Per-Capita Incomes and the Extensive Margin of Bilateral Trade: A Quantitative Ricardian Model

PRELIMINARY AND INCOMPLETE

Christian Hepenstrick*

August 11, 2009

Abstract

This paper develops a Ricardian trade model that accounts for the empirically observed positive relation between the extensive margin of a bilateral trade flow (measured as the number of goods categories with positive volumes) and per-capita incomes of trading partners. We incorporate non-homothetic consumer behavior into the Eaton and Kortum (2002) model of international trade. The central mechanism is that richer agents find it optimal to consume a broader set of varieties, which leads to a positive relation of per-capita income of the importer and the extensive margin. The positive effect of exporter per-capita income, corresponding to the standard model, comes from the fact that technologically advanced countries are the cheapest suppliers for many varieties. In the empirical part we structurally estimate the model and argue that we capture the key features of the data. For example the importer income elasticity of the extensive margin is 0.54 in the data, whereas our model produces an elasticity of 0.56 (in contrast to -0.33 in the standard model). Having estimated the model we perform a number of counterfactual experiments with a particular emphasis on per-capita income, income distribution, and the extensive margin of trade.

*University of Zurich, Institute for Empirical Research in Economics, Mühlebachstrasse 86, CH-8008 Zurich, Switzerland, phone: +41 44 634 37 25, email: hepenstrick@iew.uzh.ch.

I thank Ana Cecília Fieler for helpful comments. I gratefully acknowledge financial support by the Verein zur Förderung des Akademischen Nachwuchses (FAN) and the Ecoscientia Foundation.
1 Introduction

The extensive margin of bilateral trade flows is positively correlated with the per-capita incomes of the trading partners. Figure 1 plots for different exporters\(^1\) the per-capita incomes of importing countries against the number of HS6 categories with positive volumes in the year 2000 (both in logs). The red line represents the best linear fit and is clearly upwards sloping indicating a positive relationship between the extensive margin and importer per-capita income. Figure 2 plots for different importers the exporter per-capita income against the number of HS6-categories with positive volumes. Again there is a clear positive relation. Indeed, as can be seen from Table 1a the elasticity of the extensive margin of a bilateral trade flow with respect to importer per-capita income is 0.54 and 0.87 for exporter per-capita income (both significant at the 1% confidence level) controlling for size and the usual trade cost proxies. The elasticities coming from the Poisson pseudo maximum likelihood estimator proposed by Silva and Tenreyro (2006) are 0.35 for importer income and 0.63 for the exporter per-capita income. Further below we discuss several papers documenting these findings in depth.

The Ricardian framework proposed by Eaton and Kortum (2002) (henceforward EK) is generally used to understand the aggregate volumes of bilateral trade flows. However, the probability that a particular country is the supplier of a variety in a given market is an endogenous model outcome and a natural theoretical counterpart to the extensive margin of a bilateral trade flow.\(^2\) In the EK model a technologically advanced country gets on average good productivity draws and is therefore an active supplier for a relatively broad set of varieties. At the same time good productivity draws imply high factor productivities and therefore high per-capita incomes. I.e. the model predicts the positive correlation of per-capita income of the exporter and the extensive margin that we observe in the data. A technologically advanced importer however will produce many varieties locally, because he gets on average good productivity draws and does not have to bear transportation costs when suppling his own market. This implies a negative correlation between importer per-capita income and the extensive margin, which is at odds with the data.

The present paper introduces non-homothetic consumer behavior into the EK model. When experiencing changes in income agents then not only adjust the quantities of each variety that they already consume, but also the number of different varieties that are consumed - i.e. the extensive margin of consumption varies with income. In the general equilibrium the negative effect described above is still present, but it is dominated by the positive income effect due to the non-homothetic preferences and we thus get a positive relation of both, importer and exporter per-capita incomes, and the extensive margin of a bilateral trade flow.

\(^1\)In our sample we have 175 countries. For expositional clarity we randomly split the sample in 4 subsamples and show only the result for the subsample including the USA. The other three subsamples exhibit the same patterns as can seen from the regression including all countries reported in Tables 1a.

\(^2\)This has been noted by Debaere and Mostashari (2005).
The model has two nice features that we use in the estimation. First, we show that due to the competitive markets the aggregate volumes of bilateral trade flows are independent of the exact form of the preferences. Second, in terms of aggregate bilateral volumes our model is observationally equivalent to the standard EK model, whereas it differs with respect to the extensive margin of bilateral trade due to the non-homothetic preferences. This suggests an intuitive two stage estimation procedure. On the first stage we estimate the supply side parameters using the volumes as moment. Due to the two properties just mentioned the resulting estimates are the same one gets from estimating the standard model. On the second stage we then estimate the preference parameter taking the extensive margin of bilateral trade as the moment and using the estimates from the first stage. This approach allows us the clearly see how much we gain by our extension. In terms of volumes the performance of our model is by construction identical to the EK model, but for the extensive margin of trade flows we significantly improve. This can be seen nicely by comparing the data generated by the models to the real world data. For example the model’s elasticity of the extensive margins with respect to the importer per-capita incomes is 0.56, whereas the corresponding elasticity in the data is 0.54 - note that we do not directly target this elasticity in the estimation. In the standard model the importer income elasticity is is strongly negative with −0.33. Similar improvements follow for the income elasticity of the multilateral import margin - i.e. the number of varieties that are not locally produced - and the average number of varieties traded between rich countries vs. the average number traded between poor countries. In other words our extension to the EK model introduces only one additional parameter - the preference parameter - and gets a very good fit for the extensive margin. At the same time the model preserves the tractability and good performance of the standard model in terms of aggregate volumes.

Having estimated the structural parameters of the model we perform two of counterfactual experiments. First, we consider the effects of global decreases in transportation costs and show that the standard model strongly underestimates the rising extensive margins of trade as it does not account for the additional set of goods that agents choose consume due to their higher income. Second, we consider the rise of China and India and the effects on the structure of international trade.

In terms of theory this paper builds on the new Ricardian framework introduced by EK and the analysis of this model provided by Alvarez and Lucas (2006). Fieler (2008) extends the EK model to 2 industries with different demand elasticities. This then leads to non-homothetic consumer behavior in the sense that poor countries relatively concentrate their expenditures in the low-elasticity industry. If this industry has a low productivity variability, most varieties are produced locally and the model explains the fact that trade volume rises overproportionally with per-capita income. The preferences in her model are such that all agents consume all varieties in positive volumes, i.e. the income effects on the extensive margin effects proposed in this paper do not emerge. Matsuyama
(2000) considers hierarchic 0/1 preferences in the Dornbusch, Fischer, and Samuelson (1977) model. With this kind of preferences all consumption decisions are on the extensive margin and indeed the intuition for many of the effects emerging in our paper can be seen from his stylized framework. This paper goes further by using preferences that allow for an intensive and an extensive margin of consumption and providing a N-country model that is flexible enough to be directly estimated. More recently there are several papers such as Sauré (2009) and Foellmi et al. (2009) investigating the role of non-homothetic preferences in Krugman’s (1980) new trade theory framework. These papers find several interesting results with respect to the extensive margin. However, they mostly remain on a very stylized level as in the monopolistic competition framework the markup become endogenous with non-homothetic preferences, which yields a variety of rich effects and makes the model usually intractable for more than 2 countries (to preserve tractability for more than two countries one could of course impose symmetry, but this will fade out many of the effects).

On the empirical side this paper is related to the literature investigating the extensive margin of trade flows. Hummels and Klenow (2005) show that the extensive margin explains about 60 percent of the higher trade volumes of larger economies. Broda and Weinstein (2006) estimate the welfare effects from more varieties becoming available due to trade and show that there are significant gains. More closely related to this paper Baldwin and Harrigan (2007) analyze the determinants of non-zero US exports in HS10 categories. Among other things they find a positive effect of gdp per worker in the importing country. Similarly Kang (2004) finds positive effects of importer per capita incomes on the extensive margin. Bernasconi (2009) finally finds positive elasticities for importer and exporter per capita incomes in the cross section as well as over time using country and time fixed effects.

The remainder of this paper is structured as follows: Section 2 presents the theoretical model and describes the equilibrium. To better understand the model we will then provide a closed form example. In section 3 we estimate the model and discuss the quantitative properties of the model. In section 4 we then present the results of the counterfactual experiments. Section 5 concludes.

2 The Model

We start this section by separately describing the modelling and the corresponding optimal behavior of the demand and supply side. In a next step we will describe the general equilibrium. In order to get more intuition we will then discuss a special case for which we can derive closed form solutions.

Consumer Behavior. Agents maximize a symmetric additively separable utility function

\[ U = \int_{j=0}^{1} v(x(j)) \, dj \]
subject to their budget constraint \( E \leq \int_{p=0}^{1} p(j) x(j) dj \) and the non-negativity constraints \( x(j) \geq 0 \forall j \). The subutility function satisfies \( v'(x) > 0, v''(x) < 0, \) and \( v'(0) < \infty \). The crucial assumption is the finite marginal utility of consuming additional varieties \( v'(0) \) implying that the non-negativity constraint is potentially binding. In this case the first order conditions for a variety \( j \) are given by

\[
\begin{align*}
v'(x(j)) &= \lambda p(j) \quad \text{for } x(j) > 0 \\
v'(0) &\leq \lambda p(j) \quad \text{for } x(j) = 0
\end{align*}
\]

(1)

where \( \lambda \) is the Lagrange multiplier. As the varieties enter the utility function symmetrically, agents simply order the varieties increasing in the prices and then choose up to which price they still want to consume a positive amount. If the price distribution is represented by a cdf \( G(p) \) we can use the cdf to reindex the varieties

\[ j = G(p) \]

such that \( p(j) < p(i) \) for \( j < i \). From the first order conditions (1) we know the price \( p(M) \) at which the non-negativity constraint just becomes binding, \( p(M) = v'(0) / \lambda \). \( M \) is the index of this marginal variety and defines the extensive margin of consumption, i.e. the fraction of totally available varieties that is consumed in positive quantities. The price distribution can be used to solve for \( M \)

\[ M = G \left( \frac{v'(0)}{\lambda} \right) . \]

For the varieties \( j < M \) the first order condition defines the Marshallian demand \( x(\lambda p) = v'^{-1}(\lambda p) \). Combining above deliberations with the budget constraint and making the change of variable \( p = G^{-1}(j) \) yields

\[ E = \int_{p=0}^{v'(0)/\lambda} px(\lambda p) g(p) dp. \]

implicitly defining the marginal utility of income \( \lambda \). Knowing \( \lambda \) we can solve for the extensive margin of consumption \( M \) and the optimal quantities of the varieties \( j < M \).

**Parametrization of \( v(x) \)** In order to be able to estimate the model lateron we need to make some functional assumptions on \( v(x) \). The model is particularly tractable for the subclass of the HARA preferences \( v'(x) = (Ax - B)^D \) that satisfy \( v'(0) < \infty \). For \( B < 0, A = -1, \) and \( D = 1 \) we get quadratic preferences\(^3\), for which the model is easy to solve and estimate. The disadvantage

\(^3\)In the trade context there are two classes of models to mention that also use some form of quadratic preferences. The first class are the general oligopolistic equilibrium models (GOLE) models proposed by Neary (see for example Neary (2009) and Eckel and Neary (2009)) that feature quadratic preferences whose linear demand function make the oligopoly problem particularly tractable. However, typically the authors rule out binding non-negativity constraints as they focus on the oligopolistic interaction rather than the role of per-capita income and this then impedes the effects this paper is about. The second class of models (see for example Melitz and Ottaviano (2008)) uses quadratic preferences nested in a quasi-linear utility function as proposed Ottaviano and Thisse (1999). This framework is usually used to assess the trade related effects on markups. However, the presence of a numéraire good rules out income effects. Moreover binding non-negativity constraints are usually ruled out by assumption.
is that saturation emerges as $v'(B) = 0$, which introduces an unnecessary complication in the estimation. Therefore we use Stone-Geary preferences, $A = 1$, $B > 0$, and $D < 0$, where we do not have to worry about saturation. Note that the empirical performance of the model is very similar for quadratic and Stone-Geary preferences. Another advantage of the Stone-Geary form is that for $B = 0$ the preferences become the standard CES preferences.

The general form of Stone-Geary preferences is $v(x) = \left( (\bar{x} + x(j))^{1-\sigma} - 1 \right) / (1 - \sigma)$ with $\sigma > 0$ and $\bar{x} > 0$. However, we show in the appendix that the additional explanatory power due to Stone-Geary’s additional degree of freedom is extremely small and we therefore choose $\sigma = 1$, which will yield particularly simple expressions. Thus the subutility function is

$$v(x(j)) = \log(\bar{x} + x(j)).$$

as for this case the expressions become particularly simple. The extensive margin of consumption is then given by

$$M = G \left( \frac{1}{\bar{x} \lambda} \right)$$

and the Lagrange multiplier is defined by

$$E = \frac{M}{\lambda} - \bar{x} \int_{p=0}^{1/(\bar{x} \lambda)} pg(p) dp.$$

As a next step we look at the supply side where the country specific price distributions $G_n(p)$ will be determined.

**Production.** The production technology exhibits constant returns to scale and uses only labor\(^4\) as an input. Labor is perfectly mobile within countries, but immobile across countries, i.e. every country $i$ has a specific wage rate $w_i$. The production technology in country $i$ for variety $j$ is shifted with a country-variety specific productivity $z_i(j)$. Thus the unit costs are $w_i/z_i(j)$. Assuming perfect competition and Samuelsonian iceberg transportation costs\(^5\) ($d_{ni}$ is the amount of goods that need to be shipped in $i$ in order for one unit to arrive in country $n$) the price at which country $i$ offers variety $j$ in country $n$ is

$$p_{ni}(j) = \frac{w_i d_{ni}}{z_i(j)}. \quad (2)$$

**Technological heterogeneity.** Technological heterogeneity is modelled as in EK. We assume that country $i$’s productivity for variety $j$, $z_i(j)$, is the realization of a Fréchet distributed random

---

\(^4\)This assumption can readily be relaxed and we could allow for multiple input factors. Moreover, some (up to all) inputs can be internationally mobile. However, for the sake of expositional simplicity we consider only one input factor.

\(^5\)As standard in the literature we assume that the triangle inequality $d_{ni} \leq d_{il}d_{nl}$ holds for all countries $i, l, n$. In different economic applications in particular finance this kind of preferences is widely used. Foellmi and Zweimüller (2006) use the preferences also due to the binding non-negativity constraint.
variable

\[ \Pr[Z_i(j) \leq z] = \exp\{-T_i z^{-\theta}\}, \]

where \( T_i \) is country specific and governs absolute advantage. \( \theta \) is common to all countries and regulates the global variability in efficiency draws and therewith the gains from trade realized by comparative advantages.

We normalize the product space to one and consider a country \( n \) that consumes a measure \( M_n \leq 1 \) of all available varieties. \( M_n \) will be endogenized in the next section. Given the assumptions made up to now, a number of well known EK results directly follows (the appendix outlines the formal derivation of these results together with a quick review of an innovation process delivering Fréchet distributed productivities).

i. The price at which country \( i \) offers a variety \( j \) in country \( n \) is a Fréchet distributed random variable

\[
\Pr[P_{ni}(j) \leq p] = 1 - \exp\left\{-\left(T_i (w_i d_{ni})^{-\theta}\right) p^\theta\right\}.
\]

ii. The probability that country \( i \) is the cheapest supplier of variety \( j \) in country \( n \) is given by

\[
\pi_{ni} = \frac{T_i (w_i d_{ni})^{-\theta}}{\Phi_n}.
\]

\( \pi_{ni} \) is also the measure of varieties for which country \( i \) is the cheapest supplier in country \( n \).

iii. The minimal price on offer for variety \( j \) in country \( n \) itself is again a Fréchet distributed random variable

\[
\Pr\left[\min_{i=1}^{N} P_{ni}(j) \leq p \right] = G_n(p) = 1 - \exp\{-\Phi_n p^\theta\},
\]

where \( \Phi_n = \sum_{i=1}^{N} T_i (w_i d_{ni})^{-\theta} \). \( G_n(p) \) also represents the distribution of the lowest prices across goods in country \( n \).

iv. The measure of varieties produced in \( i \) and actually consumed in \( n \) is given by

\[
m_{ni} = M_n \pi_{ni}.
\]

We shall call \( m_{ni} \) the extensive margin of the trade flow from country \( i \) to country \( n \).

2.1 The Equilibrium

In the general equilibrium firms price according to (2) and consumers behave as implied by (1). The equilibrium then pins down a set of country specific wage rates and extensive margins of
consumption \( \{w_n, M_n\}^N_{n=1} \). We will show that the equilibrium exists and that it is unique. For now we consider homogenous agents each endowed with one unit of labor, i.e. the equilibrium income of agents in country \( n \) is given by \( w_n \).

The Volume of Bilateral Trade Flows An agent in country \( n \) observes the price distribution \( G_n(p) \) and allocates his expenditures according to (1) over the different varieties. Optimality and the assumption of non-saturation implies that he spends his total income \( w_n \). Because the productivity draws are iid across different varieties the varieties that origin from a given country \( i \) are evenly distributed over all varieties consumed. But this then implies that the share of expenditures going to country \( i \) is just equal to \( \pi_{ni} \). We show this formally in the Appendix. Denoting country \( n \)'s population size with \( L_n \) then implies that the aggregate volume of the trade flow from \( i \) to \( n \) is

\[ X_{ni} = \pi_{ni}w_nL_n, \tag{3} \]

i.e. our model delivers exactly the same prediction with respect to volume as does the standard EK model. In particular the result is independent of the preference parameters as long as the draws are iid across varieties.

Separation of Demand and Supply Side An important feature of our model is the fact that due to the competitive environment the equilibrium wage rates are independent of the demand structure. To see this, consider country \( i \)'s balance of payments

\[ \sum_{n \neq i} X_{ni} = \sum_{j \neq i} X_{ij}. \]

By adding \( X_i^i \) on both sides the right hand side becomes country \( i \)'s aggregate expenditure, which in equilibrium must be equal to total labor income \( w_iL_i \). Using (3) on the right hand side we then can write

\[ w_iL_i = \sum_{n=1}^N T_i \left( \frac{w_i d_{ni}}{\sum_{j=1}^N T_j (w_j d_{nj})} \right)^{\theta} w_n L_n. \tag{4} \]

In the Appendix we show that for given states of technologies \( \{T_i\}_{i=1}^N \), population sizes \( \{L_i\}_{i=1}^N \), transportation costs \( \{d_{ni}\}_{n=1}^N \), and comparative advantage parameter \( \theta \) there is a unique set of equilibrium wage rates that satisfies simultaneously all countries' (4). In other words the model exhibits a certain separation of the demand and supply side that is a very convenient feature of the model when it comes to estimating the model. Note that for example monopolistic competition with non-homothetic preferences would not exhibit this separation, as the optimal markup varies with the demand structure, i.e. we have feedback effects from income on the production decision and via the implied labor productivity back on income.
The Extensive Margin of Bilateral Trade Flows. However, when it comes to the extensive margin, the demand side matters of course. In the appendix we show that country $n$’s extensive margin of consumption is implicitly defined by

$$ w_n = \bar{x} (\Phi_n)^{-\frac{1}{\theta}} \left( M_n \left( -\log (1 - M_n) \right)^\frac{1}{\theta} - \gamma \left( \frac{1}{\theta} + 1; -\log (1 - M_n) \right) \right) $$

(5)

where $\gamma (z, \tilde{\theta}) = \int_0^\tilde{\theta} t^{z-1}e^{-t}dt$ represents the incomplete Gamma function. I.e. the equilibrium measure of varieties consumed in country $n$ is a function of the per capita income $w_n$, the preference parameter $\bar{x}$, and the local price distribution $G_n (\cdot)$, which in turn is driven by the global distribution of states of technologies $\{T_i\}_{i=1}^N$, the transportation costs matrix $\{d_{ni}\}_{n,i}$, the variability of the productivity draws $\theta$, and the equilibrium wage rates $\{w_i\}_{i=1}^N$. Note that $M_n$ is unique as (5) follows from maximizing a concave objective function over a convex constraint.

Using property iv. from above we get the equilibrium extensive margin of the trade flow from $i$ to $n$ as

$$ m_{ni} = \pi_{ni} M_n. $$

Here, we already see how the extensive margin of the trade flow from $i$ to $n$ depends on the technologies and therefore the per-capita incomes in these two countries. A better technology in the exporting country $i$ implies a higher $\pi_{ni}$ and a slightly higher $M_n$ as the price distribution in country $n$ shifts towards the origin. Better technology and therefore higher per-capita income in the importing country $n$ implies a higher extensive margin of consumption $M_n$ and at the same time a lower $\pi_{ni}$ as the measure of varieties that are produced cheapest in country $n$ increases. In the estimated model it turns out that the latter effect is dominated by the former, i.e. we get the positive correlation between per-capita income and extensive margin that is in the data.

### 2.2 A Closed Form Example

The novel feature of our model is the extensive margin of consumption. In what follows we set $\theta = 1$, i.e. the productivities are now exponentially distributed, which will allow us to do some comparative statics to sharpen our intuition of the model in particular with respect to the extensive margin. In the appendix we show that for $\theta = 1$ the budget restriction (5) can be written as

$$ \frac{w_n \Phi_n}{\bar{x}} = M_n - \log (1 - M_n). $$

(6)

and we see directly that the equilibrium extensive margin of consumption rises with $w_n \Phi_n / \bar{x}$.

Now for sake of simplicity we consider $N$ symmetric countries, i.e. the indices are redundant and we can normalize the income to one. $\Phi$ is then given by $\Phi = T / (1 + (N - 1) d)$. Higher states of technology $T$ or lower transportation $d$ lead to a higher $\Phi$ and thus a higher extensive margin.
of consumption. The intuition is that both changes imply that labor becomes more productive and therefore real income rises. This in turn makes the agents adjust their consumption on both the intensive and the extensive margin. In the general equilibrium trade the extensive margin of bilateral trade flows then is not only rising due to lower transportation costs or better technology respectively, but also because the set on varieties that are actually consumed increases. A higher $\bar{x}$ implies a lower equilibrium extensive margin. The reason for this is that the higher $\bar{x}$ the lower the utility from consuming additional varieties.

Finally, we consider the effect of differences in per-capita income and restrict our attention to two countries, $N = 2$. From (6) follows that the country with the higher $\Phi_n w_n$ exhibits the higher equilibrium extensive margin of consumption. Using the balance of payments we show in the appendix that $T_i > T_n$ implies $\Phi_i w_i > \Phi_n w_n$. This confirms the intuition that if we have two countries that are geographically equally remote the country with the higher state of technology has the higher per-capita income and therefore the higher extensive margin of consumption.

3 Empirical Analysis

In this section we estimate our model and quantitatively assess its performance in explaining the extensive margin of the bilateral trade flows. The theoretical model makes predictions about two key moments of bilateral trade flows - the extensive margin and the volume. For readability we repeat the respective equations. Volumes are governed by

$$X_{ni} = \pi_{ni} w_n L_n,$$

and the extensive margin is determined by

$$m_{ni} = M_n \pi_{ni},$$

where the following equations hold

$$\pi_{ni} = \frac{T_i (w_i d_{ni})^{-\theta}}{\Phi_n},$$

$$w_n = \bar{x} (\Phi_n)^{-\frac{1}{\theta}} \left( M_n (- \log (1 - M_n))^{\frac{1}{\theta}} - \gamma \left( \frac{1}{\theta} + 1; \log (1 - M_n) \right) \right),$$

$$\Phi_n = \sum_{i=1}^{N} T_i (w_i d_{ni})^{-\theta}.$$
3.1 Estimation Procedure

A natural and intuitive approach to estimating this model is to apply a two stage procedure. In the first stage we use (7) to get estimates for the technology parameters \( T_i \) and the transportation costs \( d_{ni} \) (we will see that we are not able to separately identify \( \theta \) and therefore take it from EK). This then allows us to calculate the implied \( \Phi_n \). We then can use them in the second stage to estimate the preference parameter \( \bar{x} \) by targeting the extensive margins of bilateral trade flows. In the following we describe the two stages in detail. In the results section we will briefly discuss the results from simultaneously estimating all parameters.

**First Stage.** On the first stage we want to estimate the model parameters that govern the bilateral trade volumes. Our model behaves similar to EK in terms of volumes, i.e. the results of the first stage are at the same time also estimates for the standard model. We use the estimation procedure proposed by Fieler (2009). As seen above the unique set of equilibrium wage rates is defined by

\[
w_i L_i = \sum_{n=1}^{N} \frac{T_i (w_i d_{ni})^{-\theta}}{\sum_{j=1}^{N} T_j (w_j d_{nj})^{-\theta}} w_n L_n
\]

holding for all countries \( i = 1, \ldots, N \). But this also implies that knowing the wage rates \( \{w_i\}_{i=1}^{N} \), the population sizes \( \{L_i\}_{i=1}^{N} \), \( \theta \) and the trade costs matrix \( \{d_{ni}\}_{n=1}^{N} \) would allow us to solve for the unique set of absolute advantage parameters \( \{T_i\}_{i=1}^{N} \) that are consistent with these wage rates. In the data we directly observe the population sizes. Moreover, following Fieler (2009) we take per-capita income as a proxy for wages and model the transportation costs as a function of the usual proxies

\[
d_{ni} = 1 + \left( \gamma_0 + \gamma_1 \delta_{ni} + \gamma_2 (\delta_{ni})^2 \right) S_{ni} B_{ni} F_{ni}
\]

where \( S_{ni} \) is an indicator taking the value of \( \gamma_L \) if \( n \) and \( i \) share a common language and one otherwise. \( B_{ni} \) and \( F_{ni} \) are similar indicators for common border and membership in the same free-trade agreement.\(^6\) In principle we would also like to estimate \( \theta \), however as \( \theta \) always occurs as the power to \( d_n^k \) we cannot separately identify \( \theta \) and the level of the transportation cost parameters. Therefore we choose to fix \( \theta \) at EK’s preferred level \( \theta = 8.28 \).\(^7\) So, in the first stage we seek to estimate the parameter vector \( \Upsilon = \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \gamma_L & \gamma_B & \gamma_F \end{bmatrix} \).

\(^6\)Due to the quadratic term it is possible that for high \( \delta_{ni} \) and negative \( \gamma_2 \) the transportation costs are falling in distance. We avoid this by modelling \( \delta_{ni} \) in the following way

\[
\delta_{ni} = \begin{cases} 
\text{dist}_{ni} & \text{dist}_{ni} < -\frac{\gamma_1}{2 \gamma_2} \\
-\frac{\gamma_1}{2 \gamma_2} & \text{dist}_{ni} > -\frac{\gamma_1}{2 \gamma_2}
\end{cases}
\]

if \( \gamma_2 < 0 \).

\(^7\)We will report the estimates for different \( \theta \) and it indeed turns out that whereas the estimates for the other parameters change the explanatory power of our model is not affected by the different levels of \( \theta \).
We define our point estimate for the parameter vector $\hat{\Theta}$ as

$$\hat{\Theta} = \arg \min_{\Theta} L_{\text{vol}} \left( \{ X_{ni}, X_{ni} (\Theta) \}_{i,n \neq i} \right),$$

where $X_{ni}$ are the aggregate bilateral trade flows in 2000 and $X_{ni} (\Theta)$ are the flows predicted by the model. $L (\cdot)$ is a quadratic loss function

$$L_{\text{vol}} \left( \{ X_{ni}, X_{ni} (\Theta) \}_{i,n \neq i} \right) = \left( \sum_{i=1}^{N} \sum_{n \neq i} \omega_{ni} (X_{ni} - X_{ni} (\Theta))^2 \right) \Bigg/ \left( \sum_{i=1}^{N} \sum_{n \neq i} \omega_{ni} (X_{ni})^2 \right)$$

with $\omega_{ni}$ a weight to be specified. Choosing $\omega_{ni} = 1$ would imply giving big countries particular weight as big countries trade a low and therefore we have potentially big errors for these countries. To avoid this we normalize the errors $(X_{ni} - X_{ni} (\Theta))$ with the economic mass of the trading partners and choose $\omega_{ni} = (w_i L_i + w_n L_n)^{-2}$. Note that we normalized the loss function by the uncentered second moment such that the loss function tells us how much of the total variation in the data cannot be explained with the model.

The algorithm for the numerical implementation of our estimation procedure is the following:

1. Start with an initial guess $\Theta^0$
2. Solve for the implied $\{T_i\}_{i=1}^{N}$ and use them to compute the corresponding $\{\pi_{ni}\}_{n=1}^{N}$
3. We then plug them into (7) to get the implied trade flows $\{X_{ni} (\Theta)\}_{n=1}^{N}$
4. Evaluate the loss function $L \left( \{ X_{ni}, X_{ni} (\Theta^0) \}_{i,n \neq i} \right)$
5. Adjust the guess using the complex method of numerical optimization and repeat steps 2-5 until a local minimum is found
6. Repeat steps 1-5 for different initial guesses to ensure finding a global minimum

**Second Stage.** What remains to be estimated in the second stage is the preference parameter $\hat{\xi}$. We estimate it by targeting the extensive margins of bilateral trade flows. In the first stage finding the empirical counterpart of the model moment was straightforward and intuitive - we simply took aggregate trade volume as implied by the model. However, defining the empirical counterpart to the moment of the second stage - the extensive margin - is less obvious. We will describe our measure of the extensive margin in detail in the data subsection. However, given that we have a measure for the extensive margin of a bilateral trade flow, we still need to scale this measure to make it comparable with the model. The reason is that in our model the extensive margin of bilateral trade flows is measured relative to the unit interval of potential varieties. However, there can be a set of

---

8The corresponding FORTRAN routine is available upon request from the author.
varieties that is not actually produced \((1 - \max \{M_n\})\). In the data do not observe this extensive margin relative to potential varieties, but only relative to produced and traded varieties. In order to make the data’s and the model’s extensive margins comparable we choose to normalize the bilateral trade flows with the multilateral import margin of the US \(m_{US}\), i.e. the measure of varieties that is imported in the US.

We define our point estimate for \(\tilde{x}\) as

\[
\tilde{x} = \arg \min_{\bar{x}} L^{em} \left( \left\{ \frac{m_{ni}}{m_{US}} \frac{m_{ni}}{m_{US}} \frac{m_{ni}}{m_{US}} \frac{m_{ni}}{m_{US}} \right\}_{i, n \neq i} \right),
\]

whereas \(L(\cdot)\) is again a quadratic loss function

\[
L^{em} \left( \left\{ m_{ni}, m_{ni} \frac{m_{ni}}{m_{US}} \frac{m_{ni}}{m_{US}} \frac{m_{ni}}{m_{US}} \frac{m_{ni}}{m_{US}} \right\}_{i, n \neq i} \right) = \left( \sum_{i=1}^{N} \sum_{j=1}^{N} \omega_{ni} \left( \frac{m_{ni}}{m_{US}} \frac{m_{ni}}{m_{US}} \frac{m_{ni}}{m_{US}} \frac{m_{ni}}{m_{US}} \right) \right) \left( \sum_{i=1}^{N} \sum_{j=1}^{N} \omega_{ni} \left( \frac{m_{ni}}{m_{US}} \frac{m_{ni}}{m_{US}} \frac{m_{ni}}{m_{US}} \frac{m_{ni}}{m_{US}} \right) \right)
\]

and \(\hat{\Phi}\) is the vector of parameter estimates from the first stage. \(\omega_{ni}\) are country pair specific weights. In the baseline we choose \(\omega_{ni} = 1\) for all \(k\) and \(n\), i.e. we give all country pairs equal weight. Again the loss function is normalized with the uncentered second moment of the data such that its value represents the fraction of total variation that is explained by our model.

For the implementation of this estimator we use the estimated \(\hat{\Phi}\) from the first stage to construct the implied \(\left\{ \hat{\Phi}_n \right\}_{n=1}^{N}\). Plugging them into (9) we can solve for every country’s extensive margin of consumption \(\left\{ M_n \right\}_{n=1}^{N}\) for a given preference parameters \(\bar{x}\) and \(\sigma\). Using these we can then compute the implied bilateral trade flows via (8) and get thus the value of the loss function for this particular \(\bar{x}\). We iterate over \(\bar{x}\) to find the point estimate \(\hat{x}\) that minimize the loss function.

**Simultaneous Estimation** The advantage of the two step estimator is that it is very intuitive as the first stage corresponds to the standard model and the second stage therefore gives us directly a feeling of how much we gain with our extension by considering the difference in the values of \(L^{em}\) in the standard model and our new model. However, in terms of precision a simultaneous approach might be better as the two step estimator places only very small weight on information contained in the variation in the extensive margin. For the simultaneous estimator we minimize a weighted average of the two loss functions

\[
L = \phi L^{vol} + (1 - \phi) L^{em},
\]

where \(\phi \in [0, 1]\) is a weight governing the relative importance of volumes vs. extensive margin. For \(\phi \to 1\) the simultaneous estimator converges to the two step estimator. The reason for that is
that the bilateral volumes depend only on $\Psi$. Therefore for $\phi \to 1$ the trade cost parameters $\Psi$ are chosen to minimize $L^{\text{vol}}$ and only then $\bar{x}$ is chosen to minimize $L^{en}$.

**The Data** We estimate the model with data from the year 2000. For the trade flows ($X_{ni}$) we use the COMTRADE data as provided by CEPII (Gaulier and Zignago (2008)). Per-capita income ($w_i$) and population sizes ($L_i$) are taken from the World Bank’s (2008) world development indicators (WDI). The per-capita incomes are in US-dollars at nominal exchange rates as deviations from PPP emerge endogenously noted in EK. The data used as distance proxies come all from CEPII except for the free trade agreements, which are taken from Fieeler (2008).

Measuring the extensive margin of bilateral trade is less obvious. A simple measure is to count the number of categories for which we observe positive volumes.

$$m_{ni} = \sum_{j=1}^{J} I [X_{ni} (j) > 0],$$

where $j$ represents the HS6 categories, $I [X_{ni} (j) > 0]$ is an indicator function taking the value of one if country $n$ has positive imports in category $j$ from country $i$. One might worry that this approach neglects the fact that the goods categories are designed for customs use and highly regulated industries with many exceptions feature therefore much more categories than other industries although the number of varieties might be the same in both industries. In order to control for this we should choose to weight the categories with their relative importance in world trade and define the extensive margin of bilateral trade as $m_{ni} = \sum_{j=1}^{J} I [X_{ni} (j) > 0] X (j) / X$, where $X (j) / X$ is the share of category $j$ in global trade.\footnote{Feenstra (1994) proposes a related measure for the extensive margin. However, his measure is directly derived from a homothetic utility function. Therefore his particular functional form does apply in our case and we choose the more ad-hoc measure proposed in the text.} In the context of the EK model this measure is correct if the productivity draws are iid across HS6 categories such that within each category we have the same price distribution. The simple counting measure on the other hand would be correct if the productivity draws are grouped together into the HS6 categories. The reality will lie somewhere between these two polar cases. However as the coefficient of correlation between these two polar measures for the extensive margin of trade is 0.95 we are confident that this measurement problem is not too serious. In fact the results presented below are preserved when using the weighted measure.

In our sample we have 175 countries for which all data are available. Further details on the data can be found in the appendix.
3.2 The Results

First Stage. We report the results of the first stage estimation in Table 2. The value of the loss function is 0.57 across all $\theta$, confirming the fact that $\theta$ cannot be estimated using volumes as the moment. The reason is that the transportation costs are always to the power of $\theta$ and because we estimate the transportation costs themself we cannot separate the level of the transportation costs parameter $\Upsilon$ from the level of $\theta$.

Using the first stage results we construct estimates for $\Phi_n$ that will be used in the second stage. Remember that the extensive margin of bilateral trade is given by $m_{ni} = \pi_{ni} M_n$. Because in the standard model the extensive margin of consumption is always one, we can use our first stage estimates to construct the standard model’s extensive margin of bilateral trade $m_{ni}|_{EK} = \pi_{ni} = T_i (w_i d_{ni})^{-\theta} / \Phi_n$. By incorporating non-homothetic demand our model endogenizes $M_n$. By comparing our model’s $m_{ni}$ to $m_{ni}|_{EK}$ we can then assess how much we gain by introducing the additional preference parameter $\bar{x}$.

Second Stage. The results of the second stage are reported in Table 3. Using the extensive margins implied by the standard model the value of the loss function is 2.54, i.e. the model performs much worse than a model prediction no trade at all. Introducing non-homothetic preferences and therewith one additional parameter $\bar{x}$ brings the value of the loss function down to 0.78, which is a very significant improvement. In other words our new model can explain with only one parameter 22% of the total variation in the extensive margin of bilateral trade flows.\(^\text{10}\)

Another way to assess the performance of our model is to compare moments of the data to the ones generated by our model. For that purpose we run the same regression cited in the introduction using the model data. The resulting coefficients are reported in Table 4b for $\theta = 8.28$. Table 4a compares the coefficients of per-capita income generated by our model to the coefficient in the data and the ones coming from the standard model. For the exporter per-capita gdp both models perform similar and generate elasticities of 0.67 compared to 0.87 in the data. For the importer per-capita income however, the EK model produces a strong negative elasticity ($-0.33$), whereas our model generates a positive elasticity of 0.55 that is very close to the data’s 0.54. Remember that we did not target these coefficients when estimating the model. For different $\theta$ the estimated exporter income elasticities barely change. Also the importer income elasticities of the EK model remain constant at $-0.33$, whereas the importer elasticities of our model change somewhat being closest to the data for $\theta$ between 6 and 9.

To further understand the mechanics of the model consider Figure 3 plotting the estimated extensive margin of consumption against per-capita income. The relation is increasing and slightly concave. But the relation is not perfect, because for given per-capita income countries have different

\(^{10}\)If we compare the performance of our model to simply taking teh average trade extensive margin and plugging it into the loss function it turns out that our model performs 13% better.
extensive margins of consumption due to being more or less remote. We consider further moments of the data and the model in the following section.

Simultaneous Estimations Instead of the two stage procedure proposed above we now simultaneously estimate the model parameters. Table 6 reports the parameter estimates associated with the weights ($\phi$) 0.0, 0.25, 0.5, 0.75, and 1.0.\footnote{For the weight 1.0 we actually chose $w = 0.99$ in order to get an estimate for the preference parameter.} We see that there is a trade of between matching the volumes and matching the extensive margin of bilateral trade. The more weight we place on the extensive margin, the lower $\gamma_0$, $\gamma_1$, and $\gamma_2$. The effects of the border, language, and free trade agreement dummies also vanish, whereas $\bar{x}$ becomes lower such that for $w = 0.25$ and $w = 0.0$ 10 and 27 countries consume all available varieties (for $w = 0.5$ and higher we do not have any country consuming all varieties).

Table 7 reports the elasticities of the extensive margin of bilateral trade with respect to importer and exporter income. The importer income elasticity remains very close to the empirical elasticity for $w \geq 0.5$, whereas the exporter income elasticity gets closer to the elasticity observed in the data for smaller $w$.

Another interesting moment is the ratio of the average extensive margin of trade among the richest $x\%$ of countries vs. the average among the poorest $x\%$. Table 8 reports the respective results. Whereas the EK model predicts much too small ratios, our model's ratios are somewhat too high, but get closer to the data for lower $w$.

A third interesting moment is the multilateral import margin, i.e. the number of HS6 categories for which we observe a positive import flow from at least one country. The elasticity of this moment with respect to importer per-capita gdp is 0.14 in the data. However, as can be seen from Figure 4 there seems to be top-censoring as many countries do reach the maximum of 5108 varieties. Therefore we think that the magnitude of the elasticity is less informative. However, it is clear that there is a positive relation between multilateral import margin and per-capita income. As can be seen from Table 9 our model generates this positive relation, but the elasticities are too high. Note however, that the standard model predicts a negative elasticity of $-0.18$. The reason for this negative sign is that the multilateral extensive margin of imports in the standard model is $(1 - \pi_{nn})$, whereas $\pi_{nn}$ is rising country $n$'s technology $T_n$ and thus also with per-capita income. In our model the multilateral extensive margin is $(1 - \pi_{nn}) M_n$ - the effect from the standard model is still present, but dominated by the positive correlation of per-capita income and the extensive margin of consumption $M_n$.

Finally, it is interesting to note that the relation between the estimated extensive margin of consumption and per-capita income becomes more concave for less weight on $L^{vol}$ as can be seen from Figure 5. The reason for this is the $\bar{x}$-estimate that is falling with $w$, i.e. the lower $w$ the higher the estimated marginal utility in the origin and therefore the stronger the reaction of the
extensive margin of consumption on changes in income.

4 Counterfactual Experiments

In this section we perform a number of counterfactual experiments. The objective of these experiments is not so much making quantitative predictions, but to further understand the workings of our model. In each case we start with the economy as estimated in the previous section, i.e. we take the estimations for the transportation cost parameters $\Upsilon$ and the preference parameters $\bar{x}$ and $\sigma$ together with the states of technology $\{T_i\}_{i=1}^N$ that lead to the equilibrium wage rates that correspond to the ones we observe in reality and then modify the parameters of interest. We use the results from the simultaneous estimation with $\theta = 8.28$.

4.1 Transportation Costs

A first potential interesting counterfactual experiment are global reductions in transportation costs. In the standard model the extensive margin of bilateral trade reacts due to supply side effects - lower transportation costs make trade worthwhile for more varieties, i.e. the extensive margin of trade will expand. In our model we have an additional demand side effect coming from the fact that lower transportation costs increase labor productivity and therewith wages. Because the agents have higher per-capita incomes they will expand their extensive margin of consumption which in turn affects the extensive margin of trade positively. To get a feeling of the quantitative size of this additional effect we take the estimated model and uniformly decrease transportation costs by 10, 20, 50 percent. For the 50%-reduction all countries’ transportation costs hit the lower bound of 1, i.e. we are in a perfectly integrated world. Thus we will call this case freetrade. Besides the reductions in transportation costs we also compare the status quo with autarky. Under autarky a country cannot benefit from other countries’ technologies and $\Phi_n$ reduces to $\Phi_n = T_n (w_n)^{-\theta}$. Plugging this into (9) defines us country $n$’s extensive margin of consumption under autarky $M_n|^{\text{aut}}$:

$$1 = \bar{x} (T_n)^{\frac{1}{\theta}} \left( M_n|^{\text{aut}} \left( -\log \left( 1 - M_n|^{\text{aut}} \right) \right)^{\frac{1}{\theta}} - \gamma \left( \frac{1}{\theta} + 1; -\log \left( 1 - M_n|^{\text{aut}} \right) \right) \right).$$

In the status quo world trade’s share of total world gdp is 20.5% (data: 18.6%). Lowering transportation costs increases the share to 33.2%, 49.0%, and 88.5% for the 10%- , 20%- reduction, and freetrade respectively. Figures 6 and 7 plot the absolute and the relative changes in the extensive margin of consumption due to a 20% trade costs reduction against the initial per-capita income. Whereas in relative terms the countries with the lowest initial extensive margin of consumption gain most, the relation is hump shaped for the absolute gains, as the richest countries initially already consuming most varieties. It is interesting to see that there is a considerable heterogeneity.
in the relation between change and initial per-capita income. This comes from differences in the remoteness of the countries and the associated differences in the effect of trade cost reductions on $\Phi_n$. For the 10%-reduction and free trade we observe similar qualitatively similar effects. These effects together with the effect of going from the status quo to autarky are summarized in Table 10 (relative changes) and Table 11 (absolute changes). Note that we observe the lowest changes always for rich countries such as Japan, USA, Norway, and Switzerland. In particular their extensive margin of consumption barely reacts, i.e. the only adjustments in their consumption plans occurs on the intensive margins. The countries with the strongest reactions are mostly African countries for the relative changes and small islands for the absolute changes. The median countries are generally Asian or Latin-American middle income economies.

We now turn to the reactions of the extensive margin of bilateral trade flows. Table 12 summarizes the relative changes. For the largest changes our model and the EK model make similar predictions as these are changes between rich countries that already consume almost all varieties. However, for all other country pairs the predictions of our model and the EK model largely differ. The EK model predicts consistently lower changes as it neglects the adjustments on the extensive margin of consumption.

To summarize, this counterfactual experiment shows clearly that if one considers a set of countries with large differences in technologies and therefore per-capita incomes, one needs to account for the adjustments on the extensive margin of consumption, when thinking about the reaction of the number of traded varieties to changes in transportation costs.

### 4.2 The Rise of China and India

A second interesting counterfactual experiment is the analysis of how global trade patterns, in particular the extensive margin of trade, react to a continued rise of China and India. One of the most important global trends of the past decades is the emergence of these two countries and the associated increase in per-capita income in China and India. Therefore we consider technology improvements in China and India that imply similar rises in per-capita incomes relative to the rest of the world as we observed over the last 15 years, i.e. 250% for China and 170% for India and use our estimated model to assess the global effects on the extensive margin of consumption and patterns of trade. During this period both countries’ extensive margins of imports and exports increased very substantially (on the bilateral and on the multilateral level).

Technological advances in China and India have the direct effect that these countries are the cheapest suppliers for a broader set of varieties and therefore have broader extensive margins in their exports. Our model predicts that the average extensive margin of exports rises by 73% for China and 33% for India. The numbers predicted by the standard model are very similar, albeit slightly smaller, which is due to the fact that in the standard model the other countries do not increase
their extensive margin of consumption. In our model the rest of the world slightly increases the extensive margin of consumption as the technological improvement in China and India lowers the price distribution in all countries. On average the other countries consume a 0.1% broader set of varieties. For poor countries close to India and China such as Nepal and Buthan we observe the strongest, albeit still small, reactions of around 0.5%. Whereas the extensive margins of rich countries such as the USA and Japan does not react at all. China and India themself on the other hand increase their extensive margin of consumption strongly by 93% for China and 63% for India. This in turn affects the extensive margins of their import flows, where we estimate average increases of 34% for China and 28% for India. This starkly contrasts with the prediction of the standard model of decreases of −31% and −22% for China and India respectively, which is clearly at odds with the development of China’s and India’s trade flows during the last 15 years. The reason for these contrasting predictions is that in the standard model China and India produce now more varieties locally due to the technological improvement, whereas in the new model this effect is still present, but is dominated by the non-homothetic consumer behavior and the associated increase in the extensive margin of consumption. For the trade flows between China and India our model finally predicts increases of 224% for the flow from China to India and 181% for the flow from India to China. These predicted increases are about twice as high as what is predicted by the standard model (137% and 93% respectively). In both models the trade flows between the rest of the world fall only slightly by around 1% (note that in our model they fall less due to the increase in the extensive margin of consumption in all countries).

All in all this counterfactual experiment demonstrates that the rise of China and India strongly affects the trade flows involving these two countries, but much less the other trade flows, i.e. the multilateral effects of technology shocks are relatively small in our model.

5 Conclusions

The growing interest in the extensive margin of trade flows has led several economist to observe that the extensive margin is positively related to both importer and exporter per-capita income. Whereas the positive effect of the exporter per-capita income is usually directly explained by productivity, the positive effect of the importer income has been linked to some form of non-homothetic consumer behavior. However, this intuition has usually been more a conjecture than implemented in a rigorous model (exceptions are the stylized models by Sauré (2009), Matsuyama (2000), and Foellmi et al (2009)). This paper extends the well-known Ricardian trade model proposed by Eaton and Kortum (2002) to incorporate non-homothetic preferences with potentially binding non-negativity constraints. The resulting new model is to our best knowledge to first Ricardian model featuring an extensive consumption margin that is flexible enough to be directly estimated using bilateral
trade data for 175 countries.

The empirical results show that the new model reproduces quite well the key features of the extensive margin of bilateral trade flows, whereas for the volume the model preserves the good fit already present in EK. Using the estimates of the deep model parameters we performed a number of interesting counterfactual results that helped to further understand the workings of the model.

We think that this paper is a significant contribution towards understanding the role per-capita income play for the extensive margin of bilateral trade flows. However, there are a number of questions that emerge, which suggest further research. Particularly interesting seems to be the question how the global distribution of per-capita incomes affect the incentives to innovate and the location of innovators. In order to discuss this question one needed a dynamic model with a market structure à la Eaton, Kortum, and Kramarz (2008) where firms can make positive profits and therefore can be incentivized to innovate.
References


6   Appendix

6.1 Some Properties due to Fréchet Distributed Productivity Draws

In the following we outline the proofs for the properties i. - iv. stated in the main text.

Property i. $\Pr[P_{ni}(j) \leq p] = 1 - \exp \left\{ -p^\theta T_i (w_id_{ni})^{-\theta} \right\}$

Proof. The price on offer in country $n$ for a variety $j$ produced in $i$ is a transformation of the random variable $Z_i(j)$, $P_{ni}(j) = w_id_{ni}/Z_i(j)$. Therefore we can write $\Pr[P_{ni}(j) \leq p] = \Pr[w_id_{ni}/Z_i(j) \leq p] = \Pr[w_id_{ni}/p \leq Z_i(j)] = 1 - \Pr[w_id_{ni}/p > Z_i(j)]$. But the last probability is just the probability that the productivity draw in $i$ for variety $j$ is below some $z = w_id_{ni}/p$.

In the main text we assumed that this probability is Fréchet distributed with the parameters $T_i$ and $\theta$. Combining completes the proof. For later use we define $G_{ni}(p) \equiv \Pr[P_{ni}(j) \leq p]$.

Property ii. $\Pr[P_{ni}(j) \leq \min \{P_{ns}; s \neq k\}] = T_i (w_id_{ni})^{-\theta}/\left( \sum_{i=1}^{N} T_i (w_id_{ni})^{-\theta} \right)$.

Proof. The probability $\pi_{ni}(j)$ that country $i$ supplies country $n$ with variety $j$ is the probability that the realization of the random variable $Z_i(j)$ leads to a price $P_{ni}(j)$ that is lower than the realizations of all other countries. To start note that the probability that all countries $s \neq i$ have a price higher than $p$ is

$$\Pr\left[ p \leq \min \{P_{ns}(j)\}_{s \neq i} \right] = \prod_{s \neq i} \Pr[P_{ns}(j) > p].$$

From i. we know the individual factors of the product. Inserting them yields

$$\Pr\left[ p \leq \min \{P_{ns}(j)\}_{s \neq i} \right] = \exp \left\{ -p^\theta \sum_{s \neq i} T_s (w_sd_{ns})^{-\theta} \right\}.$$  

Heuristically, the probability $\pi_{ni}(j)$ is now the weighted sum over all prices $p$, where as weights we use the probability that country $i$ supplies $j$ at a price below $p$. As we have a continuous price space the appropriate expression is the integral over the prices where we use the pdf of the prices offered by $i$ in $n$ for variety $j$ as weight

$$\pi_{ni}(j) = \int_{p=0}^{\infty} \Pr\left[ p \leq \min \{P_{ns}(j)\}_{s \neq k} \right] g_{ni}(p) \, dp.$$
By taking the first derivative of the cdf from \( i \) we get 
\[
    g_{ni}(p) = \theta p^{\theta-1} T_i(w_i d_{ni})^{-\theta} \exp \left\{ -p^\theta T_i(w_i d_{ni})^{-\theta} \right\}.
\]
Inserting this we can write
\[
    \pi_{ni}(j) = T_i(w_i d_{ni})^{-\theta} \int_0^\infty \theta p^{\theta-1} \exp \left\{ -p^\theta \sum_{s=1}^N T_s(w_s d_{ni})^{-\theta} \right\} dp.
\]
Solving the integral then completes the first part of the proof. As the draws are iid and we have a continuum of varieties this is also the measure of varieties produced in \( i \) and sold in \( n \). □

**Property iii.** 
\[
    \Pr \left[ \min_{i=1}^N \{ P_{ni}(j) \} < p \right] = 1 - \exp \left\{ -p^\theta \sum_{i=1}^N T_i(w_i d_{ni})^{-\theta} \right\}
\]

**Proof.** This cdf represents the probability the lowest price is below \( p \) or equally one minus the probability that all prices on offer are higher than \( p \). As the prices are iid distributed this is just the product of the individual probabilities
\[
    \Pr \left[ \min_{i=1}^N \{ P_{ni}(j) \} < p \right] = 1 - \prod_{i=1}^N \Pr [P_{ni}(j) > p].
\]
Substituting in the expression for the individual probabilities derived above completes the first part of the proof. As the prices are iid across varieties this is also the distribution across varieties. □

**Property iv.** 
\[
    m_{ni} = M_n \pi_{ni}
\]

**Proof.** We assume that \( M_n \leq 1 \) and we order the varieties such that the varieties with \( j > M_n \) are not demanded in positive quantities. Note that the event that country \( i \) is the supplier of variety \( j \) in country \( n \) is the realization of a Bernoulli random variable with \( f(1) = \pi_{ni} \) and \( f(0) = 1 - \pi_{ni} \) for \( j < M_n \) and \( f(0) = 1 \) for \( j > M_n \) as these varieties are not consumed at all. The Bernoulli variable is iid across the different varieties \( j < M_n \). Therefore the measure \( m_{ni} \) of varieties supplied from \( i \) to \( n \) is the expectation over the Bernoulli random variable given that \( j < M_n \), which is the expression stated in the main text. □

### 6.2 An innovation process generating our Fréchet distribution

In progress. Follows EK book

### 6.3 Derivation of the expenditure share going to a given country

Let \( H_{ni}(p) \) denote the distribution of prices of country \( i \) producers in country \( n \) that are at the same time the lowest on offer in country \( n \). The expenditures of a country \( n \) agent going to country \( k \), \( E_{ni} \), then is
\[
    E_{ni} = \int_{p=0}^{p(M_n)} p \left( \frac{1}{\lambda p - \bar{x}} \right) h_{ni}(p) dp.
\]
Now note that $H_{ni}(p)$ is given by

$$H_{ni}(p) = \int_{q=0}^{p} \prod_{s \neq i} (1 - G_{ns}(q)) \frac{\partial G_{ni}(q)}{\partial q} dq.$$ 

Inserting for $G_{ni}(p)$ and its derivative we find that $H_{ni}(p) = \pi_{ni} G_{ni}(p)$ and thus $h_{ni}(p)/\pi_{ni} = g_{n}(p)$. Now note that total expenditures of a country $n$ agent in equilibrium are

$$E_n = \int_{p=0}^{p(M_n)} p \left( \frac{1}{\lambda p - \bar{x}} \right) g_{n}(p) dp = \frac{1}{\pi_{ni}} \int_{p=0}^{p(M_n)} p \left( \frac{1}{\lambda p - \bar{x}} \right) h_{ni}(p) dp.$$ 

But this then directly implies that the share of a country $n$ agent’s expenditures going to country $i$ is given by $E_{ni}/E_{n} = \pi_{ni}$.

### 6.4 Unique solution to (4)

**Claim 1** There is a unique set of wage rates for which (4) holds for all countries $i = 1, ..., N$.

**Proof.** This proof is inspired by Alvarez and Lucas (2006). First we write (4) as an excess demand

$$Z_i \left( \{w_i\}_{i=1}^{N} \right) = \frac{1}{w_i} \sum_{n=1}^{N} \frac{T_i \left( w_i d_{ni} \right)^{-\theta}}{\sum_{j=1}^{N} T_j \left( w_j d_{nj} \right)^{-\theta}} w_n L_n - L_i,$$

where $L_i$ represents the supply of labor from country $i$ and the first term on the left hand side is the global demand for labor from country $i$. For notational ease we define $w = (w_1, ..., w_N)$

We first show that this excess demand satisfies the five properties from Proposition 17.B.2 of Mas-Colell, Whinston, and Green (1995).

i. $Z(w)$ is continuos.
   Trivial.

ii. $Z(w)$ is homogeneous of degree zero.
   Trivial.

iii. $\sum_{n=1}^{N} w_n Z_n(w) = 0$ for all $w$.
   Follows if all agent’s budget constraints hold with equality, which is the case given our non-saturation assumption.

iv. There is an $s > 0$ such that $Z_n(w) > -s$ for all $n$ and $w$.
   As the first part of the excess demand is always non-negative we have $s = \max \{ L_i \}_{i=1}^{N}$.

v. If $(w)^n \to w$, where $p \neq 0$ and $p_i = 0$ for some $i$, then $\max \{ Z_1 ((w)^n), ..., Z_N ((w)^n) \} \to \infty$.
   Follows from noting that $Z_i(w) \to \infty$ for $w_k \to 0$.

Second, note that because these five properties hold, the existence of a solution to (4) is established with Proposition 17.C.1 from Mas-Colell, Whinston, and Green (1995).
Finally, note that

$$\frac{\partial Z_i(w)}{\partial w_j} = \frac{1}{w_i} \left( \sum_{n=1}^{N} \frac{T_i(w_{i,d_{ni}})^{-\theta}}{\sum_{j=1}^{N} T_j(w_{j,d_{nj}})^{-\theta}} L_n + \theta \sum_{n=1}^{N} \frac{T_i(w_{i,d_{ni}})^{-\theta}}{\left( \sum_{j=1}^{N} T_j(w_{j,d_{nj}})^{-\theta} \right)^{2}} \right) w_n L_n > 0 \forall j \neq i,$$

i.e. the gross substitution property holds. But using Proposition 17.F.3 from Mas-Colell, Whinston, and Green (1995) this implies together with the existence proof from above that there is a unique solution $\{w_i\}_{i=1}^{N}$ that ensures that (4) holds for all countries $i = 1, \ldots, N$.  

6.5 Derivation of equation (5)

We take a country $n$ agent’s budget restriction and replace the Lagrange multiplier using

$$\lambda_n = \frac{1}{x p(M_n)}$$

as implied by optimal consumption (1) to get

$$E = \bar{x} \left( p(M_n) M_n - \int_{p=0}^{p(M_n)} p g(p) dp \right).$$

If we then insert the pdf of the Fréchet distribution we have

$$w_n = \bar{x} \left( M_n \left( -\log (1 - M_n) \right)^{\frac{1}{\theta}} - \int_{p=0}^{p(M_n)} p \Phi_n \theta p^{\theta-1} e^{-\Phi_n p^\theta} dp \right).$$

Making the change of variable $t = \Phi_n p^\theta$ (therefore $dt = \Phi_n \theta p^{\theta-1} dp$, $p = (t/\Phi_n)^{1/\theta}$, $\bar{x} = \Phi_n p(M)^\theta$) we can write

$$w_n = \bar{x} \left( \Phi_n \right)^{-\frac{1}{\theta}} \left( M_n \left( -\log (1 - M_n) \right)^{\frac{1}{\theta}} - \int_{p=0}^{p(M_n)} t^{1/\theta} e^{-t} dt \right).$$

Note that the integrals are incomplete Gamma functions, i.e.

$$w_n = \bar{x} \left( \Phi_n \right)^{-\frac{1}{\theta}} \left( M_n \left( -\log (1 - M_n) \right)^{\frac{1}{\theta}} - \gamma \left( \frac{1}{\theta} + 1; \Phi_n p(M)^\theta \right) \right).$$

Finally using the Fréchet distribution we can substitute for $\Phi_n p(M)^\theta = -\log (1 - M_n)$ to get equation (5).
6.6 Derivation of equation (6)

For \( \theta = 1 \) and by substituting in the integrals of the incomplete Gamma distributions we can write the equilibrium budget restriction (5) as

\[
\frac{w \Phi_n}{\bar{e}} = \left( M_n (-\log (1 - M_n)) - \int_{t=0}^{\log(1-M_n)^{-t}} te^{-t} dt \right)
\]

Integration by parts of the second integral yields

\[
\int_{j=0}^{\log(1-M_n)^{-t}} te^{-t} dt = (1 - M_n) \log (1 - M_n) + M_n.
\]

Inserting this into the budget restriction and solving the first integral then allows us to derive equation (6).

6.7 \( T_i > T_n \) implies \( \Phi_i w_i > \Phi_n w_n \)

We choose \( w_n \) an the numéraire and write \( w_i = w \) and thus we want to show that \( \Phi_i w > \Phi_n w \). Using the expression for the \( \Phi \) it follows that \( \Phi_i w > \Phi_n w \) if

\[
\frac{T_i}{T_n} > w.
\]

Now consider the balance of payments \( \pi_n = \pi_{in} \), which can be written as

\[
\frac{T_n}{T_n} \frac{1}{w} + \frac{1}{d} \frac{1}{\bar{d}} + \frac{T_n}{T_i} w = w.
\]

We now prove our claim by contradiction. First, we guess that \( w < 1 \). Thus the balance of payments would imply

\[
\frac{T_n}{T_n} \frac{1}{w} + \frac{1}{d} \frac{1}{\bar{d}} + \frac{T_n}{T_i} w < 1,
\]

which can be simplified to \( T_k / T_n < w < 1 \), which is a contradiction. Next we guess \( w = 1 \). For this guess the balance of payments becomes \( T_i / T_n = 1 \), which again is a contradiction. Therefore it must be that \( w > 1 \). Now note that for \( w = T_i / T_n \) the left hand side of the balance of payments becomes 1, which is a contradiction as \( w > 1 \). Finally we guess that \( w > T_i / T_n \) and therefore \( 1/w < T_n / T_i \).
implying for the balance of payments

\[
\frac{T_i}{T_n} \frac{1}{w} + \frac{1}{d} > 1
\]

which is another contradiction. Therefore the only remaining possibility is \( 1 < w < T_i/T_n \), which then implies that indeed \( \Phi_i w_i > \Phi_n w_n \).

6.8 A model with Intermediates

In progress.

6.9 Data Description

Bilateral Trade Flows  In progress.

Per-capita Incomes and Population Sizes  In progress.

Transportation Cost Proxies  In progress.

Quintiles  The main data source is the UNU WIDER (2008) world income inequality database. This database provides quintiles for 119 countries. For 112 countries we have the corresponding trade flows from the COMTRADE data. One problem is that the income concept varies widely across different countries. In particular quintiles may refer to income or expenditures. Following Dollar and Kray (2002) running the following dummy regression

\[
q_i = a_0 + a_1 D_{regio}^i + \alpha_2 D_{inc}^i + \varepsilon_i
\]

for all quintiles. \( D_{regio}^i \) are regional dummies (Eastern Asia and Pacific, Eastern Europe and Central Asia, Mid East and North Africa, Latin America and the Caribbean, South Asia, Sub Sahara, Base Category: Developed World) and \( D_{inc}^i \) a dummy taking the value one if the quintile refers to income. It turns out that \( \alpha_2 \) is not significant, which is why we decide not to adjust the quintiles.
6.10 Estimating the Model with Stone-Geary Utility

Instead of the one-parameter utility function used in the main text, one could use the more general Stone-Geary form
\[ u(x(j)) = \frac{(\tilde{x} + x(j))^{1-\sigma} - 1}{1 - \sigma}, \]
where \(\tilde{x}\) still governs the marginal utility in the origin and \(\sigma > 0\) the curvature of the utility function and therewith the substitutability of the different varieties. Note that for \(\sigma = 1\) we get the utility used in the main text. Combining the first order conditions of consumption, the budget constraint, and the country specific price distribution we get the following equation governing the extensive margin of consumption
\[ w_n = \tilde{x}(\Phi_n)^{-\frac{1}{\sigma}} \left( \frac{\gamma (\frac{\sigma-1}{\sigma}) 1 + \log (1 - M_n))}{\log (1 - M_n)} \right) - \gamma \left( \frac{1}{\theta} + 1; -\log (1 - M_n) \right). \]

As in the two stage estimation the first stage is independent of the particular preference specification, we can use the first stage estimates from the main text. On the second stage we iterate over \(\tilde{x}\) and \(\sigma\) in order to find the point estimator defined by
\[ \begin{bmatrix} \tilde{x} \\ \sigma \end{bmatrix} = \arg \min_{\tilde{x},\sigma} L \left( \left\{ \frac{m_{ni}}{m_{US}}, \frac{m_{ni}(\tilde{x},\sigma)}{m_{US}(\tilde{x},\sigma)} \right\}_{i,n} \right), \]
where \(m_{ni}(\tilde{x},\sigma) = \hat{\pi}_{ni} M(\tilde{x},\sigma)\) and \(\hat{\pi}_{ni}\) represents the first stage estimate for \(\pi_{ni}\). However, it turns out that this minimization problem seems to exhibit a continuum of \((\tilde{x},\sigma)\)-pairs as solutions (the values of the loss function vary only in the range of rounding errors), i.e. we cannot separately identify \(\tilde{x}\) and \(\sigma\). This can be seen from Figure 8 plotting the value of the loss function against different \((\tilde{x},\sigma)\)-pairs. In Figure 9 we show the grid together with the minima.
Table 1: Dependent Variable - bilateral extensive margin

**OLS**

<table>
<thead>
<tr>
<th></th>
<th>log-log</th>
<th>level-log</th>
<th>PPML</th>
</tr>
</thead>
<tbody>
<tr>
<td>i_pc_gdp</td>
<td>0.540***</td>
<td>0.016***</td>
<td>0.350***</td>
</tr>
<tr>
<td>e_pc_gdp</td>
<td>0.866***</td>
<td>0.023***</td>
<td>0.632***</td>
</tr>
<tr>
<td>i_pop</td>
<td>0.387***</td>
<td>0.010***</td>
<td>0.276***</td>
</tr>
<tr>
<td>e_pop</td>
<td>0.723***</td>
<td>0.018***</td>
<td>0.547***</td>
</tr>
<tr>
<td>distance</td>
<td>−0.881***</td>
<td>−0.016***</td>
<td>−0.448***</td>
</tr>
<tr>
<td>border</td>
<td>0.667***</td>
<td>0.112***</td>
<td>0.331***</td>
</tr>
<tr>
<td>language</td>
<td>0.756***</td>
<td>0.021***</td>
<td>0.392***</td>
</tr>
<tr>
<td>FTA</td>
<td>0.057</td>
<td>0.125***</td>
<td>−0.374***</td>
</tr>
<tr>
<td>constant</td>
<td>−11.831***</td>
<td>−0.148***</td>
<td>−10.816***</td>
</tr>
<tr>
<td>N</td>
<td>18815</td>
<td>30450</td>
<td>30450</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.338</td>
<td>0.278</td>
<td>0.465</td>
</tr>
</tbody>
</table>

*** denotes significance at the 1%-level. Standard errors are robust.
Table 2: First stage results

<table>
<thead>
<tr>
<th></th>
<th>$\theta = 8.28$</th>
<th>$\theta = 3$</th>
<th>$\theta = 6$</th>
<th>$\theta = 9$</th>
<th>$\theta = 12$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loss</td>
<td>0.57</td>
<td>0.57</td>
<td>0.57</td>
<td>0.57</td>
<td>0.57</td>
</tr>
<tr>
<td>$\gamma_0$</td>
<td>0.54</td>
<td>2.39</td>
<td>0.83</td>
<td>0.49</td>
<td>0.35</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>0.52</td>
<td>3.12</td>
<td>0.84</td>
<td>0.45</td>
<td>0.30</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>-0.12</td>
<td>-0.59</td>
<td>-0.19</td>
<td>-0.11</td>
<td>-0.07</td>
</tr>
<tr>
<td>$\gamma_B$</td>
<td>0.84</td>
<td>0.79</td>
<td>0.82</td>
<td>0.84</td>
<td>0.85</td>
</tr>
<tr>
<td>$\gamma_L$</td>
<td>0.83</td>
<td>0.75</td>
<td>0.81</td>
<td>0.83</td>
<td>0.84</td>
</tr>
<tr>
<td>$\gamma_F$</td>
<td>0.66</td>
<td>0.55</td>
<td>0.63</td>
<td>0.66</td>
<td>0.68</td>
</tr>
</tbody>
</table>

Table 3: Second stage results

<table>
<thead>
<tr>
<th></th>
<th>$\theta = 8.28$</th>
<th>$\theta = 3$</th>
<th>$\theta = 6$</th>
<th>$\theta = 9$</th>
<th>$\theta = 12$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loss</td>
<td>0.78</td>
<td>0.77</td>
<td>0.78</td>
<td>0.79</td>
<td>0.79</td>
</tr>
<tr>
<td>Loss-EK</td>
<td>2.54</td>
<td>2.49</td>
<td>2.52</td>
<td>2.55</td>
<td>2.55</td>
</tr>
<tr>
<td>$\bar{x}$</td>
<td>7.09</td>
<td>1.80</td>
<td>4.66</td>
<td>8.11</td>
<td>11.33</td>
</tr>
</tbody>
</table>
Table 4a: Elasticities in data and model ($\theta = 8.28$)

<table>
<thead>
<tr>
<th></th>
<th>Data log-log</th>
<th>PPML</th>
<th>Model log-log</th>
<th>PPML</th>
<th>EK log-log</th>
<th>PPML</th>
</tr>
</thead>
<tbody>
<tr>
<td>Importer per capita gdp</td>
<td>0.54</td>
<td>0.36</td>
<td>0.55</td>
<td>0.56</td>
<td>-0.33</td>
<td>-0.27</td>
</tr>
<tr>
<td>Exporter per capita gdp</td>
<td>0.87</td>
<td>0.63</td>
<td>0.67</td>
<td>0.56</td>
<td>0.67</td>
<td>0.56</td>
</tr>
</tbody>
</table>

Controlling for population sizes, distance, common border, common language, and membership in the same free trade agreement

Table 4b: Dependent Variable: bilateral extensive margin (model data)

<table>
<thead>
<tr>
<th></th>
<th>OLS: log-log</th>
<th>PPML</th>
</tr>
</thead>
<tbody>
<tr>
<td>i_pc_gdp</td>
<td>0.55</td>
<td>0.56</td>
</tr>
<tr>
<td>e_pc_gdp</td>
<td>0.67</td>
<td>0.56</td>
</tr>
<tr>
<td>i_pop</td>
<td>-0.27</td>
<td>-0.28</td>
</tr>
<tr>
<td>e_pop</td>
<td>0.71</td>
<td>0.56</td>
</tr>
<tr>
<td>distance</td>
<td>-0.21</td>
<td>-0.30</td>
</tr>
<tr>
<td>border</td>
<td>1.09</td>
<td>0.83</td>
</tr>
<tr>
<td>language</td>
<td>0.61</td>
<td>0.59</td>
</tr>
<tr>
<td>FTA</td>
<td>1.19</td>
<td>1.08</td>
</tr>
<tr>
<td>constant</td>
<td>-12.81</td>
<td>-10.38</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>30450</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>30450</td>
<td>0.23</td>
</tr>
</tbody>
</table>
Table 5: Elasticities for different $\theta$

<table>
<thead>
<tr>
<th></th>
<th>Model</th>
<th>EK</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Importer per capita gdp</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta = 3$</td>
<td>0.45</td>
<td>-0.33</td>
</tr>
<tr>
<td>$\theta = 6$</td>
<td>0.53</td>
<td>-0.33</td>
</tr>
<tr>
<td>$\theta = 9$</td>
<td>0.56</td>
<td>-0.33</td>
</tr>
<tr>
<td>$\theta = 12$</td>
<td>0.58</td>
<td>-0.33</td>
</tr>
<tr>
<td><strong>Exporter per capita gdp</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta = 3$</td>
<td>0.67</td>
<td>0.67</td>
</tr>
<tr>
<td>$\theta = 6$</td>
<td>0.67</td>
<td>0.67</td>
</tr>
<tr>
<td>$\theta = 9$</td>
<td>0.67</td>
<td>0.67</td>
</tr>
<tr>
<td>$\theta = 12$</td>
<td>0.67</td>
<td>0.67</td>
</tr>
</tbody>
</table>

Controlling for population size, distance, common border, common language, and membership in the same free trade agreement

Table 6: Parameter estimates for simultaneous estimation with different weights ($\theta = 8.28$)

<table>
<thead>
<tr>
<th></th>
<th>$w = 0.0$</th>
<th>$w = 0.25$</th>
<th>$w = 0.5$</th>
<th>$w = 0.75$</th>
<th>$w = 1.0$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Loss</strong></td>
<td>0.59</td>
<td>0.65</td>
<td>0.65</td>
<td>0.62</td>
<td>0.57</td>
</tr>
<tr>
<td>$L_{vol}$</td>
<td>0.97</td>
<td>0.74</td>
<td>0.62</td>
<td>0.58</td>
<td>0.57</td>
</tr>
<tr>
<td>$L_{em}$</td>
<td>0.59</td>
<td>0.62</td>
<td>0.68</td>
<td>0.74</td>
<td>0.78</td>
</tr>
<tr>
<td>$\gamma_0$</td>
<td>0.21</td>
<td>0.37</td>
<td>0.47</td>
<td>0.51</td>
<td>0.54</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>0.13</td>
<td>0.15</td>
<td>0.26</td>
<td>0.41</td>
<td>0.52</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>-0.01</td>
<td>-0.02</td>
<td>-0.04</td>
<td>-0.08</td>
<td>-0.12</td>
</tr>
<tr>
<td>$\gamma_B$</td>
<td>0.99</td>
<td>0.89</td>
<td>0.89</td>
<td>0.86</td>
<td>0.84</td>
</tr>
<tr>
<td>$\gamma_L$</td>
<td>0.99</td>
<td>0.97</td>
<td>0.91</td>
<td>0.86</td>
<td>0.83</td>
</tr>
<tr>
<td>$\gamma_F$</td>
<td>0.94</td>
<td>0.85</td>
<td>0.77</td>
<td>0.69</td>
<td>0.66</td>
</tr>
<tr>
<td>$\bar{x}$</td>
<td>0.59</td>
<td>1.02</td>
<td>2.59</td>
<td>4.48</td>
<td>7.09</td>
</tr>
</tbody>
</table>
Table 7: Income elasticities of the extensive margin of trade for different weights ($\theta = 8.28$)

<table>
<thead>
<tr>
<th></th>
<th>$w = 0.0$</th>
<th>$w = 0.25$</th>
<th>$w = 0.5$</th>
<th>$w = 0.75$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OLS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Importer per-capita gdp</td>
<td>0.35</td>
<td>0.43</td>
<td>0.53</td>
<td>0.54</td>
</tr>
<tr>
<td>Exporter per-capita gdp</td>
<td>0.85</td>
<td>0.80</td>
<td>0.74</td>
<td>0.70</td>
</tr>
<tr>
<td><strong>PPML</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Importer per-capita gdp</td>
<td>0.69</td>
<td>0.64</td>
<td>0.60</td>
<td>0.54</td>
</tr>
<tr>
<td>Exporter per-capita gdp</td>
<td>0.27</td>
<td>0.36</td>
<td>0.50</td>
<td>0.58</td>
</tr>
</tbody>
</table>

Table 8: Average extensive margin among the richest $x\%$ relative to average among poorest $x\%$

<table>
<thead>
<tr>
<th>$x%$</th>
<th>data</th>
<th>$w = 0.0$</th>
<th>$w = 0.25$</th>
<th>$w = 0.5$</th>
<th>$w = 0.75$</th>
<th>$w = 1.0$</th>
<th>EK</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>413</td>
<td>523</td>
<td>558</td>
<td>661</td>
<td>609</td>
<td>558</td>
<td>7</td>
</tr>
<tr>
<td>20</td>
<td>132</td>
<td>145</td>
<td>184</td>
<td>255</td>
<td>263</td>
<td>260</td>
<td>5</td>
</tr>
<tr>
<td>30</td>
<td>35</td>
<td>53</td>
<td>72</td>
<td>100</td>
<td>106</td>
<td>107</td>
<td>3</td>
</tr>
<tr>
<td>40</td>
<td>19</td>
<td>25</td>
<td>33</td>
<td>46</td>
<td>50</td>
<td>51</td>
<td>3</td>
</tr>
<tr>
<td>50</td>
<td>13</td>
<td>15</td>
<td>19</td>
<td>26</td>
<td>27</td>
<td>29</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 9: Importer income elasticity of the multilateral extensive margin

<table>
<thead>
<tr>
<th></th>
<th>data</th>
<th>$w = 0.0$</th>
<th>$w = 0.25$</th>
<th>$w = 0.5$</th>
<th>$w = 0.75$</th>
<th>$w = 1.0$</th>
<th>EK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Importer per capita gdp</td>
<td>0.14</td>
<td>0.44</td>
<td>0.53</td>
<td>0.65</td>
<td>0.68</td>
<td>0.70</td>
<td>-0.18</td>
</tr>
</tbody>
</table>
Table 10: Summary of the relative reaction of the extensive margin of consumption to changes in trade costs

<table>
<thead>
<tr>
<th>10%-reduction</th>
<th>20%-reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative change</td>
<td>relevant countries</td>
</tr>
<tr>
<td>max</td>
<td>14.9</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>min</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>med</td>
<td>10.2</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>freetrade</th>
<th>autarky</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative change</td>
<td>relevant countries</td>
</tr>
<tr>
<td>max</td>
<td>112.5</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>min</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>med</td>
<td>82.8</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 11: Summary of the relative reaction of the extensive margin of consumption to changes in trade costs

<table>
<thead>
<tr>
<th></th>
<th>10%-reduction</th>
<th></th>
<th>20%-reduction</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absolute change</td>
<td>relevant countries</td>
<td>Absolute change</td>
<td>relevant countries</td>
</tr>
<tr>
<td>max</td>
<td>0.06</td>
<td>Saint Kitts</td>
<td>0.12</td>
<td>New Caledonia</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Malta</td>
<td></td>
<td>Barbados</td>
</tr>
<tr>
<td></td>
<td></td>
<td>New Caledonia</td>
<td></td>
<td>Malta</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Antigua and Barbuda</td>
<td></td>
<td>Antigua and Barbuda</td>
</tr>
<tr>
<td>min</td>
<td>0.0</td>
<td>Japan</td>
<td>0.0</td>
<td>Japan</td>
</tr>
<tr>
<td></td>
<td></td>
<td>USA</td>
<td></td>
<td>USA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Norway</td>
<td></td>
<td>Norway</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Switzerland</td>
<td></td>
<td>Switzerland</td>
</tr>
<tr>
<td>med</td>
<td>0.01</td>
<td>Kazakstan</td>
<td>0.03</td>
<td>Thailand</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Djibouti</td>
<td></td>
<td>Bhutan</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Syria</td>
<td></td>
<td>Morocco</td>
</tr>
<tr>
<td></td>
<td></td>
<td>South Africa</td>
<td></td>
<td>Bolivia</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>freetrade</th>
<th>relevant countries</th>
<th></th>
<th>autarky</th>
<th>relevant countries</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absolute change</td>
<td></td>
<td></td>
<td>Absolute change</td>
<td></td>
</tr>
<tr>
<td>max</td>
<td>0.33</td>
<td>Barbados</td>
<td></td>
<td>-0.16</td>
<td>Saint Kitts and Nevis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Seychelles</td>
<td></td>
<td></td>
<td>Antigua and Barbuda</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Saint Kitts and Nevis</td>
<td></td>
<td></td>
<td>Seychelles</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Antigua and Barbuda</td>
<td></td>
<td></td>
<td>Malta</td>
</tr>
<tr>
<td>min</td>
<td>0.0</td>
<td>Japan</td>
<td></td>
<td>0.0</td>
<td>India</td>
</tr>
<tr>
<td></td>
<td></td>
<td>USA</td>
<td></td>
<td></td>
<td>Norway</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Norway</td>
<td></td>
<td></td>
<td>United States of America</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Switzerland</td>
<td></td>
<td></td>
<td>Japan</td>
</tr>
<tr>
<td>med</td>
<td>0.09</td>
<td>Bhutan</td>
<td></td>
<td>-0.02</td>
<td>Nicaragua</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bolivia</td>
<td></td>
<td></td>
<td>East Timor</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Israel</td>
<td></td>
<td></td>
<td>Argentina</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Guyana</td>
<td></td>
<td></td>
<td>Romania</td>
</tr>
</tbody>
</table>
Table 12: Summary of the reaction (in %) of the extensive margin of bilateral trade to changes in trade costs

<table>
<thead>
<tr>
<th></th>
<th>10%-reduction</th>
<th>20%-reduction</th>
<th>freetrade</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>our model</td>
<td>EK</td>
<td>our model</td>
</tr>
<tr>
<td>max</td>
<td>125.4</td>
<td>125.4</td>
<td>433.2</td>
</tr>
<tr>
<td>min</td>
<td>-27.2</td>
<td>-34.5</td>
<td>-44.9</td>
</tr>
<tr>
<td>mean</td>
<td>7.4</td>
<td>-0.1</td>
<td>10.5</td>
</tr>
<tr>
<td>median</td>
<td>2.7</td>
<td>-5.5</td>
<td>-5.0</td>
</tr>
</tbody>
</table>
Figure 1: Importer per-capita gdp vs. extensive margin for 44 randomly selected exporters (year 2000, log-log)
Figure 2: Exporter per-capita gdp vs. extensive margin for the same 44 randomly selected importers as in the previous Figure (year 2000, log-log)
Figure 3: The estimated extensive margin of consumption ($M_n$) against per-capita income (relative to US per-capita income)
Figure 4: The multilateral extensive margin of imports vs. importer per-capita income (relative to US income)
Figure 5: The extensive margin of consumption vs. importer per-capita gdp (relative to US) for different $w$. 
Figure 6: Absolute change in extensive margin of consumption due to 20%-reduction in trade costs against intial per-capita income (relative to US)
Figure 7: Relative change in extensive margin of consumption due to 20%-reduction in trade costs against initial per-capita income (relative to US)
Figure 8: The value of the loss function for different preference parameters $\sigma$ and $\bar{c}$. 
Figure 9: The grid (black) of the values of the loss function for different preference parameters. The red dots represents the minima for given $\sigma$. 