓	✓	✓	✓	✓
Estimator	OLS	OLS	OLS	OLS	OLS	OLS

Notes: The table presents regression results at the household-country level. Observations weighted using NielsenIQ household weights. Standard errors clustered two-ways at the household and origin-by-destination levels. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.2: Estimator Robustness: ζ_1 and ζ_2 as Linear Probability Estimates

	Dep. var.: $1\{X_{oh} \neq 0\}$					
	(1)	(2)	(3)	(4)	(5)	(6)
Immigrant from anywhere	.021*** (.0034)	.019*** (.0034)	.014*** (.0032)	.012*** (.0032)	.015*** (.0034)	.013*** (.0033)
Immigrant from o		.17*** (.019)		.17*** (.019)		.17*** (.021)
N	1480171	1480171	1480171	1480171	1306382	1306382
Origin-County FE	✓	✓	✓	✓		
HH Controls			✓	✓	✓	✓
Origin-County-Income FE					✓	✓
Sample Weights	✓	✓	✓	✓	✓	✓
Estimator	OLS	OLS	OLS	OLS	OLS	OLS

Notes: The table presents regression results at the household-country level. Observations weighted using NielsenIQ household weights. Standard errors clustered two-ways at the household and origin-by-destination levels. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.3: Estimates of Import Demand Preferences with Adjusted Weights

	Dep. var.: Rel. expenditure share on goods from o (\tilde{X}_{oh})					
	(1)	(2)	(3)	(4)	(5)	(6)
Immigrant from anywhere	0.32*** (0.031)	0.25*** (0.029)	0.28*** (0.034)	0.21*** (0.033)	0.28*** (0.035)	0.22*** (0.034)
Immigrant from o		0.59*** (0.080)		0.63*** (0.079)		0.62*** (0.084)
N	911,385	911,385	911,385	911,385	619,741	619,741
Origin-County FE	✓	✓	✓	✓		
HH Controls			✓	✓	✓	✓
Origin-County-Income FE					✓	✓
Sample Weights	✓	✓	✓	✓	✓	✓

Notes: The table presents regression results at the household-country level. Observations weighted using NielsenIQ household weights. Standard errors clustered two-ways at the household and origin-by-destination levels. Household controls are dummy variables indicating: race, ethnicity, household size bins, bins for the number and age of children in the household, highest education among household members, household income quintiles, marital status and age groups. The estimation sample size is less than our total sample, and falls across columns, due to fixed effects causing separation in the sense of [Correia et al. \(2019\)](#). *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.4: Household Gravity Equation with Adjusted Weights

	\tilde{X}_{oc}		\tilde{N}_{oc}	
	(1)	(2)	(3)	(4)
Immigrants/Pop. 2010	1.07*** (0.20)	1.03** (0.48)	0.86*** (0.12)	0.85* (0.49)
First-stage residuals		0.041 (0.54)		0.012 (0.54)
N	167,860	167,860	167,860	167,860
Country FE	✓	✓	✓	✓
Distance & latitude difference	✓	✓	✓	✓
1st-stage F-statistic		103.4		103.4

Notes: The table presents the results of estimating equation (13) using PPML, reweighted to match the immigrant population by origin country in the Census as described in Section A.3. In columns 1 and 2, the dependent variable is the preference-adjusted origin-specific import expenditure relative to expenditure on US goods, averaged at the county level: \tilde{X}_{oc} . In columns 3 and 4, expenditure is replaced by count of different varieties (barcodes) purchased per household, also averaged at the county level: \tilde{N}_{oc} . The first-stage residual term is taken from a first-stage regression of all the instruments on the immigrant-population share. Observations are weighted using NielsenIQ household weights. Standard errors are computed from 1,000 bootstrapped repetitions. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.5: Estimates of Import Demand Preferences with Additional Variables

	Dep. var.: Rel. expenditure share on goods from o (\tilde{X}_{oh})		
	(1)	(2)	(3)
Immigrant from anywhere	0.23*** (0.028)	0.22*** (0.032)	0.21*** (0.035)
Immigrant from o	0.54*** (0.067)	0.51*** (0.080)	0.54*** (0.078)
Hispanic ethnicity	-0.033 (0.055)	-0.030 (0.055)	-0.048 (0.053)
=1 if Hispanic and o in Latin America	0.14** (0.064)	0.13** (0.065)	0.15** (0.062)
Asian ethnicity	0.0046 (0.037)	0.0100 (0.037)	0.013 (0.037)
=1 if Asian and o in Asia	0.41*** (0.089)	0.39*** (0.087)	0.40*** (0.086)
=1 if immigrant from continent of o		0.043 (0.062)	0.017 (0.056)
=1 if common official or primary language			0.019 (0.054)
=1 if ever in colonial/dependency rel.			0.18* (0.098)
=1 if currently in colonial/dependency rel.			0.17 (0.79)
N	911,385	911,385	907,183
HH Controls	✓	✓	✓
Origin-County FE	✓	✓	✓
Sample Weights	✓	✓	✓

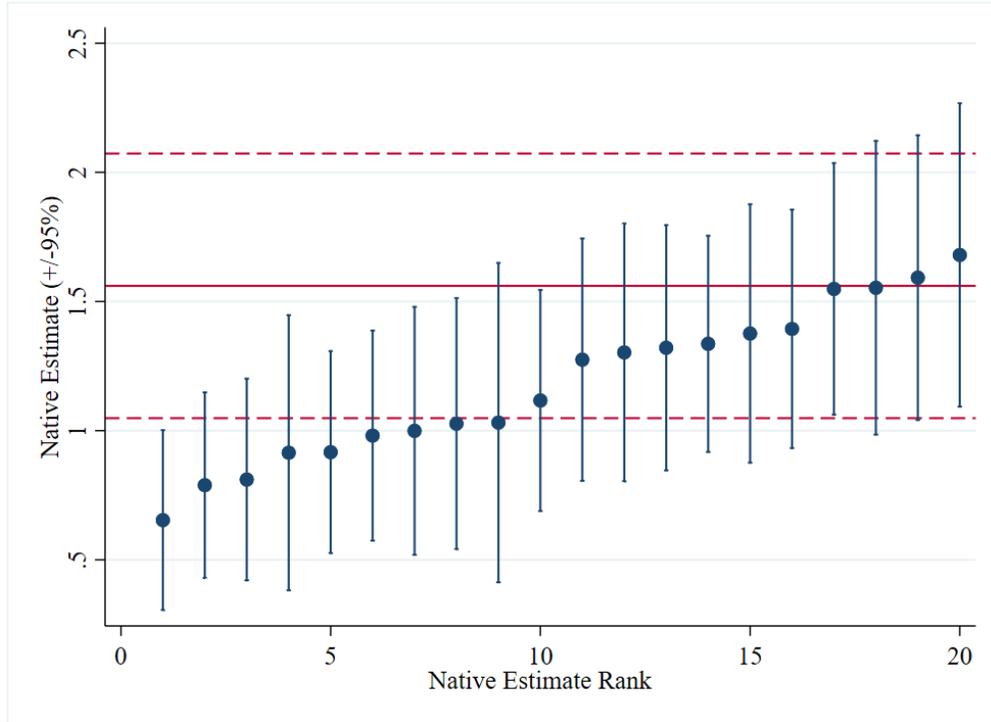
Notes: The table presents regression results at the household-country level. Observations weighted using NielsenIQ household weights. Standard errors clustered two-ways at the household and origin-by-destination levels. Household controls are dummy variables indicating: race, ethnicity, household size bins, bins for the number and age of children in the household, highest education among household members, household income quintiles, marital status and age groups. The estimation sample size is less than our total sample, and falls across columns, due to fixed effects causing separation in the sense of [Correia et al. \(2019\)](#). *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.6: Estimates of σ by Nativity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$\Delta \ln X$	$\Delta \ln p$	$\Delta \ln X$	$\Delta \ln X$				
$\Delta \ln p$	0.412*** (0.011)		-2.266*** (0.061)	-2.299*** (0.055)	-1.913*** (0.092)	-1.981*** (0.071)	-1.907*** (0.097)	-0.560** (0.261)
Price IV		0.396*** (0.008)						
Market FE	✓	✓	✓		✓		✓	✓
Origin-Market FE				✓		✓		
Balanced	✓	✓	✓	✓	✓	✓	✓	✓
Sample	All	All	All	All	Survey	Survey	Native	Immig.
First-Stage		✓						
Estimator	OLS	OLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
K-P F-Statistic			2738.70	3247.48	1860.27	2746.57	1740.51	337.73
R2	0.01	0.04						
Observations	2000840	2000834	2000834	1993558	390521	386781	348123	9112

Notes: The table presents regression results at the barcode-CBSA-bimonthly level. Column (1) and (2) provide, respectively, the OLS estimates and the first-stage estimate associated with the price instrument. Columns (3)-(4) provide 2SLS estimates for the entire sample of NielsenIQ households, columns (5)-(6) provide 2SLS estimates for the subsample of NielsenIQ households for which we have nativity information, and columns (7) and (8) provide 2SLS estimates for, respectively, native and immigrant households. All models include CBSA-period fixed effects, while Columns (4) and (6) include country-CBSA-period fixed effects. Standard errors clustered at the CBSA-period level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Figure A.1: Comparing Estimates of η for Immigrants and Natives



Notes: The figure plots estimates of η for twenty subgroups of native households, each identical in size to the total number of immigrant households present in our sample. These estimates are ranked from left to right, with 95% confidence intervals provided. The red solid line denotes the immigrant estimate of η , with the dashed red lines denoting 95% confidence intervals.

Table A.7: Estimates of the Armington Elasticity by Nativity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$\Delta \ln X$	$\Delta \ln X$	$\Delta \ln X$	$\Delta \ln X$	$\Delta \ln X$	$\Delta \ln X$	$\Delta \ln X$	$\Delta \ln X$
$\Delta \text{ Log Price}$	-3.069*** (0.534)	-3.792*** (0.977)	-4.612*** (1.302)	-9.591 (6.881)	-3.069*** (0.272)	-2.629*** (0.374)	-2.777*** (0.393)	-2.906*** (0.848)
Origin-City FE	✓	✓	✓	✓	✓	✓	✓	✓
Balanced	✓	✓	✓	✓	✓	✓	✓	✓
Sample	All	Survey	Native	Immig.	All	Survey	Native	Immig.
Estimator	2SLS	2SLS	2SLS	2SLS	OLS	OLS	OLS	OLS
Control Function					✓	✓	✓	✓
K-P F-Statistic	51.84	25.78	18.99	2.11				
Observations	33594	22552	21093	4658	33594	22552	21093	4658

Notes: This table presents regression results at the country-CBSA-bimonthly level. Columns (1)-(4) provide 2SLS estimates associated with exchange rates as instruments. In order, these estimates represent Armington elasticities estimated for the entire NielsenIQ sample (column 1), the sub-sample surveyed in 2008 about nativity (column 2), natives (column 3), and immigrants (column 4). In all cases, only country-CBSA pairs with non-zero expenditure in all 18 time periods are included for each subsample of households. Prices are calculated using purchase records for the entire NielsenIQ Homescanner dataset, with the same prices applied to all subsample. Columns (5)-(8) provide the exact same ordering of estimates but using a control function approach, in which the first-stage is estimated using the entire sample, with the residuals from this model included in the second-stage as an additional regressor. In this case, all first-stage K-P F-Statistics are identical to column 1, with column 1 and column 5 estimates of η identical by construction. All models include CBSA-by-period fixed effects. Standard errors clustered at the CBSA-period level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.8: First-stage Estimates

Dependent variable: Immigrants/Pop. 2010			
	(1)	(2)	(3)
$L_{o,-r}^{1880} \times \frac{L_{-c(o),d}^{1880}}{L_{-c(o)}^{1880}}$	0.000049** (0.000021)	-0.00046** (0.00021)	-0.00048** (0.00022)
$L_{o,-r}^{1900} \times \frac{L_{-c(o),d}^{1900}}{L_{-c(o)}^{1900}}$	-0.00012* (0.000068)	-0.0024*** (0.00092)	-0.0025*** (0.00095)
$L_{o,-r}^{1910} \times \frac{L_{-c(o),d}^{1910}}{L_{-c(o)}^{1910}}$	0.00027*** (0.000056)	-0.0022** (0.00097)	-0.0023** (0.0010)
$L_{o,-r}^{1920} \times \frac{L_{-c(o),d}^{1920}}{L_{-c(o)}^{1920}}$	0.0016*** (0.00034)	0.00057 (0.0011)	0.00050 (0.0012)
$L_{o,-r}^{1930} \times \frac{L_{-c(o),d}^{1930}}{L_{-c(o)}^{1930}}$	0.0010*** (0.00024)	0.00039 (0.0011)	0.00037 (0.0011)
$L_{o,-r}^{1970} \times \frac{L_{-c(o),d}^{1970}}{L_{-c(o)}^{1970}}$	0.00061*** (0.000068)	0.00027 (0.00058)	0.00026 (0.00061)
$L_{o,-r}^{1980} \times \frac{L_{-c(o),d}^{1980}}{L_{-c(o)}^{1980}}$	0.0025*** (0.00019)	0.0022*** (0.00072)	0.0022*** (0.00075)
$L_{o,-r}^{1990} \times \frac{L_{-c(o),d}^{1990}}{L_{-c(o)}^{1990}}$	0.0018*** (0.00020)	0.0011 (0.00097)	0.0011 (0.0010)
$L_{o,-r}^{2000} \times \frac{L_{-c(o),d}^{2000}}{L_{-c(o)}^{2000}}$	0.0011*** (0.00013)	0.0012*** (0.00036)	0.0013*** (0.00038)
N	167,860	2,280,154	2,217,377
Country FE	✓		
Barcode FE		✓	✓
County FE		✓	✓
Distance & latitude difference	✓	✓	✓
Household controls	✓		
Principal components	✓	✓	✓
F-statistic	110.0	18.3	17.4
Sample	All counties		UPC in >10 counties

Notes: Column 1 shows regression results at the county-origin level with observations weighted using aggregated NielsenIQ household weights and standard errors clustered at the county level. Columns 2 and 3 show regression results at the county-barcode level with standard errors clustered at the county level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.9: Gravity Estimates with Ancestry

	\tilde{X}_{oc}			
	(1)	(2)	(3)	(4)
Ancestry/Pop. 2010	0.62*** (0.11)	0.46 (0.47)		
Log Ancestry 2010			0.035*** (0.012)	0.030*** (0.011)
First-stage residuals		0.18 (0.49)		0.0067* (0.004)
N	163,500	163,500	163,500	163,500
Country FE	✓	✓	✓	✓
Distance & latitude difference	✓	✓	✓	✓
1st-stage F-statistic		56.3		32.4

Notes: The table presents regression results at the county-origin level using pseudo-Poisson maximum likelihood estimation. The first-stage residual term is taken from a first-stage regression of all the instruments on the ancestry variable. Observations are weighted using NielsenIQ household weights. Standard errors are computed from 1,000 bootstrapped repetitions. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.10: State-level Gravity Estimates

	Exp. share on goods from o relative to US			
	PPML + control fct		Heckman correction	
	(1)	(2)	(3)	(4)
Log Ancestry 2010	0.05*** (0.008)		-0.04 (0.043)	
Immigrants/Pop. 2010		2.40*** (0.378)		-1.46 (2.151)
N	3,773	3,773	3,083	3,083
State FE	✓	✓	✓	✓

Notes: The table presents regression results at the state-origin level. Standard errors are computed over 1,000 bootstrapped household samples. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

B Theory Appendix

B.1 Deriving Heterogeneous Firms Model Equations

Deriving equation (10). Taking the ratio of the household's first-order condition for two varieties ω_1 from country o and ω_2 from country o' , we obtain

$$\left(\frac{q_{o'h}(\omega_2)}{q_{oh}(\omega_1)}\right)^{-1/\sigma} \left(\frac{z_{o'h}}{z_{oh}}\right)^{1/\sigma} = \frac{p_{o',c(h)}(\omega_2)}{p_{o,c(h)}(\omega_1)}.$$

Define

$$P_h \equiv \left(\sum_{o \in \mathcal{O}} z_{oh} \int_{\omega \in \Omega_{o,c(h)}} p_{o,c(h)}(\omega)^{1-\sigma} d\omega \right)^{\frac{1}{1-\sigma}}, \quad (\text{B.1})$$

as the price index faced by household h for differentiated goods. Assuming the household budget is equal to X_h , we then obtain

$$(1 - \mu_0)X_h = z_{oh}^{-1} q_{oh}(\omega) p_{o,c(h)}(\omega)^\sigma P_h^{1-\sigma}. \quad (\text{B.2})$$

We rearrange to get quantity and expenditure for a variety associated with productivity φ as

$$q_{oh}(\varphi) = (1 - \mu_0)X_h z_{oh} p_{o,c(h)}(\varphi)^{-\sigma} P_h^{\sigma-1} \quad (\text{B.3})$$

$$x_{oh}(\varphi) = (1 - \mu_0)X_h z_{oh} (p_{o,c(h)}(\varphi)/P_h)^{1-\sigma}. \quad (\text{B.4})$$

From the firm's profit maximization problem, the price equation is

$$p_{o,c(h)}(\varphi) = \frac{\sigma}{\sigma - 1} \frac{w_o}{\varphi} \tau_{o,c(h)}. \quad (\text{B.5})$$

Substituting equation (B.5) into equation (B.4), summing across all households in $c(h)$, and defining $\lambda_1 \equiv (1 - \mu_0) \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}$, we obtain the following expression for county expenditure on imports from firm with productivity φ in o :

$$x_{oc}(\varphi) = \lambda_1 (w_o \tau_{oc})^{1-\sigma} \varphi^{\sigma-1} \left(\sum_{h' \in \Lambda_c} z_{oh'} X_{h'} P_{h'}^{\sigma-1} \right). \quad (\text{B.6})$$

Next, we derive variable profits earned by a firm with productivity φ selling to market c from origin o :

$$\begin{aligned}\pi_{o,c}(\varphi) &\equiv \left(p_{o,c}(\varphi) - \frac{w_o}{\varphi} \tau_{o,c} \right) \sum_{h' \in c} q_{oh}(\varphi) \\ &= (1 - \mu) \left(\frac{w_o}{\varphi} \tau_{o,c} \right)^{1-\sigma} \frac{1}{\sigma} \left(\frac{\sigma}{\sigma - 1} \right)^{1-\sigma} \sum_{h' \in \Lambda_c} z_{oh'} X_{h'} P_{h'}^{\sigma-1} \\ &= \frac{1}{\sigma} x_{oc}(\varphi).\end{aligned}$$

A firm with productivity φ only exports from o to c if it is profitable, i.e., if variable profits are at least as much as the fixed cost of exporting:

$$\pi_{oc}(\varphi) \geq f_{oc}. \quad (\text{B.7})$$

For a firm at the cutoff productivity, (B.7) holds with equality, resulting in the following equation for φ_{oc}^* , where $\lambda_2 \equiv \frac{\sigma}{\sigma-1} \left(\frac{\sigma}{1-\mu_0} \right)^{\frac{1}{\sigma-1}}$:

$$\varphi_{oc}^* = \lambda_2 w_o \tau_{oc} \left(\frac{f_{oc}}{\sum_{h' \in \Lambda_c} z_{oh'} X_{h'} P_{h'}^{\sigma-1}} \right)^{\frac{1}{\sigma-1}}. \quad (\text{B.8})$$

Returning to equation (B.1) and replacing varieties ω with productivity φ (since firms with identical productivity charge identical prices), we get:

$$P_h = \left(\sum_{o \in \mathcal{O}} z_{oh} \int_0^{+\infty} p_{o,c(h)}(\varphi)^{1-\sigma} M_{o,c(h)} g_{o,c(h)}(\varphi) d\varphi \right)^{\frac{1}{1-\sigma}},$$

where $M_{o,c(h)}$ is the measure of firms exporting from o to $c(h)$ and $g_{o,c(h)}(\varphi)$ is the (equilibrium) density of firms from o with productivity φ that export to $c(h)$.

Plugging in the equilibrium price, equation (B.5), we have

$$P_h = \frac{\sigma}{\sigma - 1} \left(\sum_{o \in \mathcal{O}} z_{oh} \left(w_o \tau_{o,c(h)} \right)^{1-\sigma} M_{o,c(h)} \int_0^{+\infty} \varphi^{\sigma-1} g_{o,c(h)}(\varphi) d\varphi \right)^{\frac{1}{1-\sigma}}. \quad (\text{B.9})$$

Turning to the gravity equation, we integrate over equation (B.4) to obtain

$$X_{oh} = \int_{\omega \in \Omega_{o,c(h)}} x_{oh}(\omega) d\omega = (1 - \mu_0) z_{oh} X_h P_h^{\sigma-1} \int_{\omega \in \Omega_{o,c(h)}} p_{o,c(h)}(\omega)^{1-\sigma} d\omega.$$

Given the equilibrium price expression in (B.5), we can substitute the last term with

$$\begin{aligned} \int_{\omega \in \Omega_{o,c(h)}} p_{o,c(h)}(\omega)^{1-\sigma} d\omega &= \left(\frac{\sigma}{\sigma-1} w_o \tau_{o,c(h)} \right)^{1-\sigma} M_{o,c(h)} \int_0^\infty \varphi^{\sigma-1} g_{o,c(h)}(\varphi) d\varphi \\ &= \left(\frac{\sigma}{\sigma-1} w_o \tau_{o,c(h)} \right)^{1-\sigma} M_o \int_{\varphi_{o,c(h)}^*}^\infty \varphi^{\sigma-1} g_{o,c(h)}(\varphi) d\varphi. \end{aligned}$$

Finally, we use the assumption that φ is Pareto distributed with shape parameter θ so that $g_{o,c(h)}(\varphi) = \theta/\varphi^{\theta+1}$ to obtain

$$X_{oh} = \lambda_1 z_{oh} X_h P_h^{\sigma-1} (w_o \tau_{o,c(h)})^{1-\sigma} M_o \frac{\theta}{\theta+1-\sigma} (\varphi_{o,c}^*)^{\sigma-\theta-1}. \quad (\text{B.10})$$

To obtain equation (10) from (B.10), we then:

- substitute equation (B.8) for $\varphi_{o,c}^*$.
- assume $M_o = \gamma Y_o$, where Y_o is the value of production in country o as in Section 4.1.
- define $S_c = \sum_{h' \in \Lambda_c} X_{h'} P_{h'}^{\sigma-1}$ and $z_{oc} = \sum_{h' \in \Lambda_c} z_{oh'} \frac{X_{h'} P_{h'}^{\sigma-1}}{S_c}$.
- define $\lambda \equiv \gamma(1 - \mu_0)^{\frac{\theta}{\sigma-1}} \sigma^{\frac{\sigma-\theta-1}{\sigma-1}} \left(\frac{\sigma}{\sigma-1} \right)^{-\theta} \frac{\theta}{\theta+1-\sigma}$.

Finally, county-level expenditure on goods from origin o is simply the summation over all household-level expenditure, and is given by the following:

$$X_{oc} = \lambda Y_o S_c (w_o \tau_{oc})^{-\theta} \left(\frac{f_{oc}}{S_c z_{oc}} \right)^{-\left(\frac{\theta}{\sigma-1}-1\right)} z_{oc}. \quad (\text{B.11})$$

B.2 Deriving County-Level Estimating Equation

This section shows how the estimating equation generated by dividing $\tilde{X}_{oh} = X_{oh}/X_{us,h}$ is equivalent to an equation aggregated to the county level, in which this aggregation is a

weighted sum across \tilde{X}_{oh} . That is, we define the weights κ_h such that

$$\tilde{X}_{o,c(h)} = \frac{\sum_{\Lambda_c} X_{oh}}{\sum_{\Lambda_c} X_{us,h}} = \sum_{\Lambda_c} \kappa_h \tilde{X}_{oh}.$$

We begin by constructing the denominator of $\tilde{X}_{o,c(h)}$, where we incorporate our normalizations $\exp(\alpha_{us}) = \tau_{us,c} = f_{us,c} = z_{us,h} = 1$ to generate:

$$\sum_h X_{us,h} = S_c^{\frac{\theta}{\sigma-1}-1} \sum_h X_h P_h^{\sigma-1} = S_c^{\frac{\theta}{\sigma-1}},$$

where we invoke the definition $S_c = \sum_h X_h P_h^{\sigma-1}$. Similarly, we construct the summation over household expenditure on origin o :

$$\sum_h X_{oh} = \exp(\alpha_o) \tau_{o,c(h)}^{-\theta} \left(\frac{f_{o,c(h)}}{S_{c(h)} z_{o,c(h)}} \right)^{-\left(\frac{\theta}{\sigma-1}-1\right)} \sum_h \left[X_h P_h^{\sigma-1} z_{oh} \right].$$

We therefore consider the weights κ_h which satisfy the following expression:

$$\begin{aligned} \exp(\alpha_o) \tau_{o,c(h)}^{-\theta} \left(\frac{f_{o,c(h)}}{z_{o,c(h)}} \right)^{-\left(\frac{\theta}{\sigma-1}-1\right)} S_{c,(h)}^{-1} \sum_h \left[X_h P_h^{\sigma-1} z_{oh} \right] \\ = \exp(\alpha_o) \tau_{o,c(h)}^{-\theta} \left(\frac{f_{o,c(h)}}{z_{o,c(h)}} \right)^{-\left(\frac{\theta}{\sigma-1}-1\right)} \sum_h \kappa_h z_{oh}. \end{aligned}$$

Note that this is true so long as:

$$\kappa_h = \frac{X_h P_h^{\sigma-1}}{S_{c,(h)}},$$

which implies that κ_h are real expenditure shares at the household level. We use the projection weights as an empirical proxy for κ_h when constructing the county average preferences z_{oc} , such that when we apply the same weights to the aggregation of expenditure across counties, we generate the following county-level estimating equation:

$$\tilde{X}_{o,c(h)} = \frac{\sum_{\Lambda_c} X_{oh}}{\sum_{\Lambda_c} X_{us,h}} = \sum_{\Lambda_c} \kappa_h \tilde{X}_{oh} = \exp(\alpha_o) \tau_{o,c(h)}^{-\theta} f_{o,c(h)}^{-\left(\frac{\theta}{\sigma-1}-1\right)} z_{o,c(h)}^{\frac{\theta}{\sigma-1}}.$$

B.3 Identification of Fixed Cost and Preference Diffusion Channels

In this section we fully differentiate equation (11) in order to arrive at two expressions relating the total import expenditure-immigrant elasticity and the extensive margin-immigrant elasticity to two parameters: β^f and β^z .^{58,59}

We begin by fully differentiating \tilde{X}_{oh} from equation (11) into terms associated with fixed costs f_{oc} , county-level preferences z_{oc} , and household-level preferences z_{oh} :

$$\begin{aligned}
 d\tilde{X}_{oh} = & \left[\int_{\varphi_{oc}^*}^{+\infty} \frac{\partial \tilde{x}_{oh}(\varphi)}{\partial f_{oc}} dG(\varphi) - \tilde{x}_{oh}(\varphi_{oc}^*) G'(\varphi_{oc}^*) \frac{\partial \varphi_{oc}^*}{\partial f_{oc}} \right] df_{oc} \\
 & + \left[\int_{\varphi_{oc}^*}^{+\infty} \frac{\partial \tilde{x}_{oh}(\varphi)}{\partial z_{oh}} dG(\varphi) - \tilde{x}_{oh}(\varphi_{oc}^*) G'(\varphi_{oc}^*) \frac{\partial \varphi_{oc}^*}{\partial z_{oh}} \right] dz_{oh} \\
 & + \left[\int_{\varphi_{oc}^*}^{+\infty} \frac{\partial \tilde{x}_{oh}(\varphi)}{\partial z_{oc}} dG(\varphi) - \tilde{x}_{oh}(\varphi_{oc}^*) G'(\varphi_{oc}^*) \frac{\partial \varphi_{oc}^*}{\partial z_{oc}} \right] dz_{oc},
 \end{aligned} \tag{B.12}$$

where we applied the Leibniz Rule to separate each term into both an intensive margin and extensive margin. Within each pair of brackets, the first term captures the intensive margin effect and the second term captures the extensive margin effect.

The expression for the intensive margin—the relative expenditure by household h on a given variety from origin o relative to its total expenditure on U.S. goods—is given by:

$$\tilde{x}_{oh}(\varphi) = (\tilde{w}_o \tau_{oc})^{1-\sigma} z_{oh} \varphi^{\sigma-1} \left(\int_{\varphi_{us,c}^*}^{+\infty} (\varphi')^{\sigma-1} dG(\varphi') \right)^{-1}, \tag{B.13}$$

whereas the productivity cut-off is equation (B.8).

It is clear from inspecting equations (B.13) and (B.8) that f_{oc} and z_{oc} only affect φ_{oc}^* , and therefore each household's extensive margin, whereas household-level preferences z_{oh} only affect the household-specific intensive margin of demand via \tilde{x}_{oh} . We can therefore apply the following restrictions: $\frac{\partial \tilde{x}_{oh}(\varphi)}{\partial f_{oc}} = 0$, $\frac{\partial \tilde{x}_{oh}(\varphi)}{\partial z_{oc}} = 0$, and $\frac{\partial \varphi_{oc}^*}{\partial z_{oh}} = 0$.

⁵⁸We assume throughout that $\beta^\tau = 0$, which implies that immigrants do not affect variables trade costs. This assumption derives from the results discussed in Table 4.

⁵⁹We refer to the extensive margin at the level of a variety.

Hence, we have an expression for the total semi-elasticity of import expenditure with respect to immigrants and an expression for the extensive margin semi-elasticity of import expenditure with respect to immigrants:

$$\frac{\partial \ln \tilde{X}_{oh}}{\partial I_{oc}} = \frac{\partial \ln \tilde{X}_{oh}}{\partial \ln f_{oc}} \frac{\partial \ln f_{oc}}{\partial I_{oc}} + \frac{\partial \ln \tilde{X}_{oh}}{\partial \ln z_{oh}} \frac{\partial \ln z_{oh}}{\partial I_{oc}} + \frac{\partial \ln \tilde{X}_{oh}}{\partial \ln z_{oc}} \frac{\partial \ln z_{oc}}{\partial I_{oc}} \quad (\text{B.14})$$

$$\frac{\partial \ln \tilde{N}_{oh}}{\partial I_{oc}} = \frac{\partial \ln \tilde{X}_{oh}}{\partial \ln f_{oc}} \frac{\partial \ln f_{oc}}{\partial I_{oc}} + \frac{\partial \ln \tilde{X}_{oh}}{\partial \ln z_{oc}} \frac{\partial \ln z_{oc}}{\partial I_{oc}}. \quad (\text{B.15})$$

Recall that when estimating β , we normalize \tilde{X}_{oh} and \tilde{N}_{oh} by $\mathcal{Z} = \bar{z}_{oh} \bar{z}_{oc}^{\frac{\theta}{\sigma-1}-1}$. That is, we normalize expenditure by the expenditure for that household which is predicted by exogenous preference terms at the household and county level. Recall further that $z_{oh} = e^{\beta z I_{oc}} \bar{z}_{oh}$ and $z_{oc} = e^{\beta z I_{oc}} \bar{z}_{oc}$. We can therefore explicitly derive our estimate of β and the extensive margin counterpart β^N as the following:

$$\begin{aligned} \beta &= \frac{\partial \ln \tilde{X}_{oh}}{\partial I_{oc}} - \frac{\partial \ln \mathcal{Z}_{oh}}{\partial I_{oc}} \\ &= \frac{\partial \ln \tilde{X}_{oh}}{\partial \ln f_{oc}} \frac{\partial \ln f_{oc}}{\partial I_{oc}} + \frac{\partial \ln \tilde{X}_{oh}}{\partial \ln z_{oh}} \frac{\partial \ln z_{oh}}{\partial I_{oc}} + \frac{\partial \ln \tilde{X}_{oh}}{\partial \ln z_{oc}} \frac{\partial \ln z_{oc}}{\partial I_{oc}} - \frac{\partial \ln \mathcal{Z}_{oh}}{\partial \ln \bar{z}_{oh}} \frac{\partial \ln \bar{z}_{oh}}{\partial I_{oc}} \\ &\quad - \frac{\partial \ln \mathcal{Z}_{oh}}{\partial \ln \bar{z}_{oc}} \frac{\partial \ln \bar{z}_{oc}}{\partial I_{oc}}, \end{aligned} \quad (\text{B.16})$$

$$\begin{aligned} \beta^N &= \frac{\partial \ln \tilde{N}_{oh}}{\partial I_{oc}} - \frac{\partial \ln \mathcal{Z}_{oh}}{\partial I_{oc}} \\ &= \frac{\partial \ln \tilde{X}_{oh}}{\partial \ln f_{oc}} \frac{\partial \ln f_{oc}}{\partial I_{oc}} + \frac{\partial \ln \tilde{X}_{oh}}{\partial \ln z_{oc}} \frac{\partial \ln z_{oc}}{\partial I_{oc}} - \frac{\partial \ln \mathcal{Z}_{oh}}{\partial \ln \bar{z}_{oh}} \frac{\partial \ln \bar{z}_{oh}}{\partial I_{oc}} - \frac{\partial \ln \mathcal{Z}_{oh}}{\partial \ln \bar{z}_{oc}} \frac{\partial \ln \bar{z}_{oc}}{\partial I_{oc}}. \end{aligned} \quad (\text{B.17})$$

We evaluate expressions (B.16) and (B.17) using the definition of \tilde{X}_{oh} provided in equation (11) and the definition of \mathcal{Z}_{oh} provided in Section 4.2. Specifically, we reduce expressions (B.16) and (B.17) in three steps:

1. Fixed costs and the extensive margin:

$$\frac{\partial \ln \tilde{X}_{oh}}{\partial \ln f_{oc}} \frac{\partial \ln f_{oc}}{\partial I_{oc}} = \beta^f.$$

2. County-level preferences and the extensive margin:

$$\frac{\partial \ln \tilde{X}_{oh}}{\partial \ln z_{oc}} \frac{\partial \ln z_{oc}}{\partial I_{oc}} - \frac{\partial \ln \mathcal{Z}_{oh}}{\partial \ln \bar{z}_{oc}} \frac{\partial \ln \bar{z}_{oc}}{\partial I_{oc}} = \left(\frac{\theta - (\sigma - 1)}{\sigma - 1} \right) \beta^z.$$

3. Household-level preferences and the intensive margin:

$$\frac{\partial \ln \tilde{X}_{oh}}{\partial \ln z_{oh}} \frac{\partial \ln z_{oh}}{\partial I_{oc}} - \frac{\partial \ln \mathcal{Z}_{oh}}{\partial \ln \bar{z}_{oh}} \frac{\partial \ln \bar{z}_{oh}}{\partial I_{oc}} = \beta^z.$$

We then derive an expression for the total import expenditure semi-elasticity with respect to the immigrant population share and the extensive margin semi-elasticity of import expenditure with respect to the immigrant population share:

$$\beta = \frac{\partial \ln \tilde{X}_{oh}}{\partial I_{oc}} = \beta^f + \left(\frac{\theta}{\sigma - 1} \right) \beta^z, \quad (\text{B.18})$$

$$\beta^N = \frac{\partial \ln \tilde{N}_{oh}}{\partial I_{oc}} = \beta^f + \left(\frac{\theta}{\sigma - 1} - 1 \right) \beta^z. \quad (\text{B.19})$$

B.4 Deriving Counterfactual Objects

Following [Dekle et al. \(2007\)](#), we denote the proportional change in a variable x as $\hat{x} = x'/x$, where an apostrophe ($'$) denotes the counterfactual value.

Immigration shocks. We start with the counterfactuals from Section 5 relating to removing immigrant effects and removing immigrant expenditures. From equation (10), we obtain the proportional change in household-origin import expenditures:

$$\hat{X}_{oh} = \hat{P}_h^{\sigma-1} \hat{f}_{o,c(h)}^{-\left(\frac{\theta}{\sigma-1}-1\right)} \left(\hat{z}_{o,c(h)} \hat{S}_{c(h)} \right)^{\frac{\theta}{\sigma-1}-1} \hat{z}_{oh}, \quad (\text{B.20})$$

where changes in household imports by origin depend on the change in the household's price level \hat{P}_h , changes in fixed costs with the origin $\hat{f}_{o,c(h)}$, changes in average household-level preferences for the origin's products $\hat{z}_{o,c(h)}$, changes in total local expenditures $\hat{S}_{c(h)}$, and changes in the household's preferences for the origin's products \hat{z}_{oh} . When o is the United States, equation (B.20) reduces to

$$\hat{X}_{us,h} = \hat{P}_h^{\sigma-1} \hat{S}_{c(h)}^{\frac{\theta}{\sigma-1}-1}. \quad (\text{B.21})$$

Hence we use equation (B.20), equation (B.21), and the fact that $\hat{f}_{o,c(h)}^{-\left(\frac{\theta}{\sigma-1}-1\right)} = e^{-\hat{\beta}^f I_{o,c(h)}}$ to obtain our counterfactual ratio as a function of observable or calibrated values:

$$\frac{X'_{oh}}{X'_{us,h}} = \frac{X_{oh}}{X_{us,h}} \left(e^{-I_{o,c(h)} \hat{\beta}^f} \right) z_{o,c(h)}^{\left(\frac{\theta}{\sigma-1}-1\right)}. \quad (\text{B.22})$$

Summing across non-U.S. origins o and holding fixed total expenditures X_h , we compute the counterfactual imports from each origin o for each household h .

Lastly, it is simple to show that under CES preferences, the change in welfare is given by the change in the price index:

$$\hat{U}_h = \hat{P}_h^{\mu_0-1}. \quad (\text{B.23})$$

The change in the welfare-relevant price index is then:

$$\hat{P}_h^{\sigma-1} = \frac{1}{\frac{X_{us,h}}{X_h} \hat{S}_{c(h)}^{\frac{\theta}{\sigma-1}-1} + \sum_{o \neq us} \frac{X_{o,h}}{X_h} \hat{f}_{oc(h)}^{-\left(\frac{\theta}{\sigma-1}-1\right)} \left(\hat{z}_{oc(h)} \hat{S}_{c(h)} \right)^{\frac{\theta}{\sigma-1}-1}}$$

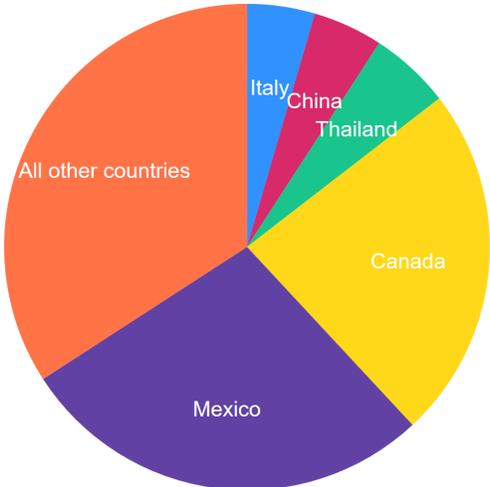
We further assume that immigrant and native households spend the same amount on grocery and personal care produces, which implies that

$$\hat{S}_{c(h)} = 1 - I_{c(h)}$$

where $I_{c(h)}$ is the share of the population who are immigrants in county $c(h)$.

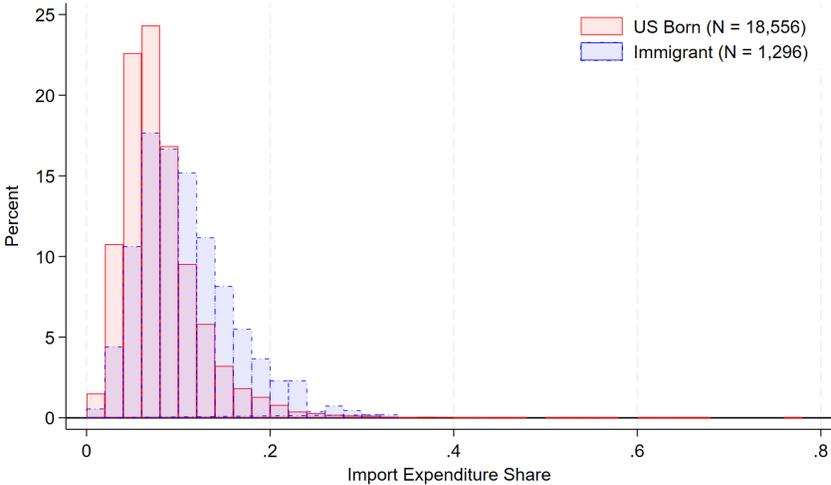
C Additional Tables and Charts

Figure C.1: Spending on Imports by Origin Country



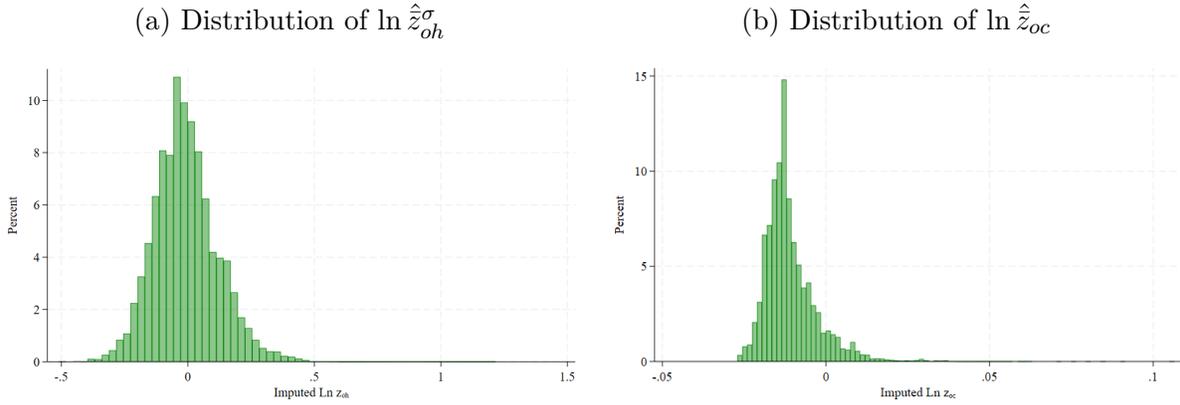
Notes: The figure shows the percent of expenditure on imports by country of origin. Data come from the NielsenIQ Household Panel 2014-2016.

Figure C.2: Distribution of Household-level Import Expenditure Share by Nativity



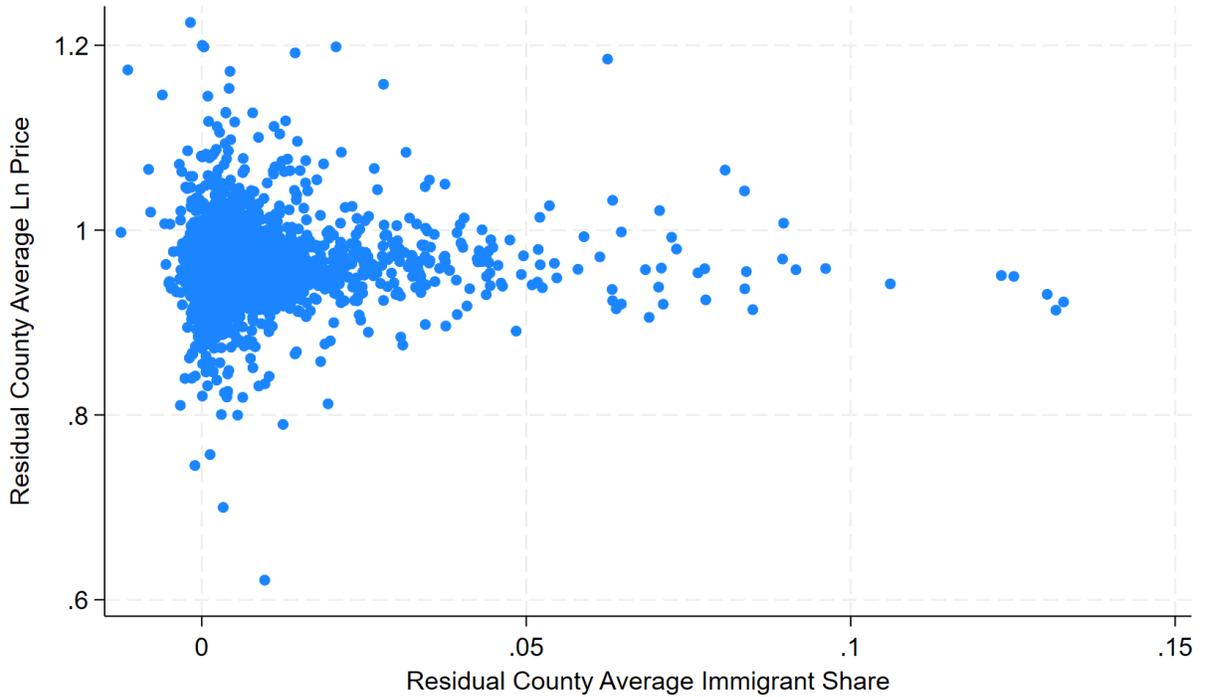
Notes: The figure shows the distribution of household’s expenditure on imported goods, split by U.S. born (in red) and foreign-born (in blue) households. Household nativity assigned as discussed in Section 2.1. Data come from the NielsenIQ Household Panel 2014-2016. We exclude households who spent less than \$1,000 over the 3 year sample period. Observations are weighted by the NielsenIQ projection factors.

Figure C.3: Distribution of Imputed Preference Terms



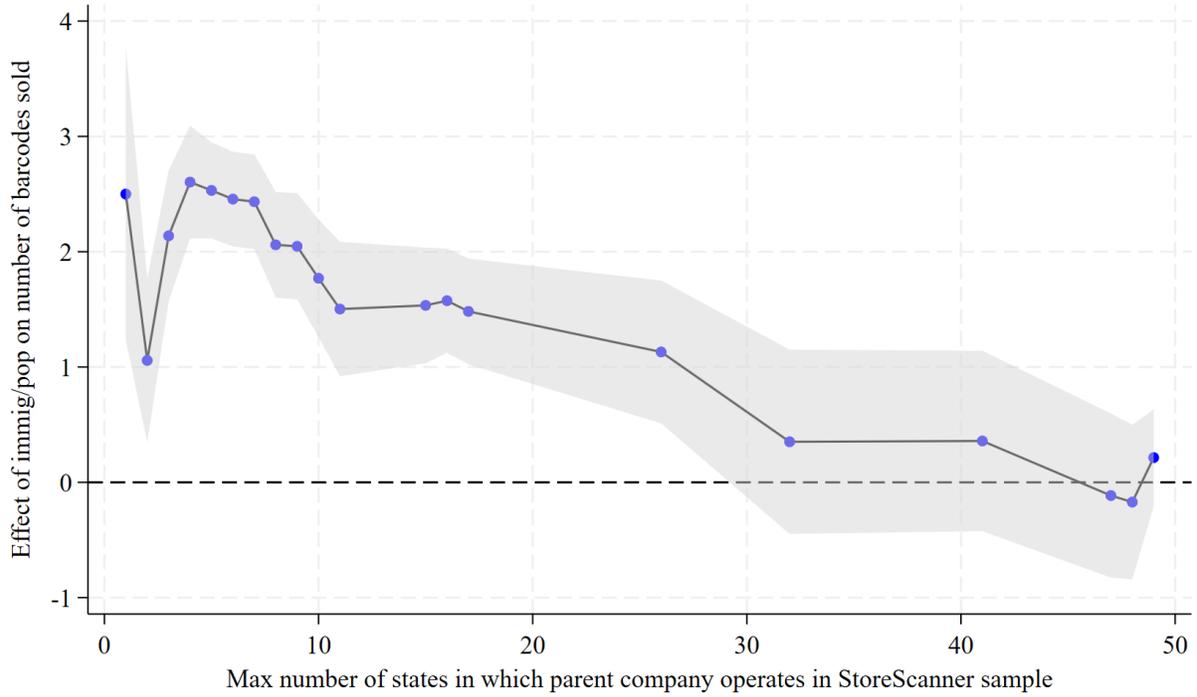
Notes: Figure (a) plots the distribution across NielsenIQ household-origin pairs of the log of $\hat{z}_{oh} = \exp(\hat{\delta}J_h + \hat{\zeta}_1\mathbf{1}[o(h) \neq US] + \hat{\zeta}_2\mathbf{1}[o(h) = o])$, where the terms $\hat{\delta}$, $\hat{\zeta}_1$, and $\hat{\zeta}_2$ are estimated from equation (5). Figure (b) plots the distribution across county-origin pairs of the log of $\hat{z}_{oc} = \sum_{h' \in \Lambda_c} \hat{z}_{oh'} \kappa_{h'}$, computed using data from the 2012-2017 American Community Survey.

Figure C.4: County Fixed Effects for Price and Immigrant Population Share



Notes: The figure plots average log price by county (y-axis) against average immigrant population share by county (x-axis). Each value is obtained by regressing the variable (log price or immigrant population share) on fixed effects for county and barcode, and taking the county fixed effect.

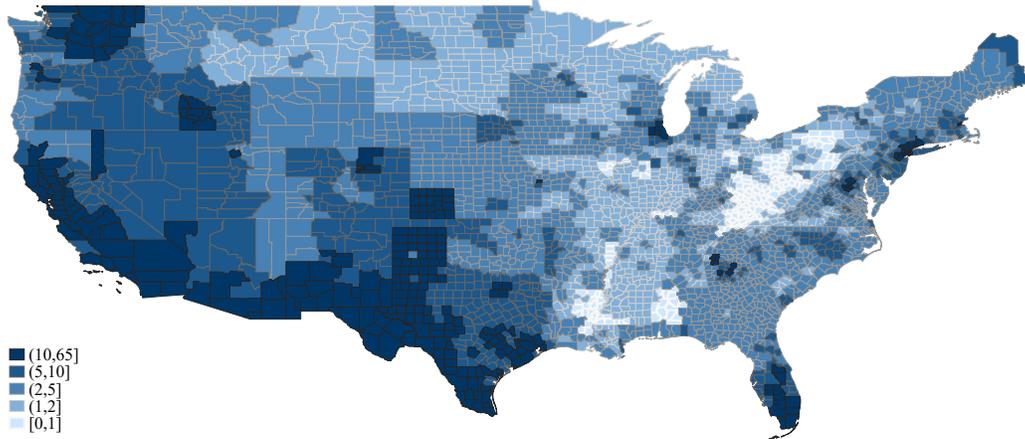
Figure C.5: Effect of Immigrants on Number of Barcodes Sold by Retail Chain Geographic Scope



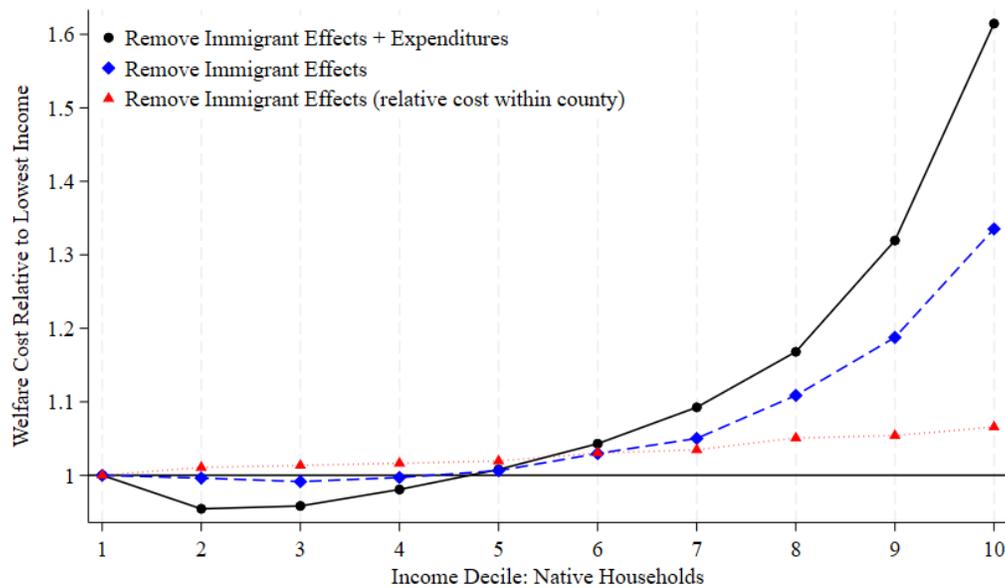
Notes: This chart plots the coefficient β^{SS} from equation (16) across different samples, except with the dependent variable being the ratio of foreign varieties from o relative to U.S. varieties sold at store s . The x-axis indicates the sample, which is determined by the maximum number of U.S. states a parent company retailer spans. 90% confidence intervals from analytic standard errors are plotted in gray.

Figure C.6: Distributional Effects of Immigrant-Induced Imports

(a) Percent Import Expenditure Attributable to Immigrant Effects



(b) Percent Change in Grocery Welfare by Income Decile



Notes: Panel (a) denotes share of county import expenditure attributable to immigrant effects, in percentages. Panel (b) depicts average native welfare costs at the income decile level of removing all immigrant expenditure (black solid line), removing immigrant effects only (blue dashed line), and removing immigrant effects while only comparing natives within the same county (red dotted line). All average welfare costs are normalized relative to the lowest income decile of natives.

Table C.1: Import Preference by Geographic Unit

	Dep. var.: Rel. expenditure share on goods from o			
	(1)	(2)	(3)	(4)
Immigrant from anywhere	0.22*** (0.028)	0.21*** (0.029)	0.23*** (0.030)	0.22*** (0.041)
Immigrant from o	0.60*** (0.065)	0.63*** (0.069)	0.57*** (0.087)	0.41*** (0.092)
N	911,385	1,002,492	684,142	398,706
Origin-County FE	✓			
Origin-3 Digit Zipcode FE		✓		
Origin-4 Digit Zipcode FE			✓	
Origin-5 Digit Zipcode FE				✓
HH Controls	✓	✓	✓	✓
Sample Weights	✓	✓	✓	✓

Notes: The table presents regression results at the household level controlling for fixed effects at different geographic granularity. Standard errors are clustered two ways at the household and county-origin pair levels. Household controls are income quintiles, household size, marital status, and household head age and gender. Sample size changes across columns because there must be at least two households per geographic unit. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table C.2: Household Characteristics and Import Preferences

Dep. var.: Rel. expenditure share on goods from o (\tilde{X}_{oh})	(1)		(2)	
Immigrant from anywhere	0.22***	(0.028)	0.23***	(0.030)
Immigrant from o	0.60***	(0.065)	0.61***	(0.077)
2nd poorest quintile (relative to poorest)	0.015	(0.027)		
median quintile (relative to poorest)	0.044	(0.027)		
2nd richest quintile (relative to poorest)	0.075***	(0.028)		
richest quintile (relative to poorest)	0.082***	(0.030)		
HH size: 2	-0.098***	(0.030)	-0.11***	(0.030)
HH size: 3	-0.14***	(0.035)	-0.17***	(0.035)
HH size: 4	-0.23***	(0.041)	-0.27***	(0.044)
HH size: >4	-0.22***	(0.082)	-0.29***	(0.090)
Children: 6-12 y.o.	-0.080	(0.088)	-0.087	(0.087)
Children: 13-17 y.o.	-0.074	(0.089)	-0.11	(0.087)
Children: <6 + 6-12	-0.094	(0.10)	-0.14	(0.11)
Children: <6 + 13-17	-0.084	(0.16)	-0.12	(0.15)
Children: 6-12 + 13-17	-0.056	(0.095)	-0.059	(0.094)
Children: All Age Groups	-0.26**	(0.13)	-0.28**	(0.14)
No Children	-0.056	(0.084)	-0.090	(0.081)
Some College	0.078***	(0.022)	0.065**	(0.026)
College Degree	0.13***	(0.022)	0.12***	(0.025)
Postgraduate Degree	0.24***	(0.024)	0.23***	(0.028)
Widowed	-0.027	(0.035)	-0.023	(0.039)
Divorced/Separated	-0.036	(0.032)	-0.058	(0.036)
Single	-0.062*	(0.034)	-0.069**	(0.034)
Black	0.052**	(0.024)	0.059**	(0.027)
Asian	0.086**	(0.038)	0.092**	(0.040)
Other	0.090**	(0.039)	0.087**	(0.044)
Hispanic	0.046	(0.035)	0.064	(0.039)
Age	-0.018	(0.031)	-0.033	(0.036)
Age ² /100	0.023	(0.052)	0.048	(0.059)
Age ³ /10000	-0.0077	(0.028)	-0.021	(0.032)
N	911,385		619,741	
Origin-County FE	✓			
Origin-County-Income FE			✓	
Sample Weights	✓		✓	
Estimator	PPML		PPML	

Notes: The table presents regression results at the household-country level. Observations weighted using NielsenIQ household weights. Standard errors clustered two-ways at the household and origin-by-destination levels. The estimation sample size is less than our total sample, and falls across columns, due to fixed effects causing separation in the sense of [Correia et al. \(2019\)](#). *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table C.3: Estimates of Immigrant Heterogeneity in Import Demand Preferences

	Dep. var.: Rel. expenditure share on goods from o			
	(1)	(2)	(3)	(4)
Immigrant from anywhere	0.17*** (0.043)	0.22*** (0.028)	0.22*** (0.030)	0.22*** (0.028)
Immigrant from o	0.70*** (0.16)	0.73*** (0.14)	0.68*** (0.069)	0.86*** (0.18)
(Immigrant from o) \times (Some College)	-0.097 (0.22)			
(Immigrant from o) \times (College Degree)	-0.13 (0.18)			
(Immigrant from o) \times (Postgraduate Degree)	-0.15 (0.21)			
(Immigrant from o) \times (2nd poorest quintile)		-0.021 (0.22)		
(Immigrant from o) \times (median income quintile)		-0.034 (0.21)		
(Immigrant from o) \times (2nd richest quintile)		-0.24 (0.16)		
(Immigrant from o) \times (richest quintile)		-0.34* (0.17)		
(Immigrant from o) \times (Cultural Distance to US)			0.22*** (0.049)	
(Immigrant from o) \times (Years in US)				-0.0055* (0.0032)
N	911,385	911,385	654,250	911,385
Origin-County FE	✓	✓	✓	✓
HH Controls	✓	✓	✓	✓
Sample Weights	✓	✓	✓	✓

Notes: The table presents results from estimating equation (5) using PPML at the household-country level. Observations are weighted using NielsenIQ household weights. Standard errors are clustered two-ways at the household and origin-by-destination levels. Household controls are dummy variables indicating: race, ethnicity, household size bins, bins for the number and age of children in the household, highest education among household members, household income quintiles, marital status and age groups. The estimation sample size is less than our total sample due to fixed effects causing separation in the sense of [Correia et al. \(2019\)](#). *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table C.4: Import Preference for Products Purchased by Both Immigrants and Natives

	Dep. var.: Rel. expenditure share on goods from o					
	(1)	(2)	(3)	(4)	(5)	(6)
Immigrant from anywhere	0.31*** (0.026)	0.28*** (0.026)	0.28*** (0.027)	0.24*** (0.027)	0.29*** (0.028)	0.25*** (0.028)
Immigrant from o		0.52*** (0.066)		0.54*** (0.068)		0.54*** (0.079)
N	856,161	856,161	855,046	855,046	585,987	585,987
Origin-County FE	✓	✓	✓	✓		
HH Controls			✓	✓	✓	✓
Origin-County-Income FE					✓	✓
Sample Weights	✓	✓	✓	✓	✓	✓

Notes: The table presents regression results at the household level for expenditures exclusively on barcodes purchased by both immigrants and natives in the NielsenIQ HomeScanner data. Standard errors are clustered two ways at the household and county-origin pair levels. Household controls are income quintiles, household size, marital status, and household head age and gender. Sample size drops across columns because there must be at least two households per geographic unit. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table C.5: Effect of Immigrants on Local Import Expenditures: Nativity vs. Full Sample, Unadjusted

	Dependent variable: Exp. share on goods from o			
	(1)	(2)	(3)	(4)
Immigrants/Pop. 2010	1.97*** (0.38)	2.84*** (0.61)	1.96*** (0.29)	3.09*** (1.00)
First-stage residuals		-1.10 (0.71)		-1.42 (1.16)
N	167,860	167,860	224,694	224,694
Country FE	✓	✓	✓	✓
Distance & latitude difference	✓	✓	✓	✓
1st-stage F-statistic		103.4		133.0
Sample	Nativity	Nativity	All	All

Notes: The table presents estimation results at the household-country level. We estimate each specification using pseudo-Poisson maximum likelihood estimation. The dependent variable in each column is the value of expenditures in county c on goods imported from o , unadjusted for local preferences: X_{oc} . The first-stage residual term is taken from a first-stage regression of all the instruments on the immigrant-population share. Observations are weighted using NielsenIQ household weights. Standard errors are computed from 1,000 bootstrapped repetitions of both first and second stage. The ‘Nativity’ sample refers to the set of households that we observe both in 2014-2016 and in the 2008 Tell Me More About Your Survey. The ‘All’ sample refers to the unrestricted set of households observed in 2014-2016. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table C.6: Effect of Immigrants on Import Expenditures, StoreScanner vs. Non-StoreScanner Stores

	Adjusted import expenditure		Adjusted import variety count	
	(1)	(2)	(3)	(4)
PANEL A: StoreScanner Stores				
Immigrants/Pop. 2010	-1.47 (1.04)	-2.16 (2.27)	-0.63 (0.49)	-1.03 (2.01)
First-stage residuals		0.89 (2.65)		0.51 (2.25)
N	152,070	152,070	152,144	152,144
Country FE	✓	✓	✓	✓
Distance & latitude difference	✓	✓	✓	✓
1st-stage F-statistic		102.3		102.3
	Adjusted import expenditure		Adjusted import variety count	
PANEL B: non-StoreScanner Stores				
Immigrants/Pop. 2010	2.28*** (0.70)	3.73* (2.10)	1.31*** (0.17)	1.50*** (0.39)
First-stage residuals		-1.87 (1.92)		-0.24 (0.43)
N	165,680	165,680	165,680	165,680
Country FE	✓	✓	✓	✓
Distance & latitude difference	✓	✓	✓	✓
1st-stage F-statistic		103.2		103.2

Notes: The table presents regression results at the county-origin level. Panel A restricts purchases to those made in stores appearing in the StoreScanner dataset, while Panel B restricts purchases to those made outside of such stores. All purchase data used in the estimates come from the HomeScanner dataset. We estimate each specification using pseudo-Poisson maximum likelihood estimation. The first-stage residual term is taken from a first-stage regression of all the instruments on the immigrant-population share. Observations are weighted using NielsenIQ household weights. Standard errors are computed over 1,000 bootstrapped repetitions of the first stage and control function estimation. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table C.7: Effect of Immigrants on Store-level Import Expenditures, StoreScanner Data

	Import expenditure		Import variety count	
	(1)	(2)	(3)	(4)
Immigrants/Pop. 2010	2.22*** (0.062)	5.07*** (0.18)	0.38*** (0.018)	0.21*** (0.052)
First-stage residuals		-3.43*** (0.19)		0.19*** (0.054)
N	2,733,744	2,733,744	2,733,744	2,733,744
Country FE	✓	✓	✓	✓
Parent Company FE	✓	✓	✓	✓
Distance & latitude difference	✓	✓	✓	✓
1st-stage F-statistic		11.4		11.4

Notes: The table presents regression results at the store-country level. We estimate each specification using pseudo-Poisson maximum likelihood estimation. The first-stage residual term is taken from a first-stage regression of all the instruments on the immigrant-population share. Standard errors are computed over 1,000 bootstrapped repetitions of the first stage and control function estimation. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table C.8: Effect of Immigrants on Prices, StoreScanner Data

	Dependent variable: $\ln(p_{wc})$			
	(1)	(2)	(3)	(4)
Immigrants/Pop. 2010	0.0242* (0.0135)	-0.0613 (0.0457)	0.0247* (0.0135)	-0.0590 (0.0451)
N	49,963,866	49,963,866	49,179,881	49,179,881
Retailer X UPC FE	✓	✓	✓	✓
County FE	✓	✓	✓	✓
Distance & latitude difference	✓	✓	✓	✓
1st-stage F-statistic		10.3		10.3
Sample	All SS UPCs		Only joint HS/SS UPCs	

Notes: The table presents regression results at the UPC-store level. Retailer X UPC FE refers to fixed effects at the store parent company-by-barcode level. Columns 1 and 2 include the full sample of barcodes, while columns 3 and 4 restrict the sample to the set of barcodes which appear in the HomeScanner dataset. Standard errors are clustered two ways at the store parent company-by-barcode level and at the county-by-origin level. Observations are weighted by the value of expenditures on the barcode in the given store. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table C.9: Effect of Immigrants on Import Expenditures, Natives only

	Adjusted import expenditure		Adjusted import variety count	
	(1)	(2)	(3)	(4)
Immigrants/Pop. 2010	1.13*** (0.21)	0.89* (0.47)	0.83*** (0.12)	0.75* (0.43)
First-stage residuals		0.29 (0.51)		0.11 (0.47)
N	166,551	166,551	166,551	166,551
Country FE	✓	✓	✓	✓
Distance & latitude difference	✓	✓	✓	✓
1st-stage F-statistic		102.1		102.1

Notes: The table presents the results of estimating equation (13) using PPML based on the sample of natives only. In columns 1 and 2, the dependent variable is the preference-adjusted origin-specific import expenditure relative to expenditure on US goods, averaged at the county level: \tilde{X}_{oc} . In columns 3 and 4, expenditure is replaced by count of different varieties (barcodes) purchased per household, also averaged at the county level: \tilde{N}_{oc} . The first-stage residual term is taken from a first-stage regression of all the instruments on the immigrant-population share. Observations are weighted using NielsenIQ household weights. Standard errors are computed from 1,000 bootstrapped repetitions. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table C.10: Counterfactual Results Summary with $\beta_f = 0.6$

Counterfactual exercise: Removing ...	(1) Change (%) grocery import expenditure	(2) Change (%) grocery welfare natives	(3) Change (\$) welfare per native HH
... all immigrant channels	-7.8	-0.086	-6.4
... trade cost channel	-0.9	-0.029	-2.1
... market size channel	-1.9	-0.058	-4.4
... demand channel	-5.3	—	—
... all immig. channels + expenditure	-26.2	-7.9	-592

Notes: This table shows the change in outcomes under various counterfactual scenarios. The first counterfactual, with results shown in the first row, removes all channels through which immigrants affect grocery import expenditures—i.e. we set $\zeta_1 = \zeta_2 = \beta_f = 0$ and recalculate \bar{z}_{oc} —but keeps total household expenditure constant. The next three rows show the results when only the following parameters change one at a time: β_f (fixed trade cost); \bar{z}_{oc} (preference-driven market size); and ζ_1 and ζ_2 (demand). In the last row, we remove all immigrant channels and the grocery expenditures made by immigrants (counterfactual 2), equivalent to removing immigrants from the U.S. population. For column 3, we assume that each household spends \$7,500 on grocery and personal care products, which is close the value computed from the Consumer Expenditure Survey.

Online Appendix References

- BONADIO, BARTHÉLÉMY (forthcoming): “Migrants, trade and market access,” *Review of Economics and Statistics*.
- BREINLICH, HOLGER, DENNIS NOVY, AND JMC SANTOS SILVA (2024): “Trade, gravity, and agglomeration,” *Review of Economics and Statistics*, 106 (5), 1418–1426.
- BURCHARDI, KONRAD B, THOMAS CHANEY, AND TAREK A HASSAN (2019): “Migrants, ancestors, and foreign investments,” *Review of Economic Studies*, 86 (4), 1448–1486.
- CARD, DAVID (2001): “Immigrant inflows, native outflows, and the local labor market impacts of higher immigration,” *Journal of Labor Economics*, 19 (1), 22–64.
- CHOI, JAERIM, JAY HYUN, AND ZIHO PARK (2024): “Bound by ancestors: Immigration, credit frictions, and global supply chain formation,” *Journal of International Economics*, 147, 103855.
- CONTE, M., P. COTTERLAZ, AND T. MAYER (2022): “The CEPII Gravity database,” CEPII working paper n°2022-05, CEPII.
- DEKLE, ROBERT, JONATHAN EATON, AND SAMUEL KORTUM (2007): “Unbalanced trade,” *American Economic Review*, 97 (2), 351–355.
- FALLY, THIBAUT (2015): “Structural gravity and fixed effects,” *Journal of International Economics*, 97 (1), 76–85.
- FEENSTRA, ROBERT C, LUCA MACEDONI, AND MINGZHI XU (2023): “Large firms, consumer heterogeneity and the rising share of profits,” Tech. rep., National Bureau of Economic Research.
- LEWBEL, ARTHUR (2019): “The identification zoo: Meanings of identification in econometrics,” *Journal of Economic Literature*, 57 (4), 835–903.
- MCCULLY, BRETT A (2024): “Immigrants, legal status, and illegal trade,” *Journal of International Economics*, 152, 104016.
- PUHANI, PATRICK (2000): “The Heckman correction for sample selection and its critique,” *Journal of Economic Surveys*, 14 (1), 53–68.
- SANTOS SILVA, JOÃO MC AND SILVANA TENREYRO (2006): “The log of gravity,” *Review of Economics and Statistics*, 641–658.
- SANTOS SILVA, JOAO MC AND SILVANA TENREYRO (2015): “Trading partners and trading volumes: implementing the Helpman–Melitz–Rubinstein model empirically,” *Oxford Bulletin of Economics and Statistics*, 77 (1), 93–105.
- SANTOS SILVA, JOAO MC AND RAINER WINKELMANN (2024): “Misspecified exponential regressions: Estimation, interpretation, and average marginal effects,” *Review of Economics and Statistics*, 1–25.
- TERRY, STEPHEN J, THOMAS CHANEY, KONRAD B BURCHARDI, LISA TARQUINIO, AND TAREK A HASSAN (2022): “Immigration, Innovation, and Growth,” .
- TYAZHELNIKOV, VLADIMIR AND XINBEI ZHOU (2024): “PPML, Gravity, and Heterogeneous Trade Elasticities,” Tech. rep., Working Paper.