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

Brett A. McCully† ⓡ Torsten Jaccard‡ ⓡ Christoph Albert§

August 5, 2024

PRELIMINARY AND INCOMPLETE: DO NOT CITE OR CIRCULATE

Abstract

Do immigrants make goods from their origin country more accessible to their non-immigrant neighbours? To answer this question, we leverage novel U.S. scanner data on grocery purchases in which we observe the country of origin for both households and products. Removing immigrants' distinctive effects on local import consumption decreases aggregate import expenditure by 7%. Three quarters of this effect is driven by immigrants' import-biased preferences, while only a quarter is associated with welfare-enhancing reductions in trade costs. A naive application of standard welfare formulas substantially over-estimates the gains to native households of immigrant-induced imports. The benefits that do accrue to natives are disproportionately concentrated among high-income and urban households.

JEL Categories: F22, J31, J61, R11.

Keywords: Price index, product variety, distributional effects

---

*We are grateful to Munseob Lee for helpful discussions. We thank Ilaria Malisan for research assistance and audience members at various seminars and conferences. 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, email: brett.mccully@carloalberto.org

‡Vancouver School of Economics, email: tjaccard@mail.ubc.ca

§Collegio Carlo Alberto, email: christoph.albert@carloalberto.org

1 Introduction

More people than ever live in a different country than the one they were born in (United Nations 2024), putting immigration at the center of contentious political debates in receiving countries. The content of these debates tends to center on how immigrants affect the nominal wages of non-immigrant households.[1] Quantifying the welfare effects of immigrants on natives, however, also requires an understanding of how immigrants affect local prices and product availability. Immigrants may, for example, lower the cost of accessing goods from their origin country by reducing information frictions between local importers and exporters in their origin.

Estimating an aggregate elasticity of import expenditure with respect to the local immigrant population, however, is insufficient for evaluating the consumption welfare effects of immigrants on natives. This elasticity cannot distinguish between the *spillover effect* of immigrants on native households' import expenditure and the *preference composition effect* associated with specific preferences of immigrants for imported goods. Furthermore, understanding the distributional effects of immigrant-induced trade requires an understanding of how natives differ in their propensity to buy imports from specific origin countries, for which the required data is exceedingly rare.

This paper overcomes these challenges by introducing a novel dataset linking household records of grocery purchases to both product-specific origin countries and household characteristics, including country of birth. We leverage these unique data to empirically and quantitatively study the impact of immigrants in the U.S. on both import volumes and native household welfare. Removing immigrants' distinctive effects on local import consumption decreases aggregate grocery import expenditures by 7%, which is roughly equivalent to reducing the prevailing tariffs applied to grocery goods by half. We provide the first direct empirical evidence for a substantial preference composition effect: almost three quarters of our estimated import-immigrant elasticity is driven by immigrant preferences for both goods from their origin country and imports in general.

Our quantitative analysis therefore yields modest welfare benefits for native households

[1] For a recent review of the academic literature on immigrants' effect on wages, see Dustmann et al. (2016). Following the literature, we use the terms "natives" and "non-immigrants" interchangeably.

as immigrants reduce bilateral trade costs to a lesser extent than one might expect, given the estimated trade-creating effect of immigrant presence. A naive county-level application of the welfare formula derived in Arkolakis et al. (2012) over-estimates the welfare benefits to native households by a factor of three. The welfare gains that do accrue to natives are highly concentrated among high-income and urban households, which is equally attributable both to positive location sorting between high-income urban natives and immigrants, and to a positive elasticity of import demand with respect to income.[2]

We are the first to document that immigrants exhibit stronger a demand than natives both for goods from their origin country and from any other import origin. Immigrants thus exhibit trade-creating effects beyond their specific origin country. This augmented import demand for all origins also serves to increase the exposure of immigrant households to all trade shocks, immigrant-induced or otherwise. A counterfactual increase in ad-valorem tariff rates applied to all imported grocery goods decreases the welfare of immigrants in the U.S. by 25% more than that of natives.[3]

This paper's central contributions are therefore twofold. First, we highlight the role of heterogeneous preferences in shaping both the substantial trade effects of immigrants and the modest welfare effects of immigrant-induced trade for native households. Second, we are the first to characterize heterogeneity in the gains from immigrant-induced trade across observable native household characteristics, documenting a strong positive bias in these gains towards high-income and urban households.

The linchpin for our analysis is a novel dataset of household grocery purchases, in which we observe the country of origin of both households and the products they purchase. The data has three key components: (i) household-level scanner data for nearly 20,000 U.S. households, (ii) detailed country-of-origin data for over half a million grocery barcodes, and (iii) survey responses eliciting the country of birth of each household. We are the first to link product and household origin countries within household-level scanner data, which we

[2]We estimate this positive income elasticity of import preference directly. "Preference" refers to household-level demand shifters for imported varieties, conditional on price, which Hottman et al. (2016) define as "appeal" when measuring firm-level market share.

[3]Increasing all import tariffs by 10 ppt is a policy Donald Trump is campaigning on in 2024. The finding that immigrants are more exposed to such a policy echoes the derivation in Borusyak and Jaravel (2021) that, to a first-order, the distribution of consumer welfare effects associated with a trade shock is approximated by the distribution of import expenditure shares across consumers.

describe in detail in Section 2.

We estimate a structural gravity model at the household-origin level in Section 3, and in doing so separately quantify the effect of immigrants on import accessibility for all households as well as the effects of specific household characteristics, such as immigrant status, on import demand. Our estimating equation nests a wide range of standard microfoundations in the trade literature (Head and Mayer 2014) and we make use of the instrumental variables from Burchardi et al. (2019) to generate exogenous variation in origin-specific immigrant population shares across U.S. counties. We estimate a significant preference composition effect: immigrants spend 28% more on imports from all origins than their within-county, non-immigrant neighbours, and 134% more on imports specifically from their origin country. Spillovers are also significant, as a percentage point increase in the share of immigrants from a given origin increases the expenditure share of all households on goods from that origin by 1.15%.

Since immigrants may causally alter the preferences of natives—what we term *preference diffusion*—it is not clear whether a positive spillover effect necessarily implies a welfare gain for native households. We therefore develop a model of trade in Section 4 which allows us to separately identify the various channels by which immigrants may increase import expenditures in their county of residence. In particular, we extend the heterogeneous-firms model of trade developed in Melitz (2003) and Chaney (2008) to allow for four possible channels through which immigrants may induce import expenditure spillovers above and beyond the preference composition effect discussed above. First, we allow for the aforementioned preference diffusion channel, in which immigrants affect import preferences of their non-immigrant neighbors. Second, immigrants' preferences and the resulting higher import demand allows more foreign firms to cover the fixed cost of exporting, thereby increasing the number of imported varieties available to non-immigrants via a *market size channel*. Finally, we allow immigrants to affect both variable and/or fixed trade costs specific to their origin country and county of residence.

The structure of the heterogeneous-firms model employed in this paper allows us to fully leverage the available data and separately identify each channel using observable moments. That is, we estimate the elasticity of barcode-specific prices to immigrant population shares, the barcode count elasticity of import expenditure to immigrants, and the aggregate elasticity of import expenditure to immigrants. Collectively, these estimates pin down the variable cost reduction channel, fixed cost reduction channel, and preference diffusion channel

Our estimates highlight the gap between the trade-creating effects of immigrants and the welfare-relevant outcomes associated with this trade: channels that are welfare-neutral from the perspective of native households constitute almost three quarters of the aggregate immigrant-import elasticity. Specifically, we find no evidence that immigrants reduce variable trade costs.[4] The fixed cost reduction channel accounts for 80% of the aggregate *spillover effect* with the preference diffusion channel accounting for the remainder. Finally, we illustrate that the market size channel depends on the ratio of two parameters: the Pareto shape parameter from which firms draw productivity, θ, and the demand substitution elasticity, $\sigma - 1$. Given our calibration of this ratio to values taken from the literature, we find that the market size channel accounts for 11% of the aggregate immigrant-import elasticity.[5] We discuss model estimation results in Section 5.

In our primary counterfactual exercise we remove the preference and trade cost effects of immigrants, which decreases aggregate import expenditure by 7%.[6] In a second counterfactual, we highlight the aggregate market size benefits of immigrant presence by removing immigrants' expenditure in addition to the preference and trade cost effects. This reduces grocery imports by a quarter with an aggregate welfare reduction for natives of nearly 1%.[7]

The aggregate welfare impact masks substantial variation, both across space and across income groups. The annual grocery dollar-equivalent benefit of all immigrant effects, including expenditure, is more than five times higher in a high immigrant county such as Queens, NY, than in the average U.S. county, with much of the Appalachian region and the Midwest receiving practically zero consumer benefits from immigrants. Across the income distribu-

[4]This finding is consistent with the assumption made by Peri and Requena-Silvente (2010), but we are the first to provide a direct empirical test of this assumption.

[5]We calibrate $\theta/(\sigma - 1) = 1.25$, which is consistent with $\{\theta, \sigma\} = \{5, 5\}$ or $\{\theta, \sigma\} = \{10, 9\}$.

[6]While this import-immigrant elasticity of 0.07 is a counterfactual result, and therefore not directly comparable to the existing empirical literature, it is worth noting that our 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).

[7]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.

tion of native households, top earners obtain a 60% higher welfare gain than households at, or below, the median income.

In a final counterfactual, we simulate an increase in variable trade costs for all imported grocery goods. This trade shock decreases the welfare of immigrant households in the U.S. by 25% more than that of natives, with college-educated immigrants facing welfare costs that are nearly 50% greater than native households without a college degree. We discuss all counterfactual results in Section 6.

The findings in this paper provide the first empirical validation of the concern voiced in Felbermayr et al. (2015): caution is needed when interpreting immigrant-induced changes in import penetration as akin to changes in welfare for native households. Low-income, less-educated, and native-born U.S. households face the lowest consumer costs associated with policies which increase barriers to either immigration or imports. This paper therefore suggests a novel factor which may contribute to the well-documented lack of political support for increased immigration among this demographic.[8]

Related literature. This paper 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 introduce and quantify a novel margin by which immigrants benefit natives: by increasing local product variety.[9] 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.

Our study is the first to leverage household-level data on import expenditures, allowing us to both quantify the contribution of a comprehensive set of mechanisms and how the impact of immigrants varies across heterogeneous households. By contrast, a vast literature on the

[8]See, for example, Card et al. (2012). For a recent review of the literature on the determinants of voter preferences on immigration policy, see Alesina and Tabellini (2024).

[9]Two prior papers have explored this margin—Mazzolari and Neumark (2012) and Chen and Jacks (2012)—but lack the data and exogenous variation to causally identify potential mechanisms nor quantify the effect on native welfare. (Iranzo and Peri 2009; Di Giovanni et al. 2015; Aubry et al. 2016) study the aggregate variety effects of immigration but with a focus on positive sorting between location productivity and migrant productivity.

immigration-trade nexus uses data on region-to-region trade flows (Gould 1994; Head and Ries 1998; Combes et al. 2005; Peri and Requena-Silvente 2010; Parsons and Vézina 2018; Burchardi et al. 2019) and, more recently, firm-level data (Ottaviano et al. 2018; Cardoso and Ramanarayanan 2022; Ariu 2022).

The closest paper to ours is Bonadio (forthcoming), who allow for a broad set of mechanisms by which immigrants may affect trade, including home-biased preferences, reduced trade costs, and increased market size. Our study advances beyond Bonadio (forthcoming) in four key ways: (i) we observe and identify home-biased immigrant preferences directly, rather than recovering these preferences indirectly via aggregate data and strong functional form assumptions; (ii) we are the first to highlight the extent to which immigrants increase imports from all origin countries; (iii) we explore the distributional effects of immigration on consumption across heterogeneous native households; and (iv) we separately identify immigrants' effect on variable trade costs, fixed trade costs, and native preferences. In sum, this paper is able to provide a more rigorous study of the welfare consequences of immigrant-induced trade due to a uniquely detailed dataset combining household and product origin countries.

Our paper also contributes to the literature on spatial variation in the local cost of living (Diamond 2016; Handbury and Weinstein 2015), local product variety (Couture 2016; Hottman 2021), and variation in the cost of living between skill groups withinin cities (Su 2022; Diamond and Moretti 2021; Handbury 2021).[10] Finally, by estimating consumer heterogeneity in exposure to trade shocks—immigrant-induced or otherwise—this paper contributes to a growing literature studying the heterogeneous consumer outcomes associated with trade shocks.[11] This paper is the first to document the extent to which import expenditure is particularly concentrated in immigrant households, thus increasing the exposure of these households to trade shocks.

[10]Existing work by Lach (2007), Cortes (2008), and Zachariadis (2012) finds immigrant effects on region-level price indices but cannot quantify (i) the mechanisms that drive these estimated effects, and (ii) heterogeneity in these effects both across and within immigrants and native households.

[11]See Fajgelbaum and Khandelwal (2016); Bai and Stumpner (2019); Amiti et al. (2020); Hottman and Monarch (2021); Borusyak and Jaravel (2021); Faber and Fally (2022); Auer et al. (2023); Jaccard (2023).

2 Data and Stylized Facts

2.1 Expenditure on tradable nondurable products

We use two datasets to 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. Detailed household demographic information and county of residence are included along with barcode-level expenditure, price, date, and store for each purchase. We restrict our analysis to the years 2014-2016 and aggregate to a single cross-section at the household level.

For a subset of NielsenIQ households, we also observe their country of birth. In 2008, NielsenIQ distributed the "Tell Me More About You" survey, which included questions about respondents' birth place, and 19,700 of these households remain in the 2014-2016 sample used here.[12] Households may have mixed nativity, and we use the following allocation rules when assigning households to an origin country. When only one member of the household was born abroad and all others were born in the U.S., we consider the household to be born in the country of the immigrant member. When a household has more than one foreign-born member, we assign the household to the larger country of origin as measured by total population.

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 extracting and organizing information found on the labels of consumer packaged goods.[13] Label Insight uses a computer vision algorithm to extract the ingredients, branding, and any other text information 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 contain some statement equivalent to "Made in ...", the Label Insight algorithm incidentally recovers a country of origin for each barcode they collect.[14] Naturally, Label Insight can only cover a segment of total consump-

[12] See Bronnenberg et al. (2012) for more details regarding this survey.

[13] See Jaccard (2023) for a more detailed discussion of this dataset.

[14] The U.S. Customs and Border Protection require that the country-of-origin printed on the label corre-

tion and their coverage is best for food and beverages, alcohol, personal care products, and cosmetics.

We therefore make use of data on the origin country for over 600,000 barcodes in these grocery product categories. Given the universality of barcodes, these data can be directly merged with the household-level purchase records from NielsenIQ. Figure C.1 documents the distribution of production origin countries in the merged scanner data with barcode origins. As expected, Mexico and Canada constitute just over half of all import expenditure, with China, Germany, Chile, France, Italy, and Spain rounding out the top 9 product origins. All other origins constitute a quarter of import expenditure, with 78 origin countries recording non-zero expenditure. The average import expenditure share is approximately 8%.[15]

Household-Level Coverage of Import Expenditure: Our final merged dataset covers $764 billion USD of expenditure and is at the household-import origin level of aggregation. When compared to estimates from the BEA Consumer Expenditure Survey (CEX), the grocery categories studied in this paper account for approximately a third of all expenditure on tradeables, with this share increasing to almost half if one excludes passenger vehicles and energy products. Within groceries, the merged household-level expenditure data used here exhibit an average expenditure per household-year of $2,200 USD, which is around 60% of the predicted expenditure on groceries in the CEX.

2.2 Immigration Data

We use the decadal Censuses from 1880–1930 and 1970–2000, as well as the pooled 2006–2010 sample of the ACS to obtain population counts of immigrants by origin.[16] We then aggregate 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. and not citizens by birth. To compute decadal migrant inflows from origin o into destination county c between two census years $t-10$ and t, denoted I^t_{oc}, we count only those respondents who

sponds to the last country in which the good underwent a "substantial transformation".

[15] Throughout this paper we make use of the projection factor weights provided by NielsenIQ when presenting aggregated statistics. These weights are not shares, but rather a population projection based on the representativeness of each household, and sum to 120 million households, which generally matches the aggregate total for the U.S.

[16] The 1940, 1950 and 1960 samples cannot be used due to missing information of the year of immigration.

migrated to the U.S. between $t-10$ and t. Following Burchardi et al. (2019), in the first sample year the measure I_{oc}^{1880} includes all those that are either first-generation immigrants from o or second-generation immigrants whose parents were born in o. These inflow measures are used in the first stage of our instrumental variables strategy outlined in Section 3.4.

Our main explanatory variable is the share of the local population who was born in country 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 providing changing geographies across waves.[17]

2.3 Stylized Facts

The combined NielsenIQ datasets described above constitute 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 household origin and motivate the analyses to follow.

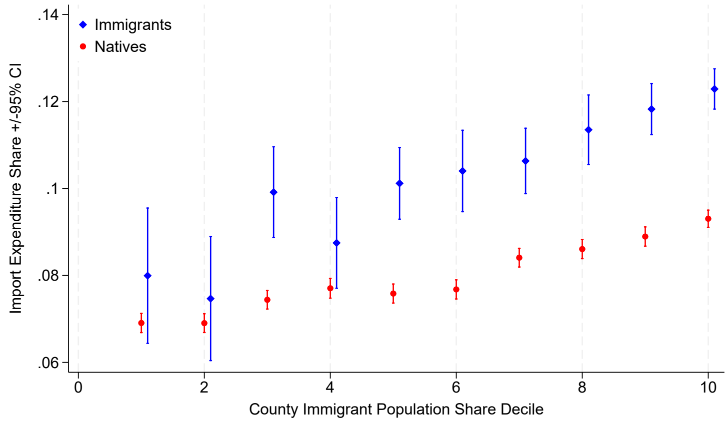
Fact 1: Household-level import expenditure is increasing in the immigrant population share of a household's county. Figure 1 plots the average import expenditure share across county deciles based on the local immigrant population share.

Both native and immigrant households exhibit a strikingly positive relationship between the presence of immigrants within a county 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% larger. For immigrant households, this differential increases to +50%. This figure represents the first direct evidence of a positive correlation between household-level import expenditure and local immigrant population shares.

Fact 2: Aggregate import expenditure shares are 38% greater for immigrant households compared to non-immigrant households. In addition to the positive correlation between import expenditure and immigrant population shares, Figure 1 provides evidence of a stark contrast in import expenditure between immigrants and native house-

[17] Historic counties until 1940; county groups in 1970/1980; and public-use micro areas (PUMAs) from 1980 to the present.

Figure 1: Immigrants, Natives, and Import Expenditure



Notes: The figure plots estimates from a linear regression at the household level in which household import expenditure shares are regressed on fixed effects at the immigrant-by-county-decile level. Counties are placed into deciles based on the immigrant population share of that county, and households are grouped into two categories: immigrant or native. 95% confidence intervals are provided, and all observations are weighted by the NielsenIQ projection factors.

holds. Within each county decile, immigrant households exhibit significantly higher import expenditure than native households. The estimates provided in Figure 1 suggest that immigrants exhibit stronger import demand than native households even within the same county.

We 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. Table C.1 provides the estimates associated with this exercise, and we 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.[18] Columns 3 to 6 of Table C.1 display results with additional controls in order to mitigate the potential bias associated with immigrants sorting into higher import counties or differing in other observable characteristics from natives, such as education or income. Even when county-level fixed effects and a suite of socio-economic household characteristics are included, the estimated differential between immigrants and natives in their average import expenditure remains highly significant and constitutes a gap of +2.8 percentage points.[19]

[18]Figure C.2 provides a raw histogram of import expenditure shares for both native and immigrant households.

[19]We add controls for income bins, household size, marital status, and head of household age and gender.

Fact 3: Immigrants spend over twice as much as natives on goods from their origin country. A key advantage of the data used here is that we are the first to simultaneously observe the specific origin country of households and products. This allows us to test directly whether the differential import expenditure associated with immigrants is driven by expenditure on all import origins, or goods specifically from that household's origin country. We turn to this analysis as our final stylized fact.

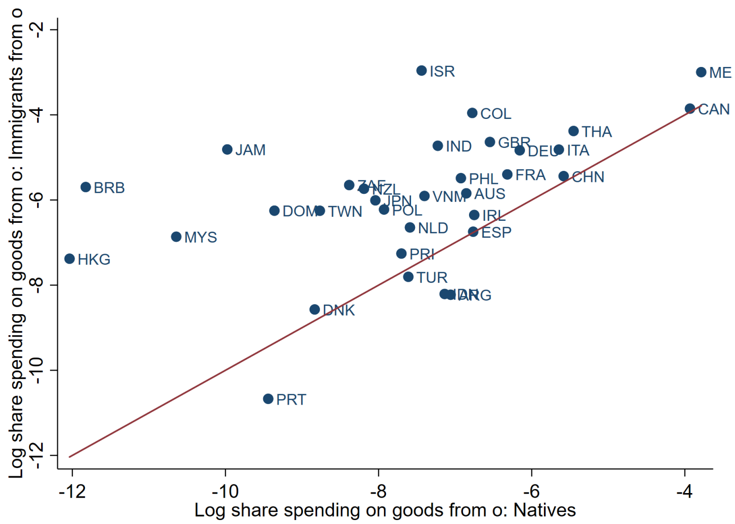
For each origin country o, we calculate the share of expenditures on goods from o by both households from o and natives. Figure 2 depicts this relationship by comparing the share of expenditures on goods from o by natives on the x-axis to that of immigrant households from o on the y-axis. The 45-degree line in red plots where natives and immigrants from o would exhibit identical expenditure on imports from o.

We find that almost all origins lie above the 45-degree line, suggesting immigrants do in fact exhibit disproportionately stronger demand for imports from their specific country of origin. For the 33 countries in our sample with non-zero expenditure by both immigrant households from that origin country and the native-born population, the median relative expenditure share on goods from origin o by immigrants from o is 2.2 times greater than the expenditure on goods from o by non-immigrant households.[20] To our knowledge, this paper is the first to provide direct evidence that the preference persistence documented in Bronnenberg et al. (2012) and Atkin (2016) exists for US immigrants with respect to demand for goods from their origin country.

The preceding three stylized facts suggest that immigrants and natives have different demand for tradables, i.e., that there exists what we call a preference composition effect. In our subsequent empirical and theoretical exercises we will quantify the importance of this channel as well as any spillovers from immigrants onto native consumption.

[20]Note that this estimate represents the weighted median relative expenditure across origins. The mean estimate is 30.9, but this is driven by outliers. When weighted by origin-specific aggregate expenditure shares, the mean difference is 3.4. Thus the median estimate of 2.2 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 one's own country (the y-axis) and spending by goods from that country by natives (x-axis). The red line is the 45-degree line, which plots when there is no preference by immigrants for goods imported from their origin country relative to natives. Household nativity assigned as discussed in Section 2.1. Data come from the NielsenIQ Household Panel 2014-2016. NielsenIQ projection factor weights used to construct expenditure shares.

3 Immigrants and Imports: the Structural Gravity Model

In this section, we estimate a structural gravity model allowing for for immigrants to reduce trade barriers and to affect local preferences. We show how, using household-level data, we can separately identify the direct effect of immigrants' home-biased preferences—what we term the preference composition effect—from the spillover effect of immigrants on native consumption within a general framework which nests a broad class of trade models.

3.1 Immigrants and Structural Gravity

We begin by considering a structural gravity model, as defined by Head and Mayer (2014), with the intent of modeling import expenditure in county c on goods from origin country o, X_{oc}. The structural gravity model associated with this flow of goods is the following:

$$X_{oc} = \alpha_o S_c \phi_{oc}$$

where α_o captures some model-adjusted size of origin o, with $\alpha_o = Y_o/\Omega_o$. Y_o measures the value of production in o and Ω_o some aggregate deflator of size in production, such as marginal cost or remoteness. S_c is a measure of real demand in county c, given by $S_c = X_c/\Phi_c$, where X_c is aggregate grocery expenditures in c and Φ_c is some price index, which we formally define below. ϕ_{oc} captures the set of bilateral factors which affect trade, such as distance, trade policy, and preference similarity.

It is important to note that the structure outlined above nests the key models used in modern quantitative studies of international trade, including Eaton and Kortum (2002), Krugman (1980), and Melitz (2003)-Chaney (2008).

The conventional interpretation of ϕ_{oc} is that it captures bilateral trade costs.[21] We generalize the standard gravity model by allowing for a bilateral affinity term, whereby consumers in c may exhibit preferences for the goods of specific origin countries. Formally, we decompose the bilateral term ϕ_{oc} into two multiplicative components: a supply component ϕ^b_{oc} capturing bilateral trade barriers, and a preference component ϕ^z_{oc} reflecting the county-specific appeal associated with goods from origin o.[22] We can then re-write our structural gravity model as:

$$X_{oc} = \alpha_o S_c \phi^b_{oc} \phi^z_{oc} \tag{1}$$

To simplify future expressions, we assume without loss of generality that for any county c: $\phi^b_{us,c} = \phi^z_{us,c} = 1$. That is, all bilateral terms are relative to the analogous term for U.S. producers selling to consumers in county c.

In this paper we aim to quantify the welfare effects of immigrants on native households' consumption of tradables. Because immigrants may affect both trade barriers ϕ^b_{oc} and bilateral affinity ϕ^z_{oc}, a gravity regression using data aggregated to the origin-by-county level will be uninformative about the degree to which immigrants separately reduce trade costs and/or increase bilateral affinity. Instead, we make use of household-level import expenditure data, allowing us to separately identify the effects of immigrants on trade costs and preferences.

[21] Head and Mayer (2014) refer to ϕ_{oc} as "bilateral accessibility", while Chaney (2008) uses "trade barriers".
[22] Introduced by Combes et al. (2005), Felbermayr et al. (2015) call ϕ^Z_{oc} "bilateral affinity".

3.2 Preference Heterogeneity and Household-Level Gravity

Each household h living in county c faces the same bilateral trade costs ϕ^b_{oc}, but households differ in their total expenditure X_h and vector of preference shifters $\boldsymbol{z_h}$. Each element $z_{oh} \in \boldsymbol{z_h}$ represents a household-origin-specific preference shifter, with the only restriction that $z_{us,h} = 1$ for all households. While we provide a micro-foundation regarding the household-level price index in Section 4, for now we simply allow for the possibility that the interaction between trade costs and household preferences may generate price indices which vary at the household level. We therefore define real expenditures by household h as X_h/Φ_h, where Φ_h denotes household h's price index. This assumption generates the following household-level gravity equation:

$$X_{oh} = \alpha_o \frac{X_h}{\Phi_h} \phi^b_{oc} z_{oh} \tag{2}$$

In order to link the household and county-level models, we note that:

$$X_{oc} = \sum_{h \in \Lambda_c} X_{oh} = \alpha_o \phi^b_{oc} \sum_{h \in \Lambda_c} \frac{X_h}{\Phi_h} z_{oh} = \alpha_o S_c \phi^b_{oc} \underbrace{\sum_{h \in \Lambda_c} \kappa_h z_{oh}}_{\phi^Z_{oc}}$$

where Λ_c is the set of households living in county c, $S_c = \sum_{h \in \Lambda_c} X_h/\Phi_h$ is real aggregate expenditure, and κ_h household-specific real expenditure weights.[23] The bilateral affinity term ϕ^Z_{oc} is therefore an expenditure-weighted average of bilateral preferences among households in c.

3.3 Estimating Spillover and Composition Effects

To render equation (2) tractable for estimation, we normalize all expenditure volumes X_{oh} by expenditure on U.S. goods at the household level. We do so to simplify our notation, dividing out county- and household-specific terms, and in anticipation of our sample having limited coverage in many U.S. counties.[24]

We define any variable $\tilde{x}_{oh}$ as the value of x for origin o divided by the equivalent value

[23] Formally, $\kappa_h = (X_h/\Phi_h)/S_c$. Notice that the definition of S_c is consistent with our county-level gravity model, since $\Phi_c = X_c/S_c$.

[24] Head and Mayer (2014) refer to this normalization when estimating gravity models as a "ratio method".

for U.S. goods. We can therefore write the household-level gravity expression as:

$$\tilde{X}_{oh} = \tilde{\alpha}_o \phi^b_{oc} z_{oh} \tag{3}$$

To estimate the supply-side effects of immigrants on county-level import expenditure from origin o, we make the following functional form assumption, in which d_{oc} is a vector of measures of distance between o and c and I_{oc} the population share of residents in county c that were born in country o:[25]

$$\phi^b_{oc} = e^{\rho d_{oc} + \beta^b I_{oc} + \eta^b_{oc}} \tag{4}$$

The parameter ρ captures the effect of distance on supply-side accessibility of county c to producers in o, and β^b measures the strength of the supply-side effects of immigrants in shaping import accessibility from their origin country. η^b_{oc} captures the unobserved component of origin-county-specific import accessibility.

Lastly, we provide a functional form for the preference vector $\boldsymbol{z_h}$. We consider two components of preferences: first, immigrants may affect the preferences of nearby households, and second, a component that relates observed socioeconomic household characteristics to import demand. For a given household and origin country, we therefore assume the following functional form for z_{oh}:

$$z_{oh} = e^{\beta^z I_{oc}} e^{[\delta J_h + \zeta_1 \mathbf{1}[o(h) \neq US] + \zeta_2 \mathbf{1}[o(h) = o] + \eta^z_{oh}]} \tag{5}$$

J_h represents a vector of observed household characteristics such as income, education, and race. ζ_1 captures the extent to which immigrant households have stronger preferences for goods from all foreign countries, and ζ_2 captures the extent to which immigrants prefer goods specifically from their origin country à la Atkin (2016) and Logan and Rhode (2010).

Household-level characteristics will not respond to changes in immigrant presence in our counterfactuals, and hence the parameters ζ_1 and ζ_2 govern the preference composition effect of changes in I_{oc}. β^z, on the other hand, captures preference diffusion in which the presence of immigrants from a given origin affects the average preference for goods from that origin

[25] d_{oc} includes the log distance between o and c and the latitude difference between o and c, as well as squared and cubed terms of that latitude difference.

across all households in the same county.[26]

Plugging these functional form assumptions into our expression for $\tilde{X}_{oh}$, we derive our estimating equation:

$$\ln \tilde{X}_{oh} = \ln \tilde{\alpha}_o + \rho d_{oc} + \beta I_{oc} + \delta J_h + \zeta_1 \mathbf{1}[o(h) \neq US] + \zeta_2 \mathbf{1}[o(h) = o] + \eta_{oh} \tag{6}$$

with $\eta_{oh} = \eta^b_{oc} + \eta^z_{oh}$ capturing idiosyncratic county and household-level deviations in import expenditure associated with origin o. The parameter $\beta = \beta^b + \beta^z$ captures spillover effects of immigrants onto import expenditure for all households, but cannot distinguish between the supply and demand effects of this spillover.

3.4 Identification and Instrumental Variables

In estimating equation (6), there may be confounders correlated with both the consumption share of a household from a specific origin and the presence of immigrants in the household's county of residence that are not captured by our baseline controls. For example, low bilateral trade costs between New York and Italy may independently expand the set of pastas available locally, which thereby draws in Italian immigrants who tend to have a strong taste for pasta. To deal with such origin-by-county specific confounders, we adopt the instrumental variable approach of Burchardi et al. (2019).[27]

The instruments work as follows. To predict the origin-by-county immigrant population in 2010, we generate a vector of exogenous immigration from the origin and into the county using 130 years of historic data. To obtain exogenous variation in origin-to-county immigration flows, we use the intuition that historic immigration flows from an origin country to a U.S. county are more likely to occur when the origin is sending many immigrants at the same time the destination county is attracting immigrants from all origins.

Concretely, 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)

[26] In 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).

[27] We 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 Section A.1 for more details.

during a given historical decade D. To deal with potential spatial correlation in confounder, we 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:

$$IV_{o,c}^{D} = \underbrace{I_{o,-r(c)}^{D}}_{\text{Push}} \times \underbrace{\frac{I_{-\mathcal{C}(o),c}^{D}}{I_{-\mathcal{C}(o)}^{D}}}_{\text{Pull}} \tag{7}$$

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

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

We use equation (7) to predict immigrant inflows into the U.S. for all decades spanning 1880 to 2000, and document these first-stage estimates in Appendix Table A.1. The push-pull instrument strongly and positively predicts the contemporary bilateral immigrant population share.

Given the prevalence of zeros in household consumption expenditure shares $\tilde{X}_{oh}$, we use pseudo-Poisson maximum likelihood (PPML) to estimate equation 6 (Silva and Tenreyro 2006). When implementing the instrument variables strategy introduced below, we account for the non-linearity of PPML by implementing a control function approach to generating exogenous variation in the immigrant population (Petrin and Train 2010; Morten and Oliveira 2024). In particular, we add the residuals from the first-stage instrumental variable regressions as controls for our main specifications.[28]

[28] Atalay et al. (2019) demonstrates that the control function approach generates consistent estimates when using PPML. They further show that the estimates are quite close to those produced by the related GMM estimation strategy developed by Wooldridge (1997) and Windmeijer (2000).

Table 1: Household Gravity Estimates

	Dependent variable: Exp. share on goods from o relative to US	
	(1)	(2)
Immigrants/Pop. 2010	1.29***	1.15***
	(0.22)	(0.24)
First-stage residuals		0.18
		(0.31)
=1 if immigrant from anywhere	0.23***	0.23***
	(0.030)	(0.030)
=1 if immigrant from origin o	0.60***	0.61***
	(0.069)	(0.071)
N	1,461,130	1,461,130
Country FE	✓	✓
Household controls	✓	✓
Distance & latitude difference	✓	✓
1st-stage F-statistic		19.5

Notes: The table presents estimation results at the household-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 in column 2. Observations are weighted using NielsenIQ household weights. Standard errors clustered two-ways at the household and county-country levels. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

3.5 Structural Gravity Results

We show the results of estimating equation (6) using PPML in Table 1. The first column shows estimates without first-stage residuals, and indicates that a higher immigrant population share corresponds to higher spending on goods from the immigrants' origin country. In column 2, we add the first-stage residuals and find that a 1 percentage point increase in the share of immigrants from a given origin increases relative expenditures on goods from that origin by 1.17 percent (SE=0.24).

Comparing the immigrant population share coefficients between columns 1 and 2, we find that the estimate falls by about 12% when adjusting for the endogenous location choices of immigrants. This is consistent with immigrants choosing their location based on where goods from their home country are more available. In terms of the preference parameter coefficients, we find that immigrants spend 26% more on imports from any origin than natives do, and

132% more on imports specifically from the immigrant's origin country.[29]

The results summarized in Table 1 provide two key takeaways. First, immigrants' preferences—the composition effect—play a significant role in shaping import expenditures. Indeed, our estimates validate the caution expressed by Felbermayr et al. (2015) in interpreting immigrants' effect on imports using aggregated data as an effect on welfare. Second, we find that spillover effects of immigrants from a given origin to the rest of the local population—captured by the immigrant-population share coefficient—are also significant.

Even controlling for immigrant preferences, the estimated spillover effect may incorporate both immigrants' effects on trade costs and on local preferences—what we call preference diffusion. While this distinction has no bearing on the trade-creating effects of immigrants, it plays a crucial role in identifying the welfare effects of immigrant-induced trade. We return to this distinction in Section 3.7 within the context of the welfare formula derived in Arkolakis et al. (2012). First, however, we discuss the robustness of our main results.

3.6 Robustness and Heterogeneity

We conduct a variety of robustness checks and heterogeneity analyses in Appendix A.2. In Section A.2.1, we re-weight households to match the distribution of immigrant country origins observed in U.S. Census data. In Section A.2.2, we show that our main results are virtually unchanged when allowing immigrants to exhibit specific preferences for countries geographically or culturally similar to their origin country. In Section A.2.3, we assess the relative importance of the extensive versus intensive margins. In Section A.2.4, we find that our baseline results are robust to controlling for the instrumental variable mean as recommended by Borusyak and Hull (2023) for instruments which combine different sources of variation according to a known formula.

[29] $\exp \hat{\zeta}_1 - 1 = \exp 0.23 - 1 = 0.26$ and $\exp \hat{\zeta}_1 + \hat{\zeta}_2 - 1 = \exp 0.23 + 0.61 - 1 = 1.32$.

3.7 Immigrants, Imports and Welfare

Arkolakis et al. (2012) (henceforth ACR) develop this simple welfare formula for a large class of trade models:

$$d\ln W_c = (d\ln X_{us,c})^{1/\nu}$$

where W_c is real expenditures or, equivalently, welfare in county c; and ν is the trade elasticity. In the formulation above, we hold constant aggregate expenditures X_c. In this subsection we show that an immigration shock which changes both local preferences and trade costs invalidates the ACR welfare formula. We also derive a multiplicative adjustment factor to $d\ln X_{us,c}$ which allows one to recover welfare effects on natives due to such an immigration shock using both aggregate data and our estimated parameters.

An immigration shock which affects trade costs deviates from ACR for two key reasons. First, if native households exhibit weaker preferences for imported goods than immigrants (ζ_1,$\zeta_2 > 0$), then natives are less sensitive to trade shocks than immigrant households. Second, immigrant-induced changes in county-level import expenditure will only translate into welfare gains for native households if driven by changes in trade barriers (i.e., ϕ^b_{oc}) rather than changes in local preferences (ϕ^z_{oc}).

With a few simplifying assumptions we derive an explicit adjustment to the standard ACR formula which allows us to express the gap between implied welfare gains from an aggregate change in import expenditure to the change in import expenditure associated with native households. For the purposes of this exercise, we assume that there exist only two regions: the United States (us) and the rest of the world. We denote native households with n and assume all households are identical except for their immigrant status. Lastly, we collapse ζ_1 and ζ_2 into a single parameter ζ which captures the relative import preference of immigrants versus native households.

We consider some change in the immigrant population share which causes aggregate county-level domestic expenditure to change by some exogenous $d\ln X_{us,c}$.[30] Given the structural gravity model described above and estimates of β^b, β^z, and ζ, one can transform the county-level change in domestic expenditure to the welfare-relevant change in domestic

[30]We assume that aggregate expenditure X_c remains constant.

expenditure of native households using the following transformation:[31]

$$d \ln X_{us,n} = d \ln X_{us,c} \left[\frac{1}{\frac{I_c}{s_{us,c}}(e^{\zeta} - 1) + 1} \right] \left[\frac{\beta^b}{\beta + \frac{e^{\zeta}-1}{I_c(e^{\zeta}-1)+1}} \right] \tag{8}$$

where I_c is the immigrant population share in county c and $s_{us,c}$ is the pre-shock domestic expenditure share in county c. With equation (8) in hand, one then follow ACR and compute native welfare as $d \ln W_n = (d \ln X_{us,n})^{1/\nu}$.

The first term associated with this transformation adjusts $s_{us,c}$ in order to recover the unobserved native household domestic expenditure share $s_{us,n}$. So long as $\zeta > 0$, and immigrants have stronger preferences for imports than native households, this term will be less than one and for any trade shock—immigrant-induced or otherwise—native households will exhibit smaller changes in welfare than those implied by the county-level aggregate $d \ln X_{us,c}$.

The second term captures the share of the aggregate change in domestic expenditure which is welfare-relevant to native households. That is, $d \ln \phi^b_{oc}/d \ln \phi_{oc}$. If at least one of β^z or ζ is positive, and both are non-negative, this second term is less than one. Therefore changes in native household welfare should be discounted when compared to the implied aggregate welfare effects. If $\beta^b = 0$, then immigrant-induced changes in domestic expenditure may be large in the aggregate, but will have zero effect on the welfare associated with native households, and $d \ln X_{us,n} = 0$.[32]

Our estimates from Table 1 imply that immigrants exhibit stronger preferences for imports from their origin country than native households, and thus $\zeta > 0$. What we cannot disentangle from our estimates is the relative magnitude of β^b and β^z: the extent to which the spillover effect of immigrants is due to supply factors (lowering trade costs) or demand factors (influencing the preference of neighbors). We therefore take these general gravity estimates as motivation for the section to follow, in which we make use of the heterogeneous firms Melitz-Chaney variant of the structural gravity class of models to run counterfactual simulations and recover the effect of immigrants on import penetration and native household welfare.

[31]The details of the derivation can be found in Appendix B.1.

[32]We also note that assuming no preference heterogeneity ($\zeta = 0$) and no effect of a shock on preferences ($\beta = \beta^b$) eliminates the multiplicative terms in equation (8).

4 Microfounding a Model of Immigration and Imports

This section uses the Melitz (2003)-Chaney (2008) micro-foundation to expand upon the structural gravity model of immigrant-induced trade in the previous section. We then leverage the equilibrium moments of this model and the detailed data available to separately identify the effect of immigrants on marginal costs, fixed costs of exporting, and household preferences, thus disentangling ϕ^b_{oc} and ϕ^z_{oc}.

We opt for the Melitz-Chaney model for two reasons. First, the increasing returns to scale nature of this model allows for market size effects, a key channel through which immigrants might affect the supply of varieties locally (Iranzo and Peri 2009; Di Giovanni et al. 2015; Aubry et al. 2016). Second, the structure of the Melitz-Chaney heterogeneous firms model allows us to fully leverage the data we possess and separately quantify the marginal cost, fixed cost, and preference spillover effects of immigrants on native households, thus identifying the supply and demand effect of immigrants on import penetration. We turn to describing this model now, as well as our estimation/calibration of the model and subsequent counterfactual exercises.

4.1 Heterogeneous Households and Firms

Households: Each household h lives in county $c(h)$ and exhibits Cobb-Douglas preferences over a homogeneous tradable good, q_0, and a differentiated good consisting of a continuum of differentiated varieties $\Omega_{o,c(h)}$ associated with each origin country $o \in \mathcal{O}$. As in the previous section, we allow for household heterogeneity in income Y_h and origin-specific preferences denoted by $z_{oh} \in \boldsymbol{z_h}$. Preferences for the differentiated sector are represented by the following CES utility function:

$$U_h = {q_0}^{\mu_0} \left[\sum_{o \in \mathcal{O}} {z_{oh}}^{\frac{1}{\sigma}} \int_{\omega \in \Omega_{o,c(h)}} q_{oh}(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right]^{\frac{\sigma}{\sigma-1}(1-\mu_0)} \tag{9}$$

with $\sigma > 1$ denoting the elasticity of substitution among differentiated varieties. The exponent μ_0 captures the expenditure share on the homogeneous good, which we assume is constant across households and therefore pins down expenditure on the differentiated sector as $X_h = (1 - \mu_0)Y_h$.

We leave the functional form of z_{oh} unchanged from the previous section (see equation 5). That is, β^z governs the spillover effect of immigrants on native household preferences for imports, δ maps exogenous household characteristics into import demand, ζ_1 governs immigrant preferences for imported goods, and ζ_2 governs immigrant preferences for goods specifically from their origin country.

Firms: Each country $o \in \mathcal{O}$ has some exogenous size Y_o and marginal cost of production w_o. Trade is characterized by county-by-origin specific 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.[33] The cost of providing q units to destination county c by a firm in origin o with productivity φ is therefore:

$$C_{oc}(q) = \frac{w_o \tau_{oc}}{\varphi} q + f_{oc} \tag{10}$$

and we assume that all entry and pricing decisions are made at the county level such that each county is an independent market.

Given the extent to which this model builds upon the structure introduced by Chaney (2008), we relegate the full derivation of the model to Appendix B.2, including all definitions of constants denoted by λ.

Equilibrium: In equilibrium, the household-specific price index is given by:

$$\Phi_h = P_h^{1-\sigma} = \lambda_3 \sum_{o \in \mathcal{O}} Y_o z_{oh} (w_o \tau_{o,c(h)})^{-\theta} \Big(\frac{f_{o,c(h)}}{S_{c(h)} z_{o,c(h)}} \Big)^{-(\frac{\theta}{\sigma-1}-1)} \tag{11}$$

in which $S_{c(h)}$ is again real aggregate expenditure in county c, as defined previously in the structural gravity model.[34] The average county-level preferences $z_{o,c(h)}$ are also the same as our definition for the bilateral affinity term introduced in Section 3.2 (ϕ^z_{oc}) and are an expenditure weighted average of the preference shifter z_{oh} across all households in c.

[33] We assume that θ is identical across all origin countries.

[34] Formally: $S_c = \sum_{h' \in \Lambda_c} X_{h'} P_{h'}^{\sigma-1}$, where Λ_c is the set of households residing in county c.

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

$$X_{oh} = \lambda_4 Y_o X_h P_h^{\sigma-1} (w_o \tau_{o,c(h)})^{-\theta} \Big(\frac{f_{o,c(h)}}{S_{c(h)} z_{o,c(h)}} \Big)^{-(\frac{\theta}{\sigma-1}-1)} z_{oh} \tag{12}$$

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

$$X_{oc} = \lambda_4 Y_o S_c (w_o \tau_{oc})^{-\theta} \Big(\frac{f_{oc}}{S_c z_{oc}} \Big)^{-(\frac{\theta}{\sigma-1}-1)} z_{oc} \equiv \alpha_o S_c^{\frac{\theta}{\sigma-1}} \phi_{oc}^b \phi_{oc}^z \tag{13}$$

Notice that we now have a micro-foundation for each term used in structural gravity model of Section 3. ϕ_{oc}^z remains unchanged, whereas real expenditures S_c is now raised by the exponent $\theta/(\sigma-1) > 1$ due to the increasing returns to scale associated with the micro-foundation of production assumed here. The real size of origin o is now defined formally as $\alpha_o = Y_o w_o^{-\theta}$.

The preference shifter z_{oc} represents the novel extension beyond the standard Chaney (2008) framework. Preferences also contribute to a market size effect associated with county-level average preferences. As preferences shift towards goods from origin o, more firms from o are able to cover the fixed costs of supplying county c, which further enhances the market penetration of imports from o to county c.

As in Section 3.3, it will be convenient when taking our main estimating equation to the data to estimate the model relative to U.S. expenditure for a given household. Using the same definition as $\tilde{x}$ from before to denote any variable relative to its U.S. equivalent, we can express the normalized household-by-origin level expenditure equation as

$$\tilde{X}_{oh} = \tilde{\alpha}_o (\tilde{\tau}_{o,c(h)})^{-\theta} \Big(\frac{\tilde{f}_{o,c(h)}}{z_{o,c(h)}} \Big)^{-(\frac{\theta}{\sigma-1}-1)} z_{oh} \tag{14}$$

It will also be useful to separate household preferences into a component that is endogenous to the local immigrant population share, $e^{\beta^z I_{oc}}$, and an exogenous component $\bar{z}_{oh}$, such that $z_{oh} = e^{\beta^z I_{oc}} \bar{z}_{oh}$. Given that the endogenous component is common to all households in a given county, we can provide the same distinction at the county level $z_{oc} = e^{\beta^z I_{oc}} \bar{z}_{oc}$, where

$\bar{z}_{oc}$ is simply an expenditure weighted average of $\bar{z}_{oh}$.

4.2 Immigrant Channels and Decomposition

We complete the model introduced in Section 4.1 by introducing functional form assumptions for variable and fixed trade costs. As in Section 3.3, we allow for the possibility that immigrants might affect these costs. We then derive our main estimating equation and highlight the various channels through which immigrants affect import penetration and welfare in this model.

Our functional form assumptions regarding the variable and fixed components of ϕ^b_{oc} closely follow the assumptions made in the structural gravity model. That is, 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:[35]

$$\tilde{\tau}_{oc} = \exp[-\frac{1}{\theta}(\rho^\tau d_{oc} + \beta^\tau I_{oc} + \eta^\tau_{oc})] \tag{15}$$

$$\tilde{f}_{oc} = \exp[-(\frac{\sigma - 1}{1 + \theta - \sigma})(\rho^f d_{oc} + \beta^f I_{oc} + \eta^f_{oc})] \tag{16}$$

where η^τ_{oc} and η^f_{oc} represent idiosyncratic deviations in trade costs across county-origin pairs that are assumed to be mean-zero. β^τ captures the strength of the the variable cost reduction channel of immigrants and β^f the fixed cost reduction channel of immigrants on import expenditure in county c.

We can now return to our expression for $\tilde{X}_{oc}$ and plug in our functional form assumptions for z_{oh}, $\tilde{\tau}_{oc}$, and $\tilde{f}_{oc}$. Taking the logarithm of this expression and differentiating yields the following decomposition of the county-level partial elasticity of import expenditure with respect to the immigrant population share:

[35]The normalization terms $\frac{1}{\theta}$ and $\frac{\sigma-1}{1+\theta-\sigma}$ simplify notation but are not necessary.

$$
\begin{aligned}
\frac{\partial \ln \tilde{X}_{oc}}{\partial I_{oc}} &= \frac{\partial \ln \phi^b_{oc}}{\partial I_{oc}} + \frac{\partial \ln \phi^z_{oc}}{\partial I_{oc}} \\
&= \underbrace{\left[\beta^\tau + \beta^f\right]}_{\substack{\text{Trade cost}\\ \text{channel}}} + \underbrace{\left[\frac{\theta}{\sigma - 1} - 1\right]\left(\beta^z + \frac{\partial \ln \bar{z}_{oc}}{\partial I_{oc}}\right)}_{\text{Market size channel}} + \underbrace{\beta^z}_{\substack{\text{Preference}\\ \text{diffusion}\\ \text{channel}}} + \underbrace{\frac{\partial \ln \bar{z}_{oc}}{\partial I_{oc}}}_{\substack{\text{Composition}\\ \text{channel}}}
\end{aligned} \tag{17}
$$

Expression (17) illustrates the channels through which immigrants affect county-level import expenditure from a given origin. The first two channels—trade costs and market size—represent changes in the supply-side effects of immigrants, or ϕ^b_{oc}. These include the variable cost reduction effect, the fixed cost reduction effect, and the market size effect associated with changes in local preferences. A shift in county-level preferences for goods from origin o will lead to greater entry by firms exporting from o, and given the CES preferences assumed in this model, this increased availability will lead to non-zero expenditure on these new varieties by non-immigrant households. The strength of this effect is governed by the ratio $\theta/(\sigma - 1) > 1$.

The final two terms capture the extent to which immigrants affect the bilateral affinity term ϕ^z_{oc}. β^z captures the effect of immigrants on preferences for goods from their origin that are common to all households in county c, whereas the preference composition channel captures the extent to which increased immigrant presence shifts the composition of households towards those with non-zero values of the parameters ζ_1 and ζ_2.

From a welfare perspective, the intuition is identical to the previous discussion regarding the structural gravity model. The only welfare relevant channels of immigrant-induced import penetration are those associated with ϕ^b_{oc}: the trade cost and market size channels.

An important feature of the micro-foundation used here is that when combined with the detailed data available, we can separately identify all parameters necessary to quantify each channel. We will again make use of our household-level data and so we return to the household-level gravity model discussed previously but accomodating the microfoundation described here:[36]

$$\ln \tilde{X}_{oh} = \alpha_o + \rho d_{o,c(h)} + \beta I_{o,c(h)} + \ln \bar{z}_{o,c(h)}^{\frac{\theta}{\sigma-1}-1} \\ + \delta J_h + \zeta_1 \mathbf{1}\left[o(h) \neq US\right] + \zeta_2 \mathbf{1}\left[o(h) = o\right] + \eta_{o,c(h)} + \eta^z_{oh} \tag{18}$$

with the following definitions:

$$\rho = \rho^\tau + \rho^f$$

$$\beta = \beta^f + \beta^\tau + (\frac{\theta}{\sigma - 1})\beta^z$$

$$\eta_{o,c(h)} = \eta^\tau_{o,c(h)} + \eta^f_{o,c(h)}$$

The specification here reveals three identification concerns. First, as discussed in Section 3.4, the unobserved component of variable costs and fixed costs η^τ and η^f are likely correlated with the immigrant population share I_{oc}, and hence we make use of the same instrumental variables strategy. Second, and perhaps more concerning, is that the preference diffusion effect of immigrants and the preference composition effect of immigrants are not separately identified: county-level preferences $\bar{z}_{oc}$ were not implied by the structural gravity model and therefore loaded on to estimates of β. Lastly, even an unbiased estimate of β would simply yield a combination of β^τ, β^f, and β^z.

4.3 Identifying the Channels of Immigrant-Induced Imports

In this section we provide a three-step identification strategy which allows us to separately identify each of the various channels by which immigrants affect imports.

Identification of Exogenous Preferences: We start by identifying how household's socioeconomic characteristics—such as race, education, household size, and income—affect the demand for imports. To do so, we control for origin-by-county factors which affect trade costs and product availability, and then project those socioeconomic characteristics onto the residual variation.

We collect all terms affected by the local immigrant population share into an origin-county fixed effect ψ_{oc} and make use of our households-level purchase data to estimate the

[36] With some abuse of notation, we define $\alpha_o = \ln \alpha_o$.

exogenous component of preferences $\bar{z}_{oh}$. Specifically, we estimate the following model, which is identical to equation (18) save for the collection of terms into ψ_{oc}:

$$\ln \tilde{X}_{oh} = \psi_{o,c(h)} + \delta J_h + \zeta_1 \mathbf{1}[o(h) \neq US] + \zeta_2 \mathbf{1}[o(h) = o] + \eta_{oh}^z \tag{19}$$

In this case it is safe to assume that the estimates $\hat{\delta}$, $\hat{\zeta}_1$, and $\hat{\zeta}_2$ are unbiased as the only error term not captured by the fixed effects is the idiosyncratic household-origin preference shock η_{oh}^z. That is, all components of the model associated with prices and firm selection are captured by the origin-county fixed effects ψ_{oc}. We estimate this specification using PPML to account for the number of zeros in $\tilde{X}_{oh}$ and recover the estimates $\hat{\delta}$, $\hat{\zeta}_1$, and $\hat{\zeta}_2$. We then construct an estimate of the exogenous household-level preference term as $\hat{\bar{z}}_{oh} = e^{(\hat{\delta} J_h + \hat{\zeta}_1 \mathbf{1}[o(h) \neq US]) + \hat{\zeta}_2 \mathbf{1}[o(h)=o])}$ and plug this estimate into the county-level average preference term $\bar{z}_{oc}$ to arrive at an estimate of $\hat{\bar{z}}_{oc} = \sum_{h' \in \Lambda_c} \hat{\bar{z}}_{oh'} \kappa_{h'}$. We make use of publicly available Census data to construct the household-level weights κ_h, and we make use of calibrated values of σ and θ taken from the literature, which we discuss in Section 5.2 to follow.

Estimating $\boldsymbol{\beta}$: With the exogenous preference estimates in hand, we can difference out both $\bar{z}_{oh}^{\sigma}$ and $\bar{z}_{oc}^{\frac{\theta}{\sigma-1}-1}$ from our main estimating equation and isolate the effect of county-level parameters on this adjusted measure of import expenditure, as shown in the following equation:

$$\ln \frac{\tilde{X}_{oh}}{\mathcal{Z}_{oh}} = \alpha_o + \rho d_{o,c(h)} + \beta I_{o,c(h)} + \eta_{o,c(h)} + \eta_{oh}^z \tag{20}$$

in which we define $\mathcal{Z}_{oh} = \hat{\bar{z}}_{oh} \hat{\bar{z}}_{o,c(h)}^{\frac{\theta}{\sigma-1}-1}$ to simplify notation.

Notice that the dependent variable represents observed household-level expenditure on imports from origin o adjusted by the household's predicted level of expenditure, given the household's observed socioeconomic characteristics but also the predicted import expenditure given the observed characteristics of all other households living in their county. Applying the deflator $\mathcal{Z}_{oh}$ to household h's expenditure on goods from origin o allows us to isolate the *spillover effect* of immigrants by adjusting for the composition effect directly.

We therefore arrive at an estimating question that is reminiscent of the structural gravity model estimated earlier (see equation 6), and we make use of the same instrumental variables strategy and again implement PPML with a control function approach.

Estimating $\boldsymbol{\beta^{\tau}}$: Even after adjusting for market size effects, we still cannot disentangle the components of β: β^{τ}, β^{f}, and β^{z}. We leverage model restrictions and data characteristics which allow us to identify this decomposition.

We have assumed throughout this section that firms price according to monopolistic competition, and thus set constant mark-ups. Specifically, the optimal pricing function for any variety ω from origin o in county c is the following:

$$p_{\omega(o),c} = \frac{\sigma}{\sigma - 1}\frac{w_o \tau_{oc}}{\varphi(\omega)} = \frac{\sigma}{\sigma - 1}\frac{w_o}{\varphi(\omega)}\tau_{us,c}\tilde{\tau}_{oc}$$

By aggregating our data to the barcode-by-county level we can estimate the price equation directly, having incorporated the functional form assumption of $\tilde{\tau}_{oc}$ from equation (15):

$$\ln p_{\omega(o),c} = \psi_c + \psi_\omega - \frac{\beta^{\tau}}{\theta} I_{oc} - \frac{\rho^{\tau}}{\theta} d_{oc} - \frac{1}{\theta}\eta^{\tau}_{oc} \tag{21}$$

where ψ_c and ψ_ω represent county and barcode-level fixed effects.[37] As with our baseline specification, we instrument for the immigrant population share to account for the likelihood that $cov[I_{oc}, \eta^{\tau}_{oc}] \neq 0$.

Estimating $\boldsymbol{\beta^{f}}$ and $\boldsymbol{\beta^{z}}$: If $\beta^{\tau} \approx 0$, we can isolate the effect of fixed costs from the effect of preference diffusion on import expenditure by comparing the intensive margin (expenditure shares, as in our baseline) with the extensive margin (counts of varieties). We show in Section 5.3 that immigrants have no discernable effect on variable trade costs.

Specifically, we follow Chaney (2008) and derive expressions for both the extensive margin elasticity of imports with respect to the immigrant population share and the total expenditure elasticity of imports with respect to the immigrant population. Since $\beta^{\tau} \approx 0$, we have two equations with two unknowns: β^{f} and β^{z}. The scanner data used in this paper provide detailed barcode count data, and so we estimate the extensive margin effect of immigrants on trade directly by replacing $\tilde{X}_{oh}$ in (20) with $\tilde{N}_{oh}$: the count of barcodes from origin o in household h's consumption basket compared to the count of barcodes from the U.S. in household h's consumption basket.

[37]Note that each barcode ω is unique to an origin country o; hence ψ_ω also captures variation in production costs w_o across origins.

While the full derivation is provided in Appendix B.3, it is simple to show that our functional form assumptions for β^f and β^z yield the following two expressions regarding the import expenditure elasticity and import variety elasticity, respectively:

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

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

The intuition is as follows. On the extensive margin, firms enter a new market if and only if they can cover their fixed costs. Immigrants can therefore facilitate firm entry by either (i) reducing fixed costs, or (ii) increasing sales-per-variety in their market, either by bringing their own preferences for imported goods or influencing their neighbors' preferences.

On the intensive margin, firms sell more in markets they already sell to if preferences for their goods increase. However, reducing fixed costs will have no impact on how much a firm sells to a given market, conditional on already selling to that market.

5 Melitz-Chaney Model Estimation Results

5.1 Estimating Preference Terms

We construct estimates for household preferences $\bar{z}_{oh}$ by estimating equation (19) using our NielsenIQ household sample. We therefore recover estimates of the parameter vector δ as well as ζ_1 and ζ_2. The detailed regression results are presented in Appendix Table C.2.

We find that import expenditure is generally increasing in income, albeit noisily, with a similar pattern of import expenditure increasing in household education.[38] We estimate that immigrant households consume more imported goods from any origin ($\hat{\zeta}_1$) with an estimated effect of 0.23 (SE=0.029), as well as more goods from their specific birth country ($\hat{\zeta}_2$), with an estimated effect of 0.64 (SE=0.069). These estimates closely match what we found in Section 3.5, and suggest that immigrant import expenditure is 1.26 times that of native households

[38]Note that these estimates combine the extensive and intensive margin of import expenditure: a higher estimate associated with a given household characteristic implies that this household has greater expenditure on the same set of origins and/or consumes imports from more origins.

for all origins, and 2.39 times greater for imports from their specific origin country.

Appendix Figure C.3a shows the distribution of $\ln \hat{\hat{z}}_{oh}$, which we recenter around zero. Household preferences for imports appear quite symmetrically and closely distributed around the mean, with only a handful of household-origin pairs exhibiting very strong or very weak preferences for imported goods.

To impute county-level preferences, $\hat{\hat{z}}_{oc}$, one must observe the characteristics of all households in a given county. Given the small sample size of the NielsenHQ data, we turn to the American Community Survey 2012-2017 sample to compute $\hat{\hat{z}}_{oc}$. We do so in two steps. First, we take the preference parameters estimated using the NielsenHQ data—$\hat{\delta}$, $\hat{\zeta}_1$, and $\hat{\zeta}_2$—and predict each ACS household's preference term $\hat{\hat{z}}_{oh}$. Then we sum across households to construct county-level preferences: $\hat{\hat{z}}_{oc} = \sum_{h' \in \Lambda_c} \hat{\hat{z}}_{oh'} \kappa_{h'}$. We compute the household weights $\kappa_{h'}$ as being the share of overall income in county c that is earned by household h'.

Appendix Figure C.3b shows the distribution of $\ln \hat{\hat{z}}_{oc}$. Due to the aggregation of household-origin expenditure shares at the county level, the distribution has a much lower range, which lies between -0.016 and 0.077. Five out of the six largest values correspond to the preference terms for products from Mexico in counties in California and Texas. Other county-origin pairs in the top 10 include preferences for Cuban products in Miami-Dade county, preferences for Chinese goods in the San Francisco and Santa Clara counties, and preferences for Indian goods in Middlesex county (NJ).

5.2 Estimates of the Immigrant Spillover Effects

Next we estimate the total effect of immigrants on imports using equation (20), in which expenditure is deflated by household and county-level preferences. Recall that in order to deflate by the appropriate market size effect, we require parameter values for σ and θ. We assume a value for the CES elasticity parameter of σ =5. In the heterogeneous firms model used here, θ is simply the elasticity of trade with respect to variable costs, and we therefore follow Costinot and Rodríguez-Clare (2014) and calibrate θ =5.[39]

[39] Recall that $\theta > \sigma - 1$ is a restriction inherent to the model. Melitz and Redding (2015) calibrate $\theta = 4.25$ when $\sigma = 4$ and Simonovska and Waugh (2014) estimate the trade elasticity as 4.10 and 4.27, depending on specification. We opt for the relatively larger value of θ =5 from Costinot and Rodríguez-Clare (2014) in order to match our larger value of σ =5.

Table 2: Estimates of Household Gravity Equation

	$\tilde{X}_{oh}/\mathcal{Z}_{oh}$		$\tilde{N}_{oh}/\mathcal{Z}_{oh}$	
	(1)	(2)	(3)	(4)
Immigrants/Pop. 2010	1.50***	1.36***	1.19***	1.14***
	(0.22)	(0.29)	(0.11)	(0.15)
First-stage residuals		0.18		0.069
		(0.38)		(0.21)
N	1,461,130	1,461,130	1,461,130	1,461,130
Country FE	✓	✓	✓	✓
Distance & latitude difference	✓	✓	✓	✓
1st-stage F-statistic		20.2		20.2

Notes: The table presents regression results at the household-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 in column 2. Observations are 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.

Columns 1 and 2 of Table 2 provide estimates of β with and without the use of the instrument from Burchardi et al. (2019). Our estimates broadly match, and are statistically indistinguishable from, those estimated using the structural gravity model in Section 3.5. In our preferred IV specification, we estimate $\hat{\beta}$ =1.36.

The modest effect of correcting for county-level preferences likely reflects two off-setting sources of bias. First, given that our initial structural gravity model omitted the spillover effect of immigrants due to the market size channel ($\bar{z}_{oc}^{\frac{\theta}{\sigma-1}-1}$), we would expect the unadjusted estimates of Section 3.5 to be biased upwards. Second, we find evidence that immigrants have stronger preferences for all origin countries via ζ_1, suggesting a downward bias in estimates which do not control for the effect, for example, of Mexican immigrants increasing the market size in county c for imports from India. That is, identifying β from unadjusted cross-origin variation is potentially mis-specified when immigrants from origin o affect native household expenditure from all other origins o'.

Table 3: Estimates of Variable Cost Parameter using Variation in Prices

	Dependent variable: Log Average Barcode Price			
	(1)	(2)	(3)	(4)
Immigrants/Pop. 2010	-0.041***	-0.017	-0.058***	-0.040
	(0.013)	(0.031)	(0.016)	(0.044)
N	2,261,777	2,261,777	1,601,674	1,601,674
Barcode FE	✓	✓	✓	✓
County FE	✓	✓	✓	✓
Distance & latitude difference	✓	✓	✓	✓
1st-stage F-statistic		17.3		17.5
Sample	All	All	>100 Counties	>100 Counties

Notes: The table presents two-stage least square regression results at the barcode-county level. The instrumental variables strategy is described in Section 3.4. Standard errors are clustered at the barcode and country level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

5.3 Decomposing Spillovers into Trade Costs and Preferences

We start by leveraging the price information that we observe in the NielsenIQ Homescanner data in order to estimate equation (21). We show our estimates in Table 3, which shows estimates of $-\frac{\beta^\tau}{\theta}$. 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 countries driving our results, we also restrict the sample to barcodes which we observe in at least 100 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 (7).

We find that the IV estimate using either sample is statistically indistinguishable from zero and very small in magnitude. The coefficient in column 2 implies that a 1 percentage point increase in the share of the local population which is born in country o raises prices by 0.01 percent, and suggests that $\hat{\beta}^\tau = -0.085$. We therefore conclude that $\hat{\beta}^\tau \approx 0$.

We then estimate equation (20) but with the relative expenditure term $\tilde{X}_{oh}$ replaced with the relative variety count share $\tilde{N}_{oh}$ in order to recover the extensive margin elasticity of immigrants on import expenditure. Columns 3 and 4 of Table 2 provides estimates of the extensive margin effect of immigrants on import expenditure.

Solving for β^f and β^z using the elasticity estimates from columns 2 and 4 of Table 2, we recover β^f =1.09 and β^z =0.22. Since our estimate of β from Table 2 is 1.36, we therefore conclude that the primary spillover channel through which immigrants affect non-immigrant households is the fixed-cost channel, which accounts for approximately 80% of the overall spillover effect implied by β.

6 Counterfactual analysis

To quantify the contribution of immigrants to trade flows and native welfare, we use our estimated model to conduct three sets of counterfactual exercises. First, we remove the channels through which immigrants affect the import expenditure of households in the U.S. That is, we set $\zeta_1 = \zeta_2 = \beta_f = \beta_z = 0$ and recalculate $\bar{z}_{oc}$ accordingly. We repeat this exercise with each individual channel successively maintained at the initial value in order to quantify the contribution of each. In a second counterfactual, we additionally remove the grocery expenditure associated with immigrant households, which corresponds to removing immigrants from the U.S. population altogether, thus capturing the market size benefits of immigrants. Finally, we simulate a 10% increase in the variable cost associated with all imported goods and quantify the relative cost of this policy on all households—immigrant and native.[40]

All derivations used in this section to calculate counterfactual changes in expenditure and welfare can be found in Appendix Section B.4.

6.1 Aggregate Effect of Immigrants on Imports & Native Welfare

To generate values which are representative of the United States as a whole, as well as meaningful counterfactual values for various U.S. cities, we leverage the American Community Survey (ACS). In particular, we use the results from estimating equation (19) with the NielsenIQ data to predict household-origin-specific expenditures for each ACS household. We further assume that each household spends $7,500 on grocery and personal care products covered by NielsenIQ, which matches estimates from the Consumer Expenditure

[40]Note that our counterfactuals exclusively allow for partial-equilibrium adjustment in the consumption space. The labor market effects of immigrants are outside the scope of our framework.

Table 4: Counterfactual Results Summary

Counterfactual exercise: Removing ...	(1) Change (%) import expenditure	(2) Change (%) welfare natives	(3) Change ($) welfare per native HH
... all immigrant channels	-7.3	-0.03	-2
... except composition	-1.9	-0.03	-2
... except homophily	-5.7	-0.03	-2
... except market size	-6.7	-0.03	-2
... except trade cost	-5.5	-0.01	-0
... all immigrant channels + expenditure	-25.7	-0.93	-69

Notes: This table shows the change in outcomes under various counterfactual scenarios. The baseline scenario in the first row removes all channels through which immigrants affect import expenditures—i.e. we set $\zeta_1 = \zeta_2 = \beta_f = 0$ and recalculate $\bar{z}_{oc}$—but keeps overall expenditure constant (we do not change β_z so that preferences stay constant). In the next rows, we keep the following parameters/variables at their estimated initial values: ζ_1 and ζ_2 (composition); ζ_2 (homophily component of composition channel); $\bar{z}_{oc}$ (market size); β_f (trade cost). In the last row, we remove all immigrant channels and the expenditures made by immigrants, equivalent to removing immigrants from the U.S. population.

Survey (CEX). Finally, we use the crosswalks provided by Burchardi et al. (2019) to generate county-specific immigrant population shares based on the PUMAs in which households are located.

We provide a summary of our results across different counterfactual scenarios in Table 4, with our baseline counterfactual scenario results appearing in the first row. Averaging across households, we find that aggregate U.S. expenditures on imports of grocery and personal care items decreases by 7.3%. We further find that removing all immigrant spillover effects yields an average welfare loss from grocery and personal care consumption of 0.03%, amounting to a welfare-equivalent fall of $2 per household.[41]

The second row of Table 4 maintains the composition effect of immigrants at their initial level. That is, the parameters ζ_1 and ζ_2, which govern immigrant import preferences relative

[41]Note that while these welfare magnitudes seem small, they are bounded above by the cost of autarkey, which is small in this setting given that the import expenditure share of US groceries is roughly 9%. Maintaining our assumption of $\theta = 5$ from earlier, autarkey carries a welfare cost of only 1.9%.

to non-immigrants, remain unchanged. The counterfactual decrease in import expenditure is reduced to 1.9%, implying that the composition channel accounts for almost three quarters of the overall trade effect of 7.3%. In row three, we keep only the homophily effect at the initial level by holding ζ_2 fixed. The reduction of the overall effect to 5.7, a fall of 22%, suggests that the preferences of immigrants for goods from their origins drives around a third of the contribution of the composition channel, while the remainder is due to their higher preferences for any imported goods relative to natives. Finally, without the market size channel, for which we keep $\bar{z}_{oc}$ fixed, and the trade cost channel, for which we keep β_f fixed, the effect of immigrants on imports is reduced by 8% and 25%, respectively.

In the last row of Table 4, we additionally remove all grocery expenditures associated with immigrants. In terms of the model, this corresponds to a reduction in $S_{c(h)}$, the real market size of county c. Accounting for this market size effect leads to an overall decline in import expenditure by 26%. The average loss in welfare for natives in this scenario is 0.93% or a welfare-equivalent fall of $69 per household. Thus, removing immigrants' expenditures in our counterfactual scenario leads to a non-negligible welfare effect on native households.

A key result highlighted in Table 4 is that the change in import expenditure associated with removing immigrant channels is generally larger than the associated change in *welfare*. This is due to two forces. On the one hand, three quarters of the effect of removing immigrants on import expenditure is due to the composition channel without welfare implications for natives. On the other hand, we assume a Pareto distribution of firm productivities, as is standard in Melitz-Chaney models for analytical tractability. This implies that a higher market share for domestic products due to a decrease in the entry of exporters to the U.S. is counteracted by higher entry of domestic firms, resulting in little overall change in varieties.

The average estimates discussed here mask considerable heterogeneity along three dimensions: heterogeneity by origin country, heterogeneity across physical geographies within the US, and heterogeneity across income groups within the US. We describe these three dimensions of heterogeneity now.

Immigrant effects by import origin country. We illustrate the unequal change in import volumes across origin countries associated with our first counterfactual exercise in

Appendix Figure C.4. The expenditure share on Mexican imports falls the most, as expected, by 12.87%. Mexico is followed by China, India, the Philippines and Vietnam with expenditure share decreases between 5.96% and 4.94%.

Within the context of a simple regression model at the origin level, we find that the trade-creating effects of immigrants is particularly pronounced for countries which are proximate to the US and large, in terms of population. Specifically, for every 10% reduction in distance between the US and a given origin o, the counterfactual trade-creating effect of immigrants increases in magnitude by 1.6%. For every 10% increase in population of an origin country, this same effect increases in magnitude by 0.6%.[42] When estimated alongside distance and population, the GDP per capita of origin countries has no significant correlation with the recovered trade-creating effect of immigrants.

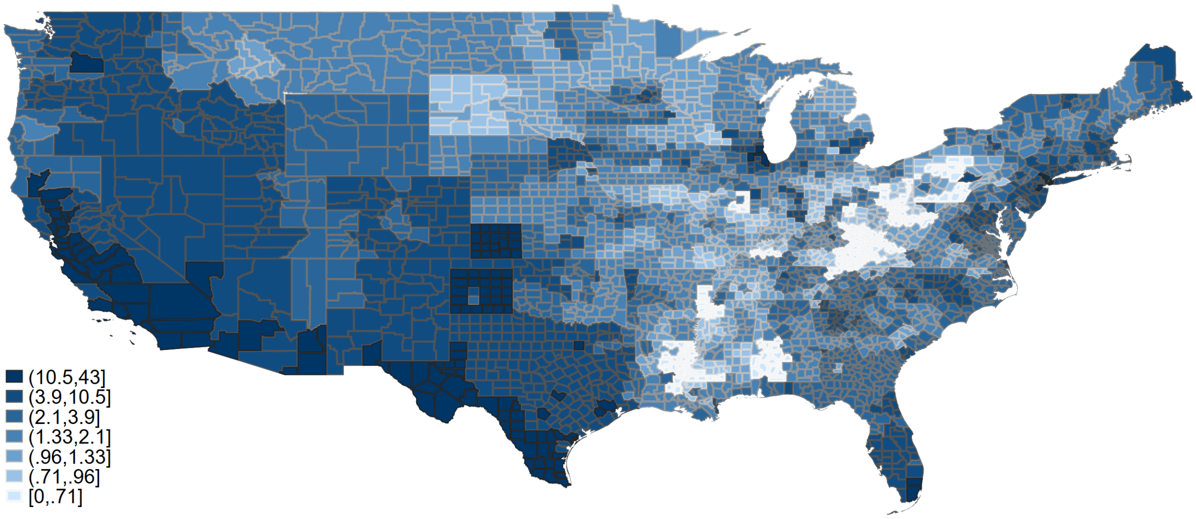
Immigrant effects across US counties. We graphically depict the substantial variation across U.S. counties in terms of the fall in import volumes associated with removing immigrant effects. Figure 3 maps the change in import expenditure share associated with our first counterfactual—removing immigrant effects only—at the county level while, in the Appendix, Figure C.5 maps the average dollar-equivalent change in utility associated with our second counterfactual: removing both immigrant effects and expenditure. In both cases, the impact of immigrants on imports and welfare is remarkably concentrated in the Southwest, West Coast, and East Coast of the United States, as well as almost all major cities.

The most affected counties with respect to imports are El Paso, TX (-34%); Los Angeles, CA (-21%); Santa Clara, CA (-21%); Kern, CA (-18%); and Riverside, CA (-18%). Assuming an initial MFN tariff rate of 2.5% applied to the grocery goods studied here, as well as our calibration of $\theta = 5$ from earlier, these estimates are equivalent to Los Angeles CA facing a county-specific increase in tariff rates to 7.4% - a three-fold increase.

In terms of annual dollar-equivalent welfare effects for large counties, the most affected are Queens, NY ($385); Dade, FL ($356); Hudson, NJ ($308); Santa Clara, CA ($291); and Los Angeles, CA ($275).

[42] While these are of course only correlations, both estimates are significant at levels of 99% within a sample of 72 countries.

Figure 3: Spatial Distribution of the Effect of Removing Immigrant Effects on Imports



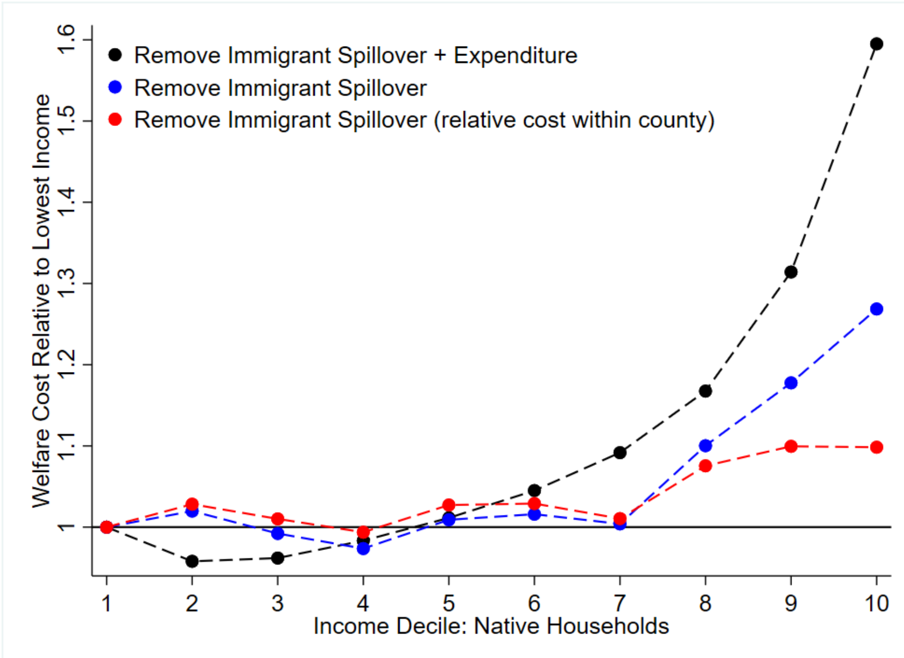
Notes: This chart plots the percent decrease in the value of grocery and personal care imports when the trade-creating effect of immigrants are removed following the procedure outlined in Appendix Section B.4.

Immigrant effects by native household income. While prior literature has emphasized the distributional consequences of immigrants in the labor market (e.g., Dustmann et al. 2013), we are the first to do so looking at the consumption side, enabled by our highly detailed household-level data and approach.

Figure 4 depicts the average relative magnitude of welfare costs associated with various counterfactual exercises across income deciles. For each counterfactual, all welfare costs are relative to the lowest income decile. The blue line depicts our first counterfactual: removing the effects of immigrants but not their expenditure. The black line depicts our second counterfactual in which we additional remove immigrant expenditure. The red line depicts the relative welfare effects of our first counterfactual but within counties, thus isolating the role of preferences in shaping the cost of this counterfactual across deciles, rather than geographic sorting into counties with more immigrants.

In all cases, there is very little difference in welfare effects associated with the first six income deciles. Almost all of the variation across income deciles occurs in the top four deciles, in which the costs of removing immigrant effects and expenditure are monotonically increasing. In our primary counterfactual, households in the wealthiest income decile face costs of losing access to immigrant-induced imports that are 25% larger than households at or below

Figure 4: Percent Change in Grocery Welfare by Income Decile



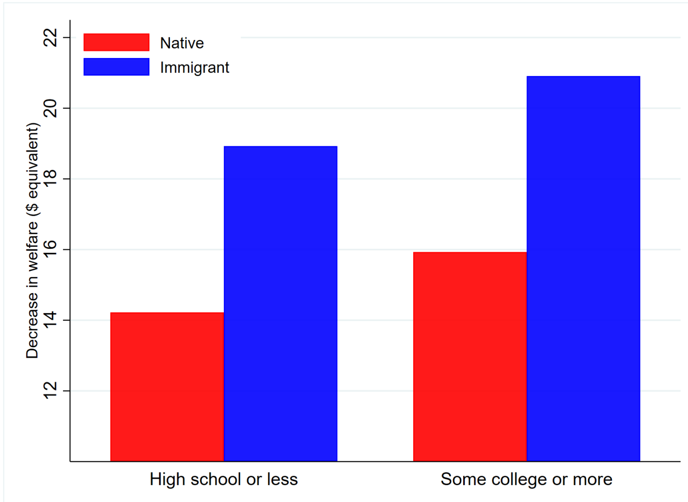
Notes: The chart depicts output from three regressions at the household level: welfare costs of removing immigrant spillovers and expenditure on income decile dummy variables (black), welfare costs of removing immigrant spillovers on income decile dummy variables (blue), and the same regression but including county-level fixed effects (red). All estimates are relative to the lowest income decile.

the seventh income decile. The red series highlights that just under half of this differential is driven by greater preferences for imported goods exhibited by high-income households, with positive geographic sorting between immigrants and high-income households explaining the remaining half.

The black line highlights that when we remove immigrants' trade effects and expenditure, the reduction in market size has far larger distributional consequences, with households in the wealthiest decile facing costs that are 60% greater than households at or below the median income level.

While the welfare estimates presented here are surely a lower-bound in that they are only relevant for grocery products, they do shed light on the remarkable variation in the consumption gains from immigrant populations across cities, counties, and income groups within the U.S. These disparities may shed light on the strong polarization across U.S. households in their attitudes towards immigrants, immigration policy, and globalization in general. We finish by discussing this point in further detail.

Figure 5: Effect of Variable Trade Cost Shock by Demographic Group



Notes: The chart depicts the counterfactual effect of a 10 percentage point increase in ad-valorem tariffs applied to τ_{od} for all o, d pairs on natives (in red) and immigrants (in blue). Effects are grouped by education, with the effect on those with high school or less on the left and those with some college or more on the right.

6.2 Differential Impact of Trade Shocks by Nativity, Education

In our final counterfactual, we consider the unequal effects of a trade shock. Inspired by the 2024 Trump Campaign's proposal to raise tariffs on all imported goods to 10 percentage points, we simulate the effect on households of an increase in the ad-valorem tariff rate applied to all imports from a base of 2.5% to 12.5%. We show the results of such a trade shock stratified by education and nativity in Figure 5. Two key findings emerge.

First, immigrants suffer much greater welfare losses than natives, regardless of how much education they have. On average, immigrants would lose the welfare-equivalent of about $20 compared to about $15 for natives. This result is driven by immigrants greater preference for imports in their consumption baskets, as documented in Section 2 and via the estimates of ζ_1 and ζ_2 discussed previously.

Second, more highly educated households lose out more from a trade shock. In particular, high-educated immigrant households are clearly the most negatively affected demographic, with welfare costs that are ∼11% greater than those faced by similarly educated native households, and almost 50% greater than less-educated native households.

Our results demonstrate that a seemingly nativity-neutral policy such as a universal tariff increase results in highly disparate impacts depending on education and nativity. Low-income, less-educated, and native-born U.S. households face the lowest consumer costs associated with policies which increase barriers to either immigration or imports. This paper therefore suggests a novel factor which may contribute to the well-documented lack of support for increased immigration among this demographic.

As suggestive evidence in support of this claim, we regress the county-average percentage of votes that the Republican party received with Donald Trump as candidate in the presidential elections of 2016 and 2020 on the immigrant-induced welfare increase based on our counterfactual in the last row of Table 4. The results in Appendix Table C.3 show that for every dollar increase in immigrant-induced welfare, the unconditional vote share received by the Republican party in 2016 decreased by 0.16 percentage points. This estimate is invariant to including state fixed effects, highlighting the usefulness of the household-level data used here. Adding further controls (log county population, native unemployment rate and immigrant population share) decreases the coefficient to 0.10, while it remains highly significant. The corresponding coefficient for the 2020 election is around 0.07 and thus somewhat lower, possibly because the immigration debate was less salient during the 2020 election campaigns.

7 Conclusion

This paper provides the first detailed decomposition of immigrant-induced import expenditure into a welfare-enhancing component and a *composition effect*. We find that the *composition effect* dominates and that within the welfare-enhancing component of immigrant-induced trade, wealthy households in large urban areas accrue the vast majority of the benefits.

A core contribution of this paper lies in separately identifying the various channels and mechanisms through which immigrants affect import expenditure. In estimating the spillover effects of immigrants on import expenditure of non-immigrant households, we make use of the leave-out push-pull instrumental variable introduced by Burchardi et al. (2019) to generate exogenous variation in origin-specific immigrant population shares across U.S. counties. By leveraging the structure inherent in the Chaney (2008) framework alongside detailed price data at the barcode level and a robust identification strategy, we are able to separately identify the effect of immigrants on price (and therefore variable trade costs), variety availability (fixed costs), and the intensive margin of demand (preferences). We are the first to provide direct evidence that among these spillover effects, the fixed cost reductions associated with immigrants are the dominant source of immigrant-induced trade.

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.

Alesina, Alberto and Marco Tabellini, "The political effects of immigration: Culture or economics?," *Journal of Economic Literature*, 2024, *62* (1), 5–46.

Amiti, Mary, Mi Dai, Robert C Feenstra, and John Romalis, "How did China's WTO entry affect US prices?," *Journal of International Economics*, 2020, *126*, 103339.

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.

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.

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*, 2023, p. rdad098.

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.

Borusyak, Kirill and Peter Hull, "Nonrandom exposure to exogenous shocks," *Econometrica*, 2023, *91* (6), 2155–2185.

_ **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.

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.

Card, David, "Immigrant inflows, native outflows, and the local labor market impacts of higher immigration," *Journal of Labor Economics*, 2001, *19* (1), 22–64.

_ , Christian Dustmann, and Ian Preston, "Immigration, wages, and compositional amenities," *Journal of the European Economic Association*, 2012, *10* (1), 78–119.

Cardoso, Miguel and Ananth Ramanarayanan, "Immigrants and exports: Firm-level evidence from Canada," *Canadian Journal of Economics*, 2022, *55* (3), 1250–1293.

Chaney, Thomas, "Distorted gravity: the intensive and extensive margins of international trade," *American Economic Review*, 2008, *98* (4), 1707–21.

Chen, Bo and David S. Jacks, "Trade, variety, and immigration," *Economics Letters*, 2012, *117* (1), 243–246.

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.

Cortes, Patricia, "The effect of low-skilled immigration on US prices: evidence from CPI data," *Journal of Political Economy*, 2008, *116* (3), 381–422.

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.

Couture, Victor, "Valuing the consumption benefits of urban density," *University of California, Berkeley, Working Paper*, 2016.

Diamond, Rebecca, "The determinants and welfare implications of US workers' diverging location choices by skill: 1980–2000," *American Economic Review*, 2016, *106* (3), 479–524.

_ and Enrico Moretti, "Where is standard of living the highest? Local prices and the geography of consumption," Technical Report, National Bureau of Economic Research 2021.

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, "The impact of immigration: Why do studies reach such different results?," *Journal of Economic Perspectives*, 2016, *30* (4), 31–56.

_ , _ , and _ , "Labor supply shocks, native wages, and the adjustment of local employment," *Quarterly Journal of Economics*, 2017, *132* (1), 435–483.

Eaton, Jonathan and Samuel Kortum, "Technology, geography, and trade," *Econometrica*, 2002, *70* (5), 1741–1779.

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.

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.

Giovanni, Julian Di, Andrei A Levchenko, and Francesc Ortega, "A global view of cross-border migration," *Journal of the European Economic Association*, 2015, *13* (1), 168–202.

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.

Handbury, Jessie, "Are poor cities cheap for everyone? Non-homotheticity and the cost of living across US cities," *Econometrica*, 2021, *89* (6), 2679–2715.

_ **and David E Weinstein**, "Goods prices and availability in cities," *Review of Economic Studies*, 2015, *82* (1), 258–296.

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

Hottman, Colin J, "Retail markups, misallocation, and store variety across US cities," 2021.

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

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," *Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4619541*, 2023.

Krugman, Paul, "Scale Economies, Product Differentiation, and the Pattern of Trade," *American Economic Review*, 1980, *70* (5), 950–59.

Lach, Saul, "Immigration and prices," *Journal of Political Economy*, 2007, *115* (4), 548–587.

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.

Mazzolari, Francesca and David Neumark, "Immigration and product diversity," *Journal of Population Economics*, 2012, *25* (3), 1107–1137.

Melitz, Marc J, "The impact of trade on intra-industry reallocations and aggregate industry productivity," *Econometrica*, 2003, *71* (6), 1695–1725.

_ **and Stephen J Redding**, "New trade models, new welfare implications," *American Economic Review*, 2015, *105* (3), 1105–1146.

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.

Ottaviano, Gianmarco IP and Giovanni Peri, "Rethinking the effect of immigration on wages," *Journal of the European Economic Association*, 2012, *10* (1), 152–197.

Ottaviano, Gianmarco I.P., Giovanni Peri, and Greg C. Wright, "Immigration, trade and productivity in services: Evidence from U.K. firms," *Journal of International Economics*, 2018, *112*, 88–108.

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.

Silva, JMC Santos and Silvana Tenreyro, "The log of gravity," *Review of Economics and Statistics*, 2006, *88* (4), 641–658.

Simonovska, Ina and Michael E Waugh, "The elasticity of trade: Estimates and evidence," *Journal of International Economics*, 2014, *92* (1), 34–50.

Su, Yichen, "Measuring the value of urban consumption amenities: A time-use approach," *Journal of Urban Economics*, 2022, *132*, 103495.

United Nations, "Global Issues: International Migration," 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.

Zachariadis, Marios, "Immigration and international prices," *Journal of International Economics*, 2012, *87* (2), 298–311.

Online Appendix

Contents

A Empirical Appendix

A.1 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 Bonadio (forthcoming), McCully (2024), and Choi et al. (2024).

The immigration leave-out push-pull instrument interacts the arrival into 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

$$IV_{o,c}^{D} = I_{o}^{D} \times \frac{I_{c}^{D}}{I^{D}},$$

where I_{o}^{D} is the number of immigrants from origin o coming to the U.S. in decade D, and I_{c}^{D}/I^{D} is the fraction of immigrants to the U.S. who choose to settle in county c in decade D.

There may still exist threats to the exogeneity of the instrument as defined thus far. These threats include 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 $I_{o,c}^{D}$ flows when constructing the instrument for the county-country pair oc.

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 migrants 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 (7).

A violation of the identification assumption may occur if, say, immigrants skilled at importing goods from France tend to settle in Chicago and immigrants skilled in importing goods from South Korea settle in 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 (7) 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 extensive robustness checks for the instrument in the same setting and find that it holds up to a battery of tests. Following Burchardi et al. (2019), we include five principal component terms which capture the variation of interactions of the instruments within county-country pairs and across decades.[43]

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 household level in Table A.1. We find that the push-pull instrument strongly and positively predicts the contemporary bilateral immigrant population.

We estimate the first-stage four ways. In columns 1 and 2, the specification is at the household-by-origin level. Since we cluster standard errors at the level of the instrumental variables—the origin-by-county level—the estimates are equivalent to a specification at the origin-by-county level but each county weighted based on the location of Nielsen households within the U.S. In column 1, we predict immigrant population shares without using information on household nativity. In column 2, we include household nativity variables. In both cases, the first-stage F-statistic is about 20 and surpasses conventional thresholds. Coefficients are always positive and typically statistically significant, with the exception of the early 20th century.

Columns 3 and 4 show estimates from data at the barcode-by-county level. We again cluster standard errors at the origin-by-county level. We again estimate an F-statistic near 20, with most coefficients 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 new cohorts of immigrant groups likely change their location choices over time.

[43]We compute 1,013 higher-order interaction terms, defined as $I^{D'}_{o,-r(c)} \times \cdots \times I^{D}_{o,-r(c)} I^{D}_{-\mathcal{C}(o),c} / I^{D}_{-\mathcal{C}(o)}$ for each $D' < D \leq 2000$. We then compute five principal components which capture the variation contained within those 1,013 terms.

Table A.1: First stage regression

	Dependent variable: Immigrants/Pop. 2010			
	(1)	(2)	(3)	(4)
$I^{1880}_{o,-r(d)} \times \frac{I^{1880}_{-c(o),d}}{I^{1880}_{-c(o)}}$	0.000063***	0.000057***	-0.00015	-0.00015
	(0.000021)	(0.000020)	(0.00015)	(0.00016)
$I^{1900}_{o,-r(d)} \times \frac{I^{1900}_{-c(o),d}}{I^{1900}_{-c(o)}}$	0.000033	0.000017	-0.00058	-0.00072
	(0.00013)	(0.00013)	(0.00072)	(0.00087)
$I^{1910}_{o,-r(d)} \times \frac{I^{1910}_{-c(o),d}}{I^{1910}_{-c(o)}}$	0.00026	0.00024	-0.00046	-0.00078
	(0.00020)	(0.00020)	(0.00048)	(0.00063)
$I^{1920}_{o,-r(d)} \times \frac{I^{1920}_{-c(o),d}}{I^{1920}_{-c(o)}}$	0.0018***	0.0018***	0.00056	0.00036
	(0.00025)	(0.00025)	(0.00070)	(0.00088)
$I^{1930}_{o,-r(d)} \times \frac{I^{1930}_{-c(o),d}}{I^{1930}_{-c(o)}}$	0.0016***	0.0016***	0.0029***	0.0031***
	(0.00017)	(0.00017)	(0.00058)	(0.00069)
$I^{1970}_{o,-r(d)} \times \frac{I^{1970}_{-c(o),d}}{I^{1970}_{-c(o)}}$	0.00086***	0.00084***	0.00084***	0.00092***
	(0.000081)	(0.000080)	(0.00023)	(0.00030)
$I^{1980}_{o,-r(d)} \times \frac{I^{1980}_{-c(o),d}}{I^{1980}_{-c(o)}}$	0.0032***	0.0032***	0.0042***	0.0047***
	(0.00028)	(0.00028)	(0.00058)	(0.00071)
$I^{1990}_{o,-r(d)} \times \frac{I^{1990}_{-c(o),d}}{I^{1990}_{-c(o)}}$	0.0023***	0.0022***	0.00093	0.0012
	(0.00025)	(0.00025)	(0.00075)	(0.00090)
$I^{2000}_{o,-r(d)} \times \frac{I^{2000}_{-c(o),d}}{I^{2000}_{-c(o)}}$	0.0015***	0.0015***	0.0015***	0.0016***
	(0.00019)	(0.00019)	(0.00029)	(0.00034)
=1 if immigrant from anywhere		0.000022		
		(0.000072)		
=1 if immigrant from origin o		0.013***		
		(0.0032)		
N	1,461,130	1,461,130	2,261,777	1,601,674
Country FE	✓	✓		
Barcode FE			✓	✓
County FE			✓	✓
Distance & latitude difference	✓	✓	✓	✓
Principal components	✓	✓	✓	✓
F-statistic	20.2	19.5	17.3	17.5
Sample	All counties	All counties	All counties	UPC in >100 counties

Notes: Columns 1 and 2 show regression results at the household-origin level with observations weighted using Nielsen household weights and standard errors clustered two-ways at the household and origin-by-county levels. Columns 3 and 4 show regression results at the barcode-county level with standard errors clustered two-ways at the barcode and origin-by-county levels. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

A.2 Robustness of Gravity Results

In this appendix we test the robustness of our main estimates in several ways.

A.2.1 Sample weights

As discussed by Feenstra et al. (2023), the Nielsen household sample may not be perfectly representative of the U.S. population in terms of income and price sensitivity. We also find that the Nielsen data is not representative of the distribution of immigrant origin countries. Mexican-born immigrants, for example, make up 15% of Nielsen households but 30% of households in the ACS. This may be driven by a combination of two factors. First, cross-sectionally Nielsen 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 2014-16.[44]

To gauge the importance of Nielsen's lack of representativeness 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 2013-2017 ACS sample. We show the results in Table A.2. Similar to our baseline results, immigrants have a positive and statistically significant effect on consumption of others on goods from the immigrants' origin. Moreover, the magnitude of the estimated coefficient increases by almost 36% to 1.59. We therefore conclude that our baseline estimates are not driven by Nielsen's lack of representativeness.

A.2.2 Alternative measures of origin country connectedness

In our baseline specification, equation 6, we allow households to have specific preferences for (i) all imports, and (ii) imports specifically from the household's origin country. Households may additionally exhibit specific preferences towards goods from countries close—geographically or culturally—to their origin country.

We test the importance of such specific households preferences in Table A.3. We start

[44]The Nielsen survey is voluntary so households may drop out at any time.

Table A.2: Gravity regressions with adjusted weights

	Dependent variable: Exp. share on goods from o relative to US (1)	(2)
Immigrants/Pop. 2010	1.40***	1.59***
	(0.25)	(0.28)
First-stage residuals		-0.26
		(0.29)
=1 if immigrant from anywhere	0.26***	0.26***
	(0.037)	(0.037)
=1 if immigrant from origin o	0.60***	0.60***
	(0.086)	(0.087)
N	1,461,130	1,461,130
Country FE	✓	✓
Household controls	✓	✓
Distance & latitude difference	✓	✓
1st-stage F-statistic		20.2

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 county-country levels. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

by exploring whether immigrant households exhibit a specific preference for goods from their continent of origin (in addition to their country of origin) in column 1. We find that immigrant households spend 12% more on goods imported from their origin continent. The coefficient on the country-of-origin dummy falls slightly to 0.54 from 0.6.

We also consider whether non-immigrants with ancestry from a given origin region exhibit a specific preference for goods from that region. While we cannot observe ancestry for every household in the Nielsen HomeScanner data, we do observe Hispanic ethnicity. We leverage this variable to assess whether Hispanic households exhibit greater demand for foods imported from Latin America in column 2. We find a positive but statistically insignificant relationship between Hispanic background and demand for imports from Latin America.

Immigrant households may prefer goods from similar cultures, or from countries with similar colonial backgrounds, not merely countries geographically proximate to their origin. We test whether such cultural or colonial history characteristics affect immigrant product demand in column 3. We proxy for cultural similarity between an immigrant's origin country and an import origin country with a dummy for the two foreign countries sharing the same

Table A.3: Gravity regressions with additional preference terms

	Dependent variable: Exp. share on goods from o relative to US			
	(1)	(2)	(3)	(4)
Immigrants/Pop. 2010	1.19***	1.14***	1.19***	1.17***
	(0.24)	(0.24)	(0.24)	(0.24)
First-stage residuals	0.17	0.17	0.17	0.15
	(0.31)	(0.31)	(0.31)	(0.31)
=1 if immigrant from anywhere	0.20***	0.23***	0.22***	0.20***
	(0.035)	(0.030)	(0.033)	(0.037)
=1 if immigrant from origin o	0.52***	0.59***	0.62***	0.54***
	(0.085)	(0.069)	(0.072)	(0.085)
=1 if immig. from continent of o	0.12*			0.092
	(0.066)			(0.063)
=1 if hispanic and o in Latin America		0.073		0.075
		(0.059)		(0.057)
=1 if common official or primary language			0.026	0.0051
			(0.057)	(0.058)
=1 if ever in colonial or dependency relationship			0.11	0.14
			(0.10)	(0.10)
=1 if currently in colonial or dependency relationship			0.35	0.37
			(0.84)	(0.83)
N	1,461,130	1,461,130	1,460,552	1,460,552
Country FE	✓	✓	✓	✓
Household controls	✓	✓	✓	✓
Distance & latitude difference	✓	✓	✓	✓
1st-stage F-statistic	19.5	19.5	19.5	19.4

Notes: The table presents regression results at the household-country level. Observations weighted using NielsenIQ household weights. The dummies indicating common language or colonial relationship are taken from the CEPII Gravity Database (Conte et al. 2022) with country pairs being based on household origin and import expenditure origin country. Standard errors clustered two-ways at the household and county-country levels. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

language as measured by Conte et al. (2022). We measure current or past colonial relationships also using the database of Conte et al. (2022). We find no statistically significant effect of a shared language or colonial relationships on product demand.

Finally, we include all aforementioned variables in a single specification in column 4 and find no statistically significant effect of any variable on imported product demand.

A.2.3 Extensive and intensive margin

Do immigrants induce natives to purchase more intensively from immigrants' origin countries or to drive diversification of product origins for native households? We answer this question in Table A.4. We first estimate equation (6), replacing the dependent variable with a dummy for whether household h makes any purchases of imports from o in columns 1 and 2. We separately estimate equation (6) while dropping all observations with household-by-origin expenditures equal to 0 in columns 3 and 4.

Table A.4: Extensive and Intensive Margin of Household Import Expenditures

	=1 if $X_{oh} > 0$		$\ln(\tilde{X}_{oh})$	
	(1)	(2)	(3)	(4)
Immigrants/Pop. 2010	0.12	-0.41**	1.35***	0.60**
	(0.079)	(0.20)	(0.23)	(0.27)
=1 if immigrant from anywhere	0.012***	0.013***	0.20***	0.20***
	(0.0037)	(0.0037)	(0.025)	(0.025)
=1 if immigrant from origin o	0.18***	0.19***	0.59***	0.61***
	(0.020)	(0.021)	(0.070)	(0.071)
N	1,540,110	1,540,110	270,869	270,869
Country FE	✓	✓	✓	✓
Household controls	✓	✓	✓	✓
Distance & latitude difference	✓	✓	✓	✓
1st-stage F-statistic		19.8		11.9

Notes: The table presents regression results at the household-country level. The dependent variable in columns 1 and 2 is a dummy for whether the household spends a positive amount on goods imported from o. The dependent variable in columns 3 and 4 is the log of the relative expenditure share $\tilde{X}_{oh}$, dropping all 0s. Columns 1 and 3 display OLS results, while columns 2 and 4 display two-stage least squares results. Observations weighted using Nielsen household weights. Standard errors clustered two-ways at the household and county-country levels. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

We obtain two results. First, we estimate that the spillover effect of immigrants onto natives is driven primarily by natives more intensively purchasing goods from local immigrants' origin countries, as indicated by the first row. In column 2, we surprisingly find that a 1 percentage point higher share of immigrants from a given origin living locally actually reduces the likelihood of households purchasing goods from that origin by 0.4 percentage points. In contrast, an identical increase in the immigrant population share from a given origin country increases the expenditure share on goods from that origin by 0.6 percentage points, conditional on exhibiting positive expenditures otherwise (column 4).

Our second main finding from Table A.4 is that immigrants' preferences for imports (rows 2 and 3) affect both the likelihood and intensity of purchasing imported products.

A.2.4 Controlling for instrumental variable mean

In a recent paper, Borusyak and Hull (2023) argue that in a formula instrument combining variation from different sources, some units may be more exposed to exogenous variation than other units due to, for example, fixed geographic features. In our context, combining the push and pull shocks to construct our push-pull instrumental variable, one might worry that some origins typically send large numbers of immigrants to the U.S. and some counties typically receive large numbers of immigrants, both due to factors such as geography. To address this issue, we follow the advice of Borusyak and Hull (2023) and compute the average instrument value across each permutation of push and pull shocks within country-county pair.[45]

When controlling for the instrument mean, the coefficient on the immigrant population share falls from 1.17 to 0.56 while remaining statistically significantly different from 0. We note, however, that while the mean instrument term may capture potentially endogenous geographic factors as outlined above, it may also absorb some desirable, exogenous variation in immigration across county-country pairs. For this reason we do not control for the mean instrument in our baseline estimation. In particular, immigrants may be consistently drawn to a particular metro area due to factors that do not affect grocery sector imports.

B Theory Appendix

B.1 Deriving Adjusted ACR Welfare Formula

Begin by noting that given our assumption of constant expenditure, we can characterize the relationship between changes in domestic and import expenditure at the county level. We

[45]In practice, we randomly reshuffle push and pull shocks across years according to a uniform distribution. Given the 10 decades for which we can construct our instrument, we therefore generate 100 potential instruments (10 of them factual and 90 counterfactual) for each county-country pair and take the arithmetic mean.

denote the rest of the world with the subscript m.

$$d\ln X_c = 0 \quad \implies \quad s_{us,c} d\ln X_{us,c} = -s_{m,c} d\ln X_{m,c} = -(1 - s_{us,c}) d\ln X_{m,c}$$

The derivation of this result begins with the following observation:

$$X_c = X_{us,c} + X_{m,c} \implies d\ln X_c = d\ln(X_{us,c} + X_{m,c})$$

Notice that for any variable x, $d\ln x = dx\frac{1}{x}$. We can therefore re-write the final expression above as:

$$d\ln X_c = \frac{d(X_{us,c} + X_{m,c})}{X_{us,c} + X_{m,c}} = \frac{d(X_{us,c} + X_{m,c})}{X_c} = \frac{dX_{us,c}}{X_c} + \frac{dX_{m,c}}{X_c}$$

Invoking again the transformation between $d\ln x$ and dx, we can write this expression as:

$$d\ln X_c = \frac{X_{us,c} d\ln X_{us,c}}{X_c} + \frac{X_{m,c} d\ln X_{m,c}}{X_c} = s_{us,c} d\ln X_{us,c} + s_{m,c} d\ln X_{m,c}$$

The last step is to assume that $d\ln X_c = 0$ and $s_{us,c} + s_{m,c} = 1$. Notice that the exact same steps can be used to derive the same expression for native households, although in this case we want to focus on welfare-relevant changes in domestic expenditure. Formally, this would be the component of changes in domestic expenditure which are due to changes in supply-side accessibility of imports associated with immigrants: $d\ln \phi^d_{mc}$.

The welfare relevant change in domestic expenditure for native households is therefore characterized by the following:

$$s_{us,n} d\ln X_{us,n} = -s_{m,n} d\ln X_{m,n} \frac{d\ln \phi^d_{mc}}{d\ln X_{m,n}} = -(1 - s_{us,n}) d\ln X_{m,n} \frac{d\ln \phi^d_{mc}}{d\ln X_{m,n}}$$

We can use the county-level aggregate expression to relate welfare-relevant changes in native household domestic expenditure to the county-level aggregate change in the following way:

$$d\ln X_{us,n} = d\ln X_{us,c} \left(\frac{1 - s_{us,n}}{s_{us,n}}\right) \left(\frac{s_{us,c}}{1 - s_{us,c}}\right) \left(\frac{d\ln X_{m,n}}{d\ln X_{m,c}}\right) \left(\frac{d\ln \phi^d_{mc}}{d\ln X_{m,n}}\right) \tag{B.1}$$

Deriving the Expenditure Share Adjustment Term. We begin by noting the following identities for county-level expenditure shares:

$$s_{us,c} = (1 - I_c)s_{us,n} + I_c s_{us,f} \qquad s_{m,c} = (1 - I_c)s_{m,n} + I_c s_{m,f}$$

We also make note that due to our assumption regarding the structure of preferences: $s_{m,f}/s_{m,n} = e^{\zeta}$. Given that expenditure shares must sum to one, it is trivially true that for immigrants $s_{us,f} = 1 - s_{m,f} = 1 - e^{\zeta}s_{m,n}$. For native households: $s_{us,n} = 1 - s_{m,n}$. Combining these two expressions, we derive:

$$e^{\zeta}(1 - s_{us,n}) = 1 - s_{us,f} = 1 - \Big[\frac{s_{us,c} - s_{us,n}(1 - I_c)}{I_c}\Big]$$

We can therefore solve for $s_{us,n}$—the native domestic expenditure share—as a function of $s_{us,c}$, I_c, and ζ by re-arranging the previous expression:

$$s_{us,n} = \frac{s_{us,c} + I_c(e^{\zeta} - 1)}{I_c(e^{\zeta} - 1) + 1}$$

Plugging this definition into the term of interest in our main equation of interest ($d \ln X_{us,n}$), we can derive our final expression:

$$\left(\frac{1 - s_{us,n}}{s_{us,n}}\right)\left(\frac{s_{us,c}}{1 - s_{us,c}}\right) = \frac{1}{\frac{I_c}{s_{us,c}}(e^{\zeta} - 1) + 1} \tag{B.2}$$

Deriving the Welfare-Relevant Component of Trade Shocks. Notice that the final two terms of the equation derived above reduce to the ratio: $\frac{d \ln \phi^d_{mc}}{d \ln X_{m,c}}$. We have an explicit expression for this ratio from the model:

$$\frac{d \ln \phi^d_{mc}}{d \ln X_{m,c}} = \frac{\beta^d}{\beta^d + \beta^z + \frac{e^{\zeta}-1}{I_c(e^{\zeta}-1)+1}} = \frac{\beta^d}{\beta + \frac{e^{\zeta}-1}{I_c(e^{\zeta}-1)+1}}$$

where the fraction in the denominator derives from evaluating:

$$\frac{d \ln[\sum \kappa_h z_{oh}]}{dI_c} = \frac{e^{\zeta} - 1}{I_c(e^{\zeta} - 1) + 1}$$

This derivation relies on the assumption that all households are identical except for their immigrant-derived preferences governed by ζ, so therefore $\kappa_h = \kappa$, and $J_h = J$ for all households. We also assume that with enough households, the average idiosyncratic component of preferences η^z_{oh} is equal to zero in expectation.

We can therefore express the average preference term $\sum \kappa_h z_{oh}$ as:

$$\sum_{h\in\Lambda_c} \kappa_h z_{oh} = \kappa e^{\delta J} \sum_{h\in\Lambda_c} e^{\zeta}$$

Since $\zeta = 0$ for any household that is not an immigrant, this expression further reduces to:

$$\sum_{h\in\Lambda_c} \kappa_h z_{oh} = \kappa e^{\delta J}[(1 - I_c) + I_c e^{\zeta}] = \kappa e^{\delta J}[I_c(e^{\zeta} - 1) + 1]$$

As a final step, we can derive the partial elasticity of this term with respect to I_c in order to show that:

$$\frac{d\ln[\sum \kappa_h z_{oh}]}{dI_c} = \frac{d[\sum \kappa_h z_{oh}]}{dI_c} \frac{1}{\sum \kappa_h z_{oh}} = \frac{\kappa e^{J}(e^{\zeta} - 1)}{\kappa e^{\delta J}[I_c(e^{\zeta} - 1) + 1]} = \frac{e^{\zeta} - 1}{I_c(e^{\zeta} - 1) + 1}$$

Notice that so long as $\zeta > 0$, this expression is positive and the composition effect of immigrants has a positive effect on county-level import expenditure. As $\zeta \uparrow$, this effect intensifies, as the derivative of the composition effect with respect to ζ is simply e^{ζ}.

B.2 Deriving Heterogeneous Firms Model Equations

Deriving equations (11) and (12). We start by deriving county-level expenditures on a variety supplied by a firm with productivity φ and imported from country o, $x_{oc}(\varphi)$.

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

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

Define

$$P_h \equiv \left(\sum_{o\in\mathcal{O}} (z_{oh})^{\sigma} \int_{\omega\in\Omega_{o,c(h)}} p_{o,c(h)}(\omega)^{1-\sigma} d\omega\right)^{\frac{1}{1-\sigma}} \tag{B.3}$$

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

$$(1-\mu_0)X_h = z_{oh}^{-\sigma} q_{oh}(\omega) p_{o,c(h)}(\omega)^{\sigma} P_h^{1-\sigma} \tag{B.4}$$

Solving for q_{oh}, we get quantity and expenditure for a variety associated with productivity φ as

$$q_{oh}(\varphi) = (1-\mu_0)X_h z_{oh}^{\sigma} p_{o,c(h)}(\varphi)^{-\sigma} P_h^{\sigma-1} \tag{B.5}$$

$$x_{oh}(\varphi) = (1-\mu_0)X_h z_{oh}^{\sigma} (p_{o,c(h)}(\varphi)/P_h)^{1-\sigma} \tag{B.6}$$

From the firm's profit maximization problem, we obtain the price equation

$$p_{o,c(h)}(\varphi) = \frac{\sigma}{\sigma-1}\frac{w_0}{\varphi}\tau_{oc(h)} \tag{B.7}$$

Substituting this expression in the equation for $x_{oh}(\varphi)$, 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 expression for expenditure $x_{oc}(\varphi)$ given by (B.8):

$$x_{oc}(\varphi) = \lambda_1 (w_o\tau_{oc})^{1-\sigma}\varphi^{\sigma-1}\Big(\sum_{h'\in\Lambda_c} z_{oh'}^{\sigma} X_{h'} P_{h'}^{\sigma-1}\Big) \tag{B.8}$$

To derive the productivity cutoff term φ^*_{oc}, we start by deriving variable profits earned by a firm with productivity φ selling to market c from origin o:

$$\begin{aligned}
\pi_{o,c}(\varphi) &\equiv \left(p_{o,c}(\varphi) - \frac{w_o}{\varphi}\tau_{o,c}\right)\sum_{h'\in c} q_{oh}(\omega(\varphi)) \\
&= (1-\mu_0)\left(\frac{w_o}{\varphi}\tau_{o,c}\right)^{1-\sigma}\frac{1}{\sigma}\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}\sum_{h'\in c}(z_{oh'})^{\sigma} W_{h'}(P_{h'})^{\sigma-1} \\
&= \frac{1}{\sigma}x_{oc}(\varphi)
\end{aligned}$$

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:

$$\pi_{oc}(\varphi) \geq f_{oc}$$

At the cutoff productivity, this holds with inequality, resulting in equation (B.9) for φ^*_{oc}, where $\lambda_2 = \frac{\sigma}{\sigma-1}\left(\frac{\sigma}{1-\mu_0}\right)^{\frac{1}{\sigma-1}}$:

$$\varphi^*_{oc} = \lambda_2 w_o \tau_{oc} \left(\frac{f_{oc}}{\sum\limits_{h' \in \Lambda_c} z^{\sigma}_{oh'} X_{h'} P^{\sigma-1}_{h'}} \right)^{\frac{1}{\sigma-1}} \tag{B.9}$$

Returning to equation (B.3) and replacing varieties ω with productivity φ (since firms with identical productivity charge identical prices), we get:

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

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

Plugging in our equilibrium price function $p_{o,c(h)}(\omega)$, we have

$$P_h = \frac{\sigma}{\sigma-1} \left(\sum_{o \in \mathcal{O}} (z_{oh})^{\sigma} \left(w_o \tau_{o,c(h)} \right)^{1-\sigma} M_{o,c(h)} \int_0^{+\infty} \varphi^{\sigma-1} g_{o,c(h)}(\varphi) d\varphi \right)^{\frac{1}{1-\sigma}} \tag{B.10}$$

We derive the gravity equation using the expression for $x_{oh}(\varphi)$ as

$$X_{oh} = \int_{\omega \in \Omega_{o,c(h)}} x_{oh}(\omega) d\omega = (1-\mu_0) z^{\sigma}_{oh} X_h P^{\sigma-1}_h \int_{\omega \in \Omega_{o,c(h)}} p_{o,c(h)}(\omega)^{1-\sigma} d\omega$$

Given the equilibrium price (B.7), 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 $g_o(\varphi) = \theta/\varphi^{\theta+1}$ to obtain

$$X_{oh} = \lambda_1 z_{oh}^{\sigma} X_h P_h^{\sigma-1} (w_o \tau_{o,c(h)})^{1-\sigma} M_o \frac{\theta}{\theta + 1 - \sigma} (\varphi_{o,c}^{*})^{\sigma-\theta-1} \tag{B.11}$$

To obtain equation (11) from (B.10) and equation (12) from (B.11), we perform the following operations:

- substitute (B.9) for $\varphi_{o,c}^{*}$
- assume $M_o = \gamma Y_o$
- define $S_c = \sum_{h' \in \Lambda_c} X_{h'} P_{h'}^{\sigma-1}$ and $z_{oc} = \sum_{h' \in \Lambda_c} z_{oh'}^{\sigma} \frac{X_{h'} P_{h'}^{\sigma-1}}{S_c}$
- define $\lambda_3 \equiv \gamma \left(\frac{\sigma}{1-\mu_0}\right)^{\frac{\sigma-\theta-1}{\sigma-1}} \left(\frac{\sigma}{\sigma-1}\right)^{\sigma-\theta-1} \frac{\theta}{\theta+1-\sigma}$
- define $\lambda_4 \equiv \gamma(1-\mu_0)^{\frac{\theta}{\sigma-1}} \sigma^{\frac{\sigma-\theta-1}{\sigma-1}} \left(\frac{\sigma}{\sigma-1}\right)^{-\theta} \frac{\theta}{\theta+1-\sigma}$

B.3 Deriving Expenditure and Extensive Margin Immigrant Elasticity

In this section we fully differentiate Equation (14) 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.

We begin by assuming that $\beta^\tau \approx 0$, which implies that immigrants do not affect variables trade costs. This assumption derives from the results discussed in Table 3.

We follow Chaney (2008) and fully differentiate $\tilde{X}_{oh}$ from Equation (14) into terms associated with fixed costs f_{oc}, county-level preferences z_{oc}, and household-level preferences z_{oh}. By applying Leibniz Rule, we can separate each term into both an intensive margin and extensive margin, and the full differentiation is given by the following:

$$
\begin{aligned}
d\tilde{X}_{oh} = & \Bigg[\int\limits_{\bar{\varphi}_{oc}}^{+\infty} \frac{\partial \tilde{x}_{oh}(\varphi)}{\partial \tilde{f}_{oc}} dG(\varphi) - \tilde{x}_{oh}(\bar{\varphi}_{oc}) G'(\bar{\varphi}_{oc}) \frac{\partial \bar{\varphi}_{oc}}{\partial \tilde{f}_{oc}}\Bigg] d\tilde{f}_{oc} \\
& + \Bigg[\int\limits_{\bar{\varphi}_{oc}}^{+\infty} \frac{\partial \tilde{x}_{oh}(\varphi)}{\partial z_{oh}} dG(\varphi) - \tilde{x}_{oh}(\bar{\varphi}_{oc}) G'(\bar{\varphi}_{oc}) \frac{\partial \bar{\varphi}_{oc}}{\partial z_{oh}}\Bigg] dz_{oh} \\
& + \Bigg[\int\limits_{\bar{\varphi}_{oc}}^{+\infty} \frac{\partial \tilde{x}_{oh}(\varphi)}{\partial z_{oc}} dG(\varphi) - \tilde{x}_{oh}(\bar{\varphi}_{oc}) G'(\bar{\varphi}_{oc}) \frac{\partial \bar{\varphi}_{oc}}{\partial z_{oc}}\Bigg] dz_{oc}
\end{aligned} \tag{B.12}
$$

In all cases, 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 aggregate expenditure on U.S. goods by h - is given by the following:

$$
\tilde{x}_{oh}(\varphi) = (\tilde{\omega}_o \tau_{oc})^{1-\sigma} z_{oh} \varphi^{\sigma-1} \Big(\int\limits_{\bar{\varphi}_{us,c}}^{+\infty} \varphi^{\sigma-1} dG(\varphi) \Big)^{-1} \tag{B.13}
$$

whereas the productivity cut-off expression is the following:

$$
\bar{\varphi}_{oc} = \lambda_2 \omega_o \tau_{oc} \left(\frac{f_{oc}}{S_c z_{oc}} \right)^{\frac{1}{\sigma-1}} \tag{B.14}
$$

It is clear from inspecting Equation B.13 and Equation B.14 that f_{oc} and z_{oc} only affect $\bar{\phi}_{oc}$, and therefore the 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: $\frac{\partial \tilde{x}_{oh}(\varphi)}{\partial \tilde{f}_{oc}} = 0$; $\frac{\partial \tilde{x}_{oh}(\varphi)}{\partial z_{oc}} = 0$; and $\frac{\partial \bar{\phi}_{oc}}{\partial z_{oh}} = 0$.

We therefore have an expression for the aggregate 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. These expressions are given by, respectively:

$$
\frac{\partial \ln \tilde{X}_{oh}}{\partial I_{oc}} = \frac{\partial \ln \tilde{X}_{oh}}{\partial \ln f_{oc}} \frac{\partial \ln f_{oc}}{\partial I_{oc}} + \frac{\partial \ln \tilde{X}_{oh}}{\partial \ln z_{oh}} \frac{\partial \ln z_{oh}}{\partial I_{oc}} + \frac{\partial \ln \tilde{X}_{oh}}{\partial \ln z_{oc}} \frac{\partial \ln z_{oc}}{\partial I_{oc}} \tag{B.15}
$$

$$
\frac{\partial \ln \tilde{N}_{oh}}{\partial I_{oc}} = \frac{\partial \ln \tilde{X}_{oh}}{\partial \ln f_{oc}} \frac{\partial \ln f_{oc}}{\partial I_{oc}} + \frac{\partial \ln \tilde{X}_{oh}}{\partial \ln z_{oc}} \frac{\partial \ln z_{oc}}{\partial I_{oc}} \tag{B.16}
$$

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:

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

$$\begin{aligned}\beta^N &= \frac{\partial \ln \tilde{N}_{oh}}{\partial I_{oc}} - \frac{\partial \ln \mathcal{Z}_{oh}}{\partial I_{oc}} \\ &= \frac{\partial \ln \tilde{X}_{oh}}{\partial \ln f_{oc}}\frac{\partial \ln f_{oc}}{\partial I_{oc}} + \frac{\partial \ln \tilde{X}_{oh}}{\partial \ln z_{oc}}\frac{\partial \ln z_{oc}}{\partial I_{oc}} - \frac{\partial \ln \mathcal{Z}_{oh}}{\partial \ln \bar{z}_{oh}}\frac{\partial \ln z_{oh}}{\partial I_{oc}} - \frac{\partial \ln \mathcal{Z}_{oh}}{\partial \ln \bar{z}_{oc}}\frac{\partial \ln \bar{z}_{oc}}{\partial I_{oc}}\end{aligned} \tag{B.18}$$

While these expressions seem daunting, they are trivial to evaluate given the definition of $\tilde{X}_{oh}$ provided in Equation 14 and the definition of $\mathcal{Z}_{oh}$ provided earlier.

Specifically, we can solve these three expression in three parts:

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

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

$$\beta^N = \frac{\partial \ln \tilde{N}_{oh}}{\partial I_{oc}} = \beta^f + \left(\frac{\theta}{\sigma - 1} - 1\right)\beta^z \tag{B.20}$$

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.

To obtain an expression for the change in household-origin import expenditures, we start with equation (12), and express the ratio of counterfactual to observed household-level imports from o as

$$\hat{X}_{oh} = \hat{P}_h^{\sigma-1} \hat{f}_{o,c(h)}^{-\left(\frac{\theta}{\sigma-1}-1\right)} \left(\hat{z}_{o,c(h)}\hat{S}_{c(h)}\right)^{\frac{\theta}{\sigma-1}-1} \hat{z}_{oh} \tag{B.21}$$

where changes in household imports by origin depend on the change in overall price level, changes in fixed costs with the origin, changes in local market demand for the origin's products, and changes in average household-level preferences. When o is the United States, equation (B.21) reduces to

$$\hat{X}_{us,h} = \hat{P}_h^{\sigma-1} \hat{S}_{c(h)}^{\frac{\theta}{\sigma-1}-1} \tag{B.22}$$

Hence we use equations (B.21) and (B.22) as well as the fact that $\hat{f}_{o,c(h)}^{-\left(\frac{\theta}{\sigma-1}-1\right)} = e^{-\hat{\beta}^f I_{o,c(h)}}$ and $\hat{z}_{oh} = e^{-\hat{\beta}^z I_{o,c(h)}}$ to obtain our counterfactual ratio as a function of observable or calibrated values:

$$\frac{X'_{oh}}{X'_{us,h}} = \frac{X_{oh}}{X_{us,h}} \left(e^{-I_{o,c(h)}(\hat{\beta}^f + \hat{\beta}^z)}\right) z_{o,c(h)}^{\left(\frac{\theta}{\sigma-1}-1\right)} \tag{B.23}$$

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

Lastly, while it is simple to show that under CES preferences, the change in welfare is given by the change in the price index, we must be careful to distinguish between welfare-relevant components of the price index and components of the price index which are not welfare relevant. To begin with, changes in utility are generally given by:

$$\hat{U}_h = \hat{P}_h^{\mu_0 - 1} \tag{B.24}$$

Notice, however, that P_h includes changes in preferences associated with β^z, which are not welfare-relevant. We therefore compute the change in the welfare-relevant price index as follows:

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

in which we purge the price index relevant for firms, and therefore containing β^z, in order to isolate the welfare-relevant component of changes in the price index.

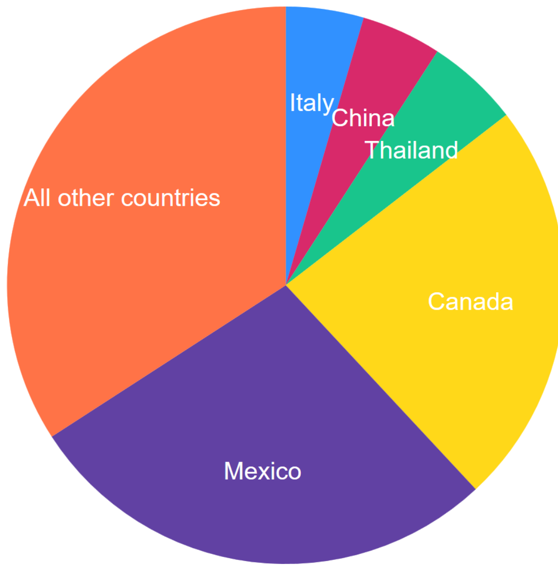
We further assume that immigrant and native households spend the same amount on grocery and personal care produces, which implies that

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

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

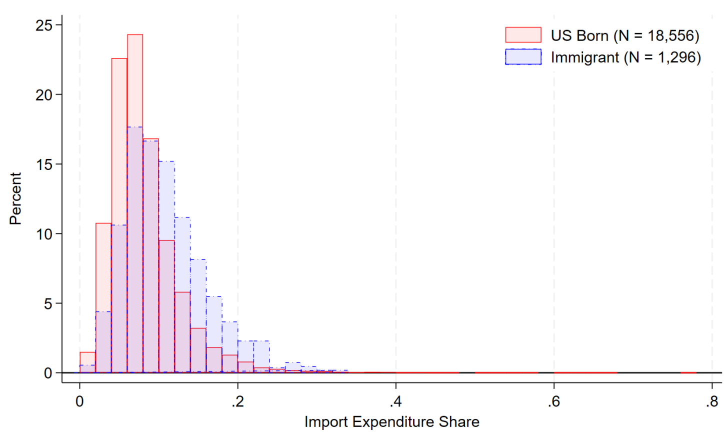
C Additional Tables and Charts

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.

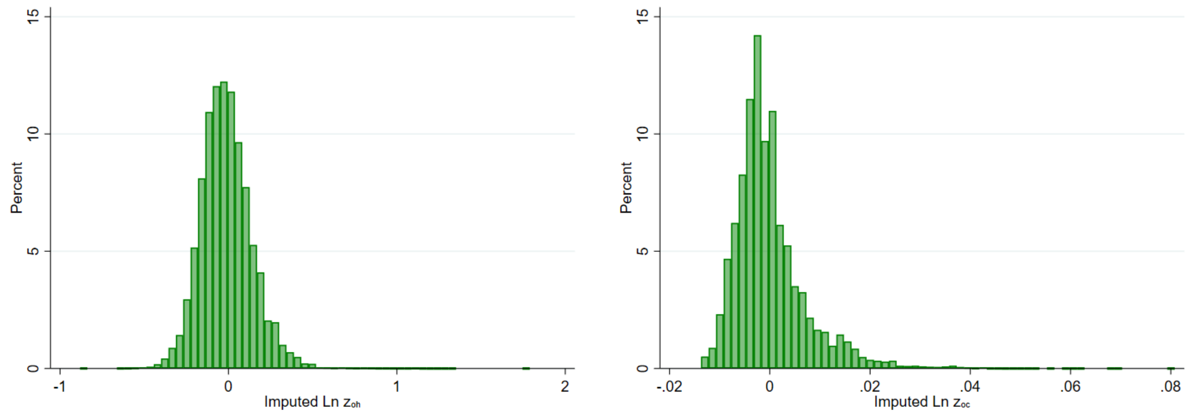
Figure C.2: Distribution of Household-level Import Expenditure Share by Nativity



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.

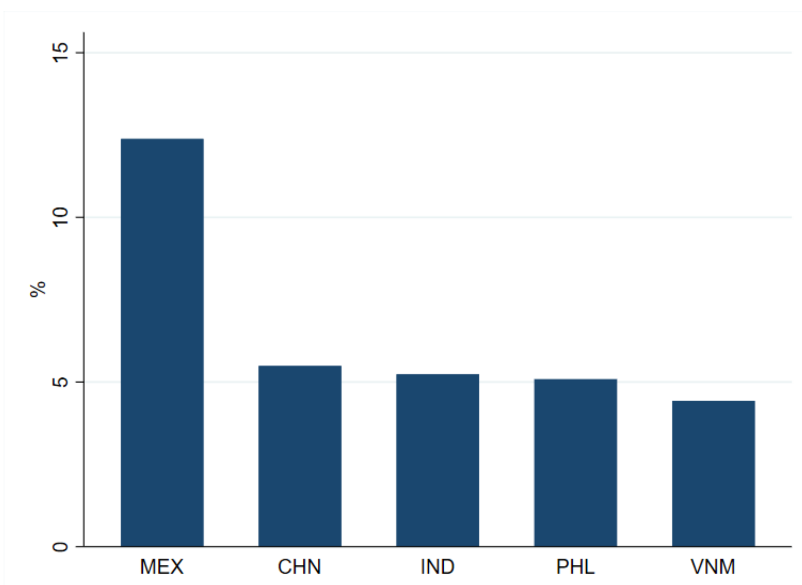
Figure C.3: Distribution of Imputed Preference Terms

(a) Distribution of $\ln \hat{\hat{z}}^{\sigma}_{oh}$ (b) Distribution of $\ln \hat{\hat{z}}_{oc}$



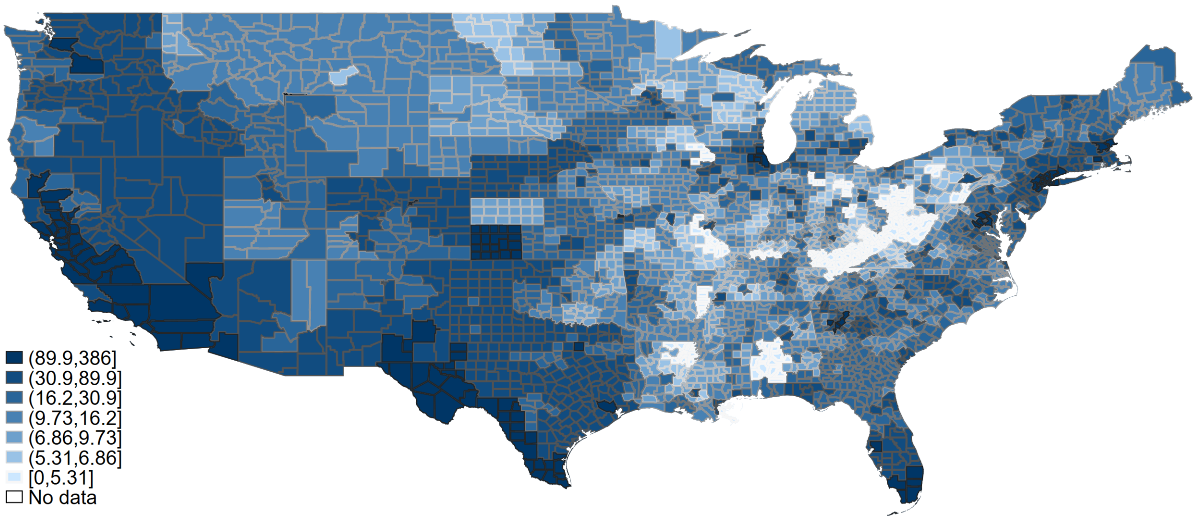
Notes: Figure (a) plots the distribution across NielsenIQ household-origin pairs of the log of $\hat{\hat{z}}_{oh} = \exp\left(\hat{\delta} J_h + \hat{\zeta}_1 \mathbf{1}[o(h) \neq US] + \hat{\zeta}_2 \mathbf{1}[o(h) = o]\right)$, where the terms $\hat{\delta}$, $\hat{\zeta}_1$, and $\hat{\zeta}_2$ are estimated from equation (19). Figure (b) plots the distribution across county-origin pairs of the log of $\hat{\hat{z}}_{oc} = \sum_{h' \in \Lambda_c} \hat{\hat{z}}_{oh'} \kappa_{h'}$, computed using data from the 2012-2017 American Community Survey.

Figure C.4: Most Impacted Origins under Baseline Counterfactual



Notes: This chart shows the percent increase in imports by origin attributable to the presence of immigrants. We compute imports under our counterfactual scenario as discussed in Appendix Section B.4.

Figure C.5: Spatial Distribution of the Effect of Removing Immigrants on Welfare



Notes: This chart plots the dollar decrease in the dollar-equivalent grocery welfare the trade-creating effect of immigrants and immigrant expenditure are removed following the procedure outlined in Appendix Section B.4.

Table C.1: Relationship between Import Expenditure Shares and Immigrant Status

	Dependent variable: Import expenditure share					
	(1)	(2)	(3)	(4)	(5)	(6)
=1 if immigrant	0.028***	0.031***	0.023***	0.027***	0.024***	0.028***
	(0.0018)	(0.0027)	(0.0017)	(0.0026)	(0.0017)	(0.0026)
N	19,700	19,700	19,107	19,107	19,107	19,107
County fixed effects			✓	✓	✓	✓
Household controls					✓	✓
Weighted		✓		✓		✓

Notes: The table presents regression results at the household level. Standard errors are clustered at the county level. Household controls are income bins, household size, marital status, and household head age and gender. Sample drops when including county fixed effects due to the 593 households living in a county with no other Nielsen panelists in our sample. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table C.2: Effect of Household Characteristics on Import expenditure

	Dep. var.: Rel. expenditure share on goods from o (1)	
Immigrant from o	0.64***	(0.069)
Immigrant from anywhere	0.23***	(0.029)
Income: 10k-30k	0.031	(0.042)
Income: 30k-50k	0.011	(0.040)
Income: 50k-70k	0.074*	(0.042)
Income: 70k-100k	0.063	(0.042)
Income: >100k	0.18***	(0.043)
HH size: 2	-0.073**	(0.029)
HH size: 3	-0.10***	(0.033)
HH size: 4	-0.19***	(0.041)
HH size: >4	-0.19**	(0.085)
Children: 6-12 y.o.	-0.087	(0.088)
Children: 13-17 y.o.	-0.10	(0.092)
Children: <6 + 6-12	-0.11	(0.10)
Children: <6 + 13-17	-0.051	(0.16)
Children: 6-12 + 13-17	-0.056	(0.096)
Children: All Age Groups	-0.26**	(0.12)
No Children	-0.070	(0.084)
Some College	0.064***	(0.023)
College Degree	0.097***	(0.024)
Postgraduate Degree	0.18***	(0.027)
Widowed	0.0043	(0.036)
Divorced/Separated	-0.0026	(0.034)
Single	-0.021	(0.034)
Black	0.058**	(0.024)
Asian	0.075**	(0.035)
Other	0.097**	(0.040)
Hispanic	-0.036	(0.034)
Age	-0.018	(0.032)
Age^2	0.00025	(0.00054)
Age^3	-0.00000087	(0.0000029)
N	868,261	
County-origin FE	✓	

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 C.3: Immigrant-induced welfare increase and voting outcomes in 2016 and 2020

	Dependent variable: Republican votes (%)			
	(1)	(2)	(3)	(4)
Welfare change ($)	-0.16***	-0.17***	-0.10***	-0.069***
	(0.0085)	(0.0091)	(0.018)	(0.018)
Log population			-4.33***	-4.96***
			(0.18)	(0.18)
Native unemployment rate			-101.5***	-80.5***
			(12.0)	(12.0)
Immigrant share			-8.48	-14.7
			(9.07)	(9.03)
N	3,038	3,038	3,038	3,038
Election year	2016	2016	2016	2020
State FE		✓	✓	✓

Notes: The table presents regression results at the county level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Online Appendix References

Bonadio, Barthélémy, "Migrants, trade and market access," *Review of Economics and Statistics*, forthcoming.

Borusyak, Kirill and Peter Hull, "Nonrandom exposure to exogenous shocks," *Econometrica*, 2023, *91* (6), 2155–2185.

Burchardi, Konrad B, Thomas Chaney, and Tarek A Hassan, "Migrants, ancestors, and foreign investments," *Review of Economic Studies*, 2019, *86* (4), 1448–1486.

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.

Conte, M., P. Cotterlaz, and T. Mayer, "The CEPII Gravity database," CEPII Working Paper N°2022-05, CEPII 2022.

Dekle, Robert, Jonathan Eaton, and Samuel Kortum, "Unbalanced trade," *American Economic Review*, 2007, *97* (2), 351–355.

Feenstra, Robert C, Luca Macedoni, and Mingzhi Xu, "Large firms, consumer heterogeneity and the rising share of profits," Technical Report, National Bureau of Economic Research 2023.

McCully, Brett, "Immigrants, legal status, and illegal trade," Technical Report 2024.