NBER WORKING PAPER SERIES

PRODUCT VARIETY, THE COST OF LIVING AND WELFARE ACROSS COUNTRIES

Alberto Cavallo
Robert C. Feenstra
Robert Inklaar

Working Paper 28711
http://www.nber.org/papers/w28711

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2021

We thank seminar participants at the NBER, Georgetown University, the University of Groningen and Yale University for helpful comments. We also thank Ninghui Li, Anna Ignatenko and Mingzhi Xu for their excellent research assistance. This research is supported by the NSF (US), SSHRC (Canada), and NWO (Netherlands) under a Digging into Data grant of the Trans-Atlantic Platform (NSF Award #1724649). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Product Variety, the Cost of Living and Welfare Across Countries
Alberto Cavallo, Robert C. Feenstra, and Robert Inklaar
NBER Working Paper No. 28711
April 2021
JEL No. E01,F12

ABSTRACT

We use the structure of the Melitz (2003) model to compare the cost of living and welfare across countries, while incorporating product variety measured by the count of barcodes or firms. For 47 countries, we compare welfare relative to the United States to conventional measures of real consumption. Relative welfare is similar to or higher than that indicated by real consumption for a select group of nations in Europe and some large countries like China and Russia, but lower in most other countries. This qualitative pattern has some similarities to that found in Jones and Klenow (2016), but for very different reasons.

Alberto Cavallo
Harvard Business School
Morgan Hall 287
Soldiers Field
Boston, MA 02163
and NBER
acavallo@hbs.edu

Robert Inklaar
University of Groningen
Faculty of Economics
PO Box 800
9700 AV Groningen
The Netherlands
R.C.Inklaar@rug.nl

Robert C. Feenstra
Department of Economics
University of California, Davis
One Shields Avenue
Davis, CA 95616
and NBER
rcfeenstra@ucdavis.edu

1. Introduction

Liberalizing trade is well understood to improve country welfare. Melitz (2003) is among the models that generate a very simple formula – based on openness – for the gains from trade (Arkolakis, Costinot and Rodríguez-Clare, ACR, 2012). But while the Melitz model is well-suited to compute the gains from trade *within* a country, can it be also used to compare welfare *between* countries? We will demonstrate that it can, provided that product variety and other domestic variables are measured and taken into account. We compare the theoretical cost-of-living predictions from the Melitz model to *consumption price levels* from the International Comparisons Project (ICP), focusing on cross-country prices for tradable goods. We will find that the theoretical cost of living relative to the US is similar to or less than indicated by consumption price levels for a select group of nations in Europe and some large countries like China and Russia, which implies a higher welfare level. But the cost of living in many other countries is higher than indicated by consumption price levels, reflecting lower levels of product variety and thus lower welfare in those countries.

This paper contributes to the literature in several ways. ACR (2012) allowed for a *foreign shock* to impact a country, by which they mean a change in foreign variables such as iceberg trade costs, foreign country size, etc. That allowed them to compare equilibria within a country. We shall expand on that literature by incorporating a *domestic shock*, meaning a change in domestic trade costs, country size, productivity, etc. That will enable us to compare equilibria between countries: in particular, it is the domestic shock that leads to differences in product variety across countries, as we show in section 2.

Second, we demonstrate the feasibility of using online barcode data to measure product variety across countries. Barcode-level data has previously been used to compare *prices* across countries (Cavallo et al. 2018), and *product variety* has been compared across cities in the United States (Handbury and Weinstein, 2015) and in China (Feenstra, Xu and Antoniades, 2020). But it is challenging to compare product variety across countries due to differing classification systems for barcodes, so that identical products cannot be identified.[1] We overcome this difficulty by relying on simple counts of barcodes in certain sectors for major retailers, and when that country information could not be collected, then we use the much cruder count of firms within each sector and country to proxy for variety.

Third, in order to compare the theoretical predictions from the Melitz model to country price levels from ICP, we also need to control for differences in country productivity. The "next generation" of Penn World Table (PWT, see Feenstra, Inklaar and Timmer, 2015) calculates productivity using a measure of *real GDP on the output-side,* i.e. GDP deflated with *aggregate output prices* that can be compared across countries. Here we use similar techniques to obtain *output prices at the sectoral level*, which differ from the sectoral consumption price levels in the ICP. So this paper becomes an evaluation of how both the ICP and PWT datasets compare to theoretical predictions from the Melitz model.

Finally, this paper contributes to the broader literature on measures of welfare that are "beyond GDP", to use the phrase of Jones and Klenow (2016). They propose a welfare concept that combines cross-country differences in consumption, leisure, mortality and inequality into a single consumption-equivalent measure. Our goal here is much less ambitious: to incorporate

[1] An exception is Argente, Hsieh and Lee (2020), who use the fact that the US and Mexico share a barcode system to compare identical products across these two countries. Because they also have consumption data by barcode in both countries, exact price indexes as in Feenstra (1994) can be constructed. We do not have barcode consumption data to construct exact price indexes, so we rely on the count of barcodes and on the theoretical structure of the Melitz model to measure welfare.

product variety to measure the cost of living across countries while focusing on tradable goods. But there are some broad similarities in our methods and results. Just as we use micro-level barcode counts for product variety in some countries, while relying on the macro-level count of firms in other countries, Jones and Klenow likewise compare results obtained with micro and macro data. They find that most Western European countries have welfare relative to the US that is comparable to or higher than indicated by conventional measures of real consumption, whereas we find that result for a more select group of European nations and also some large countries like China, India and Russia. Conversely, many middle-income and smaller countries are moved farther below the US benchmark in both their analysis and in ours.

In section 2, we examine the cross-country comparison of equilibria in the Melitz (2003) model, which depends on: the share of expenditure on domestic goods (reflecting openness); domestic and foreign trade costs; productivity; the terms of trade; and the extent of product variety available to consumers. Implementing the theoretical expression for the cost of living between countries requires data on all these variables, as well as model parameters. In section 3, we describe the data used to measure these variables and parameters. Our results are presented in section 4 and section 5 concludes, while the proofs of Propositions and other details are in the Online Appendix.

2. Modeling Welfare Between Countries

Extending the model of Melitz (2003) and Chaney (2008) to measure welfare between countries means that we must allow for differences their populations, productivities and domestic trade costs. The latter are modeled as iceberg costs, meaning that $\tau_d \geq 1$ units must be sent from the domestic firms for one unit to reach the consumer. Like the foreign trade costs in Melitz and Chaney, these iceberg costs use up resources. That is a plausible description of resources used in domestic transportation and in the wholesale and retail sectors, which we use to measure τ_d.[2] We consider two equilibria that can experience a *domestic shock,* by which we mean a change in domestic iceberg costs τ_d, a change in mean productivity, or a change in population. In addition, like in ACR (2012), the two equilibria can experience a *foreign shock,* defined as changes in the iceberg costs of international trade and in the foreign values of fixed costs and population (though we will introduce below a restriction on the extent to which fixed costs can differ across countries).

The rest of the model is familiar from Melitz (2003), so our exposition will be brief. We assume a CES utility function with elasticity of substitution $\sigma > 1$. A mass M_e of domestic firms pay an entry cost f_e to receive a draw of productivity from the distribution $g(\varphi)$. With trade, the CES price index for the home consumer is defined over domestic and foreign goods as:

$$P = \left[M_d \int_{\varphi_d}^{\infty} p_d(\varphi)^{1-\sigma} \frac{g(\varphi)}{[1-G(\varphi_d)]} d\varphi + M_x^* \int_{\varphi_x^*}^{\infty} p_x^*(\varphi)^{1-\sigma} \frac{g(\varphi)}{[1-G(\varphi_x)]} d\varphi \right]^{1/(1-\sigma)}, \tag{1}$$

where the first integral reflects the consumer prices $p_d(\varphi)$ of the mass $M_d = M_e[1-G(\varphi_d)]$ of domestic firms with productivity $\varphi \geq \varphi_{\mathrm{d}}$, and the second integral reflects the import prices $p_x^*(\varphi)$ of the mass $M_x^* = M_e^*[1-G(\varphi_x^*)]$ of *foreign* firms with productivity $\varphi \geq \varphi_x^*$. We follow Chaney (2008) in assuming that the density of home and foreign productivities is Pareto distributed:

$$G(\varphi) = 1-(\varphi / A)^{-\theta} \text{ for } \varphi \geq A \text{ and } \theta > (\sigma - 1) > 1. \tag{2}$$

Note that the mean productivity is then $\int_A^{\infty} \varphi g(\varphi) d\varphi = \left(\frac{\theta}{\theta-1}\right) A.$ It follows that the lower-bound for

[2] Our theoretical analysis could also be extended to include differences in excise taxes across countries, which are part of our empirical analysis.

productivity, *A*, is proportional to the mean productivity.

To obtain the share of expenditure on domestic goods, which we denote by λ_d, we take the ratio of the first term on the right of (1) to the whole term in brackets,

$$\lambda_d = \left[M_d \int_{\varphi_d}^{\infty} p_d(\varphi)^{1-\sigma} \frac{g(\varphi)}{[1-G(\varphi_d)]} d\varphi \right] \Big/ P^{(1-\sigma)} . \tag{3}$$

This expression can be simplified by solving for domestic prices. The marginal costs of production at home are w/φ, so that with the usual CES markup the consumer price is $p_d(\varphi) = [\sigma/(\sigma-1)]\,(\tau_d w/\varphi)$, where $\tau_d \geq 1$ are the domestic iceberg costs. Substituting these prices into the numerator of (3) and computing the integral, we obtain:

$$M_d \int_{\varphi_d}^{\infty} \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} \left(\frac{w\tau_d}{\varphi}\right)^{1-\sigma} \frac{g(\varphi)}{[1-G(\varphi_d)]} d\varphi = M_d \left(\frac{\theta}{\theta-\sigma+1}\right)\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} \left(\frac{w\tau_d}{\varphi_d}\right)^{1-\sigma} . \tag{4}$$

Substituting (4) into (3), the share of expenditure on domestic goods is:

$$\lambda_d = \left(\frac{\theta}{\theta-\sigma+1}\right)\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} \left(\frac{w\tau_d}{\varphi_d}\right)^{1-\sigma} \frac{M_d}{P^{1-\sigma}} . \tag{5}$$

Now consider two equilibria, with the shocked equilibrium denoted by a prime. The ratio of CES price indexes is denoted by P'/P, and it measures the change in the cost of living between the two equilibria, i.e. the inverse of the change in welfare. Then the ratio P'/P is readily obtained by re-arranging (5) as:

$$\begin{aligned} \frac{P'}{P} &= \left[\frac{(M'_d)^{\frac{1}{1-\sigma}} w'\tau'_d / \varphi'_d}{M_d^{\frac{1}{1-\sigma}} w\tau_d / \varphi_d} \right] \left(\frac{\lambda'_d}{\lambda_d}\right)^{\frac{1}{\sigma-1}} \\ &= \left(\frac{w'\tau'_d / \varphi'_d}{w\tau_d / \varphi_d}\right)\left(\frac{M'_d / \lambda'_d}{M_d / \lambda_d}\right)^{\frac{1}{1-\sigma}} . \end{aligned} \tag{6}$$

The first term on the right of (6) in brackets is the ratio of the *CES price index of domestic goods*, where the variety term $M_d^{\frac{1}{1-\sigma}}$ (and, likewise, in the prime equilibrium) is the welfare effect of any change in the mass of *domestic* varieties, while $w\tau_d / \varphi_d$ is proportional to the *average price* of these domestic varieties (using equation (4)). The second term on the right of (6) is the ratio of the share of spending on domestic goods, or one minus the share of spending on new imported varieties, which adjusts the price index for import varieties as in Feenstra (1994). Either an increase in domestic variety ($M_d > M'_d$) or in import variety ($\lambda_d < \lambda'_d$) reduces the price index. By rewriting the price index as in the second line of (6), we can see that the term M_d / λ_d (and, likewise, in the prime equilibrium) measures the "overall" product variety taking into account both domestic and import varieties, with an increase in overall variety lowering the overall price index according to the exponent $1/(1-\sigma) < 0$.

We do not describe the rest of the equilibrium conditions here, but they are outlined in the Appendix. The zero-cutoff-profit (ZCP) condition determines the threshold productivity φ_d for home domestic firms, which is:

$$\varphi_d^{\sigma-1} = \frac{wf_d\sigma^\sigma}{X}\left(\frac{w\tau_d}{P(\sigma-1)}\right)^{\sigma-1}, \tag{7}$$

where *X* denotes total expenditure on the differentiated good at home. The mass of entering firms is determined by the free entry and full employment conditions.

We consider two equilibria that can experience both a domestic and a foreign shock, meaning different values of the iceberg costs, fixed costs, population, and productivity *A*. In this way, we can examine the impact on one country from a change in the foreign variables (following ACR), or we can compare the equilibria between two countries that have differing values for both the home and foreign shock variables. The equilibrium conditions that we have described above are enough to determine the sources of welfare differences between the two equilibria. We take the ratio of the ZCP productivity in (7) between the two equilibria, and substitute that into (6) to obtain,

$$\left(\frac{X'/w'f_d'}{X/wf_d}\right)^{\frac{1}{\sigma-1}} = \left(\frac{M_d'/\lambda_d'}{M_d/\lambda_d}\right)^{\frac{1}{\sigma-1}}. \qquad (8)$$

The expression on the right of (8) is the inverse of the domestic variety and share terms appearing on the right of (6). Expression (8) therefore measures the welfare gain between the two equilibria due to any expansion of "overall" variety.[3] Comparing two equilibria with the *same* values of expenditure *X* relative to fixed costs wf_d, then *there will be no welfare change due to variety*: equation (8) shows that $M_d'/\lambda_d' = M_d/\lambda_d$ when $X'/w'f_d' = X/wf_d$, which means that there is no change due to variety in the relative price indexes in (6).[4] That is the case in the one-sector Melitz-Chaney model in ACR (2012), for example, where trade balance ensures that expenditure equals labor income, $X = wL$, and changes in *L* and f_d are ruled out, so that the left of (8) is unity. It follows from the right of (8) that $M_d'/\lambda_d' = M_d/\lambda_d$ so there is no welfare change due to variety. By allowing for domestic shocks, however, we can have welfare changes due to variety, either within or between countries.

Expression (8) shows us how to interpret the variety terms appearing in (6), but we still need to solve for ZCP productivity levels appearing there. As mentioned, we assume a Pareto distribution for firm productivity given by (2). The mass of operating domestic firms equals $M_d = M_e[1-G(\varphi_d)] = M_e(\varphi_d/A)^{-\theta}$ where M_e is the mass of entering firms. Then using this

---

[3] In particular, welfare rises with increased openness, resulting in $\lambda_d' < \lambda_d$, provided that this impact exceeds any reduction in domestic varieties, $M_d' < M_d$, so that $M_d'/\lambda_d' > M_d/\lambda_d$.

[4] Note that M_d can change due to a domestic shock, but with an equal and offsetting change in λ_d.

in (8), we obtain,

$$\frac{\lambda_d'}{\lambda_d} = \left(\frac{X'/w'f_d'}{X/wf_d}\right)^{-1}\left(\frac{M_e'}{M_e}\right)\left(\frac{\varphi_d'/A'}{\varphi_d/A}\right)^{-\theta} = \left(\frac{X'/w'L'}{X/wL}\right)^{-1}\left(\frac{f_d'/f_e'}{f_d/f_e}\right)\left(\frac{\varphi_d'/A'}{\varphi_d/A}\right)^{-\theta}, \qquad (9)$$

where the final equality uses the fact that the mass of entering firms is inversely proportional to the effective population size, $M_e \propto L/f_e$, as shown in the Appendix. The ratio of the fixed costs of production and entry that appears in (9) is difficult to identify from the data, so we simplify our model by assuming that it is the same across countries. We state this assumption formally by adding a country superscript $i = 1,\ldots,C$:

Assumption 1:

The fixed costs of production and entry for the home market are proportional, $f_d/f_e = f_d'/f_e' = f_d^i/f_e^i$ for all countries $i = 1,\ldots,C$.

Assumption 1 ensures that the ratio $(f_d'/f_e')/(f_d/f_e)$ vanishes in (9). We will also consider the following stronger version, which implies Assumption 1:

Assumption 1$'$:

The fixed costs of production and entry for the home market are proportional to L^α, $0 \le \alpha \le 1$, so $f_d/L^\alpha = f_d'/L'^\alpha = f_d^i/L^{i\alpha}$ and $f_e/L^\alpha = f_e'/L'^\alpha = f_e^i/L^{i\alpha}$ for all countries $i = 1,\ldots,C$.

This stronger version is motivated by the fixed market penetration costs discussed by Arkolakis (2010), which in a simplified version of his model are L^α. With these assumptions, we readily obtain the following result by computing real wages from (6) and substituting for the ZCP productivity levels (φ_d'/φ_d) from (9):

Proposition 1:

(a) Under Assumption 1, the ratio of real wages between two equilibria is:

$$\left(\frac{w'/P'}{w/P}\right)=\left(\frac{A'}{A}\right)\left(\frac{\lambda_d'}{\lambda_d}\right)^{-\frac{1}{\theta}}\left(\frac{\tau_d'}{\tau_d}\right)^{-1}\left(\frac{M_d'/\lambda_d'}{M_d/\lambda_d}\right)^{\frac{1}{\sigma-1}}\left(\frac{X'/w'L'}{X/wL}\right)^{-\frac{1}{\theta}}. \tag{10}$$

(b) Under Assumption $1'$, this expression becomes:

$$\left(\frac{w'/P'}{w/P}\right)=\left(\frac{A'}{A}\right)\left(\frac{\lambda_d'}{\lambda_d}\right)^{-\frac{1}{\theta}}\left(\frac{\tau_d'}{\tau_d}\right)^{-1}\left(\frac{M_d'/\lambda_d'}{M_d/\lambda_d}\right)^{\frac{1}{\sigma-1}-\frac{1}{\theta}}\left(\frac{L'}{L}\right)^{\frac{(1-\alpha)}{\theta}}, \tag{11}$$

and product variety is determined by,

$$\left(\frac{M_d'/\lambda_d'}{M_d/\lambda_d}\right)=\left(\frac{L'}{L}\right)^{1-\alpha}\left(\frac{X'/w'L'}{X/wL}\right). \tag{12}$$

To interpret these results, the first term on the right of (10) and (11) is the ratio of overall productivity levels. The second term is the ratio of the share of expenditure on domestic goods, or an inverse measure of openness but with a negative exponent: when that share falls— indicating that more varieties are available from abroad— then the gains from trade are higher. This is the sufficient statistic identified by ACR for the gains from a foreign shock.

The third term on the right of (10) and (11) is the inverse ratio of domestic trade costs, so that a country with higher domestic trade costs will have correspondingly lower welfare. It is surprising that the domestic trade costs do not involve an exponent reflecting the share of expenditure on domestic goods. To explain this, consider two countries where the only difference between them is that one has higher domestic trade costs, $\tau_d' > \tau_d$. That country will have higher domestic prices and therefore lower real wages and welfare, depending on its consumption of the domestic good. But that country will also have lower expenditure on its domestic goods, $\lambda_d' < \lambda_d$, due to the higher prices. So, from (10) and (11), the higher domestic trade costs are *partially offset* by the lower domestic share, meaning that country welfare does not fall in direct proportion to the higher domestic trade costs.[5]

The fourth term appearing on the right of (10) and (11) is the "overall" welfare gain from domestic and import varieties available to consumers, as discussed just after (8), but this term appears with differing exponents in (10) and (11): the exponent on overall variety in (10) exceeds that in (11), which would seem to give a greater impact of product variety on welfare in (10). But that seemingly greater impact in (10) is offset by the final term in each equation. The final term in (10) is an adjustment for trade imbalance $(X' / w'L') / (X / wL)$, while the final term in (11) is an adjustment for the size of the labor force, which adds a further impact of product variety in (11) and (12) due to this "scale effect" from country size.

We stress that the formulations in (10) and (11) are theoretically equivalent because (12) holds. Nevertheless, we prefer to work with (11) while estimating the exponent $(1-\alpha)$ on the labor force from the regression shown by (12). When running this regression at the sectoral level, (X_s / wL) measures the *expenditure share* of each sector *s* in GDP. We will find that this sectoral expenditure share is a rather poor predictor of product variety in the regression (12), but that the relative population in (12) is a robust regressor that adjusts for the scale of economies, and therefore predicts variety. For this reason, we rely on equations (11) and (12). In contrast, the formulation in (10) would be very sensitive to the inclusion of the expenditure share at the sectoral level, which we do not find is a reliable predictor of variety, and is therefore not a reliable predictor of the cost of living.

[5] There is one parameterization, however, where the welfare will fall in direct proportion to the domestic trade costs, and that is where the domestic costs of transport and retail trade, *apply equally to domestic and imported goods*. This simple case is assumed below and in the Appendix to derive (15).

3. From Theory to the Data

3.1 The Cost of Living

We shall use Proposition 1(b) to compare the cost of living across countries. To achieve that, we invert (11) to obtain the cost of living between countries *i* and *j*:

$$\left(\frac{P^i}{P^j}\right)=\left(\frac{w^i/A^i}{w^j/A^j}\right)\left(\frac{\lambda^{ii}}{\lambda^{jj}}\right)^{\frac{1}{\theta}}\left(\frac{\tau^{ii}}{\tau^{jj}}\right)\left(\frac{M^i/\lambda^{ii}}{M^j/\lambda^{jj}}\right)^{\frac{1}{\theta}-\frac{1}{\sigma-1}}\left(\frac{L^i}{L^j}\right)^{-\frac{(1-\alpha)}{\theta}}. \tag{13}$$

where we have dropped the subscript *d*, so that λ^{ii} is understood as the share of expenditure in country *i* coming from domestic production, τ^{ii} are domestic trade costs in country *i*, and M^i is the mass of domestic products available. We can compare this theoretical cost-of-living index across countries to the *price level of traded consumption*, which we denote by PC^{Ti} in country *i*. The price level of traded consumption is computed from the 2011 round of the ICP (World Bank, 2014) and is measured as the observed prices of tradable consumption goods in each country, converted to US$ using the nominal exchange rate and measured relative to the US prices of the same goods. By construction, then, PC^{Ti} in country *i* is measured relative to the United States as country *j* (i.e., $PC^{T,US}\equiv 1$).

Several adjustments to (13) are needed to bridge the gap between our stylized model and the data we shall apply to it. First, our model has only labor, while in reality there are many factors of production. This feature is readily incorporated by consideration of the terms w^i/A^i and likewise for country *j* (i.e., the United States). Let w^i denote a weighted average of factor prices used in production. The term A^i is the lower bound to productivity in (7), and as such it also reflects the mean productivity in country *i* (as discussed just below equation (2)). Suppose we measure country productivity using a dual approach, which would equal the ratio of the weighted average of factor prices to the aggregate output price. Then the ratio w^i/A^i would equal the output price level, which we denote by PY^i. In contrast to the price level of consumption, the price level of output reflects the prices of *produced goods* in each country, relative to the US. In particular, export prices are included in the price level of output, whereas import prices are included in the price level of consumption. The "next generation" of PWT (Feenstra, Inklaar and Timmer, 2015) measured the *aggregate* output price by correcting for the terms of trade in this fashion. We shall use the same approach to measure *sectoral* output prices in traded sectors, which we denote by PY_s^{Ti} (see the Appendix and section 4.1).

Second, we shall apply formula (13) at the sectoral level, denoted by the subscript *s*. We use seven sectors of consumption shown in Table 1, defined at the two-digit level of the "classification of individual consumption by purpose" (COICOP). Third, we restrict ourselves to potentially traded products for household consumption – meaning goods rather than services – and these goods shares vary between 25% (Other goods) and 100% (Food, Beverages and Tobacco), as shown in Table 1.[6] For example, expenditure in the transportation sector includes taxi services, which we omit from traded products. The domestic expenditures shares λ_s^{ii}, in particular, are measured for manufactured goods in each sector *s*. Domestic trade costs τ_s^{ii} in sector s include the margin earned in transportation and retail trade and also taxes on products, notably sales tax, VAT and excise taxes. Information on the construction of both these terms is provided in the Appendix, and their average values over our 47 countries are reported in Table 1.

Let ω_s^{Ti} equal the Sato-Vartia weight of traded goods in sector *s* relative to the US (see the Appendix). Then averaging across the sectors, (13) is re-written as the cost of living in

[6] Four other sectors of consumption are omitted because the products in those sectors are either all nontraded (education, hotels and restaurants) or have very few traded products (housing and utilities, communication).

Table 1: Consumption sectors, goods share in each sector, and variable means

Sector	Code	Goods share (%)	Mean λ_s^{ii}	Mean τ_s^{ii}	Product variety
Total traded consumption		47			
Food, beverages & tobacco	01-02	100	0.73	1.78	Yes
Clothing & footwear	03	97	0.27	2.05	Yes
Furnishing, household equipment	05	89	0.44	2.06	Yes
Health	06	24	0.30	1.80	No
Transport	07	57	0.52	1.82	No
Recreation and culture	09	45	0.58	1.65	Yes*
Other goods	12	17	0.53	1.94	Yes

Notes: Code is the COICOP code for the sector; goods share is the share of total sectoral expenditure on goods rather than services, averaged over the 47 countries.
* Barcode data for newspapers and books are not available within the Recreation and Culture sector.

traded goods for country *i*, CoL^{Ti} / CoL^{Tj}, relative to the US as country *j*:

$$\frac{CoL^{Ti}}{CoL^{Tj}} \equiv \prod_{s=1}^{S} \left(\frac{PY_s^{Ti}}{PY_s^{Tj}}\right)^{\omega_s^{Ti}} \left(\frac{\lambda_s^{ii}}{\lambda_s^{jj}}\right)^{\frac{\omega_s^{Ti}}{\theta_s}} \left(\frac{\tau_s^{ii}}{\tau_s^{jj}}\right)^{\omega_s^{Ti}} \left(\frac{M_s^i / \lambda_s^{ii}}{M_s^j / \lambda_s^{jj}}\right)^{\frac{\omega_s^{Ti}}{\theta_s} - \frac{\omega_s^{Ti}}{\sigma_s - 1}} \left(\frac{L^i}{L^j}\right)^{-\frac{(1-\alpha)\omega_s^{Ti}}{\theta_s}}, \qquad (14)$$

where the weights of traded goods across sectors sum to unity, $\sum_{s=1}^{S} \omega_s^{Ti} = 1$. The last term appearing on the right of (14) is the population of each country *i*, relative to the US.

3.2 Product Variety

A final variable needed to measure (14) is the number of domestic product varieties in each sector, M_s^i. Our first measure of domestic varieties is based on the *number of domestic firms* active in each sector, taken from the Bureau van Dijk's ORBIS global dataset, which in turn, is based on business registers in different countries. We eliminate duplicate names and drop firms with zero employees to eliminate shell companies. As a verification exercise, we also collected data on the number of firms from national enterprise statistics, primarily from the OECD Structural Business Statistics and Eurostat Enterprise Statistics, supplemented by national reports. For most countries, the correspondence between the two sources is close; the correlation of the log number of firms between both sources is 0.75, rising to 0.90 when excluding India and Indonesia. Both of those countries have very large numbers of informal firms, which would skew upwards their variety count.

The number of firms in each sector is obtained for 46 countries, and in Table 2 we show the results for 24 countries where we also have an alternative measure of variety from barcode data.[7] The number of firms is a very crude measure of the number of products because of multi-product firms. At one extreme, large firms produce very many products that are not counted. At the other extreme, certain low-income countries like India have a very large number of informal firms, as just mentioned, which include firms perhaps serving only a single city or neighborhood. We might think of these firms as producing less than a single (national) product.

To obtain a more accurate count of product variety, we rely on a count of barcodes for goods sold within each of the sectors as shown in Table 1, except for Health and for Transport.[8] In these sectors without barcode data, we use the relative variety estimates from the firm-count data. The barcode counts are obtained from micro data available at the Billion Prices Project (BPP), for all products sold by some of the largest multi-channel retailers in in 24 countries: Australia, Brazil, Canada, Chile, China, Colombia, France, Germany, Greece, India, Ireland, Italy, Japan, Mexico, the Netherlands, New Zealand, Poland, Russia, South Africa, South Korea, Spain, Turkey, the United Kingdom, and the United States. To compute the barcode counts at the

[7] Barcode count data cover 26 countries including Argentina and Uruguay, but the we do not have data on the share of domestic expenditure for these two countries. We also lack data on the number of firms from ORBIS for Colombia, which means our analysis based on this variety measures covers 46 countries. See the Appendix Table A1 for the firm counts in 46 countries and 7 sectors and Appendix Table A2 for the barcode counts in the 24 countries that we cover in this paper.

[8] Barcodes for the Recreation and Culture sector are not measured for all products in that sector (see the notes to Table 1), so λ_s^{ii} is adjusted to match this coverage based on WIOD/TiVA data.

sector level, we first take the modal daily barcode count for each retailer in the BPP sample during 2018, computed at a 4-digit COICOP level of aggregation (e.g., "Coffee, Tea, and Cocoa"). To avoid double counting varieties sold in multiple retailers, we use the largest barcode count for each 4-digit category available across retailers, and then add up all the barcodes at the sectoral level. The number of barcodes in each sector is denoted by N_s^i, shown in Table 2.

Table 2: Number of Firms and Domestic Varieties

			Food & Beverages		Recreation & Culture (Electronics)	
Country	Orbis Firm Count (7 Sectors)	# Barcodes (5 Sectors, N)	# Barcodes (N)	Share Domestic Barcodes (B)	# Barcodes (N)	Share Domestic Barcodes (B)
Australia	63,482	117,992	9,738	0.61	29,217	0.01
Brazil	140,822	237,604	7,721	0.90	70,128	0.48
Canada	38,117	110,006	13,502	0.56	30,910	0.10
Chile	7,761	61,122	3,680	-	6,810	-
China	282,499	211,779	22,123	0.77	23,065	0.99
France	55,284	248,857	11,235	0.83	23,793	0.11
Germany	76,061	250,767	15,860	0.87	98,334	0.32
Greece	5,607	59,449	4,454	0.81	11,678	0.13
India	57,678	50,438	4,039	0.99	2,019	-
Ireland	5,045	40,088	9,162	0.63	3,005	0.13
Italy	162,167	80,063	7,819	0.89	14,348	-
Japan	65,994	563,973	16,163	0.88	165,692	0.16
Korea	78,893	254,510	41,641	-	42,891	-
Mexico	32,253	57,096	7,789	0.80	7,275	0.11
Netherlands	34,534	122,835	12,038	0.67	17,533	0.18
New Zealand	10,591	74,351	7,006	-	20,800	-
Poland	57,567	60,790	7,927	0.53	28,590	0.34
Russia	195,887	84,725	7,821	0.71	21,049	0.11
South Africa	37,498	50,238	9,493	-	14,512	-
Spain	64,970	175,981	12,741	0.67	35,568	0.20
Turkey	66,240	97,158	6,753	0.78	8,910	0.38
United Kingdom	91,124	110,509	11,996	0.78	19,880	0.19
United States	359,967	375,943	22,386	0.89	80,598	0.24

When collecting the count of barcodes in these sectors, we are mixing both domestically produced and imported goods. For just two sectors – Food, Beverages and Tobacco and Recreation and Culture –we further collected the country-of-origin information for a random sample of the total number of barcodes.[9] Specifically, we hired freelancers to manually check

[9] Within Recreation and Culture, we collected country of origin information for barcodes in Electronics and certain other consumer products such as bicycles.

500 randomly sampled barcode items per sector in each country. Using a custom mobile phone application, each freelancer visited one of the retailers in the BPP sample, scanned the barcode of each product, took a photo of the country-of-origin label, and determined if the product was domestic or imported. When no country of origin is listed, then the product is treated as domestically made (see the Appendix). For these two sectors, we therefore have the *barcode domestic ratio*, i.e., the ratio of domestically produced to the total number of sampled barcodes, which is denoted by B_s^i.[10] The number of domestically produced barcodes is therefore $M_s^i = N_s^i B_s^i$. In addition, for these sectors we also have the expenditure domestic ratio, which we have denoted by λ_s^{ii}. The overall measure of product variety is therefore,

$$\left(\frac{M_s^i}{\lambda_s^{ii}}\right) = \left(\frac{N_s^i B_s^i}{\lambda_s^{ii}}\right).$$

Outside of Food, Beverages and Tobacco, and Recreation and Culture (and for some countries within those sectors),[11] we do not have information on the share of domestically produced barcodes. In these cases, we make the simple assumption that $B_s^i \approx \lambda_s^{ii}$.[12] That is, we are assuming that the barcode domestic share is approximately equal to the expenditure domestic share. In that case, the overall measure of product variety is approximated by the total count of barcodes (including domestic and imported goods) for each sector:

$$M_s^i / \lambda_s^{ii} \approx N_s^i.$$

[10] To test the validity of our estimates for the barcode domestic ratios, we also computed an alternative metric using country of origin information collected online for individual products in a subset of 9 countries. This information was scraped from the website of a single retailer in each country. The correlation between the online and offline barcode domestic ratios is 0.76. More details are provided in the Appendix.

[11] No data for either sector could be collected for Chile, New Zealand, South Africa and South Korea. No data for Recreation and Culture could be collected for India and Italy.

[12] For those countries where we have the barcode domestic share (see Table 2), the median value of that share in Food, Beverages and Tobacco is 0.78 as compared to the median value of λ_s^{ii} which is 0.73; while in Recreation and Culture (Electronics) those medians are 0.18 and 0.14. So our simple assumption is plausible.

3.3 Parameter Values

Also needed in (14) are the elasticity of substitution σ and the Pareto parameter θ, which we treat as the same across sectors. For the Pareto parameter, we obtained an estimate for the Melitz-Chaney model by relying on the simulated method of moments from Simonovska and Waugh (2014), who also use cross-country data on the prices of goods collected by the International Comparisons Project (ICP).[13] Using ICP data from 2011, we obtain the pooled estimate of $\theta = 5.1$ across all sectors.[14]

For the elasticities of substitution, we considered the recent work of Redding and Weinstein (2020), who estimate elasticities of substitution across barcode varieties using the Nielsen Homescan data. These barcodes are for grocery store items (many of their barcodes are within our Food, Beverage and Tobacco sector), and they obtain a median elasticity of $\sigma = 6.5$. This estimate is incompatible with $\theta = 5.1$, however, because it violates $\theta > \sigma - 1$. This finding is not surprising, since we are comparing a median estimate of σ obtained from barcode data within narrow modules to a Pareto parameter θ obtained from ICP price data using the "basic heading" price data pooled across all sectors.[15]

To reconcile the estimates of σ and θ, we proceed in two way. First, $\sigma = 6.5$ can be compared with an initial estimate of $\theta = 8.3$ from Eaton and Kortum (2002, p. 1754) that is not too model-dependent. The ratio of these parameters gives $\theta/(\sigma - 1) = 1.5$, which is an acceptable spread between the parameters.[16] Conversely, we can start with the ratio $\theta/(\sigma - 1) = 1.5$ and

[13] We are grateful to Mike Waugh for providing us with the programs required to run these estimates, see https://github.com/mwaugh0328.

[14] We were not successful in obtaining sectoral estimates for the Melitz-Chaney model.

[15] Broda and Weinstein (2008) argue that the price differences across countries are much less for goods with identical barcodes than observed in datasets without barcode classifications.

[16] Eaton, Kortum and Kramarz (2011, p. 1472) find that $\theta/(\sigma - 1) = 1.75$ from an initial calculation of this ratio from French data that is, once again, not too model-dependent.

apply this to the estimate of θ = 5.1 obtained from the model-based methods of Simonovska and Waugh (2014) to obtain σ = 4.4. That estimate is very close to the median value of the elasticity of substitution obtained by Broda and Weinstein (2006), once we concord their results to products within our Food, Beverage and Tobacco sector (resulting in σ = 4.2). So using these two approaches, we obtain low estimates of (σ, θ) = (4.4, 5.1) and high estimates of (σ, θ) = (6.5, 8.3). We have made all our calculations using both sets of estimates and find that the results are remarkably similar. While the scale of the cost of living is about 1.5–2 times greater when the low parameter estimates are obtained, the relative position of countries is much the same. We present in the text the results obtained with the high estimates (σ, θ) = (6.5, 8.3), and report in the Appendix the results obtained using the low estimates (σ, θ) = (4.4, 5.1).

4. Empirical Results

4.1 Openness, Domestic Trade Costs and the Terms of Trade

We refer to the lambda-ratio that appears in (14), $(\lambda_s^{ii} / \lambda_s^{jj})^{\frac{1}{\theta_s}}$, as "inverse openness". It and the ratio of domestic trade costs, $(\tau_s^{ii} / \tau_s^{jj})$, are both constructed at the sectoral level, with *j* = *USA*. For convenience in graphing these, however, we take the weighted average across sector within each country, and plot $\prod_{s=1}^{S} (\lambda_s^{ii} / \lambda_s^{jj})^{\frac{\omega_s^{Ti}}{\theta_s}}$ versus $\prod_{s=1}^{S} (\tau_s^{ii} / \tau_s^{jj})^{\omega_s^{Ti}}$, as shown in Figure 1. There is a negative relationship between these two variables, so countries that are more open than the United States tend to have higher domestic trade costs and vice versa. This negative relationship is due, in part, to the generally higher rates of indirect taxation in European countries who also are more open. At the other end of the scale are countries such as Indonesia (IDN), India (IND), China, and Colombia, where domestic trade costs are notably lower than in

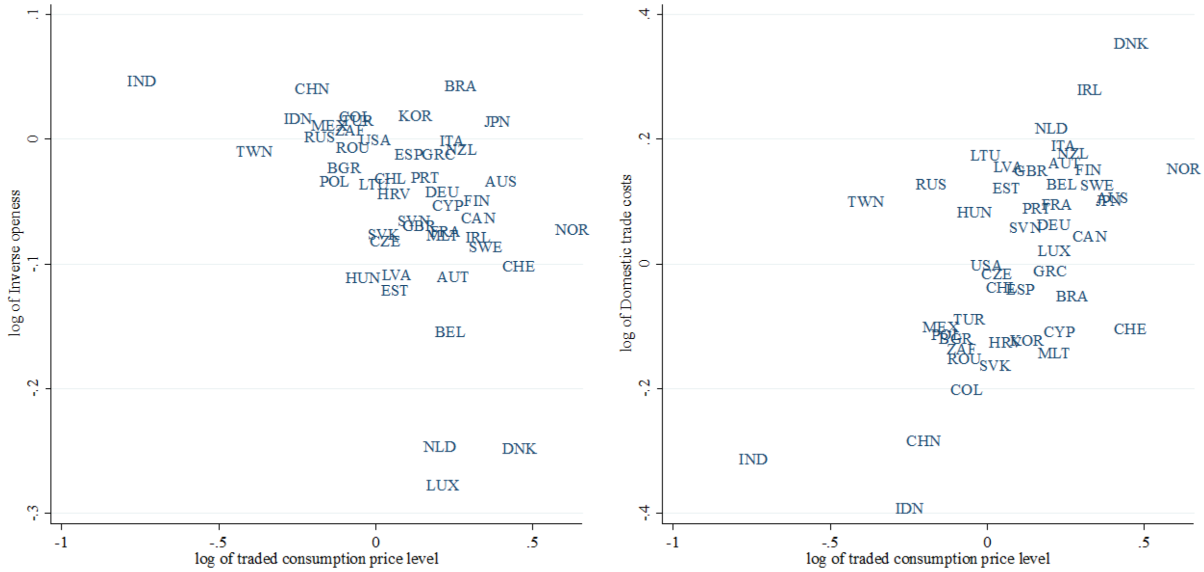
Figure 1. Inverse openness versus domestic trade costs by country (USA=1)

the US, but these countries are less open to trade, as shown by higher inverse openness.

It is instructive to graph overall openness and the domestic trade costs against the overall price level of tradable consumption goods, PC^{Ti}, as shown in Figure 2, in log terms with the US at point (0,0) in both panels. As noted earlier, we construct the level of consumption prices by aggregating the prices of tradable consumption goods within each country (from the 2011 round of the ICP). So in contrast to the *theoretical cost of living*, which we are measuring using the Metliz model, the *consumption price level* simply reflect price data collected across countries and expressed relative to the US. From the second panel of Figure 2, we can see that traded consumption prices are strongly correlated with domestic trade costs, which is no surprise: excise taxes and retail margins increase country prices. Many of the high-domestic-trade cost countries are also more open than the US, however, as shown in the first panel.

Next, we calculate all the factors in the cost of living in (14) – including the output price, domestic trade costs, and openness – *except* for variety in the final two terms. In Figure 3, the

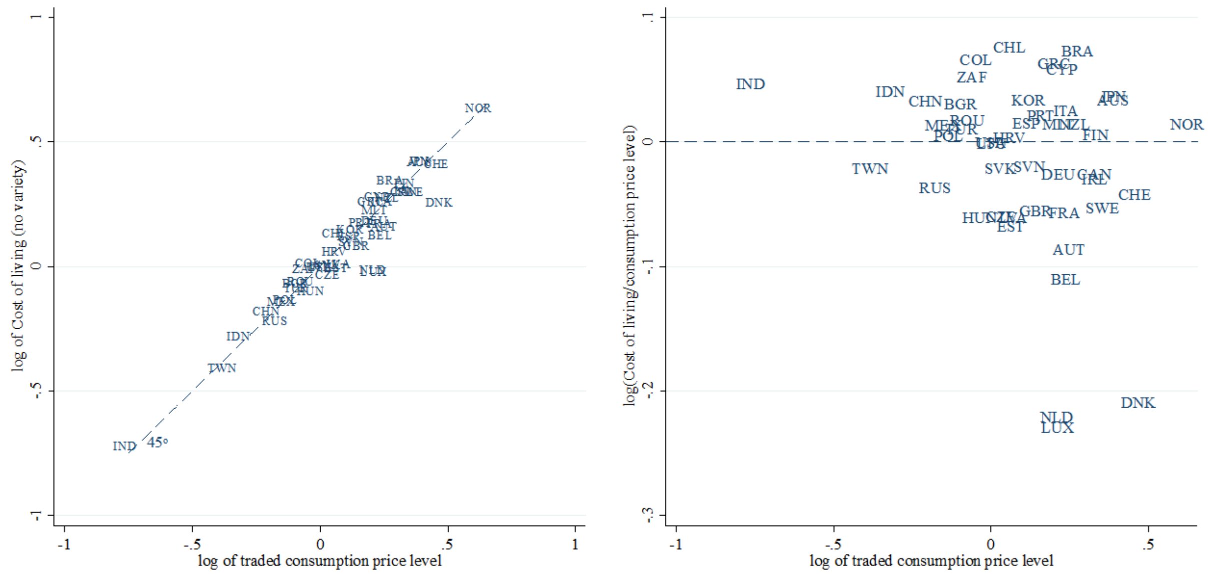
Figure 2: Inverse openness and domestic trade costs versus traded consumption price level (log scale, USA=0)



first panel plots the log of $\prod_{s=1}^{S}\left(PY_s^{Ti}\right)^{W_s^{Ti}} (\lambda_s^{ii} / \lambda_s^{jj})^{\frac{W_s^{Ti}}{\theta_s}} (\tau_s^{ii} / \tau_s^{jj})^{W_s^{Ti}}$ against the consumption prices of traded goods, PC^{Ti}. The cost of living (without variety) and the consumption prices are tightly clustered around the 45-degree line in the first panel. From the second panel, we see that roughly an equal number of countries have a cost of living (without variety) relative to the US that is higher versus lower than indicated by their consumption price level. Those with a cost of living below their relative consumption price include Taiwan and Russia, some countries in Eastern Europe (Czech Republic, Estonia, Hungary, Slovakia and Slovenia) and many in Western Europe (Austria, Belgium, Denmark, France, Germany, Ireland, Luxembourg, the Netherlands, Sweden, Switzerland, CHE and the United Kingdom, GBR), as well as Canada. Denmark, Luxembourg and the Netherlands are extreme cases where the cost of living is more than 20% below their relative consumption prices.

To understand why Denmark and the Netherlands have such a low cost of living (without variety) relative to the US in Figure 3, and conversely why some other countries have higher cost

Figure 3. Cost of living (without variety) due to productivity, openness and the domestic trade costs versus traded consumption price level (log scale, USA=0)



Notes: The left-hand figure plots log CoL^{Ti} versus log PC^{Ti} for the 47 countries in our analysis, with log CoL^{Ti} as defined in equation (14) (excluding the variety effects) and log PC_c^{Ti} computed as the price level of traded consumption, normalized to USA=1. The right-hand figure plots $\log\left(CoL^{Ti}/PC^{Ti}\right)$ versus log PC^{Ti}.

of living despite being more open than the US, we need to explore the terms influencing the cost of living: in particular, the output price levels PY_s^{Ti}. These output prices differ from the consumption prices PC_s^{Ti} because consumption prices include imports, whereas output prices include exports. These two variables therefore differ by the *terms of trade*. The terms of trade are constructed from the quality-adjusted export and import prices estimated by Feenstra and Romalis (2014). As already mentioned, these prices are used in PWT to construct an aggregate output price, and here we follow much the same procedure to construct sectoral output prices.

Specifically, to obtain the output prices we start with the price level of consumption for traded goods and net out the domestic trade costs (which we assume are identical for domestically produced and imported goods); then we add export prices; and finally, we net out tariff-inclusive import prices. This calculation gives (see the Appendix):

$$\frac{PY_s^{Ti}}{PY_s^{Tj}} = \left(\frac{PC_s^{Ti} / \tau_s^{ii}}{PC_s^{Tj} / \tau_s^{jj}} \right)^{\frac{(1-\omega_s^{Xi})}{(1-\omega_s^{Mi})}} \left(\frac{\tilde{P}_s^{Xi}}{\tilde{P}_s^{Xj}} \right)^{\omega_s^{Xi}} \left(\frac{\tilde{P}_s^{Mi}}{\tilde{P}_s^{Mj}} \right)^{-\frac{\omega_s^{Mi}(1-\omega_s^{Xi})}{(1-\omega_s^{Mi})}}, \quad (15)$$

where $\tilde{P}_s^{Xi}$ and $\tilde{P}_s^{Mi}$ are quality-adjusted prices for exports and imports,[17] while ω_s^{Xi} and ω_s^{Mi} are the associated Sato-Vartia weights. Substituting (15) into (14) we obtain:[18]

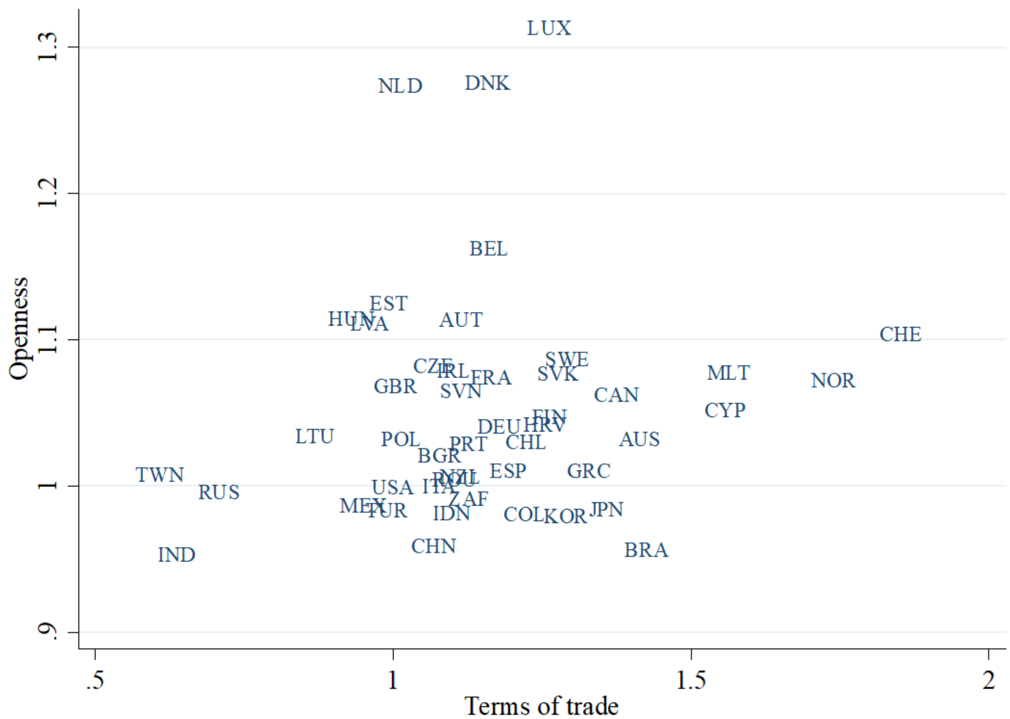
$$\begin{aligned} \frac{CoL^{Ti}}{CoL^{Tj}} \equiv \prod_{s=1}^{S} & \left(\frac{PC_s^{Ti}}{PC_s^{Tj}} \right)^{\frac{\omega_s^{Ti}(1-\omega_s^{Xi})}{(1-\omega_s^{Mi})}} \left(\frac{\tau_s^{ii}}{\tau_s^{jj}} \right)^{\frac{\omega_s^{Ti}(\omega_s^{Xi}-\omega_s^{Mi})}{(1-\omega_s^{Mi})}} \left(\frac{\tilde{P}_s^{Xi}}{\tilde{P}_s^{Xj}} \right)^{\omega_s^{Ti}\omega_s^{Xi}} \left(\frac{\tau_s^{Mi}\tilde{P}_s^{Mi}}{\tau_s^{Mj}\tilde{P}_s^{Mj}} \right)^{-\frac{\omega_s^{Ti}\omega_s^{Mi}(1-\omega_s^{Xi})}{(1-\omega_s^{Mi})}} \\ \times \prod_{s=1}^{S} & \left(\frac{\lambda_s^{ii}}{\lambda_s^{jj}} \right)^{\frac{\omega_s^{Ti}}{\theta_s}} \left(\frac{M_s^i / \lambda_s^{ii}}{M_s^j / \lambda_s^{jj}} \right)^{\frac{\omega_s^{Ti}}{\theta_s} - \frac{\omega_s^{Ti}}{\sigma_s - 1}} \left(\frac{L^i}{L^j} \right)^{-\frac{(1-\alpha)\omega_s^{Ti}}{\theta_s}}. \end{aligned} \quad (16)$$

To interpret the first line of (16), consider the simplified case where trade is balanced sector-by-sector, with $\omega_s^{Xi} = \omega_s^{Mi}$. In that case the first line starts with the weighted consumption price level. The next term, which is domestic trade costs, disappears when $\omega_s^{Xi} = \omega_s^{Mi}$ because it *equally impacts* the cost of living (on the left) and the consumption price level (on the right). The remaining terms on the first line are interpeted as the *terms of trade*, i.e. the price of exports relative to imports. In the Metliz model, the beneficial impact of trade is measured by openness, which lowers the cost of living by appearing inversely on the second line of (16); but in ICP or PWT data, the beneficial impact of trade is measured by the terms of trade, which lowers the consumption price level as compared to the output price level. In Figure 4, we plot openness (i.e. the *inverse* of the lambda-ratio) against the terms of trade (i.e. the relative price of exports). We see that there is a positive correlation between the two, though with some outliers.

[17] In Feenstra and Romalis (2014) the quality-adjusted import prices are measured net of tariffs, since they are used to deflate duty-free imports in GDP, but here we measure them inclusive of tariffs.

[18] In Appendix Tables A3 and A4, we show the log values of the terms in (16) to provide a decomposition of the cost of living to relative to the consumption price level.

Figure 4. Openness versus the terms of trade by country (USA=1)



Consider Denmark, Luxembourg and the Netherlands in Figure 4, which are very open but have terms of trade that are not much higher than for the United States, i.e., close to unity. Their openness contributes to a low cost of living relative to the US, while having terms of trade close to unity *does not* contribute to low consumption prices. As a result, these three countries have the lowest costs of living (without variety) as compared to consumption prices in Figure 3. Similarly for Belgium, which is less open than these three countries in Figure 4 but still more than the US, and so its cost of living relative to consumption price is higher than for these three countries in Figure 3. Then consider Switzerland (CHE), which is somewhat more open that the US but has the highest terms of trade in Figure 4, which contributes to low consumption prices. As a result, its cost of living as compared to its consumption price is higher than for Belgium in Figure 3. We conclude that openness versus the terms of trade contributes meaningful variation to the cost of living (without variety) relative to the consumption price level.

4.2 Product Variety

The main novelty of our approach is to incorporate product variety. Our theory implies that variety is related to country size according to (12), where *X* is expenditure and *L* is population. Rather than thinking of aggregate expenditure for the economy, we instead we think of it as *sector-level expenditure* X_s, in the same sectors for which we measure variety. We then rewrite (12) as:

$$\left(\frac{M_s^i / \lambda_s^{ii}}{M_s^j / \lambda_s^{jj}}\right) = \left(\frac{L^j}{L^i}\right)^{1-\alpha} \left(\frac{X_s^i / w^j L^j}{X_s^j / w^j L^j}\right), \qquad (12')$$

where $X_s^i / w^i L^i$ denotes expenditure on sector *s* relative to GDP of country *j*. Then we estimate this relationship as a regression where we initially do not difference with respect to country *j* and we move λ_s^{ii} to the right:

$$\ln M_s^i = \beta_0 + \beta_1 \ln \lambda_s^{ii} + (1-\alpha)\ln L^i + \beta_2 \ln\left(\frac{X_s^i}{GDP^i}\right) + \varepsilon_s^i. \qquad (17)$$

We also consider constraining $\beta_1 = 1$ and moving λ_s^{ii} back to the left:

$$\ln\left(\frac{M_s^i}{\lambda_s^{ii}}\right) = \beta_0 + (1-\alpha)\ln L^i + \beta_2 \ln\left(\frac{X_s^i}{GDP^i}\right) + \varepsilon_s^i. \qquad (17')$$

The regressions in (17) and (17') can be estimated using either barcode or firm counts to measure product variety. Before estimating this regression using firm-count data, however, we should recognize that the count of firms is potentially a *proxy* for "true" product variety. Specifically, suppose that the firm-count measure of product variety $\tilde{M}_s$, by sector *s*, is related to "true" variety M_s measured using the barcode count according to:

$$\ln \tilde{M}_s^i = \mu_0 + \mu \ln M_s^i + u_s^i.$$

Differencing with respect to the United States as country *j*, and taking the weighted average across sectors to reduce the errors, we obtain:

$$\sum_{s=1}^{S} W_s^{Ti} \ln\left(\tilde{M}_s^i / \tilde{M}_s^j\right) = \mu \sum_{s=1}^{S} W_s^{Ti} \ln\left(M_s^i / M_s^j\right) + U^{ij}. \qquad (18)$$

The ordinary least squares estimate of (18) is OLS estimate of $\hat{\mu} = 2.03$ (s.e. = 0.24). We conclude that taking approximately the *square root* of the firm count gives an estimate of variety that is reasonably close to that obtained from barcode count (for those countries where we have both sources of data). Accordingly, we take $\mu \approx 2$ to the left of (18) and the variety effect in (16) becomes $\sum_{s=1}^{S}\left(\frac{\omega_s^{Ti}}{\theta_s} - \frac{\omega_s^{Ti}}{\sigma_s - 1}\right)\ln\left[\left(\sqrt{\tilde{M}_s^i} / \lambda_s^{ii}\right) / \sqrt{\tilde{M}_s^j} / \lambda_s^{jj}\right]$, with $\left(\frac{\omega_s^{Ti}}{\theta_s} - \frac{\omega_s^{Ti}}{\sigma_s - 1}\right) < 0.$ This proxy from the firm-count is compared to the variety effect using the barcode count in Figure 5. It is clear that these two measures are highly correlated with very similar scales. From now on, we take approximately the square root of the firm count when measuring variety with those data.[19]

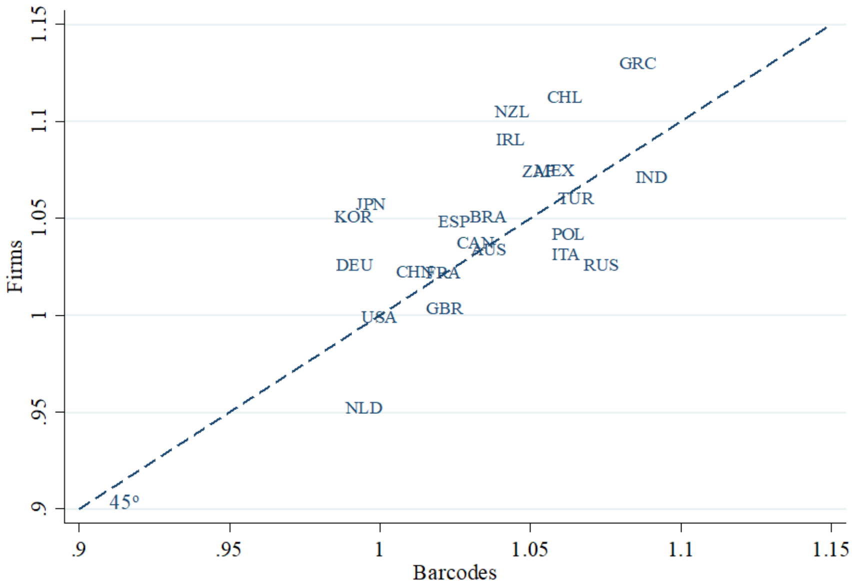
We now estimate the regressions (17) and (17') on both the barcode data and the square root of firm counts, with the results shown in Table 3. From columns (1) and (2), the variable $\ln \lambda_s^{ii}$ performs reasonably well, with a coefficient less than its predicted value of unity but highly significant. We find a coefficient $(1-\hat{\alpha})$ on $\ln L^i$ that is between 0.1 and 0.3, though the sector share of GDP has a coefficient far below unity and marginally significant. When we move λ_s^{ii} to the left of the equation and estimate (17'), the results in columns (3) and (4) are quite sensitive to whether we include India or not, since both the barcode for that country are unusually small as illustrated in Figure 6.[20] When omitting India from the barcode data, we find that $(1-\hat{\alpha}) = 0.2$.

---

[19] To be precise, we use $1/\hat{\mu} = 0.49$ to transform the firm count, and we also take into account the standard error of this estimate to compute the standard errors shown in Table 3, columns (4) and (6) and in Figures 7 and 9.

[20] India's online retail sector was relatively undeveloped when these barcodes were counted in 2018. The country ranked last in the UN "Ecommerce Index" among those included in our barcode sample (UNCTAD 2017) and the World Bank estimated that only about 1.6% of sales took place online that year (Kathuria et al., 2019).

Figure 5. Variety effects by country – firm count versus barcode count (USA=1)



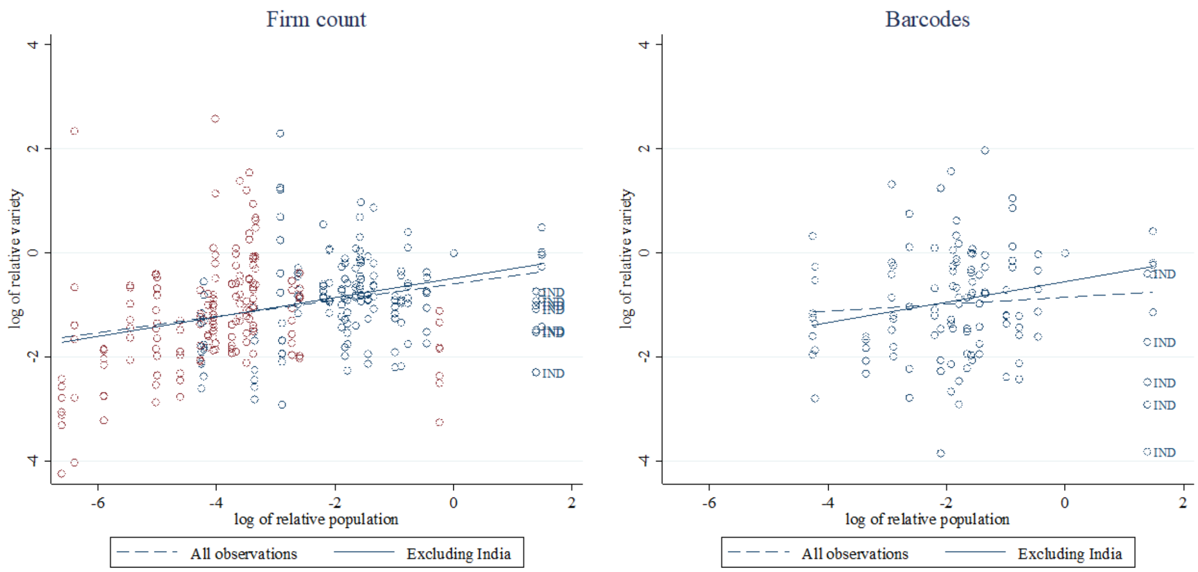
Note: Figure plots the aggregate variety effect relative to the USA

Table 3. Product Variety Regressions

Variety:	(1) Barcode count	(2) Barcode count	(3) Barcode count	(4) Barcode count	(5) Firm count	(6) Firm count
Dependent variable:	$\ln M_s^i$	$\ln M_s^i$	$\ln \frac{M_s^i}{\lambda_s^{ii}}$	$\ln \frac{M_s^i}{\lambda_s^{ii}}$	$\ln \frac{\sqrt{M_s^i}}{\lambda_s^{ii}}$	$\ln \frac{\sqrt{M_s^i}}{\lambda_s^{ii}}$
$\ln \lambda_s^{ii}$	0.604*** (0.134)	0.542*** (0.133)				
$\ln L^i$	0.112 (0.102)	0.298*** (0.076)	0.034 (0.096)	0.200** (0.077)	0.171*** (0.063)	0.203*** (0.066)
$\ln\left(\frac{X_s^i}{GDP^i}\right)$	0.267* (0.143)	0.247* (0.136)	0.080 (0.126)	0.035 (0.121)	-0.156 (0.195)	-0.153 (0.188)
Include India:	Yes	No	Yes	No	Yes	No
Observations	100	95	100	95	299	292

Notes: * p≤ 0.1, ** p ≤ 0.05, *** p ≤ 0.01

Figure 6. Variety and size at the sectoral level



Note: Observations in blue are covered by firm count data and barcode data, while observations in red are only in the firm count data.

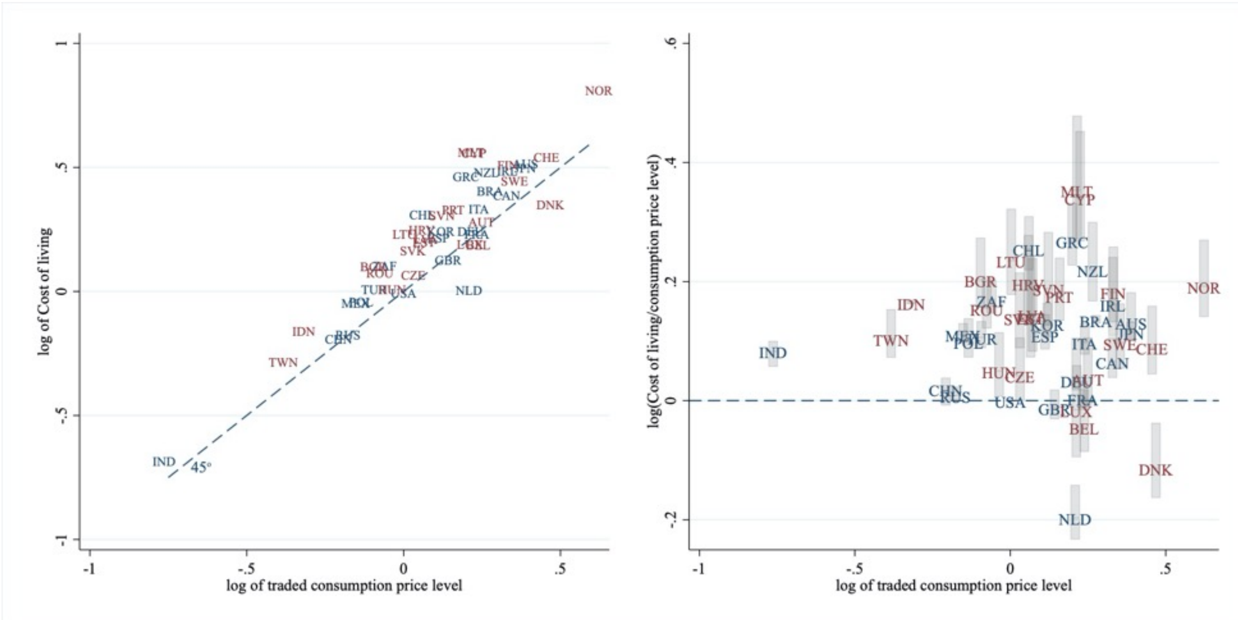
Turning to the results with firm counts, columns (5) and (6) give results quite similar to $(1-\hat{\alpha})=0.2$, especially when India is omitted. Accordingly, we shall use $(1-\hat{\alpha})$ from columns (4) and (6), and its standard error, to compute the cost of living according to (16) and its associated 95% confidence interval.

Figure 7 shows the estimates of the cost of living for 46 countries based on the firm count data. The variety effect increases the cost of living in all countries relative to the United States, which has nearly the greatest variety.[21] As a result, most countries have a greater cost of living relative to the US than indicated by their relative consumption prices (second panel). Only Belgium, Denmark, Luxembourg, the Netherlands and the United Kingdom (GBR) have

---

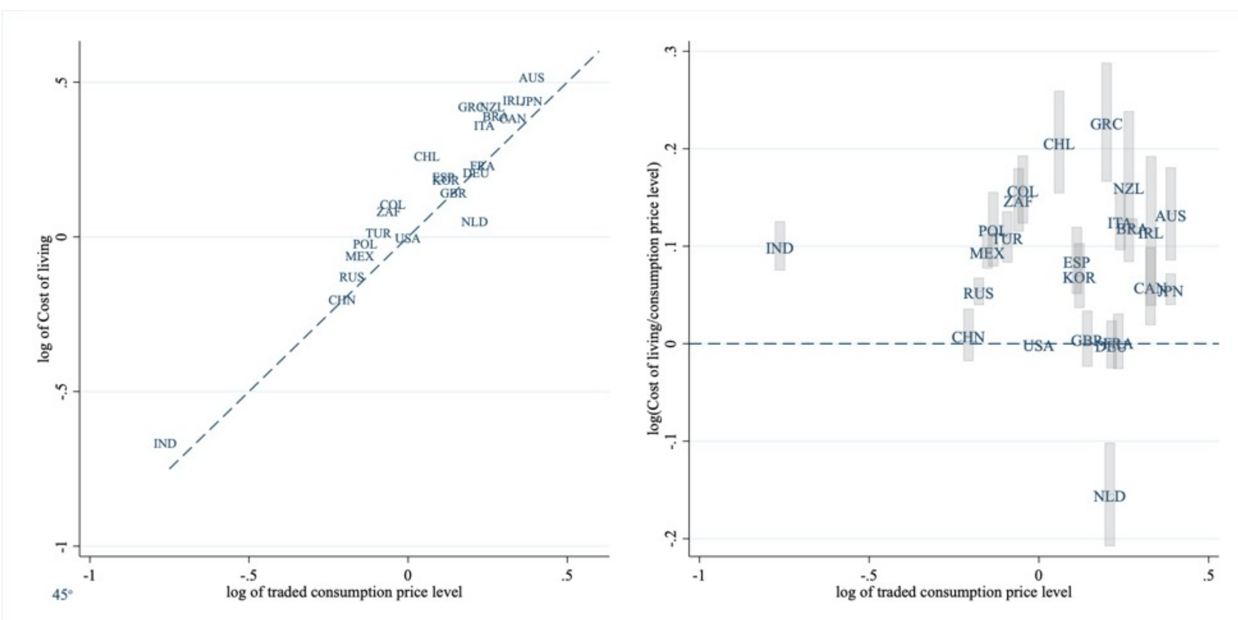
[21] Japan exceeds the United States in the raw barcode count across all sectors shown in Table 2, while Germany and South Korea each exceed the US in specific sectors. Our overall measure of product variety is computed as $[(M_s^i/\lambda_s^{ii})/(M_s^j/\lambda_s^{jj})]$, then including its negative exponent as in (16) and weighting across sectors. Figure 5 shows that the Netherland has greater overall variety than the US using barcode counts, as does Germany when using the square root of the firm counts, due to their high openness (which increases effective import variety).

Figure 7. Cost of living versus the traded consumption price level – firm count data



Notes: The left-hand figure plots log CoL^{Ti} versus log PC_c^{Ti} for the 46 countries in our analysis with firm count data, with log CoL^{Ti} as defined in equation (14) and log PC_c^{Ti} computed as the price level of traded consumption, with PC^{Ti} and CoL^{Ti} normalized to USA=1. The right-hand figure plots $\log(CoL^{Ti}/PC^{Ti})$ versus log PC^{Ti}. Countries in blue are covered by firm count data and barcode data, while observations in red are only in the firm count data.

Figure 8. Cost of Living versus the Traded Consumption Price Level – barcode data



Notes: The left-hand figure plots log CoL^{Ti} versus log PC_c^{Ti} for the 23 countries in our analysis with barcode data, with log CoL^{Ti} as defined in equation (14) and log PC_c^{Ti} computed as the price level of traded consumption, with PC^{Ti} and CoL^{Ti} normalized to USA=1. The right-hand figure plots $\log(CoL^{Ti}/PC^{Ti})$ versus log PC^{Ti}.

a cost of living relative to the US that is lower than their traded consumption price level. A collection of other countries have relative costs of living that are insignificantly different from their relative consumption prices, based on the 95% confidence intervals for the cost of living shown in the second panel. This group includes China and Russia, and several countries in Europe: Austria, Czech Republic, France, Hungary, and nearly Germany. Figure 8 shows the results based on the barcode data for the 24 countries with available data. Here, too, most countries have a higher cost of living relative to the US than their consumption price level, with only the Netherlands at a lower level; while China, France, Germany and the United Kingdom are insignificantly different in the two measures.

To further examine the relationship between CoL^{Ti} and PC^{Ti} and to understand how the different factors contribute to the their difference, we perform a decomposition analysis similar to Eaton, Kortum and Kramarz (2004). We take the difference between the “true” cost of living in (16) and the price of consumption,

$$\Delta \ln CoL^{Ti} \equiv \ln CoL^{Ti} - \ln PC^{Ti}, \qquad (19)$$

corresponding to the second panels in Figures 7 and 8. The log of all the terms appearing on the first line of (16) are denoted by Z_1^i, which we refer to as “trade costs plus the terms of trade”, since they include tariffs (in the import prices), domestic trade costs (when $\omega_s^{Xi} \neq \omega_s^{Mi}$) and the terms of trade. The other terms appearing on the second line of (16) are denoted by $\ln Z_k^i$, $k = 2,3,4,$ which refer to inverse openness, variety, and the scale effect of population We define $\Delta \ln Z_k^i \equiv \ln Z_k^i - \ln PC^{Ti}$ as the difference with the consumption price level, and we run the regressions:

$$\Delta \ln Z_k^i = \gamma_{0k} + \gamma_k \Delta \ln CoL^i \ , \quad k = 1,...,4. \qquad (20)$$

Table 4. Difference between the cost of living and the traded consumption price level

Explanatory variable: $\ln(CoL^{Ti}/PC^{Ti})$	(1) No variety	(2) Firm count	(3) Barcode count	(4) Firm count	(5) Barcode count
Dependent variables:					
Trade costs and terms of trade	0.067	0.097	0.124	0.164	0.209
	(0.053)	(0.046)	(0.088)	(0.044)	(0.095)
Inverse openness	0.933	0.559	0.624	0.312	0.420
	(0.053)	(0.085)	(0.133)	(0.120)	(0.225)
Variety		0.343	0.252	0.369	0.244
		(0.050)	(0.062)	(0.024)	(0.038)
Scale				0.154	0.128
				(0.068)	(0.115)
Number of countries	47	46	24	46	24

Note: Each line in the table corresponds to a γ_k coefficient estimated from equation (20). Robust standard errors are in parentheses.

These regressions aim to account for the cross-country variation in the *difference* between the relative cost of living and the consumption price level. Table 4 presents the results. By construction, the regression coefficients shown in Table 4 sum to unity, so we can interpret them as the portion of the variation in the cost-of-living difference relative to the consumption prices, $\Delta \ln CoL^{Ti} \equiv \ln CoL^{Ti} - \ln PC^{Ti}$, that is explained by the dependent variable in each regression. We focus on the results shown in column (4), representing the broadest sample of countries using firm counts, though the results in column (5) are very similar using barcode counts.

Trade costs and the terms of trade combined have a positive and significant regression coefficient of 0.29 as shown in the first row of Table 4, column (4), indicating that 29% of the variation in the cost-of living as compared to relative consumptions price is explained by those terms. Inverse openness accounts for 24% of the cost-of-living difference, as shown in the second row. The variety and scale terms account for 33% and 14%, respectively, of the variation in the cost-of-living index as compared to the consumption prices. Summing these two terms, we see that the combined effect of product variety and scale account for nearly one-half (47%) of the cost-of-living differences across countries, as compared to their consumption prices. That impact of product variety is rather large, and it suggests that conventional price indexes could miss this source of welfare variation across countries. To explore this issue further, we convert our results for the cost of living into real consumption and compare them to the alternative measure of real consumption developed by Jones and Klenow (2016).

4.3 Comparison with Jones-Klenow

Jones and Klenow (2016) propose a measure of welfare across nations that is meant to be much more inclusive that consumption, by also incorporating leisure, mortality and inequality into a single consumption-equivalent measure. Our analysis, in contrast, is a more restrictive measure of welfare from the Melitz model that incorporates openness and product variety. Despite the differences in our approaches, it is worth asking whether the cross-country variation in welfare – as compared to a conventional measure of real consumption – has any similarity in their analysis and in ours. We shall find that they do.

Up until now, our paper has relied on sectors of tradable goods in household consumption (see Table 1). For a broader comparison, in particular to Jones and Klenow (2016), we now consider including non-traded consumption, for which countries differ substantially in who pays for these products. For healthcare and education, in particular, how much of these are purchased directly by consumers and how much by the government varies considerably across countries. Yet regardless of who pays for them—be it households, non-profit organizations or the government—these services are consumed. We thus use a measure of total consumption that includes these services, corresponding to the statistical concept of “actual individual consumption" (AIC).[22]

We rely on ICP data to calculate the price level of nontraded consumption that is included in AIC, denoted by PC^{Ni} in country *i*. We add these nontraded prices into our previous calculation of the cost of living for traded goods consumption:

$$\frac{CoL^i}{CoL^j} \equiv \left(\frac{CoL^{Ti}}{CoL^{Tj}}\right)^{\omega^{Ti}} \left(\frac{PC^{Ni}}{PC^{Nj}}\right)^{\omega^{Ni}}, \qquad (21)$$

where the Sato-Vartia weights satisfy $\omega^{Ti} + \omega^{Ni} = 1$ (see the Appendix). Likewise the price level of AIC, inclusive of the nontraded services, is constructed as:

$$\frac{PC^i}{PC^j} \equiv \left(\frac{PC^{Ti}}{PC^{Tj}}\right)^{\omega^{Ti}} \left(\frac{PC^{Ni}}{PC^{Nj}}\right)^{\omega^{Ni}}. \qquad (22)$$

We are adding the same nontraded prices to both the cost of living and to the price of consumption goods, so that procedure will tend to reduce any differences in these two measures. Our final step is to use (21) or (22) to deflate nominal AIC in US$ for each country relative to the United States, to obtain theory-based real consumption (or welfare) as compared to actual real consumption using ICP prices.

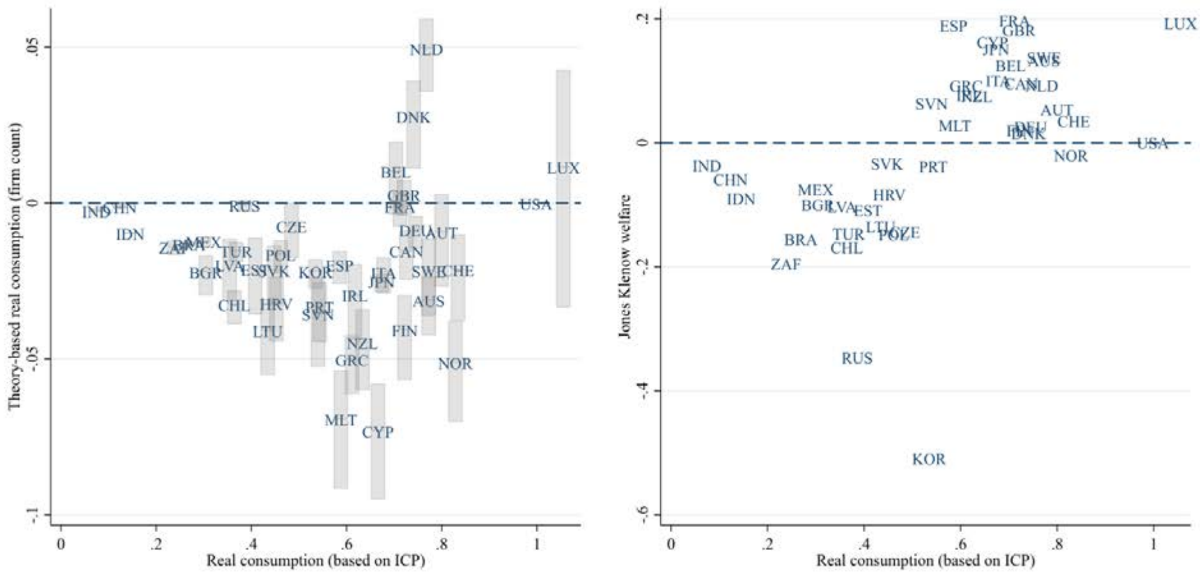
The results when using the square root of the firm count to measure product variety are shown in Figure 9, where our theory-based real consumption measured relative to actual consumption is shown in the first panel, and the Jones-Klenow measure of welfare relative to actual real consumption for a matching set of countries is shown in the second panel.[23] The most obvious difference between the two panels is in the vertical scale of each: welfare in Jones and Klenow differs by ±20% of actual real consumption for all countries except Russia and South

---

[22] We only exclude net purchases of households abroad, which cannot be allocated to a type of products. In their macro-level comparison Jones and Klenow (2016) focus on an even broader measure of consumption that also includes expenditure on collective goods and services. Collective services make up 10–12 percent of the Jones-Klenow consumption measure for most of our set of countries.

[23] Not included are Hungary, Romania and Taiwan, which were not covered by Jones and Klenow (2016).

Figure 9. Ratio of theory-based real consumption to actual real consumption using ICP prices, versus actual real consumption – firm count data



Notes: The left-hand figure plots the log of the ratio of theory-based real consumption (using firm counts) to real consumption (based on ICP prices), against log real consumption, for 43 countries in our sample. The right-hand panel plots the log of ratio of welfare from Jones and Klenow (2016) to actual real consumption (based on ICP prices), against log real consumption, for the matching 43 countries in their sample.

Korea, whereas our measure of real consumption differs by only ±5% from actual real consumption in all countries except Cyprus and Malta. The smaller scale in our case is not surprisingly in view of the more limited scope of our welfare measure.[24] In other respects, however, there are some similarities in the results.

Following countries from the left to the right, India (IND) and China appear first and have theory-based real consumption from the Melitz model that is very close to actual real consumption.[25] Similarly, welfare in these countries is not too far below actual consumption for

[24] Recall that Figure 9 is computed with the high elasticities (σ, θ) = (6.5, 8.3). In Appendix Figure A9 we instead use the low estimates (σ, θ) = (4.4, 5.1), and in that case theory-based real consumption in the first panel is between ±10% of actual real consumption for all countries (including Cyprus and Malta), or still less than one-half of the range for Jones-Klenow welfare.

[25] India had a higher cost of living relative to the US than its traded consumption price level in Figure 7. In Figure 9, first panel, India has theory-based consumption very close to actual real consumption because of its high share of nontraded goods, which move (20) and (21) closer together.

Jones and Klenow. As we move to the right, welfare relative to actual consumption falls for most countries in our sample and also for Jones-Klenow, and then this ratio rises again. The key difference between the two panels in Figure 9 is that welfare relative to the US in nearly all the Western European countries *exceeds* real consumption for Jones and Klenow; whereas in our case, welfare is higher only for Belgium, Denmark, Luxembourg, the Netherlands and the United Kingdom. In addition, for much the same group of European nations with relative costs of living similar to their relative consumption prices in Figure 7 (including Austria, Czech Republic and France), welfare from the Melitz model is insignificantly different from actual real consumption. For these combined European nations along with China, India and Russia, welfare is thus above or comparable to actual real consumption relative to the US in our analysis; but for all other countries, welfare relative to the US falls short of actual real consumption, as also occurs for a reduced set of middle-income countries in Jones and Klenow.

5. Conclusions

The monopolistic competition model suggests that product variety is an important determinant of welfare. There are two challenges with evaluating this hypothesis. First, the most disaggregate data for measuring product variety – which is barcode data – is not typically available across multiple countries with the same classification system. When it is available, as for Mexico and the United States (Argente, Hsieh, and Lee, 2020), then it becomes possible to construct exact consumer price indexes as in Feenstra (1994).[26] In the absence of a common classification system across many countries, we have relied on the count of barcode items from micro-data in the Billion Prices Project; and when those data are not available, on the simple

[26] Note however, that the number of barcodes that is found in both countries is 8.5 percent of the number of total Mexican barcodes and 1.5 percent of US barcodes, so even in this case the number of identical barcodes between the countries is quite limited and could understate the number of identical products.

count of firms as a proxy for variety.

Second, literature on the gains from trade under monopolistic competition (Arkolakis, Costinot and Rodríguez-Clare, 2012) has emphasized the gains within a country when it is shocked by a change in foreign variables, such as trade costs. To evaluate welfare between countries, however, we also need to include shocks to domestic variables, and that is what we have incorporated here. As country size, fixed costs or productivities change, we find that product variety responds endogenously. We develop a parsimonious expression for the "true" costs of living in the Melitz (2003) model that incorporates changes to all these variables, and therefore changes in product variety. We model fixed costs by relying on a simplified version of Arkolakis (2010), where such costs depend on the mass of customers. So our model includes a scale effect of population on variety, with an estimated elasticity of about 0.2. Because we compare the theoretical cost of living with the price level of consumption as measured from ICP data, we also end up comparing the openness of a country (which lowers the theoretical cost of living) with the terms of trade (which lowers the price level of consumption relative to the price of output). Differences between openness and the terms of trade lead to commensurate differences between the cost of living and the consumption price level.

Before adjusting for product variety, roughly an equal number of countries in our sample have a cost of living from the Melitz model that is above versus below their consumption price relative to the US. Those differences are principally explained by the countries' openness as compared to their terms of trade. The United States, however, has higher product variety than nearly all other countries.[27] Once we incorporate variety, then, the relative cost of living is raised in many countries, and we find that only five countries – Belgium, Denmark, Luxembourg,

[27] See note 21.

Netherlands and the United Kingdom – have a cost of living relative to the US that is significantly below its consumption price, while China, France and Russia are very close. There is a collection of other European countries – including Austria, Czech Republic, Germany and Hungary – whose relative costs of living are not significantly different from their relative consumption prices.

We have also used cost of living and consumption price levels to compute theoretical measure of welfare and real actual individual consumption (AIC) across countries. By construction, our theoretical measure of welfare varies inversely with the theoretical cost of living: Belgium, Denmark, Luxembourg, Netherlands and the United Kingdom now have welfare relative to the US above actual consumption, while a further group of European countries, along with China, India, and Russia, have relative welfare that is not substantially different from actual consumption; the remaining set of countries have lower relative welfare. That pattern is more pronounced in Jones and Klenow (2016), where nearly all Western European countries have welfare relative to the US exceeding real actual consumption, whereas a reduced set of middle-income countries have lower welfare. It is surprisingly but perhaps reassuring that our narrow focus on the determinants of welfare in the Melitz model leads to a pattern of welfare across countries that has similarities to Jones and Klenow (2016), even though they focus on much broader determinants of welfare.

References

Argente, David, Chang-Tai Hsieh, Munseob Lee. 2020. "Measuring the Cost of Living in Mexico and the US," NBER Working Paper 27806,

Arkolakis, Costas, Arnaud Costinot and Andrés Rodríguez-Clare. 2012. "New Trade Models, Same Old Gains?" *American Economic Review* 102(1): 94-130.

Arkolakis, Costas. 2010. "Market Penetration Costs and the New Consumers Margin in International Trade," *Journal of Political Economy*, 118(8): 1151 – 1199.

Cavallo, Alberto, W. Erwin Diewert, Robert Inklaar, and Marcel P. Timmer. 2018. "Using Online Prices for Measuring Real Consumption across Countries," *American Economic Review*, 108, May, 483-487.

Chaney, Thomas. 2008. "Distorted Gravity: the Intensive and Extensive Margins of International Trade," *American Economic Review* 98(4): September, 1707-1721.

Broda, Christian and David Weinstein. 2006. "Globalization and the Gains from Product Variety." *Quarterly Journal of Economics,* 121(2), May, 541-585.

Broda, Christian and David Weinstein. 2008. "Understanding International Price Differences Using Barcode Data." NBER Working Paper 14017.

Eaton, Jonathan and Samuel Kortum. 2002. "Technology, Geography and Trade." *Econometrica*, 70(7): September, 1741-1780.

Eaton, Jonathan, Samuel Kortum and Francis Kramarz. 2004. "Dissecting Trade: Firms, Industries, and Export Destinations." *American Economic Review*, 94(2): May, 150-154.

Eaton, Jonathan, Samuel Kortum and Francis Kramarz. 2011. "An Anatomy of International Trade: Evidence from French Firms." *Econometrica*, 79(5): 1453–98.

Feenstra, Robert C. 1994. "New Product Varieties and the Measurement of International Prices," *American Economic Review*, 84(1): March, 157-177

Feenstra, Robert C., Robert Inklaar and Marcel Timmer. 2015. "The Next Generation of the Penn World Table," *American Economic Review*, 105(10): October, 3150-3182.

Feenstra, Robert C., Mingzhi Xu and Alexis Antoniades. 2020. "What is the Price of Tea in China? Towards the Relative Cost of Living in Chinese and US Cities," *The Economic Journal*, 130(632), November, 2438–2467.

Handbury, Jessie and David E. Weinstein. 2015. "Goods Prices and Availability in Cities." *Review of Economic Studies*, 82(1): 258-296.

Jones, Charles I. and Peter J. Klenow. 2016. "Beyond GDP? Welfare across Countries and Time." *American Economic Review,* 106(9): 2426-2457.

Kathuria, Sanjay, Arti Grover, Viviana Maria Eugenia Perego, Aaditya Mattoo, and Pritam Banerjee. 2019. *Unleashing E-Commerce for South Asian Integration*. The World Bank.

Melitz, Marc. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica*, 71(6): November, 1695-1725.

Redding, Stephen J. and David Weinstein. 2020. "Measuring Aggregate Price Indexes with Taste Shocks: Theory and Evidence for CES Preferences." *Quarterly Journal of Economics*, 135(1): 503-560.

Simonovska, Ina and Michael E. Waugh. 2014. "Trade Models, Trade Elasticities, and the Gains from Trade," NBER Working Paper 20495.

UNCTAD 2017. "B2C E-Commerce Index 2017." *UNCTAD Technical Notes on ICT for Development*, 2017. https://unctad.org/system/files/official-document/tn_unctad_ict4d09_en.pdf.

World Bank. 2014. *Purchasing Power Parities and the Real Size of World Economies. A Comprehensive report of the 2011 International Comparison Program.* World Bank: Washington: DC.

Product Variety, the Cost of Living and Welfare Across Countries

Online Appendix: Not for Publication

Proof of Proposition 1:

The final equality in (9) uses $M_e \propto L / f_e$. To prove this condition, we complete the description of the model in part (a) below, and then we prove Proposition 1 in part (b).

a) With CES demand using the consumer price $p_d(\varphi) = [\sigma / (\sigma - 1)]\,(w\tau_d / \varphi)$, and total home expenditure of X, the home demand for a firm with productivity φ is:

$$y_d(\varphi) = \frac{X}{P^{1-\sigma}} \left[\frac{w\tau_d \sigma}{\varphi(\sigma - 1)} \right]^{-\sigma}.$$

Multiplying by price minus variable cost, $p_d(\varphi) - (w\tau_d / \varphi) = [1/(\sigma - 1)]\,(w\tau_d / \varphi)$, profits in the home market are,

$$\pi_d(\varphi) = \underbrace{\left[\frac{X}{\sigma^\sigma} \left(\frac{w\tau_d}{P(\sigma - 1)} \right)^{1-\sigma} \right]}_{B_d} \varphi^{\sigma-1} - wf_d,$$

where f_d are the fixed costs that use labor in the domestic market. It follows that the zero-cutoff-profit (ZCP) condition in the domestic market is:

$$\pi_d(\varphi_d) = B_d \varphi_d^{\sigma-1} - wf_d = 0 \quad \Rightarrow \quad \varphi_d^{\sigma-1} = \frac{wf_d}{B_d} = \frac{wf_d \sigma^\sigma}{X} \left(\frac{w\tau_d}{P(\sigma - 1)} \right)^{\sigma-1}, \qquad \text{(A1)}$$

which also appears as condition (7) in the main text.

With iceberg costs of exporting τ_x, export prices are $p_x(\varphi) = [\sigma / (\sigma - 1)]\,(w\tau_x / \varphi)$ and so demand for the home firm with productivity φ is:

$$y_x(\varphi) = \frac{X^*}{P^{*1-\sigma}} \left[\frac{w\tau_x \sigma}{\varphi(\sigma - 1)} \right]^{-\sigma}.$$

Multiplying by price minus variable cost, $p_x(\varphi) - (w\tau_x / \varphi) = [1/(\sigma - 1)]\,(w\tau_x / \varphi)$, profits in the export market are,

$$\pi_x(\varphi) = \underbrace{\left[\frac{X^*}{\sigma^\sigma} \left(\frac{w\tau_x}{P^*(\sigma-1)} \right)^{1-\sigma} \right]}_{B_x} \varphi^{\sigma-1} - wf_x ,$$

where f_x are the fixed costs for exporting. It follows that the zero-cutoff-profit condition in the export market is:

$$\pi_x(\varphi_x) = B_x \varphi_x^{\sigma-1} - wf_x = 0 \quad \Rightarrow \quad \varphi_x^{\sigma-1} = \frac{wf_x}{B_x} = \frac{wf_x \sigma^\sigma}{X^*} \left(\frac{w\tau_x}{P^*(\sigma-1)} \right)^{\sigma-1} . \tag{A2}$$

Total employment at home for domestic and export sales equals:

$$L = M_e f_e + M_d \int_{\varphi_d}^{\infty} \left[\frac{\tau_d y_d(\varphi)}{\varphi} + f_d \right] \frac{g(\varphi)}{[1-G(\varphi_d)]} d\varphi + M_x \int_{\varphi_x}^{\infty} \left[\frac{\tau_x y_x(\varphi)}{\varphi} + f_x \right] \frac{g(\varphi)}{[1-G(\varphi_x)]} d\varphi . \tag{A3}$$

Notice that we have multiplied the quantity delivered to home and foreign consumers by their respective iceberg costs, τ_d and τ_x, to obtain the quantity produced by the firm. Multiply the entire expression by wages *w*, and then multiply and divide the production terms by $\sigma/(\sigma-1)$ to obtain prices $p_d(\varphi) = (\tau_d/\varphi)[\sigma/(\sigma-1)]$ and $p_x(\varphi) = (\tau_x/\varphi)[\sigma/(\sigma-1)]$, so that:

$$\begin{aligned} wL &= w\left(M_e f_e + M_d f_d + M_x f_x\right) + \left(\frac{\sigma-1}{\sigma} \right) \left[M_d \int_{\varphi_d}^{\infty} \frac{p_d(\varphi) y_d(\varphi) g(\varphi)}{[1-G(\varphi_d)]} d\varphi + M_x \int_{\varphi_x}^{\infty} \frac{p_x(\varphi) y_x(\varphi) g(\varphi)}{[1-G(\varphi_x)]} d\varphi \right] \\ &= w\left(M_e f_e + M_d f_d + M_x f_x\right) + \left(\frac{\sigma-1}{\sigma} \right) wL, \end{aligned}$$

where the bracketed term on the first line is total revenue earned by firms. With zero expected profits, total revenue equals the payment to labor *wL*, so then $L = \sigma\left(M_e f_e + M_d f_d + M_x f_x\right)$ is obtained. It follows that the full employment condition (A3) is simplified as:

$$\left(\frac{\sigma-1}{\sigma} \right) L = M_d \int_{\varphi_d}^{\infty} \left[\frac{\tau_d y_d(\varphi)}{\varphi} \right] \frac{g(\varphi)}{[1-G(\varphi_d)]} d\varphi + M_x \int_{\varphi_x}^{\infty} \left[\frac{\tau_x y_x(\varphi)}{\varphi} \right] \frac{g(\varphi)}{[1-G(\varphi_x)]} d\varphi . \tag{A4}$$

CES demand with prices $p_d(\varphi) = (\tau_d/\varphi)[\sigma/(\sigma-1)]$ implies that $y_d(\varphi) = (\varphi/\varphi_d)^\sigma y_d(\varphi_d)$.

Using the Pareto distribution for productivity, the first integral in (A4) is then:

$$\int_{\varphi_d}^{\infty}\left[\frac{\tau_d y_d(\varphi)}{\varphi}\right]\frac{g(\varphi)}{[1-G(\varphi_d)]}d\varphi=\int_{\varphi_d}^{\infty}\frac{\tau_d y_d(\varphi_d)}{\varphi}\left(\frac{\varphi}{\varphi_d}\right)^{\sigma}\frac{g(\varphi)}{[1-G(\varphi_d)]}d\varphi$$

$$=\frac{\tau_d y_d(\varphi_d)}{\varphi_d}\int_{\phi_d}^{\infty}\left(\frac{\varphi}{\varphi_d}\right)^{\sigma-1}\frac{\theta\varphi^{-\theta-1}}{(\varphi_d)^{-\theta}}d\varphi$$

$$=\frac{\tau_d y_d(\varphi_d)}{\varphi_d}\frac{\theta}{(\sigma-\theta-1)}\left(\frac{\varphi}{\varphi_d}\right)^{\sigma-\theta-1}\Bigg|_{\varphi_d}^{\infty}$$

$$=f_d\,\frac{(\sigma-1)\theta}{(\theta-\sigma+1)},$$

where the last line uses $\tau_d y_d(\varphi_d)/\varphi_d=(\sigma-1)f_d$, as follows from (A1). Likewise we have $\tau_x y_x(\varphi_x)/\varphi_x=(\sigma-1)f_x$ from (A2), and so the second integral in (A4) is evaluated as:

$$\int_{\varphi_x}^{\infty}\left[\frac{\tau_x y_x(\varphi)}{\varphi}\right]\frac{g(\varphi)}{[1-G(\varphi_x)]}d\varphi=f_x\,\frac{(\sigma-1)\theta}{(\theta-\sigma+1)}.$$

Substituting these back into (A4) we arrive at:

$$L=\frac{\sigma\theta}{(\theta-\sigma+1)}\left(M_d f_d+M_x f_x\right).$$

Using $L=\sigma\left(M_e f_e+M_d f_d+M_x f_x\right)$ we obtain $M_e=L(\sigma-1)/\sigma\theta f_e$, so that $M_e\propto L/f_e$.

b) Now completing the proof of Proposition 1, from (6) we have:

$$\frac{w'/P'}{w/P}=\left(\frac{M_d'/\lambda_d'}{M_d/\lambda_d}\right)^{\frac{1}{\sigma-1}}\left(\frac{\tau_d'}{\tau_d}\right)^{-1}\left(\frac{\varphi_d'}{\varphi_d}\right).\tag{A5}$$

The final ratio on the right of (A5) is solved using (9) as,

$$\frac{\varphi_d'}{\varphi_d}=\frac{A'}{A}\left(\frac{\lambda_d'}{\lambda_d}\right)^{-1/\theta}\left(\frac{X'/w'L'}{X/wL}\right)^{-1/\theta},\tag{A6}$$

where $f_d'/f_e'=f_d/f_e$ from Assumption 1. Substituting (A6) into (A5), we obtain (10).

Under Assumption 1′ and (8), the final term in (10) becomes:

$$\left(\frac{X'/w'L'}{X/wL}\right)^{-\frac{1}{\theta}} = \left(\frac{X'/w'}{X/w}\right)^{\frac{1-\alpha}{\alpha\theta}} \left(\frac{X'/w'f'_d}{X/wf_d}\right)^{-\frac{1}{\alpha\theta}} = \left(\frac{X'/w'}{X/w}\right)^{\frac{1-\alpha}{\alpha\theta}} \left(\frac{M'_d/\lambda'_d}{M_d/\lambda_d}\right)^{-\frac{1}{\alpha\theta}}. \qquad \text{(A7)}$$

We can use (A7) to solve for product variety as in (12). To obtain real wages, we use (12) to solve for $(X'/w'L')/(X/wL)$ and substitute that into (10) to obtain (11). QED

Sato-Vartia weights:

We consider the general case of a nested CES function, where the expenditure across traded goods is aggregated using a CES function, the expenditure across the various traded sectors is aggregated using a second CES function, and then nontraded goods included within "actual individual consumption" (AIC) are added with a third CES function.

At the lowest level, the traded goods price index P_s^{Ti} is obtained from the prices of goods purchased from home, P_s^{Tii}, and those that are purchased from abroad, P_s^{Tji}, $j \neq i$:

$$P_s^{Ti} = \left[\sum_{j=1}^{C} b_s^j (P_s^{Tji})^{1-\sigma}\right]^{1/(1-\sigma)}, \ \sigma > 1. \qquad \text{(A8)}$$

This price index is comparable to what appears in (1) in our model, where the mass of products from each country in (1) is captured above by the (constant) parameter b_s^j. Above this level, the price index of traded goods P^{Ti} for country *i* is given by:

$$P^{Ti} = \left[\sum_{s=1}^{S} b_s (P_s^{Ti})^{1-\eta}\right]^{1/(1-\eta)}, \ \eta > 0.$$

Finally, we denote the price of nontraded goods included in AIC by P^{Ni}, which could be an aggregate over multiple sectors, and construct the overall price index in country *i* as:

$$P^i = \left[b_N^i (P^{Ni})^{1-\delta} + b_T^i (P^{Ti})^{1-\delta}\right]^{1/(1-\delta)}, \ \delta > 0.$$

Choose country *j* (i.e. the United States) as the base country. Then the traded goods price index in country *i* relative to *j* can be measured by the Sato-Vartia price index:

$$\frac{P^{Ti}}{P^{Tj}} = \prod_{s=1}^{S} \left(\frac{P_s^{Ti}}{P_s^{Tj}} \right)^{\omega_s^{Ti}}, \tag{A9}$$

where the Sato-Vartia weights, $\sum_{s=1}^{S} \omega_s^{Ti} = 1$, are defined over the expenditure shares on traded goods. Since we have already used the variable *X* to denote expenditures and *s* to denote sectors, we will use *x* to denote expenditure shares. So $x_s^{Ti} \equiv X_s^{Ti} / X^{Ti}$ is the share of expenditure on traded goods of sector *s* relative to total expenditure, $X^{Ti} = \sum_{s=1}^{S} X_s^{Ti}$, in country i. Then the Sato-Vartia weights used in (A9) are:

$$\omega_s^{Ti} \equiv \frac{(x_s^{Ti} - x_s^{Tj})}{(\ln x_s^{Ti} - \ln x_s^{Tj})} \Bigg/ \left[\sum_{s=1}^{S} \frac{(x_s^{Ti} - x_s^{Tj})}{(\ln x_s^{Ti} - \ln x_s^{Tj})} \right].$$

These are the Sato-Vartia weights that appear in (14), (15) and (16) in the main text.

When we include the nontraded services that are part of AIC, the overall price index is:

$$\frac{P^i}{P^j} = \left(\frac{P^{Ti}}{P^{Tj}} \right)^{\omega_s^{Ti}} \left(\frac{P^{Ni}}{P^{Nj}} \right)^{\omega_s^{Ni}}, \tag{A10}$$

where total AIC expenditure in country *i* is $X^i \equiv (X^{Ti} + X^{Ni})$ with the expenditure shares $x^{Ti} \equiv X^{Ti} / X^i$ and $x^{Ni} \equiv X^{Ni} / X^i$, and so the Sato-Vartia weights used in (A10) are:

$$\omega^{Ti} \equiv \frac{(x^{Ti} - x^{Tj})}{(\ln x^{Ti} - \ln x^{Tj})} \Bigg/ \left[\frac{(x^{Ti} - x^{Tj})}{(\ln x^{Ti} - \ln x^{Tj})} + \frac{(x^{Ni} - x^{Nj})}{(\ln x^{Ni} - \ln x^{Nj})} \right], \quad \omega_s^{Ni} \equiv 1 - \omega_s^{Ti}.$$

These weights appear in (21) and (22) in the main text.

Output prices:

To construct a measure of output prices used in (14), we use the above equations and also follow the framework of Inklaar and Timmer (2014). The price index P_s^{Ti} in (A8) combines domestically produced goods, with price P_s^{Tii} and imports with price P_s^{Tji} for $j \neq i$. We define P_s^{Mi} as the import price index,

$$P_s^{Mi} = \left[\sum_{j \neq i} b_s^j (P_s^{Tji})^{1-\sigma} \right]^{1/(1-\sigma)},$$

so that the overall traded goods price index in sector *s* can be constructed as,

$$\frac{P_s^{Ti}}{P_s^{Tj}} = \left(\frac{P_s^{Tii}}{P_s^{Tjj}} \right)^{1-\omega_s^{Mi}} \left(\frac{P_s^{Mi}}{P_s^{Mj}} \right)^{\omega_s^{Mi}},$$

where $x^{Tii} \equiv X^{Tii} / X^{Ti}$ and $x^{Mi} \equiv \sum_{j \neq i} X^{Tji} / X^{Ti}$ are the expenditure shares on domestic goods and imports, respectively, and the Sato-Vartia weights on imports is:

$$\omega_s^{Mi} \equiv \frac{(x_s^{Mi} - x_s^{Mj})}{(\ln x_s^{Mi} - \ln x_s^{Mj})} \Bigg/ \left[\frac{(x_s^{Tji} - x_s^{Tjj})}{(\ln x_s^{Tii} - \ln x_s^{Tjj})} + \frac{(x_s^{Mi} - x^{Mj})}{(\ln x^{Mi} - \ln x^{Mj})} \right]. \quad \text{(A11)}$$

Notice that from (A11) we construct the domestic price of tradable goods as:

$$\frac{P_s^{Tii}}{P_s^{Tjj}} = \left(\frac{P_s^{Ti}}{P_s^{Tj}} \right)^{1/(1-\omega_s^{Mi})} \left(\frac{P_s^{Mi}}{P_s^{Mj}} \right)^{-\omega_s^{Mi}/(1-\omega_s^{Mi})}. \quad \text{(A12)}$$

All these prices are inclusive of the domestic trade costs τ_s^{ii} needed to deliver a good to consumers, while import prices also include foreign trade costs . τ_s^{ji} [28] We let $\tilde{P}_s^{Tii} \equiv P_s^{Tii} / \tau_s^{ii}$ denote the prices *net* of the domestic trade costs – or what is called a "basic" price – which is the price that home producers face for domestic sales. Home firms also export, so the total value of home production Y_s^{Ti} on tradable goods equals:

$$Y_s^{Ti} \equiv P_s^{Yi} Q_s^{Yi} = \tilde{P}_s^{Tii} Q_s^{Tii} + \tilde{P}_s^{Xi} Q_s^{Xi} = (X_s^{Ti} / \tau_s^{ii}) + \tilde{P}_s^{Xi} Q_s^{Xi},$$

[28] For simplicity, we assume that domestic trade costs are identical for domestically produced and imported goods.

where $\tilde{P}_s^{Xi}$ is the export price index, the *Q's* denote the associated quantities, and sales to home consumers net of trade costs is $\tilde{P}_s^{Tii}Q_s^{Tii} = X_s^{Tii} / \tau_s^{ii}$. The export price index is defined using f.o.b. (free-on-board) prices net of any trade costs (i.e. net of transport costs and tariffs),

$$\tilde{P}_s^{Xi} = \left[\sum\nolimits_{j \neq i} b_s^i (P_s^{Tij} / \tau_s^{ij})^{1-\sigma}\right]^{1/(1-\sigma)}.$$

Assuming a CES production function for domestic consumption and exports, the output price is constructed as a Sato-Vartia index:

$$\frac{P_s^{Yi}}{P_s^{Yj}} = \left(\frac{\tilde{P}_s^{Tii}}{\tilde{P}_s^{Tjj}}\right)^{1-\omega_s^{Xi}} \left(\frac{\tilde{P}_s^{Xi}}{\tilde{P}_s^{Xj}}\right)^{\omega_s^{Xi}}, \qquad \text{(A13)}$$

where $y_s^{Tii} \equiv (X_s^{Tii} / \tau_s^{ii}) / Y_s^{Ti}$ and $y^{Xi} \equiv 1 - y_s^{Tii}$ are the production shares on domestic goods and exports, respectively, and the associated Sato-Vartia weight on exports is:

$$\omega_s^{Xi} \equiv \frac{(y_s^{Xi} - y_s^{Xj})}{(\ln y_s^{Xi} - \ln y_s^{Xj})} \Bigg/ \left[\frac{(y_s^{Tii} - y_s^{Tjj})}{(\ln y_s^{Tii} - \ln y_s^{Tjj})} + \frac{(y_s^{Xi} - y_s^{Xj})}{(\ln y_s^{Xi} - \ln y_s^{Xj})}\right]. \qquad \text{(A14)}$$

Substituting (A12) into (A13) we obtain the price of output:

$$\begin{aligned}
\frac{P_s^{Yi}}{P_s^{Yj}} &= \left(\frac{P_s^{Tii} / \tau_s^{ii}}{P_s^{Tjj} / \tau_s^{jj}}\right)^{1-\omega_s^{Xi}} \left(\frac{\tilde{P}_s^{Xi}}{\tilde{P}_s^{Xj}}\right)^{\omega_s^{Xi}} \\
&= \left(\frac{P_s^{Ti}}{P_s^{Tj}}\right)^{\frac{(1-\omega_s^{Xi})}{(1-\omega_s^{Mi})}} \left(\frac{\tau_s^{ii}}{\tau_s^{jj}}\right)^{-(1-\omega_s^{Xi})} \left(\frac{\tilde{P}_s^{Mi}\tau_s^{ii}}{\tilde{P}_s^{Mj}\tau_s^{ii}}\right)^{-\frac{\omega_s^{Mi}(1-\omega_s^{Xi})}{(1-\omega_s^{Mi})}} \left(\frac{\tilde{P}_s^{Xi}}{\tilde{P}_s^{Xj}}\right)^{\omega_s^{Xi}} \qquad \text{(A15)} \\
&= \left(\frac{P_s^{Ti} / \tau_s^{ii}}{P_s^{Tj} / \tau_s^{jj}}\right)^{\frac{(1-\omega_s^{Xi})}{(1-\omega_s^{Mi})}} \left(\frac{\tilde{P}_s^{Xi}}{\tilde{P}_s^{Xj}}\right)^{\omega_s^{Xi}} \left(\frac{\tilde{P}_s^{Mi}}{\tilde{P}_s^{Mj}}\right)^{-\frac{\omega_s^{Mi}(1-\omega_s^{Xi})}{(1-\omega_s^{Mi})}},
\end{aligned}$$

where the first line comes from using $\tilde{P}_s^{Tii} \equiv P_s^{Tii} / \tau_s^{ii}$ in (A13), the second line comes from substituting (A12) and using $\tilde{P}_s^{Mi} \equiv P_s^{Mi} / \tau_s^{ii}$ to denote the c.i.f. and tariff-inclusive prices of imports, but net of *domestic* trade costs, and the third line from simplification.

The price indexes that we have constructed so far are the theoretically correct CES indexes. To relate these to the price level that we construct from ICP and PWT data, let us start with (A10). The price ratio on the left is what we measure as the price level of consumption for traded goods, so we replace P_s^{Ti} / P_s^{Tj} with PC_s^{Ti} / PC_s^{Tj}. This price level of consumption also appears first on the second and third lines of (A13). In that case, the price of output P_s^{Yi} / P_s^{Yj} appearing on the left is replaced with PY_s^{Ti} / PY_s^{Tj}. Finally, the export and import prices $\tilde{P}_s^{Xi} / \tilde{P}_s^{Xj}$ and $\tilde{P}_s^{Mi} / \tilde{P}_s^{Mj}$ are measured by the quality-adjusted export and imports prices from Feenstra and Romalis (2014), for $j=USA$.

To implement the resulting equations which appears as (15) and (16) in the main text, we draw on the World Input-Output Tables (Timmer et al., 2015, 2016) for calculating the Sato-Vartia weights for import and export shares, as shown in (A11) and (A14). For Colombia, Chile, New Zealand and South Africa, we use the data from the OECD TiVA tables. The traded consumption prices are the same as discussed in the main text, aggregated from the (revised) ICP 2011 PPPs and consumption expenditure data using GEKS indexes. The import and export price data are organized by SITC rev. 2, so first we use the concordance to the Broad Economic Category (BEC-4) classification to select only traded products consumed by households.[29] Second, we use the concordance between 4-digit SITC rev. 2 and 3-digit ISIC rev. 2 constructed by Marc Muendler,[30] and bridge that to ISIC rev. 4, the industry classification used in WIOD and OECD TiVA. We aggregate to ISIC rev. 4 industries using export values from Comtrade and GEKS indexes. In the final step, we use export values by ISIC rev. 4 industry from WIOD and

[29] We select food and beverages, mainly for household consumption, primary (BEC code 112) and processed (122); processed fuels and lubricants (32), transport equipment, passenger motor cars (51) and consumption goods (6). This selection means that products used by industry, as supplies or capital goods, are omitted.
[30] https://econweb.ucsd.edu/muendler/docs/conc/sitc2isic.pdf.

OECD TiVA to aggregate to the traded consumption sectors.

Other data:

Other data used in (14) is obtained as follows. The share of consumption expenditure on domestic products, λ_s^{ii}, is computed based on WIOD, as are the trade flows for the gravity equation estimation. Colombia, Chile, New Zealand and South Africa are not in WIOD, so we use the inter-country input-output tables of OECD TiVA to compute λ_s^{ii} for those countries. Domestic trade costs τ_s^{ii} in sector s are measured as consumption expenditure at purchaser's prices divided by consumption expenditure at basic prices, which excludes the margin earned in transportation and retail trade and excludes taxes on products, notably sales tax, VAT and excise taxes. For most countries, we rely on the margins and tax tables (sometimes also referred to as valuation tables) provided by Eurostat and the OECD, which report consumption at purchaser's prices and at basic prices. For the remainder of countries, we use data from national input-output tables, from Eurostat's Structural Business Statistics for retail trade, or WIOD to approximate trade margins.[31] To estimate consumption taxes by sector, we use information on total taxes on products by sector and ensure that the tax rate (taxes as a share of consumption expenditure at purchaser's prices) does not exceed that country's indirect tax rates.[32]

[31] We rely on national input-output data for China, Japan, Indonesia, New Zealand, Russia, South Africa, Taiwan; Eurostat retail survey data for Germany, Spain and Switzerland; and WIOD data for India. The retail survey data abstracts from transportation margins, but most transportation costs are registered as intermediate inputs rather than as margins.

[32] Country-level indirect tax rates are from the OECD Consumption Tax Trends 2018 publication. On average across European countries with the requisite data, only 60 percent of taxes on products are borne directly by consumers, so scaling is important. Excise taxes on alcoholic beverages, tobacco and fuel lead to higher tax rates in the food and transport sectors so in those sectors, the tax rate is allowed to exceed the national indirect tax rate, though not by more than the maximum excess rate observed in other European countries. In Japan, a uniform VAT rate of 5 percent is applied to all sectors, which is increased by an additional 5.8 percent in the food and transport sectors based on estimates of the revenue from excise taxes relative to VAT in the OECD Consumption Tax Trends 2018 publication.

Appendix Table A1. Orbis Firm Counts in 46 Countries

Sector	Food & beverages	Clothing & footwear	Furnishing & household eq.	Health	Transportation	Recreation & culture	Other goods & services
COICOP	01-02	03	05	06	07	09	12
Australia	18,315	7,954	10,152	926	3,947	21,352	836
Austria	5,416	3,316	5,841	315	725	3,096	199
Belgium	15,361	6,910	7,274	2,040	1,326	14,775	395
Brazil	31,524	30,127	29,188	2,271	5,616	38,407	3,689
Bulgaria	12,144	10,692	4,652	90	184	2,679	810
Canada	5,595	3,458	10,927	1,071	1,822	14,037	1,207
Chile	3,676	1,290	1,153	58	97	1,369	118
China	61,415	56,998	47,803	6,684	19,681	74,328	15,590
Croatia	4,165	2,332	1,969	66	161	2,584	321
Cyprus	784	280	323	13	45	264	37
Czechia	21,975	23,053	36,497	111	1,078	13,023	441
Denmark	2,330	804	1,491	155	163	1,545	140
Estonia	1,262	1,700	1,373	27	105	995	90
Finland	3,534	5,199	3,207	75	445	4,062	218
France	28,419	5,561	7,384	447	1,649	10,800	1,024
Germany	19,856	5,191	17,666	1,776	2,991	26,788	1,793
Greece	2,503	992	769	130	56	873	284
Hungary	8,518	5,910	5,717	153	653	6,841	1,672
India	15,945	11,074	6,927	9,829	3,242	9,224	1,437
Indonesia	1,717	1,425	1,404	376	471	2,647	373
Ireland	1,608	436	869	237	209	1,635	51
Italy	45,299	62,181	27,315	719	2,690	21,045	2,918
Japan	21,175	5,714	11,724	698	4,040	18,549	4,094
Korea	16,679	8,928	11,185	1,155	10,147	27,288	3,511
Latvia	1,300	1,203	1,119	49	74	899	113
Lithuania	1,305	1,171	1,403	31	58	632	129
Luxembourg	229	39	86	10	20	162	7
Malta	90	22	66	23	6	70	7
Mexico	9,914	6,347	5,947	793	2,358	5,364	1,530
Netherlands	8,173	5,196	11,810	428	927	7,599	401
New Zealand	5,079	1,001	1,477	245	507	2,170	112
Norway	3,605	3,548	2,554	69	170	2,753	77
Poland	14,740	13,183	15,117	498	1,505	10,034	2,490
Portugal	8,477	8,875	4,911	178	551	2,781	434
Romania	18,805	12,261	7,706	158	532	4,338	1,003
Russia	60,918	50,262	49,429	1,162	2,074	29,265	2,777
Slovakia	4,525	4,763	4,859	36	522	3,537	307
Slovenia	3,075	1,397	2,175	38	230	2,028	187
South Africa	17,921	4,269	4,941	788	707	7,579	1,293
Spain	21,845	12,811	12,563	510	1,881	14,150	1,210
Sweden	2,681	699	2,502	160	657	2,896	188
Switzerland	5,920	2,945	2,867	544	317	6,150	186
Taiwan	7,768	3,903	12,912	459	1,650	15,698	2,236
Turkey	13,550	17,559	18,552	838	3,023	11,165	1,553
United Kingdom	22,919	13,892	18,259	1,704	4,786	26,748	2,816
United States	56,689	28,022	95,313	10,917	19,661	140,302	9,063

Appendix Table A2. BPP Barcode Counts by Sector in 24 Countries

Sector COICOP	Food & beverages 01-02	Clothing & footwear 03	Furnishing & household eq. 05	Recreation & culture 09	Other goods & services 12
Australia	9,738	64,319	11,513	29,217	3,205
Brazil	7,721	11,493	133,418	70,128	14,844
Canada	13,502	17,224	38,401	30,910	9,969
Chile	3,680	16,205	25,516	6,810	8,911
China	22,123	87,193	59,736	23,065	19,662
Colombia	5,707	15,975	13,003	5,694	3,515
Germany	15,860	26,219	87,676	98,334	22,678
Spain	12,741	60,832	43,763	35,568	23,077
France	11,235	26,766	183,281	23,793	3,782
United Kingdom	11,996	39,254	26,142	19,880	13,237
Greece	4,454	7,236	30,092	11,678	5,989
India	4,039	38,675	4,091	2,019	1,614
Ireland	9,162	8,896	11,389	3,005	7,636
Italy	7,819	13,434	41,214	14,348	3,248
Japan	16,163	136,015	160,810	165,692	85,293
Korea	41,641	47,999	95,512	42,891	26,467
Mexico	7,789	17,137	17,269	7,275	7,626
Netherlands	12,038	38,104	42,526	17,533	12,634
New Zealand	7,006	11,613	26,341	20,800	8,591
Poland	7,927	1,221	19,268	28,590	3,784
Russia	7,821	13,755	38,533	21,049	3,567
Turkey	6,753	37,719	32,532	8,910	11,244
United States	22,386	57,305	185,983	80,598	29,671
South Africa	9,493	4,901	10,182	14,152	11,150

In Appendix Tables A3 and A4, we show the log values of the terms in (16) to provide a decomposition of the ratio of the cost of living to the consumption price. Table A3 uses the firm counts, and Table A4 use the barcode counts.

Appendix Table A3. Traded consumption prices, the cost of living and a decomposition, Orbis firm counts

Country	ISO-code	PC	CoL	ln(CoL/PC)	Due to:			
					Trade costs and terms of trade	Inverse openness	Variety	Scale
India	IND	0.466	0.507	0.083	0.000	0.047	0.070	-0.034
				[0.057, 0.100]				[-0.060, -0.017]
Taiwan	TWN	0.682	0.757	0.105	-0.011	-0.009	0.062	0.063
				[0.072, 0.154]				[0.031, 0.112]
Indonesia	IDN	0.727	0.858	0.165	0.024	0.017	0.118	0.006
				[0.162, 0.169]				[0.003, 0.010]
China	CHN	0.813	0.829	0.020	-0.007	0.041	0.023	-0.036
				[-0.008, 0.038]				[-0.064, -0.018]
Russia	RUS	0.838	0.846	0.009	-0.039	0.003	0.026	0.019
				[-0.001, 0.024]				[0.009, 0.033]
Mexico	MEX	0.859	0.960	0.111	0.002	0.012	0.073	0.024
				[0.099, 0.130]				[0.012, 0.043]
Poland	POL	0.874	0.965	0.099	0.038	-0.032	0.042	0.051
				[0.073, 0.138]				[0.025, 0.091]
Bulgaria	BGR	0.909	1.114	0.203	0.054	-0.022	0.080	0.091
				[0.156, 0.274]				[0.045, 0.162]
Turkey	TUR	0.911	1.013	0.106	-0.004	0.016	0.059	0.035
				[0.088, 0.134]				[0.017, 0.063]
Romania	ROU	0.927	1.083	0.155	0.024	-0.006	0.070	0.067
				[0.121, 0.207]				[0.033, 0.118]
South Africa	ZAF	0.942	1.115	0.169	0.045	0.008	0.073	0.044
				[0.147, 0.203]				[0.021, 0.077]
Hungary	HUN	0.965	1.014	0.050	0.049	-0.109	0.026	0.084
				[0.007, 0.115]				[0.041, 0.149]
United States	USA	1.000	1.000	0.000	0.000	0.000	0.000	0.000
Lithuania	LTU	1.003	1.269	0.235	0.034	-0.035	0.123	0.113
				[0.178, 0.322]				[0.056, 0.200]
Slovakia	SVK	1.031	1.184	0.139	0.054	-0.075	0.061	0.099
				[0.089, 0.215]				[0.049, 0.175]
Czechia	CZE	1.031	1.076	0.042	0.021	-0.080	0.019	0.083
				[0.000, 0.106]				[0.041, 0.146]
Croatia	HRV	1.060	1.291	0.198	0.047	-0.042	0.089	0.104
				[0.145, 0.278]				[0.052, 0.185]

Country	ISO-code	PC	CoL	ln(CoL/PC)	Due to:			
					Trade costs and terms of trade	Inverse openness	Variety	Scale
Chile	CHL	1.061	1.370	0.255	0.108	-0.031	0.107	0.071
				[0.219, 0.310]				[0.035, 0.125]
Estonia	EST	1.068	1.229	0.141	0.052	-0.119	0.075	0.133
				[0.073, 0.244]				[0.066, 0.236]
Latvia	LVA	1.074	1.242	0.145	0.047	-0.106	0.083	0.122
				[0.083, 0.239]				[0.060, 0.216]
Spain	ESP	1.118	1.248	0.110	0.027	-0.011	0.048	0.046
				[0.086, 0.145]				[0.023, 0.082]
South Korea	KOR	1.127	1.283	0.130	0.015	0.019	0.050	0.045
				[0.107, 0.164]				[0.022, 0.079]
Slovenia	SVN	1.130	1.365	0.189	0.046	-0.064	0.084	0.123
				[0.126, 0.283]				[0.060, 0.217]
United Kingdom	GBR	1.153	1.140	-0.011	0.013	-0.067	0.004	0.039
				[-0.031, 0.019]				[0.019, 0.068]
Portugal	PRT	1.172	1.398	0.177	0.052	-0.029	0.072	0.083
				[0.135, 0.240]				[0.041, 0.146]
Greece	GRC	1.221	1.597	0.269	0.075	-0.011	0.123	0.082
				[0.227, 0.332]				[0.040, 0.145]
Netherlands	NLD	1.232	1.012	-0.197	0.023	-0.243	-0.048	0.071
				[-0.233, -0.141]				[0.035, 0.126]
Luxembourg	LUX	1.236	1.218	-0.015	0.045	-0.274	0.057	0.156
				[-0.094, 0.106]				[0.077, 0.277]
Germany	DEU	1.238	1.282	0.034	0.016	-0.041	0.026	0.033
				[0.018, 0.060]				[0.016, 0.058]
Malta	MLT	1.240	1.766	0.354	0.091	-0.076	0.177	0.161
				[0.272, 0.479]				[0.080, 0.286]
Cyprus	CYP	1.252	1.761	0.341	0.111	-0.052	0.138	0.144
				[0.268, 0.453]				[0.071, 0.255]
France	FRA	1.264	1.269	0.004	0.016	-0.072	0.022	0.038
				[-0.015, 0.034]				[0.019, 0.068]
Belgium	BEL	1.269	1.214	-0.044	0.043	-0.152	-0.017	0.082
				[-0.085, 0.019]				[0.040, 0.145]
Italy	ITA	1.271	1.402	0.098	0.027	-0.001	0.032	0.040
				[0.078, 0.129]				[0.020, 0.072]
Austria	AUT	1.283	1.333	0.038	0.023	-0.109	0.036	0.088
				[-0.007, 0.106]				[0.043, 0.156]

Country	ISO-code	PC	CoL	ln(CoL/PC)	Due to: Trade costs and terms of trade	Inverse openness	Variety	Scale
New Zealand	NZL	1.303	1.624	0.220	0.023	-0.008	0.101	0.104
				[0.167, 0.300]				[0.051, 0.184]
Brazil	BRA	1.317	1.508	0.135	0.030	0.044	0.050	0.011
				[0.129, 0.144]				[0.005, 0.020]
Canada	CAN	1.389	1.483	0.066	0.036	-0.062	0.037	0.054
				[0.038, 0.107]				[0.026, 0.095]
Ireland	IRL	1.392	1.636	0.162	0.048	-0.077	0.087	0.103
				[0.110, 0.241]				[0.050, 0.182]
Finland	FIN	1.393	1.672	0.182	0.054	-0.047	0.076	0.099
				[0.132, 0.259]				[0.049, 0.175]
Sweden	SWE	1.424	1.570	0.098	0.032	-0.084	0.064	0.085
				[0.054, 0.163]				[0.042, 0.151]
Japan	JPN	1.474	1.655	0.116	0.022	0.015	0.057	0.022
				[0.105, 0.132]				[0.011, 0.038]
Australia	AUS	1.475	1.684	0.133	0.067	-0.033	0.034	0.064
				[0.100, 0.182]				[0.032, 0.114]
Switzerland	CHE	1.578	1.726	0.090	0.059	-0.100	0.042	0.090
				[0.045, 0.159]				[0.044, 0.159]
Denmark	DNK	1.597	1.427	-0.113	0.036	-0.245	-0.003	0.098
				[-0.163, -0.037]				[0.048, 0.174]
Norway	NOR	1.865	2.260	0.192	0.086	-0.071	0.076	0.101
				[0.141, 0.270]				[0.050, 0.179]

Appendix Table A4. Traded consumption prices, the cost of living and a decomposition, BPP barcode counts

Country	ISO-code	PC	CoL	ln(CoL/PC)	Due to: Trade costs and terms of trade	Inverse openness	Variety	Scale
India	IND	0.466	0.515	0.100	0.000	0.047	0.086	-0.034
				[0.075, 0.126]				[-0.059, -0.008]
China	CHN	0.813	0.820	0.009	-0.007	0.041	0.012	-0.036
				[-0.018, 0.036]				[-0.063, -0.009]
Russia	RUS	0.838	0.884	0.054	-0.039	0.003	0.071	0.019
				[0.040, 0.068]				[0.005, 0.033]
Mexico	MEX	0.859	0.945	0.095	0.002	0.012	0.056	0.024
				[0.077, 0.113]				[0.006, 0.042]
Poland	POL	0.874	0.983	0.117	0.038	-0.032	0.061	0.051

				[0.079, 0.156]				[0.013, 0.089]
Turkey	TUR	0.911	1.016	0.109	-0.004	0.016	0.063	0.035
				[0.083, 0.136]				[0.009, 0.061]
South Africa	ZAF	0.942	1.092	0.148	0.045	0.008	0.051	0.043
				[0.115, 0.180]				[0.011, 0.076]
Columbia	COL	0.954	1.117	0.158	0.048	0.019	0.045	0.046
				[0.123, 0.193]				[0.011, 0.081]
United States	USA	1.000	1.000	0.000	0.000	0.000	0.000	0.000
Chile	CHL	1.061	1.305	0.207	0.108	-0.031	0.060	0.070
				[0.154, 0.259]				[0.017, 0.122]
Spain	ESP	1.118	1.218	0.086	0.027	-0.011	0.024	0.046
				[0.051, 0.120]				[0.011, 0.080]
South Korea	KOR	1.127	1.208	0.070	0.015	0.019	-0.009	0.044
				[0.036, 0.103]				[0.011, 0.078]
United Kingdom	GBR	1.153	1.159	0.005	0.013	-0.067	0.021	0.038
				[-0.024, 0.034]				[0.009, 0.067]
Greece	GRC	1.221	1.532	0.227	0.075	-0.011	0.082	0.081
				[0.166, 0.288]				[0.020, 0.142]
Netherlands	NLD	1.232	1.056	-0.155	0.023	-0.243	-0.005	0.071
				[-0.208, -0.101]				[0.017, 0.124]
Germany	DEU	1.238	1.238	-0.001	0.016	-0.041	-0.008	0.033
				[-0.025, 0.024]				[0.008, 0.057]
France	FRA	1.264	1.267	0.002	0.016	-0.072	0.021	0.038
				[-0.026, 0.031]				[0.009, 0.066]
Italy	ITA	1.271	1.442	0.126	0.027	-0.001	0.060	0.040
				[0.096, 0.156]				[0.010, 0.070]
New Zealand	NZL	1.303	1.532	0.161	0.023	-0.008	0.043	0.103
				[0.084, 0.239]				[0.025, 0.180]
Brazil	BRA	1.317	1.486	0.120	0.030	0.044	0.035	0.011
				[0.112, 0.129]				[0.003, 0.019]
Canada	CAN	1.389	1.473	0.059	0.036	-0.062	0.031	0.053
				[0.019, 0.099]				[0.013, 0.093]
Ireland	IRL	1.392	1.562	0.116	0.048	-0.077	0.043	0.102
				[0.039, 0.192]				[0.025, 0.178]
Japan	JPN	1.474	1.560	0.056	0.022	0.015	-0.003	0.021
				[0.040, 0.072]				[0.005, 0.037]
Australia	AUS	1.475	1.685	0.133	0.067	-0.033	0.036	0.063
				[0.085, 0.181]				[0.016, 0.111]

Notes: Decomposition by showing the log values of the terms in (16).

Data Collection for the Barcode Domestic Ratio:

Our main measure of the *barcode domestic share* uses data collected by freelancers in physical stores of large retailer. The first columns in Table A5 provide details of this data collection effort in 19 countries. In some countries we hired multiple freelancers to collect data from several large companies. The freelancers took photos of the product labels (see example shown in Figure A1), which we then used to monitor and validate their work. A more detailed description of the mobile-phone app used by the freelancers can be found on Cavallo (2017).

Appendix Figure A1: Example of a Crowdsourced Product Image

.

Note: Freelancers were instructed to take a photograph of the package's country-of-origin information. In this example taken inside a German electronics retailer, the product is made in China.

The crowdsourcing method makes it possible to collect country-of-origin data from many locations, but it also limited us to a relatively small sample of about 1000 products in each country (500 food products and 500 electronic product). As a robustness check, we were also able to collected web-scraped data from the websites of retailers that show the country-of-origin information for individual goods. These online estimates are only available for 9 food and 2 electronics retailers (covering 10 countries), but the product samples are much larger because they include all goods available for sale in these companies. The last columns of Table A5 show the number of domestic and imported varieties using this online scraped data. The barcode domestic ratios are very similar, with a correlation 0.76 between the benchmark offline (crowdsourced) and online estimates.

As a final robustness check, we also estimated the domestic barcode ratio for food in the US using Nielsen' Scanner data, shown in the last column of Table A5. Reassuringly, the barcode domestic ratio is 0.86 with scanner data, 0.90 with online scraped data, and 0.89 with the crowdsourced data.

Appendix References

Cavallo, Alberto. 2017. "Are Online and Offline Prices Similar? Evidence from Large Multi-Channel Retailers." *American Economic Review* 107, no. 1 : 283–303.

Inklaar, Robert and Marcel Timmer. 2014. "The Relative Price of Services." *Review of Income and Wealth*, 60(4): 727-746

Timmer, M. P., Dietzenbacher, E., Los, B., Stehrer, R. and de Vries, G. J. 2015. "An Illustrated User Guide to the World Input–Output Database: The Case of Global Automotive Production." *Review of International Economics,* 23: 575–605.

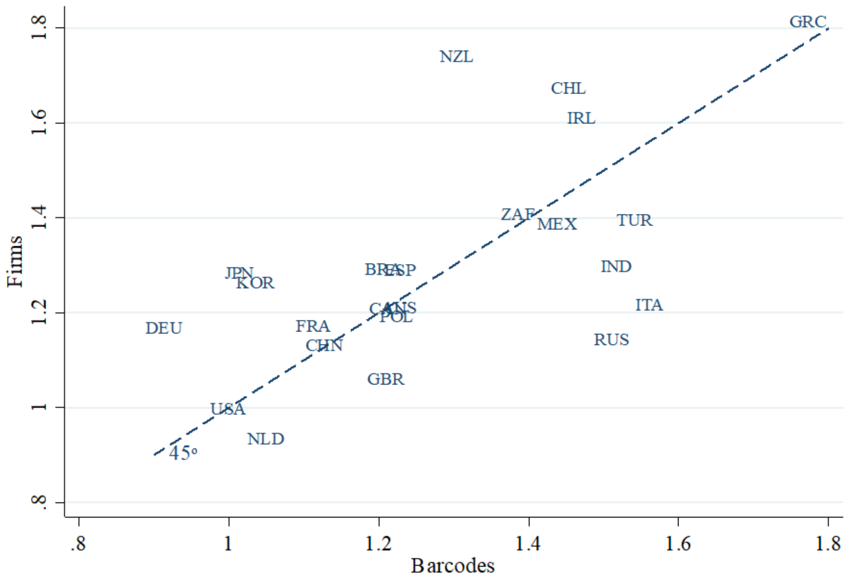
Timmer, M. P., Los, B., Stehrer, R. and de Vries, G. J. 2016. "An Anatomy of the Global Trade Slowdown based on the WIOD 2016 Release." *GGDC research memorandum* number 162, University of Groningen.

Appendix Table A5. Offline and Online Data Collection for Estimating the Share of Domestic Barcodes

		Mobile Phone Data Collection						Online Scraped Data				Scanner Data
Country	Type	Workers	Retailers	Barcodes	Domestic	Imported	Share Domestic Barcode	Barcode Online	Domestic Online	Imported Online	Share Domestic Online	Share Domestic (Nielsen)
AUS	Food	1	2	482	294	188	0.61	14,829	11,249	3,580	0.76	
BRA	Food	4	4	478	430	48	0.90	12,293	10,392	1,901	0.85	
CAN	Food	1	2	408	230	178	0.56					
CHN	Food	2	2	517	397	120	0.77	21,843	17,932	3,911	0.82	
DEU	Food	2	3	513	448	65	0.87	26,808	21,157	5,651	0.79	
ESP	Food	4	5	419	280	139	0.67					
FRA	Food	4	7	472	392	80	0.83					
GBR	Food	3	5	547	426	121	0.78					
GRC	Food	3	6	585	475	110	0.81					
IND	Food	3	4	206	204	2	0.99					
IRL	Food	3	4	423	268	155	0.63	9,219	5,926	3,293	0.64	
ITA	Food	4	3	420	374	46	0.89	7,881	6,620	1,261	0.84	
JPN	Food	1	8	508	447	61	0.88					
MEX	Food	4	7	346	276	70	0.80					
NLD	Food	1	5	501	337	164	0.67	5,593	1,891	3,702	0.34	
POL	Food	1	2	384	205	179	0.53					
RUS	Food	3	3	533	377	156	0.71	21,900	12,538	9,362	0.57	
TUR	Food	2	3	513	402	111	0.78					
USA	Food	3	3	509	454	55	0.89	23,259	21,035	2,224	0.90	0.86
AUS	Electronics	1	6	1035	15	1020	0.01	4,474	92	4,382	0.02	
BRA	Electronics	4	7	487	235	252	0.48					
CAN	Electronics	1	3	435	44	391	0.10					
CHN	Electronics	1	3	516	510	6	0.99					
DEU	Electronics	3	7	502	162	340	0.32					
ESP	Electronics	4	5	382	78	304	0.20					
FRA	Electronics	3	5	502	53	449	0.11					
GBR	Electronics	1	2	308	60	248	0.19					
GRC	Electronics	3	3	411	52	359	0.13					
IRL	Electronics	3	2	150	20	130	0.13					
JPN	Electronics	1	4	420	66	354	0.16					
MEX	Electronics	3	8	428	46	382	0.11					
NLD	Electronics	1	1	505	89	416	0.18					
POL	Electronics	3	3	467	158	309	0.34					
RUS	Electronics	2	3	391	42	349	0.11					
TUR	Electronics	5	7	445	168	277	0.38					
USA	Electronics	4	6	410	97	313	0.24	6,596	2,477	4,119	0.38	

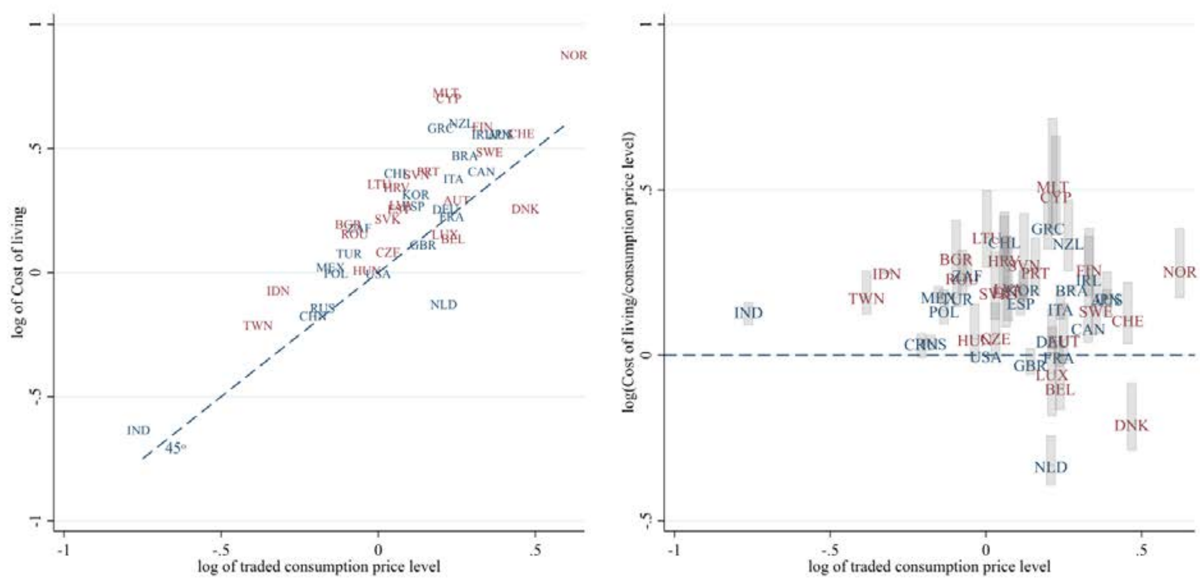
Appendix Figures A5, A7, A8, and A9 using low estimates ($(\sigma, \theta) = (4.4, 5.1)$):

Figure A5: Variety effects by country – firm count versus barcode count (USA=1)



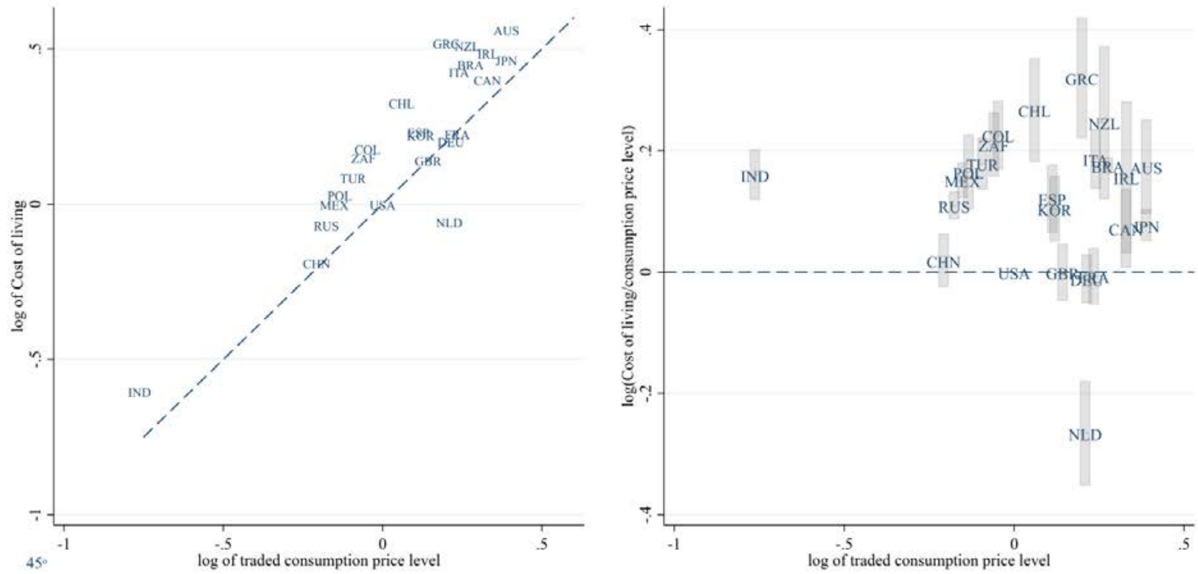
Note: Figure plots the aggregate variety effect relative to the USA

Figure A7. Cost of living versus the traded consumption price level – firm count data



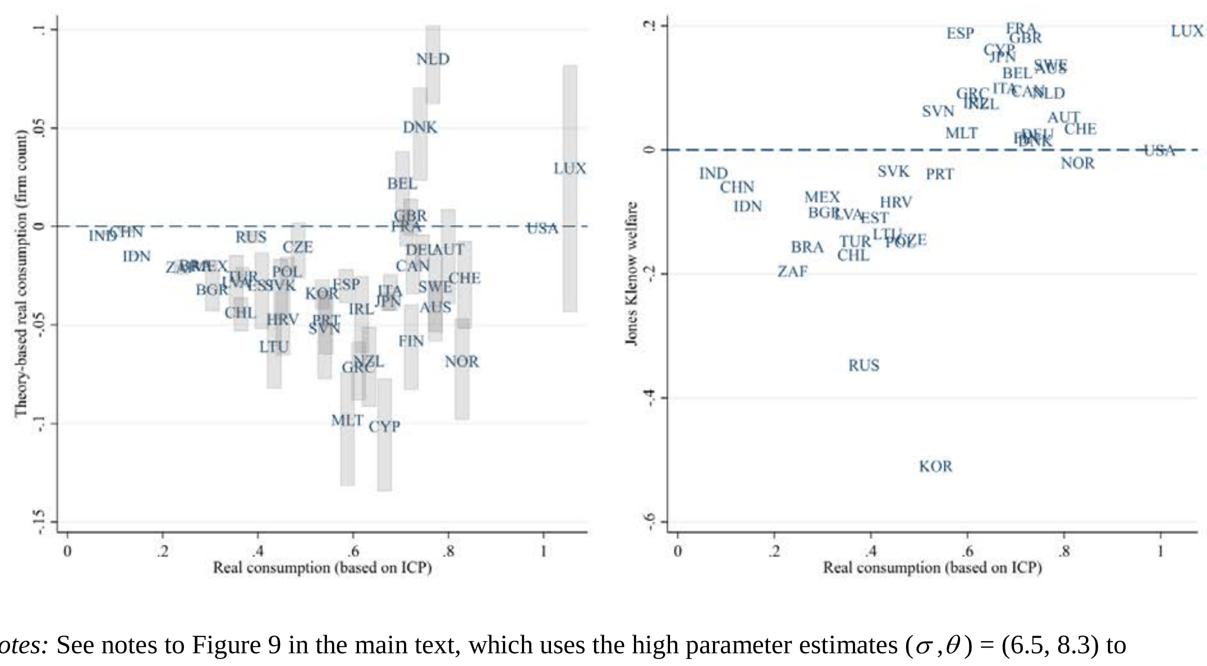
Notes: See notes to Figure 7 in the main text, which uses the high parameter estimates $(\sigma, \theta) = (6.5, 8.3)$ to obtain the cost of living. This Appendix Figure A7 uses the low estimates $(\sigma, \theta) = (4.4, 5.1)$.

Figure A8. Cost of Living versus the Traded Consumption Price Level – barcode data



Notes: See notes to Figure 8 in the main text, which uses the high parameter estimates (σ, θ) = (6.5, 8.3) to obtain the cost of living. This Appendix Figure A8 uses the low estimates (σ, θ) = (4.4, 5.1).

Figure A9. Ratio of theory-based real consumption to real consumption using ICP prices, versus real consumption – firm count data



Notes: See notes to Figure 9 in the main text, which uses the high parameter estimates (σ, θ) = (6.5, 8.3) to obtain the theory-based real consumption. This Appendix Figure A9 uses the low estimates (σ, θ) = (4.4, 5.1).