Inattentive Importers *

KUNAL DASGUPTA † JORDI MONDRIA ‡

PRELIMINARY AND INCOMPLETE

Abstract

Importers rarely observe the price of every good in every market because of information frictions. In this paper, we aim to explain how the presence of such frictions shape the pattern of trade across countries. To this end, we introduce rationally inattentive importers in a multi-country, multi-good Ricardian trade model. We derive a gravity-like equation linking bilateral trade flows and the cost of processing information faced by importers. In this setting, a reduction in conventional trade costs has large effects on trade flows as importers re-optimize information processing across countries. The model explains a number of findings in the literature related to the response of trade flows to various trade barriers.

*We would like to thank Victor Aguirregabiria, Martin Burda, Davin Chor, Arnaud Costinot, Michal Fabinger, Peter Morrow, Ananth Ramanarayanan and Kim Ruhl for helpful discussion.

†University of Toronto.

‡University of Toronto.

1 Introduction

Trade costs are high. Our understanding of the determinants of these costs, however, remains limited (Head and Mayer, 2013). One possible source of such costs is information frictions. Imperfect information about prices plagues markets. An importer rarely observes the price of a given product in every market. How does imperfect information affect international trade flows? In this paper, we take a first step towards answering this question.

Despite a widespread agreement among economists that imperfect information could and do create significant barriers to trade, we lack a framework that formalizes the link between information and trade. The difficulty in developing such a framework is due to the absence of a standard way of modeling information. In the international trade literature, the informational structure is usually treated as exogenous. But as Anderson and van Wincoop (2004) point out in their definitive survey on trade costs, we need more careful modeling of information costs. We believe that by treating information as a parameter that is determined outside the model, we are missing out on several interesting insights, some of which could have potentially important implications for how countries trade. Borrowing tools from the rational inattention literature (Sims, 2003, 2006), we develop a model of information and trade.

Our decision to model information as in the rational inattention theory is guided by three considerations. *First*, the central premise of this theory is that although information is freely available, agents have a limited capacity to process information. Faced with a capacity constraint, agents must decide how much information they want to process about their variable(s) of interest. We believe this to be a quite accurate description of the real world. *Second*, rational inattention has an appealing feature that the process through which agents process information is endogenous. In the context of our model, this causes importers to have a different level of information about different source countries in equilibrium, without taking recourse to ad-hoc differences in such information across countries. And *finally*, by using rational inattention, we can appeal to powerful results from the multinomial choice literature that allows us to derive elegant, closed-form solutions for bilateral trade flows as functions of information costs and other primitives. This allows us to perform simple comparative static exercises, despite a complex underlying problem.

We introduce rational inattention in a N-country, multi-good Ricardian model of trade. As in Eaton and Kortum (2002), we model productivity of a good as stochastic. The key object of interest in our model is the unconditional probability that importers in country j buy a good k from producers in country i, $\pi_{ij}(k)$. Following Matejka and McKay (2012), we derive a system of equations involving the $\pi_{ij}(k)$s, $i = 1, ...N$, for a particular j. Solving these unconditional probabilities requires solving N-dimensional integrals. We show that when the productivity distributions have a particular form, one can obtain closed-form solutions for the $\pi_{ij}(k)$s. Under a mild condition, we obtain a gravity-like equation linking bilateral trade with the cost of processing information. We believe this to be one of the main contributions of the paper.

The endogenous processing of information affects the response of trade flows to a change in conventional trade frictions between trading partners. As the trade cost between importing country j and exporting country i declines, country j importers start to purchase more from country i because the average price offered by country i producers is now lower. This is the standard effect of trade costs on trade flows present in any trade model. In our model, there is an additional effect however. Faced with a cost of processing information, importers in country j choose how much information to process about every source country, including country j. A lowering of expected price in country i raises the expected benefit of processing information about country i. Country j importers respond by paying more attention to country i and less attention to every other country, thereby boosting the volume of trade between j and i further. Thus, when importers are rationally inattentive, small differences in conventional trade costs could have large effect on trade flows - there is a *magnification effect*. This is the key insight of our model of inattentive importers. We show that this mechanism can explain a number of findings in the literature related to the response of trade flows to various trade barriers.

A number of papers have provided evidence of informational asymmetry in international trades. Using data on trade in agricultural goods in Philippines, Allen (2012) shows that most of the reduction in trade can be attributed to informational frictions, thereby highlighting their significance. In a highly influential paper, Rauch (1999) showed that proximity, common language and colonial ties are more important for trade in differentiated products, which is more dependent on information, than for products traded on organized exchanges. Gould (1994) shows that immigrant links to the home country have a strong positive effect on both exports and imports for the U.S. while Head and Ries (1998) find the same for Canada. Although higher import from the home country could simply reflect greater demand from the immigrants, the same cannot be said of exports. The latter, the argument goes, probably reflects better information possessed by the immigrants about their home markets. Rauch and Trindade (2002) find that for differentiated goods, the presence of ethnic Chinese networks in both the trading partners increases trade.

One of the few papers to explicitly use proxies for information cost in explaining trade flows is Portes and Rey (2005). They run a standard gravity equation and find that informational flows, captured by telephone call traffic and multinational bank branches, have significant explanatory power for bilateral trade flows. Morales et al. (2011) show that the entry of a exporter in a particular market increases the likelihood of his entry into other similar markets. Their finding seems to suggest that we may not be in a full information world, and when a firm enters a market, it gets new information about similar markets. An absence of perfect information about foreign markets also features in the exporting models of Eaton et al. (2010) and Albornoz et al. (2012). Chaney (2013) incorporates exporter networks into a model of trade. Among other things, he shows that his network model can explain the distribution of foreign markets accessed by individual exporters. The importance of networks in trade is suggestive of the presence of informational barriers.

The paper that is closest in spirit to our paper is Allen (2012). He considers producers sequentially searching for the lowest price across markets which makes information about prices endogenous. As in our paper, Allen derives bilateral trade flows as a function of information costs. Our model, however, differs from Allen (2012) in two important ways. First, the information friction in our model is on the side of the buyers, rather than sellers. Producers in Allen (2012) search for the most attractive market while importers in our model search for the most attractive source. Second, in Allen (2012) if a producer searches N markets, he exactly knows the price in those N markets and has no information about prices in the remaining markets. In contrast, importers in our model have varying degrees of information about every market, but never observe prices in any market perfectly.

In a related paper, Arkolakis et al. (2012) introduce staggered adjustment in the Eaton-Kortum model of trade. They assume that in each period, consumers continue to buy from the same supplier with some probability - consumers are inattentive. Accordingly, with some probability, consumers do not respond to price shocks that hit other suppliers. Arkolakis et al. takes the inattention as given, however, and is therefore silent on how the degree of inattention itself could respond to trade costs.

Recent work in the areas of macroeconomics and finance has used information theoretic ideas. In macroeconomics, Sims (2003, 2006), Luo (2008), Luo and Young (2009) and Tutino (2013) have applied rational inattention to study the consumption and savings behavior of households. Mackowiak and Wiederholt (2009), Woodford (2009) and Matejka (2012) have used information theoretic ideas to analyze the price setting behavior of firms. In finance, Peng and Xiong (2006), Van Nieuwerburgh and Veldkamp (2009, 2010) and Mondria (2010) have applied rational inattention in asset pricing and portfolio choice models. To the best of our knowledge, we are the first ones to apply rational inattention to the study of international trade.

The rest of the paper is organized as follows. In Section 2, we specify preferences, production structure and the market clearing conditions. Section 3 introduces inattentive importers and derives the equilibrium. We also perform meaningful comparative statics in this section, for a given distribution of wages and input costs across countries. In Section 4, we numerically solve the full general equilibrium of the model and perform counterfactuals. Section 5 concludes. All the proofs are in the Appendix.

2 Model

We consider a Ricardian model of trade with N countries. Country i is populated by L_i individuals. Each individual consumes a final good and supplies a unit of labor inelastically. For simplicity, we assume that labor is the only factor of production. Labor is perfectly mobile within a country but immobile across countries. There are two sectors that add value domestically - an intermediate goods sector that produces intermediate inputs that can be traded, and a final goods sector that produces a non-traded good for consumption. Henceforth, we shall express all the variables in per capita terms.

Trade costs. The "standard" trade cost between exporting country i and importing country j is captured by τ_{ij}. We do not take a stand on the interpretation of τ_{ij} (i.e., whether it is an iceberg cost) until later, except noting that a higher value corresponds to higher trade costs. τ_{ij} includes all types of costs that are typically suggested by the gravity literature like transportation costs, border costs, policy barriers, etc. Importantly, this does not include information costs. For simplicity, we assume that trade costs are symmetric.

Intermediate good sector. There is a continuum of intermediate inputs indexed by $k \in [0, 1]$. Production of each intermediate requires labor and an aggregate of all the intermediates. Good k can be produced in country i and made available to country j using the following technology:

$$q_{ij}(k) = f\left(l_i, Q_i, A_i, z(k), \frac{1}{\tau_{ij}}\right),$$

where l_i and Q_i denotes the amount of labor and the aggregate intermediate good. A_i is a location parameter common to all goods k produced in country i, while $z(k)$ is a random shock drawn independently for each good k from a cumulative distribution function $F(z)$, with a zero location parameter, which is assumed to be common across countries.[1] We separate the location parameter, A_i, from the random productivity realizations for tractability reasons. At this point we simply assume that f' is positive with respect to each of its arguments. We shall have more to say about the functional form for f later on.

[1] All the results in the paper are robust to making A_i good-specific and $z(k)$ country-specific.

Each intermediate good is combined using a CES aggregator into Q_i:

$$Q_i = [\int_0^1 q_i(k)^{1-\frac{1}{\sigma}} dk]^{\frac{\sigma}{\sigma-1}},$$

where $\sigma > 1$ is the elasticity of substitution. The CES aggregate is produced by competitive firms simply by combining all the intermediates. The producers of the aggregate good (henceforth importers) in country i purchase intermediate goods from all over the world; the corresponding amounts are given by $q_i(k)$. The price of one unit of the aggregate is then given by

$$P_i = [\int_0^1 p_i(k)^{1-\sigma} dk]^{\frac{1}{\sigma-1}}, \tag{1}$$

where $p_i(k)$ is the price of good k paid by importers in country i.

Final good sector. Perfectly competitive firms combine labor and intermediate goods to produce a final good:

$$y_i = \left(\frac{l_i}{\alpha}\right)^{\alpha} \left(\frac{Q_i}{1-\alpha}\right)^{1-\alpha}.$$

We assume that y_i is also the utility of each individual in country i. The share of labor in final good production, α, is common across countries. The price of the final good is then

$$s_i = w_i^{\alpha} P_i^{1-\alpha}, \tag{2}$$

where w_i is the wage in country i.

Cost of intermediates. The cost of importing one unit of good k from country i to country j is given by $1/z_{ij}(k)$, where $z_{ij}(k)$ is the *adjusted* productivity of country i (adjusted for trade costs). In this paper, we use two alternative definitions of $z_{ij}(k)$:

ASSUMPTION 1A: $z_{ij}(k) = \frac{A_i + z(k)}{(w_i^{\beta} P_i^{1-\beta})\tau_{ij}}$.

ASSUMPTION 1B: $z_{ij}(k) = \frac{A_i}{(w_i^{\beta} P_i^{1-\beta})\tau_{ij}} + z(k)$.

Under Assumption 1A, trade costs and input costs in country i affect both the location parameter and the shape of the productivity distribution in country i in producing good k for country j, as in Eaton and Kortum (2002). Such importing costs costs can be generated if $f = C[A_i + z_i(k)] l_i^{\beta} Q_i^{1-\beta} / \tau_{ij}$, where C is some constant. Under this assumption, our model cannot be solved analytically. But because this is the more standard assumption in the literature, we study the properties of the model under this assumption numerically.

Under assumption 1B, trade costs and input costs only affect the location parameter in country i. This assumption has two advantages. First, this specification for costs allows us to obtain a closed-form solution for trade shares and gain intuition about the new friction introduced in the paper. Second, this specification introduces a smaller error when computing trade flows if the true trade costs are of a per unit nature rather than ad-valorem.[2]

Importers in j want to pay the lowest price for each good k. They would ideally like to import good k from the most efficient country, that is the country with $\max[z_{ij}(k); i = 1, ..., N]$. Prices and productivity realizations, however, are not perfectly observed when importers choose the country from where to purchase the good. We assume though, that the price of good k in country i becomes fully observable once country i has been chosen to supply a good. This assumption of perfect observability *ex-post*, combined with perfect competition in the market for intermediate inputs, implies that the producers of the same intermediate good in any country do not engage in strategic price setting.[3] The price at which producers in country i are willing to sell good k to importers in country j is then given by

$$p_{ij}(k) = \frac{1}{z_{ij}(k)}. \tag{3}$$

i.e., producers are willing to sell their goods at marginal cost. It must be emphasized that $p_{ij}(k)$ is the price that is *actually paid* by country j importers if they choose to purchase good k from country i. The un-observability of prices *ex-ante*, however, implies that $p_{ij}(k)$ may not be the lowest price for good k faced by the importers in country j.

Market clearing. Let the set of intermediate goods that country j purchases from country i be denoted by $\Omega_{ij} \subset [0, 1]$. The share of expenditure by country j on goods imported from country i is given by

$$s_{ij} = \int_{\Omega_{ij}} p_{ij}(k) q_{ij}(k) dk / (L_j P_j Q_j),$$

where $L_j P_j Q_j$ is the total expenditure on intermediates in country j. The total imports in country

[2]To see this, suppose the true cost of intermediate good k produced by country i in country j is $\frac{1}{\bar{z}_i + z(k)} + \tau_{ij}$, where $\tau_{ij} > 0$. This can be re-written as $[1 + \hat{\tau}_{ij}]\frac{1}{\bar{z}_i + z(k)}$, where $\hat{\tau}_{ij} = \tau_{ij}(\bar{z}_i + z(k))$. The ad-valorem trade cost is increasing in productivity. Hence, if we use an ad-valorem cost that is invariant across source countries, we end up predicting more trade with less productive source countries and less trade with more productive source countries relative to what the data might suggest. Under our assumption 1B, the error would still be there, but would be smaller because the ad-valorem cost is unaffected by $z(k)$.

[3]Essentially, in the presence of information frictions, firms selling a homogeneous good might choose to charge a price greater than marginal cost even with free entry.

i are given by

$$\text{Imports} = L_i P_i Q_i \sum_{j \neq i} s_{ji}.$$

while total exports are given by

$$\text{Exports} = \sum_{j \neq i} s_{ij} L_j P_j Q_j.$$

Balanced trade requires that

$$L_i P_i Q_i \sum_{j=1}^{N} s_{ji} = \sum_{j=1}^{N} s_{ij} L_j P_j Q_j,$$

where exports and imports are now been broadly defined to include country i's sale to itself. Following Alvarez and Lucas (2007), it can be shown that the trade balance equation reduces to a system of equations involving the wages:

$$L_i w_i = \sum_{j=1}^{N} s_{ij} L_j w_j. \tag{4}$$

As we show in the next section, the s_{ij}-s are functions of w_i and other parameters. Hence, given the endowment of labor in each country, (4) represents N non-linear equations in N unknowns - w_i. As shown by Alvarez and Lucas (2007), this system of equations will typically have at least one set of solution.

In the next section, we introduce inattentive importers and solve for the equilibrium of the model. All the results in this section are conditional on wages and input prices.

3 Rationally Inattentive Importers

We assume that importers are rationally inattentive. Importers choose how to process information about each country, given a limited capacity to process information. In equilibrium, this yields the amount of information they have about each country. The innovation provided by rational inattention is that importers are not constrained to learn about each country's productivity draws with a particular signal structure, but rather, are allowed to choose the optimal mechanism to process information. As argued by Sims (2003) and Matejka and McKay (2012), there is no need to model the signal structure explicitly - it is enough to solve for the optimal joint distribution of the variable of interest and actions. In our model, importers choose the (i) joint distribution of trade cost adjusted productivity in each country, and (ii) the country from where to import a particular good, subject to a capacity constraint for processing information.

3.1 Information

Following Sims (2003), we use tools from information theory to model the limited information processing capabilities of importers. We define $\lambda_j(k)$ as the unit cost of information of country j importers about good k and $\kappa_j(k)$ as the total amount of information processed by country j about suppliers of good k. Let $\tilde{Z}(k)$ be the vector of adjusted productivity realizations of all countries for good k. Information theory measures information as the reduction in uncertainty. By paying a cost $\lambda_j(k)\kappa_j(k)$, country j importers can reduce their uncertainty about the realization of $\tilde{Z}(k)$ by $\kappa_j(k)$. We use *entropy* as the measure of uncertainty about $\tilde{Z}(k)$ and *mutual information* as the measure of uncertainty reduction.

Definition. The *entropy* $H(X)$ of a discrete random variable X that takes values x in $\mathbb{X}$ is

$$H(X) = -\sum_{x\in\mathbb{X}} p(x)\log p(x),$$

where $p(x)$ is the probability mass function of X.

Definition. The *mutual information* of two random variables X and Y (taking values y in $\mathbb{Y}$) is given by

$$I(X;Y) = \sum_{x\in\mathbb{X}}\sum_{y\in\mathbb{Y}} p(x,y)\log\frac{p(x,y)}{p(x)p(y)},$$

where $p(x,y)$ is the joint probability mass function of X and Y, while $p(y)$ is the marginal probability mass function of Y.

In other words, entropy is the expectation of $\log(\frac{1}{p(x)})$. As an example, consider a random variable that takes only two values: x_1 with probability p and x_2 with probability $1-p$. The entropy of this random variable is $-p\log p-(1-p)\log(1-p)$. Figure 1 plots the entropy as a function of p. As the figure suggests, entropy is a hump-shaped function, attaining a maximum at $p=\frac{1}{2}$ and a minimum of zero for both $p=0$ and $p=1$. These properties of entropy are actually quite general. When $p=\{0,1\}$, the random variable is not "random" any more. Accordingly, there is no uncertainty - entropy is zero. But as p rises above zero (or falls below one), uncertainty is introduced and consequently entropy becomes positive. Entropy is maximum when all realizations of p are equally likely.

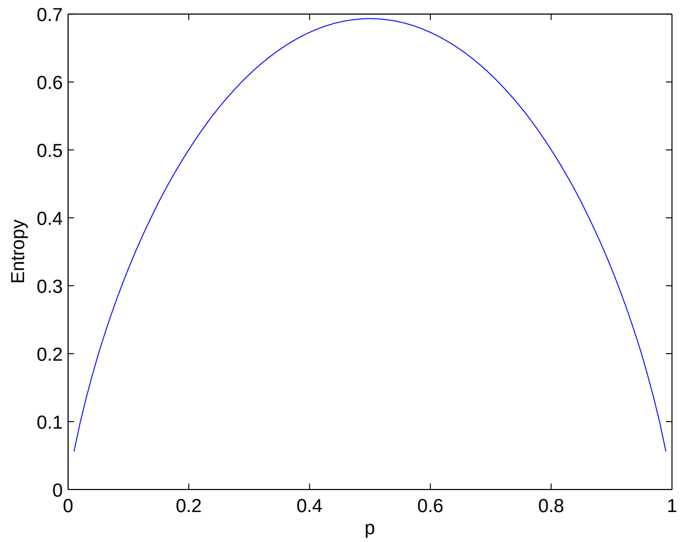

Figure 1: Entropy when the random variable is binary

It is straight-forward to show that mutual information can be re-written as

$$I(X;Y) = H(X) - E_y[H(X|Y)],$$

where $H(X|Y) = -\sum_x p(x|y) \log p(x|y)$ is the entropy of X conditional on Y. The following properties of mutual information (Cover and Thomas, 1991) will be useful later on:

PROPERTY 1: $I(X;Y) \geq 0$.

PROPERTY 2: $H(X) - E_y[H(X|Y)] = H(Y) - E_x[H(Y|X)]$.

Because of Property 1, mutual information can be interpreted as the reduction in uncertainty in X caused by the knowledge of Y. Property 2 suggests that the role of X and Y in the definition can be reversed.

3.2 Results for the General Case

In our model, importers face uncertainty regarding which country has the highest adjusted productivity, and hence, the lowest cost of delivering good k. The object of interest is the likelihood of importing a good from a particular country. Let us define $f_{ij}(k)$ as the probability that country j buys good k from country i conditional on the realization of $\tilde{Z}(k)$, and $\pi_{ij}(k)$ as the unconditional probability that country j buys good k from country i. Then,

$$\pi_{ij}(k) = \int_{\tilde{Z}(k)} f_{ij}(k) d\tilde{F}(\tilde{Z}(k)). \tag{5}$$

Importers in country j process information about $\tilde{Z}(k)$ to reduce its entropy $H(\tilde{Z}(k))$. Instead of explicitly modeling the optimal signal structure, rational inattention allows us to measure uncertainty reduction as the mutual information between $\tilde{Z}(k)$ and the country i chosen by country j:

$$\begin{aligned}
\kappa_j(k) &= H\big(\tilde{Z}(k)\big) - E_i\big[H\big(\tilde{Z}(k)|ij\big)\big] \\
&= H(ij) - E_{\tilde{Z}(k)}\big[H\big(ij|\tilde{Z}(k)\big)\big] \\
&= -\sum_{i=1}^{N} \pi_{ij}(k)\log\pi_{ij}(k) + \int_{Z(k)} \Big(\sum_{i=1}^{N} f_{ij}(k)\log f_{ij}(k)\Big) dF(\tilde{Z}(k)).
\end{aligned} \tag{6}$$

where the equality between the first and the second line follows from Property 2. $H(ij)$, in the above equation, captures the *ex-ante* uncertainty of country j's importers about which country i to buy good k from. Once the importers observe $\tilde{Z}(k)$, albeit imperfectly, their uncertainty is reduced. The resulting difference is the information that country j importers have about the productivity of exporters across the world in good k.

If information could be processed freely, an importer would find out the true realization of $\tilde{Z}(k)$. There are, however, a multitude of costs, captured by $\lambda_j(k)$, involved in processing information about the true productivity of a supplier. Importers in country j choose to consume good k from the country that has the highest expected productivity, taking into account the information processing costs. Therefore, importers in country j solve the following optimization problem:

$$\max_{[f_{ij}(k)]_{i=1}^{N}} \sum_{i=1}^{N} \int_{\tilde{Z}(k)} z_{ij}(k) dF(\tilde{Z}(k)) - \lambda_j(k)\kappa_j(k), \tag{7}$$

subject to

$$f_{ij}(k) \geq 0 \quad \forall i, \tag{8}$$

$$\sum_{i=1}^{N} f_{ij}(k) = 1, \tag{9}$$

where $\kappa_j(k)$ is given by (6), $f_{ij}(k)$ is given by (5) and $1/z_{ij}(k)$ is the trade cost adjusted price. Rationally inattentive importers choose a probability distribution over where to buy good k from. Following Matejka and McKay (2012), the next proposition derives the equilibrium conditional probabilities.

Proposition 1. *If $\lambda_j(k) > 0$, then conditional on the realization of $\tilde{Z}(k)$, the probability of importers in country j choosing to import good k from country $i \in \{1, ..., N\}$ is given by*

$$f_{ij}(k) = \frac{\pi_{ij}(k) \exp\left(\frac{z_{ij}(k)}{\lambda_j(k)}\right)}{\sum_{h=1}^{N} \pi_{hj}(k) \exp\left(\frac{z_{hj}(k)}{\lambda_j(k)}\right)}. \tag{10}$$

If the countries are *a priori* identical, then $\pi_{ij}(k) = 1/N$ for all i and the posterior probability that country j buys good k from country i follows a multinomial logit (REFERENCE). Hence, in this model, the demand for goods is a modified multinomial logit that takes into account the priors, $\pi_{ij}(k)$. The following corollary makes an important observation:

Corollary 1. *If $\nexists i$ such that $\pi_{ij}(k) = 1$ and $\pi_{hj}(k) = 0$ for all $h \neq i$, then $\nexists i$ such that $f_{ij}(k) = 1$.*

If the importers in country j attach positive prior probabilities of importing good k from at least two countries, then conditional on the productivity draws, they will never buy good k from only one country. Notice that the corollary contrasts sharply with the result in Eaton and Kortum (2002). In their paper, even though the prior (unconditional) probabilities of importing good k by country j is positive for every exporting country i, this probability drops to zero for every exporting country but one after the productivity draws are realized. Under full information, once the productivities are drawn, importers in country j purchase good k almost surely from one country, the country which has the lowest cost of delivering good k to country j. In our model, this is not true any more. If country j is populated by a large number of importers for each good k, by applying a Law of Large Numbers we can conclude that a fraction $f_{ij}(k)$ of them will import the good from country i.[4] In the literature, when a narrowly defined good is imported from many countries, it is customary to assume that they represent different varieties (REFERENCE). In our model, a good that is identical in every respect could still be imported from multiple countries because of information frictions.

In the next few propositions, we characterize the unconditional probability that country j chooses to import good k from country i.

Proposition 2. *$\pi_{ij}(k)$ is increasing in the location parameter A_i and decreasing in input costs w_i, P_i, and trade costs τ_{ij}.*

Proposition 2 states that *a priori*, importers in country j are less likely to purchase goods from countries that are farther away or have higher input costs, other things remaining equal. A decrease in A_i lowers the the average productivity of good k in country i and reduces the probability that the price of good k in i is the lowest expected price. In a full information model, this results in a lower probability of purchasing good k from country i. In our model, there is an additional effect. The rationally inattentive importer in country j compares the expected marginal benefit of processing information about country i's productivity with the marginal cost of information. As the probability of getting a lower price in country i declines, so does the information processed by country j importers about country i. Consequently, $\pi_{ij}(k)$ drops further - the presence of information costs creates a *magnification effect*. The same mechanism reduces $\pi_{ij}(k)$ following an increase in input costs or trade costs.

Proposition 3. *If countries are identical but $\tau_{ij} > \tau_{jj}$ for all $i \neq j$, $\pi_{jj}(k)$ is increasing in $\lambda_j(k)$.*

Proposition 3 states that all else equal, an increase in the cost of processing information increases the probability of purchasing good k from the home country. Intuitively, the higher are the information processing costs, the less information importers are able to incorporate into their decision making and the greater is the weight attached to the initial priors. Because the expected adjusted productivity in the home country is the highest among all the countries because of trade costs, increased importance of the prior raises the likelihood of buying the good from the home country. This result is related to findings of home bias in consumption. DISCUSS THEIR FINDINGS. It must be emphasized that the home bias in our model arises not due to exogenous differences in the information structure, but endogenous differences in information across countries being processed by importers.

Proposition 4. *If countries are identical but $\tau_{ij} > \tau_{jj}$ for all $i \neq j$, then as $\lambda_j(k) \to \infty$, $\pi_{ij}(k) \to 0$ for all $i \neq j$ and $\pi_{jj}(j) \to 1$.*

[4] Because of perfect competition among the importers, the number of importers is not determined.

Proposition 4 goes one step further than the previous propositions and establishes that all else equal, if the information processing cost of a good becomes extremely large, consumers stop importing this good completely. Figure 2 is helpful in illustrating the intuition of the model.

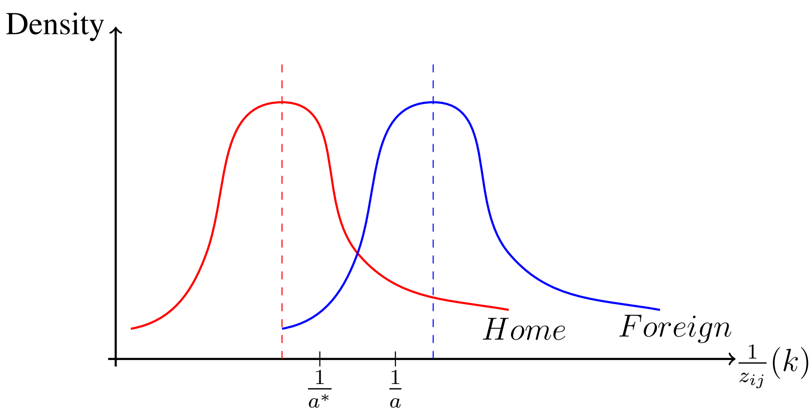

Figure 2: Distribution of adjusted productivity

Imagine for simplicity that there are two countries in the world with the same productivity distribution and input costs. The only difference between countries is that, for a given productivity realization, the foreign country needs to charge higher prices in the home market because of trade costs. Figure 2 plots the price (inverse of adjusted productivity) distribution for the home and foreign countries. The price distribution of the home country lies to the left of the foreign country because of trade costs. The dotted lines correspond to the mean of the distributions. If the cost of processing information is zero, then consumers have perfect information and choose to consume from the country with the lowest price. In particular, if the home country draws a while the foreign country draws a^*, then home importers should buy the good from the foreign country. Of course, given the shape of the distribution, home importers will buy goods from home more often than not.

If information is imperfect, however, consumers will choose to purchase goods from the country they *believe* offers the lowest price. In the extreme case, where the information costs tend to infinity and no information is processed, consumers will choose to buy all goods from the home country. Intuitively, for any good, consumers will choose to purchase the good from the country with lowest expected price. Since the two countries are identical, except that goods from the foreign country face trade costs, it will be always the case that goods produced at home have the lowest expected price. Hence, using a continuity argument, it is easy to see that the higher are the information processing costs, the lower the likelihood of purchasing a good from the foreign country.

Observe that while deriving Propositions 1, 2, 3 and 4, we did not specify a particular functional form for $z_{ij}(k)$ or for the distribution of random productivity draws $F(z)$. In particular, these results are satisfied under both assumptions 1A and 1B and for *any* $F(z)$. At this level of generality, however, we cannot explicitly solve for the $\pi_{ij}(k)$-s. Rather, we have to solve for the fixed point of (5) where $f_{ij}(k)$ is given by (10).

3.3 Closed-form solution for a particular $F(z)$

Here we use Assumption 1B and a specific form for $F(z)$ to derive of a closed-form solution for $\pi_{ij}(k)$. This facilitates a more detailed analysis of the model. Following Cardell (1997), we define a distribution $C(\lambda)$.

Definition. For $0 < \lambda < 1$, $C(\lambda)$ is a distribution with a probability density function given by

$$g_\lambda(z) = \frac{1}{\lambda} \sum_{n=0}^{\infty} \left[\frac{(-1)^n e^{-nz}}{n! \Gamma(-\lambda n)} \right] \tag{11}$$

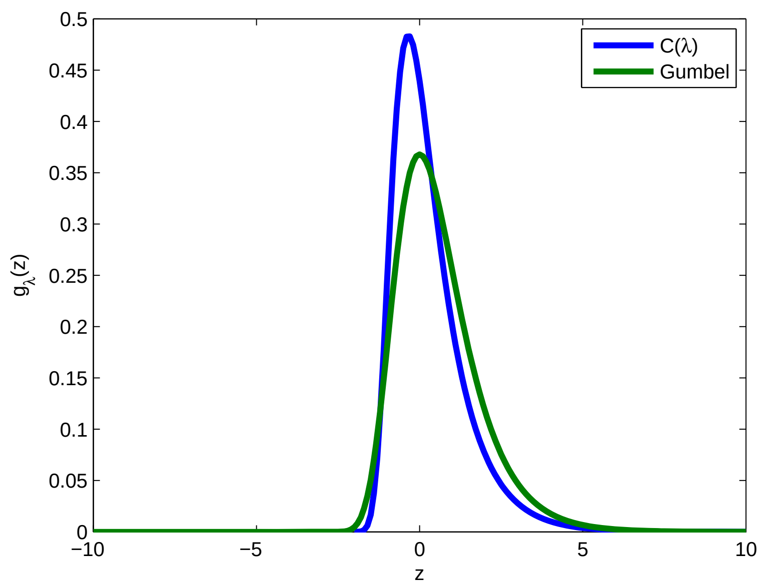

Figure 3: Distribution $C(\lambda)$

The main property of the $C(\lambda)$ distribution is that if a random variable ϵ is drawn from a Type I extreme value (Gumbel) distribution and another random variable ν is drawn from $C(\lambda)$, then $\nu + \lambda\epsilon$ is a random variable distributed as Type I extreme value. The relation between $C(\lambda)$ and a Gumbel distribution is shown in Figure 3.[5] It is clear that qualitatively, the two distributions are very similar. The next proposition shows that when Assumption 1B is satisfied and the random productivities are drawn from a $C(\lambda)$ distribution, there exists a closed-form solution for $\pi_{ij}(k)$.

Proposition 5. *Under assumption 1B, if the $z(k)$-s are drawn independently from a cumulative distribution $C\big(\lambda(k)\big)$ where $0 < \lambda(k) < 1$, then $\pi_{ij}(k)$ is given by*

$$\pi_{ij}(k) = \frac{\exp\left(\frac{\bar{z}_{ij}}{1-\lambda(k)}\right)}{\sum_{h=1}^{N} \exp\left(\frac{\bar{z}_{hj}}{1-\lambda(k)}\right)},$$

where $\bar{z}_{ij} = \frac{A_i}{(w_i^{\beta} P_i^{1-\beta})\tau_{ij}}$.

The unconditional probability that country j buys good k from country i follows a multinomial logit. This result contributes to the literature on rational inattention by finding a closed-form solution to a problem with asymmetric prior probabilities. Eaton and Kortum (2002) obtain a similar expression for $\pi_{ij}(k)$. In their paper, this is also the probability that country i offers the lowest price for good k to country j. It is derived entirely from the primitive productivity distributions in each country - it is a feature of technology.[6] In contrast, the unconditional probability $\pi_{ij}(k)$ in our model is derived from the conditional probabilities $f_{ij}(k)$ that are *chosen* by inattentive importers.[7]

The derivation of closed form-solutions for $\pi_{ij}(k)$'s relies on a few key assumptions. First, $\lambda(k)$ must lie in a bounded interval $[0, 1]$. It is straightforward to show that unlike in the more general model, as $\lambda(k) \to 1$, $\pi_{j'j} \to 1$, where $\bar{z}_{j'j} = max_i[\bar{z}_{ij}]$. When $\lambda(k)$ gets arbitrarily close to 1, country j importers tend to import almost exclusively from the country with the highest average productivity. In other words, a value of $\lambda(k) = 1$ corresponds to prohibitively high information processing costs. Second, we assume that $z(k)$ is drawn from a cumulative distribution $C(\lambda)$ which has a support of $(-\infty, \infty)$. Accordingly, $z(k)$ and hence, $z_{ij}(k)$ could be negative. Although $\pi_{ij}(k)$ is defined irrespective of the sign of $z_{ij}(k)$, a negative productivity does not make sense. As a result, we truncate any negative draw of $z_{ij}(k)$ at zero. This obviously introduces an error in our derivation of $\pi_{ij}(k)$. We can, however, make this error as small as possible by choosing $\bar{z}_{ij}$ appropriately. For example, if $\bar{z}_{ij}$ equals 2 in Figure 3, the probability that $\bar{z}_{ij} + z(k)$ is

[5] For this plot, we choose $\lambda = 0.5$.

[6] A Fréchet distribution for the random productivity draws generates a Weibull distribution for the price distributions, resulting in a closed-form expression for $\pi_{ij}(k)$.

[7] For this derivation, we need $\lambda_j(k) = \lambda(k) \forall j$. Because $\lambda(k)$ governs the productivity distribution for each country, we set it equal across countries to ensure that they draw from a distribution that has the same shape and differs only in terms of the mean.

negative is less than 0.00004. And finally, the parameter $\lambda(k)$ plays two different roles. It is not only the cost of processing information but is a parameter of the distribution $C(\lambda)$ as well. In particular, λ affects the shape of the productivity distribution.[8]

The prior probabilities are not observed in the data. They are, however, related to the share of country j's expenditure on good k bought from country i, $s_{ij}(k)$. The following lemma formalizes this relation:

Lemma 1. *For σ large, $s_{ij}(k)$ can be approximated by $\pi_{ij}(k)$.*

Therefore, for σ large, we can use the terms $\pi_{ij}(k)$ and $s_{ij}(k)$ interchangeably. Let $X_{ij}(k)$ be the value of good k imported by country j from country i, and $E_j(k) = \sum_i X_{ij}(k)$ be the total expenditure by country j on good k. Proposition 5 and Lemma 1 then implies that

$$X_{ij}(k) = \frac{\exp\left(\frac{\bar{z}_{ij}}{1-\lambda(k)}\right)}{\sum_{h=1}^{N}\exp\left(\frac{\bar{z}_{hj}}{1-\lambda(k)}\right)} E_j. \qquad (12)$$

Equation (12) brings out the exact relationship between the share of country j's expenditure on good k that goes to country i, and the cost of processing information. Observe that (12) is similar to a gravity equation that connects bilateral trade flows to trade barriers, geographic and otherwise. Hence, our model provides a micro-foundation for the assumption in the literature that bilateral trade flows depend on information frictions. The trade flow equation also bears a close resemblance to the corresponding equation in Eaton and Kortum (2002), the difference being that instead of a parameter that captures the variance of the productivity distribution, we have a parameter that captures information friction. To see how information costs could distort trade flows, consider the relative imports of country j from countries i and i':

$$\frac{X_{ij}(k)}{X_{i'j}(k)} = \exp\left(\frac{\bar{z}_{ij} - \bar{z}_{i'j}}{1-\lambda(k)}\right).$$

In the presence of information costs, the differences between countries are magnified. Even a small difference in adjusted average productivities could have a large effect on relative trade shares if the cost of processing information is high. The intuition is the same as before: when processing information is costly, importers tend to place a greater weight on their prior beliefs. As a result, source countries that are slightly better (in terms of providing goods cheaply) capture a disproportionately large share of the market for *any* good.

[8] As a result, we must be careful in interpreting results when performing comparative statics with λ.

The next proposition describes the properties of $\pi_{ij}(k)$. Without loss of generality, we assume that for a given country j, the exporting countries are ordered with respect to $\bar{z}_{ij}$ with $\bar{z}_{1j}$ being the largest and $\bar{z}_{Nj}$ being the smallest.

Proposition 6. *$\pi_{ij}(k)$ $\big($ and $s_{ij}(k)\big)$ has the following properties:*

1. *$\frac{\partial \pi_{ij}(k)}{\partial \bar{z}_{ij}} > 0$ and $\frac{\partial \pi_{hj}(k)}{\partial \bar{z}_{ij}} < 0$ for all $h \neq i$.*
2. *$\frac{\partial \log \pi_{ij}(k)}{\partial \log \tau_{ij}}$ (the trade elasticity) is increasing in $\lambda(k)$.*
3. *$\frac{\partial \log \pi_{ij}(k)}{\partial \log \tau_{ij}}$ is decreasing in τ_{ij}.*
4. *$\pi_{1j}(k)$ is monotone increasing while $\pi_{Nj}(k)$ is monotone decreasing in $\lambda(k)$. For any other i, $\pi_{ij}(k)$ has a hump shape. Let $\lambda_{ij}(k)$ be defined implicitly by $\frac{\partial \pi_{ij}(k)}{\partial \lambda(k)} = 0$. Then $\lambda_{ij}(k)$ is decreasing in i.*

Part 1 of Proposition 6 indicates that conditional on the cost of intermediate inputs and wages in countries other than i, any increase in the location parameter of country i raises the likelihood of importers in country j buying good k from country i. As discussed earlier, as the expected productivity of good k produced in country i rises and the expected price falls, country j importers optimally choose to pay more attention to country i. This has an additional effect on the probability of importing from country i, over and above the usual effect. This result contributes to the literature on border effects. McCallum (1995) had shown that the border between U.S. and Canada has a disproportionately large effect on international and intra-national trade: trade between Canadian provinces seem to be 22 times higher than trade between a Canadian province and a U.S. state. Although this number has been shown to be much lower in subsequent research by Anderson and van Wincoop (2003), it still seems too large to be explained by conventional costs that might be involved in border-crossing. Our model suggests a possible explanation: even if the international border imposes a small cost on trade between Canadian provinces and U.S. states, inattentive importers in Canada, for example, could choose to purchase much more from Canadian provinces relative to U.S. states. We use a numerical exercise in Section 3.4 to show how very small differences in trade costs between importing region j and exporting regions i and h could lead to large differences in bilateral trade flows between j and i, and j and h. DISCUSS ALTERNATIVE EXPLANATIONS OF BORDER PUZZLE.

Our model also provides a possible explanation for the distance elasticity puzzle. Disdier and Head (2008) study thousands of measures of distance elasticity of trade in the data and find that it is about $0.9 (\approx 1)$ on average. That means, if distance falls by 10 percent, bilateral trade rises by about 10 percent. Now, transport costs are roughly 20 percent of total trade costs (Anderson and van Wincoop, 2004). Therefore, a 10 percent fall in distance should reduce transport costs by about 2 percent. So it seems that a 2 percent decline in transport costs increases bilateral trade by 10 percent. As pointed out by Grossman (1998), this number is implausibly large. One possible explanation is that distance proxies for other barriers like informational frictions. In our model, a 2 percent fall in transport costs can easily lead to a 10 percent increase in trade if information costs are large enough. As we show in Section 3.4, this result is obtained even with more familiar productivity distributions such as the Fréchet and trade cost specifications such as iceberg trade costs.

Part 2 of Proposition 6 discusses the properties of the trade elasticity. The trade elasticity in our model is endogenous and depends on equilibrium values of wages and input costs. If $\lambda(k)$ varies across goods, trade elasticity varies too. One might expect that the cost of processing information about differentiated goods could be higher relative to homogeneous goods. Differentiated goods, such as electronic goods, have many attributes that might be harder to assess compared to a homogeneous good like steel or cement; this makes it costlier to reduce uncertainty about differentiated goods. In a seminal paper, Rauch (1999) showed that the distance elasticity is higher for differentiated goods. Rauch's explanation was that because differentiated goods involve greater uncertainty, trade in such goods is more dependent on information. If an increase in distance also reduces the flow of information between two countries, this will have a relatively bigger impact on trade in differentiated goods. Our model formalizes this intuition. When distance with an exporting country i increases, importers optimally choose to process less information about every good in country i, and relatively less information about goods that have a high information cost. Hence, trade involving the latter goods is affected more.

The trade elasticity is also decreasing with the level of trade costs, as suggested by Part 3 of Proposition 6. Although standard gravity models do not have this feature, Eaton and Kortum (2002) provides evidence that seems to suggest that distance elasticity of trade is falling in distance. Table 1 below collects estimates from their paper.

As shown in the table, when we move from the first to the second distance interval, the (absolute value of) log change in trade rises from 3.1 to 3.66. This is a difference of 0.56. But when we move from the second to the third interval, the difference is only 0.37. This difference falls further to 0.19 when we move from the third to the fourth interval. Because we are doubling the distance as we move from one interval to the next, the table suggests that the distance elasticity is falling with distance.

Part 4 of Proposition 6 sheds light on a property of the model that again highlights a novel

Distance intervals	Log change in bilateral trade
$[0, 375)$	-3.1
$[375, 750)$	-3.66
$[750, 1500)$	-4.03
$[1500, 3000)$	-4.22

Table 1: The effect of distance on trade
*Source:*Eaton and Kortum (2002)

insight from rational inattention theory. As the cost of information rises, the share of imports coming from every country except the most attractive (z_{1j}) and the least attractive (z_{Nj}) displays non-monotonicity. This contrasts with the response of import shares to a change in any other cost. As part 1 of Proposition 6 suggests, a decrease in τ_{1j} reduces π_{ij} for every country other than country 1. Now, a reduction in τ_{1j} can be interpreted as an increase in the cost of accessing market j by every country relative to country 1. In particular, if $j = 1$, then an uniform import tariff imposed by country j on every country will reduce the share of every country in country j's imports. But if the cost of information rises, the import share of every country other than j does not decline. Rather, when $\lambda(k)$ is small, an increase in $\lambda(k)$ raises the import shares from countries that have high $\bar{z}_{ij}$. This happens due to endogenous information processing. When information becomes costly, importers in country j relocate attention to countries that have lower expected prices resulting in the share of imports coming from some of these countries actually going up. This suggests that information costs differ from more traditional trade costs in important ways.

Before using data to calibrate the location parameters and trade costs in Section 4, we consider the benchmark case of symmetric countries, i.e., we set $\frac{A_i}{w_i^{\beta} P_i^{1-\beta}} = 1 \forall i, j$ and $\tau_{ij} = \tau$. We also assume that the goods are symmetric and that countries face the same information cost, i.e. $\lambda_j(k) = \lambda \forall j, k$. In this case, imports of a country as a share of its GDP are given by

$$m_j = (1 - \alpha)(1 - s_{jj}),$$

where

$$s_{jj} = \frac{(N-1)\exp(\frac{1-\tau}{\tau(1-\lambda)})}{1 + (N-1)\exp(\frac{1-\tau}{\tau(1-\lambda)})}.$$

One way of ascertaining how country size might affect trade flows in a symmetric world is to look at how s_{jj} varies with N. Because the size of each country is $\frac{1}{N}$ of the size of the world, a bigger country corresponds to a smaller N. The relationship between import share and size is shown in Figure 4.

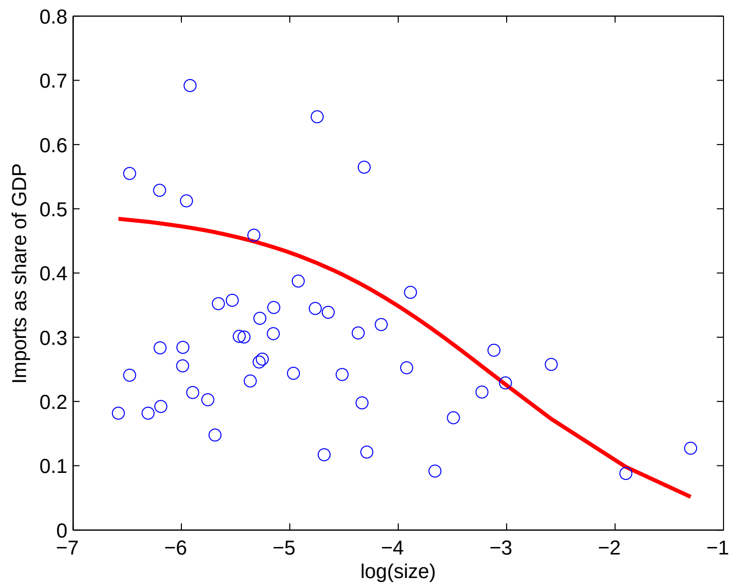

Figure 4: Import shares as a function of size

Each circle in Figure 4 corresponds to the import share of a country whose size is given by its share in world GDP. We use data for the 50 largest countries (in terms of GDP) that account for more than 95 percent of world GDP. The import share data is averaged over 1995-95 while the GDP data is for 1999. In the plot, we leave out two countries: Singapore and Hong Kong. Both of these countries have import shares greater than one. Although there is a lot of variation, the plot shows a clear negative relation between size and import share. The biggest economies like the U.S. and Japan purchase about 10 percent of their output from the rest of the world, while for smaller economies like the Czech Republic or Hungary, the corresponding figure is greater than 50 percent. The data for imports, as well as GDP is from World Bank's *World Development Indicators*.

On this curve, we super-impose data for trade shares generated by our model. The red curve in Figure 4 captures the relationship between m_j and $\log \frac{1}{N}$ as predicted by the model. Notice that the variation in N does not really capture cross country variation in size, because each N corresponds to a different world. Nevertheless, one could think of the red curve as showing the import share of country j, where j's trading partners are symmetric. Following Alvarez and Lucas (2007), we choose $\alpha = 0.75$. We choose τ and λ so as to match the average import share (weighted by GDP) of 0.21. Because we have one degree of freedom, various combinations of (τ, λ) are consistent with an average share of 0.21. The curve in Figure 4 is drawn for $\tau = 2.7$. This is an ad-valorem trade cost of 170 percent which is the average trade cost for developed countries as reported by Anderson and van Wincoop (2004). This corresponds to a λ of 0.8. Notice that the import share generated by the model displays a clear negative relation with respect to size. The fit of the curve is not good however, as suggested by a correlation between the model and the data of 0.35. Given how parsimonious the model is, this is to be expected.

3.4 When productivity draws follow a Fréchet

The analytical results in Section 3.3 were derived under a special technology and productivity distribution. Under assumption 1A, we have a Ricardian model with the more familiar iceberg trade costs. The following proposition states that the model by Eaton and Kortum (2002) is a limiting case of our model with rationally inattentive importers when the cost of information processing equals zero.

Proposition 7. *Under Assumption 1A, if $\lambda_j(k) \to 0$, then our model is equivalent to Eaton and Kortum (2002).*

With unlimited capacity to process information, importers will be in a full information world all the time and will purchase a good k from the country that offers the lowest price. As mentioned earlier, in the presence of information costs, one cannot analytically derive the unconditional probabilities any more. To obtain $\pi_{ij}(k)$s, we compute the integrals defined in (5) using Monte-Carlo methods.

We simulate a model with four countries and with one hundred realizations of productivity for each country. For this numerical exercise, we choose a Fréchet distribution for productivity realizations so that $F(z) = e^{-z^{-\theta}}$ as in Eaton and Kortum (2002). We assume that $\theta = 8.28$. To illustrate the role of information frictions, we also assume that all countries have the same input costs and are only distinguished by trade costs. Let the home country be denoted by 1. We number countries by "distance" from the home economy and assume that $\tau_{11} = 1; \tau_{21} = 1.01; \tau_{31} = 1.02; \tau_{41} = 1.03$. This implies that each country faces a cost of selling to country 1 that is one percent higher than its neighboring country. Our object of interest is $\pi_{i1}(k)$, the probability that country 1 importers buy good k from country i. By assuming that all goods have the same information processing costs, $\lambda_1(k) = \lambda_1$, one can interpret $\pi_{i1}(k)$ as the fraction of goods that country 1 importers buy from country i, π_{i1}.

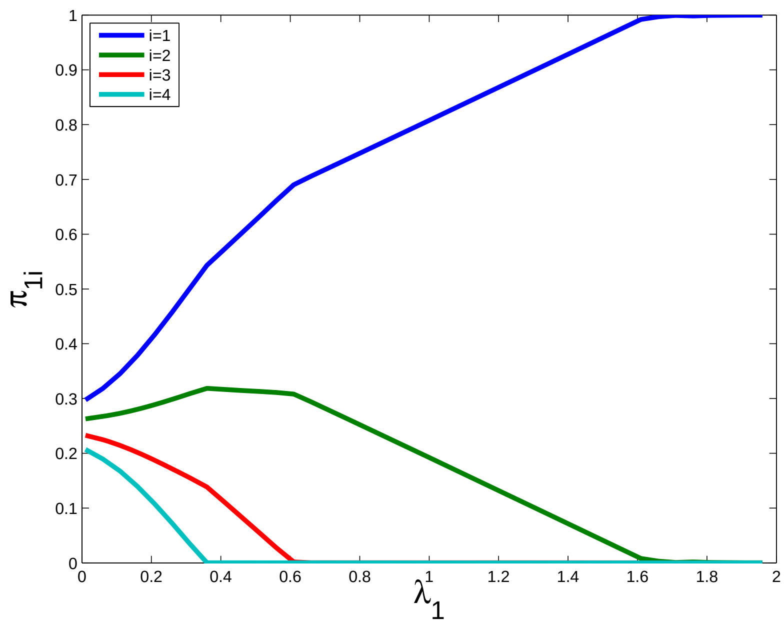

Figure 5: $\pi_{ij}(k)$ under iceberg trade cost

Figure 5 shows the $\pi_{i1}(k)$s for different levels of the information processing cost λ_j. We see, as we proved in part 3 of Proposition 6, that π_{11} is increasing with λ_1 while π_{41} is decreasing with λ_1. The fraction of goods bought from country 2 displays non-monotonicity. Observe that the fraction of goods purchased from countries 2, 3 and 4 decreases towards zero as the information processing costs increase. These probabilities are never equal to zero, but they get very close to zero for large enough λ_1. Intuitively, if information costs are too high, importers in country 1 can only process a limited amount of information. Hence, they choose to process very little information about countries that are more distant because the likelihood of getting a good adjusted productivity draw from these countries is very small - it is optimal for the importers to process information mostly about a few close countries. If trade flows are truncated below, our model provides a possible explanation for observing zero trade flows when information costs are high Helpman et al. (2008).

Figure 5 also shows that for a given λ_1, small differences in trade costs could have large implications for trade flows. First, for a λ_1 close to 0.36, home importers purchase a negligible fraction of goods from country 4. The latter is not extremely far away from the home country; its trade costs are only 3 percent higher than the costs at home. In fact, in a full information world, home importers import almost 20 percent of their goods from country 4. Second, for $\lambda_1 = 0.36$, a one percent difference in trade costs between country 2 and country 3 generates a difference in the fraction of goods purchased by home importers of 18 percentage points between the two economies. Assuming that π_{i1} is approximately equal to s_{i1}, the trade share of country i, this number implies that bilateral trade between country 2 and home is almost 2.3 times higher than bilateral trade between country 3 and home. Thus, this simple numerical exercise suggests that in the presence of information costs, small differences in distance can have large effects on trade flows.

4 General Equilibrium

THIS SECTION IS INCOMPLETE.

In this section, we solve for the full general equilibrium of the model. We set β, the share of labor in the production of intermediate inputs, to one. This simplifies the analysis as we do not have to solve for the prices. We continue to assume that σ, the elasticity of substitution, is large enough so that the share of trade in good k between exporting country i and importing country j can be approximated by π_{ij}. Initially, we also set $\lambda_j(k) = 1$ for all j, k. This means that $\pi_{ij}(k) = \pi_{ij}$ is also the share of expenditure of country j on imports from country i.

To solve the model, we need data on trade costs τ_{ij} and the location parameters for the productivity distributions, A_i. We set $\tau_{ij} = \tau d_{ij}^{0.3}$, where d_{ij} is the great circle distance between capital cities of countries i and j, and τ is chosen so as to match the average bilateral trade flow between countries. We use R&D expenditures (share of GDP) for the year 2005 as a proxy for A_i. We use the same set of countries as in Section 3.3 except for Venezuela and Bangladesh due to a lack of R&D data. For the following exercise, we choose $\lambda = 0.8$.

Figure 6 shows the correlation between the nominal wages generated by the model and the data. For wage data, we use per capita GDP for the year 1999. The plot shows a clear positive correlation between wages in the model and the data. The exact correlation is about 0.72. The relation between the import shares generated by the model and corresponding shares in the data is shown in Figure 7 for the entire sample. As is clear, there is apparently a weak correlation between the model and the data. The exact correlation is about 0.19. What could be driving this result? Looking carefully at the import shares, we see that the deviation between the import shares in the data and those generated by the model is systematic - the deviations are much bigger for lower income countries. In particular, the model predicts that these countries should import much more than they actually do.

To examine this further, we look at the relation between the import shares (model versus data)

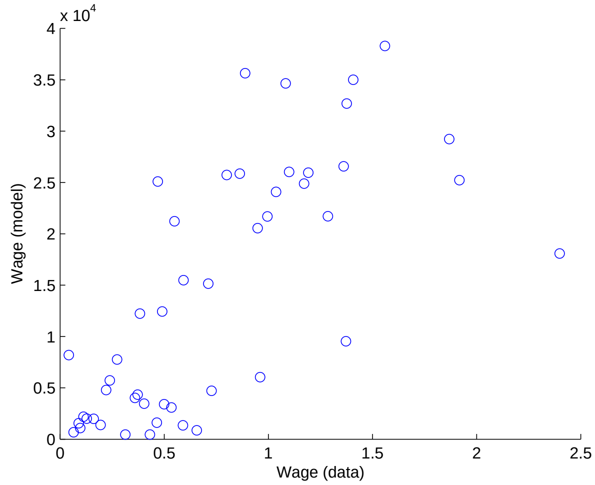

Figure 6: Relation between wages in the model and the data

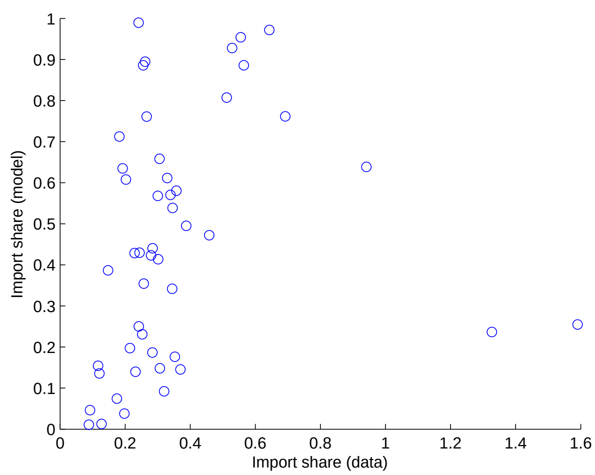

Figure 7: Import shares for all countries (model vs data)

of only those countries that have a per capita GDP greater than 20,000 (measured in 1990 U.S. dollars). From this sample, we also leave out Singapore ad Hong Kong. The resultant relation is shown in Figure 8. Now, there is a clear positive relation between the model and the data. The exact correlation jumps to 0.87.

5 Conclusion

TO BE WRITTEN

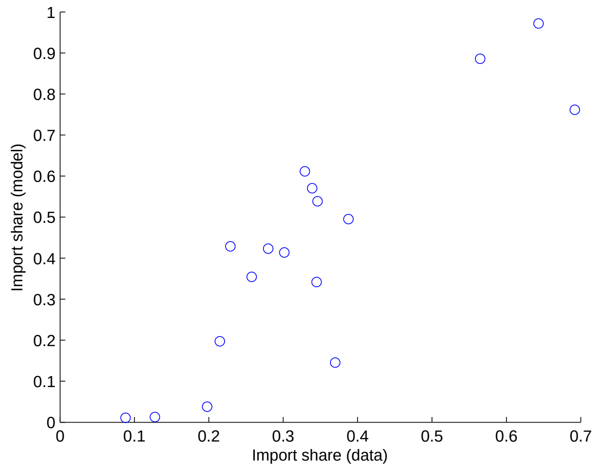

Figure 8: Import shares for rich countries (model vs data)

References

Albornoz, F., Pardo, H. F. C., Corcos, G. and Ornelas, E. (2012), 'Sequential exporting', *Journal of International Economics* **88**(1), 17 – 31.

Allen, T. (2012), Information frictions in trade, mimeo, Northwestern University.

Alvarez, F. and Lucas, R. E. (2007), 'General equilibrium analysis of the eaton–kortum model of international trade', *Journal of monetary Economics* **54**(6), 1726–1768.

Anderson, J. E. and van Wincoop, E. (2003), 'Gravity with gravitas: A solution to the border puzzle', *The American Economic Review* **93**(1), 170–192.

Anderson, J. E. and van Wincoop, E. (2004), 'Trade costs', *Journal of Economic Literature* **42**, 691–751.

Arkolakis, C., Eaton, J. and Kortum, S. (2012), Staggered adjustment and trade dynamics, mimeo, Yale University.

Cardell, N. S. (1997), 'Variance components structures for the extreme-value and logistic distributions with application to models of heterogeneity', *Econometric Theory* **13**(02), 185–213.

Chaney, T. (2013), A network structure of international trade, NBER Working Paper 19285.

Cover, T. M. and Thomas, J. A. (1991), *Elements of information theory*, Wiley-interscience.

Disdier, A.-C. and Head, K. (2008), 'The puzzling persistence of the distance effect on bilateral trade', *The Review of Economics and Statistics* **90**(1), 37–48.

Eaton, J., Eslava, M., Kugler, M. and Tybout, J. (2010), A search and learning model of export dynamics, mimeo, pennsylvania state university.

Eaton, J. and Kortum, S. (2002), 'Technology, geography, and trade', *Econometrica* **70**(5), 1741–1779.

Gould, D. M. (1994), 'Immigrant links to the home country: Empirical implications for U.S. bilateral trade flows', *The Review of Economics and Statistics* **76**(2), 302–16.

Grossman, G. (1998), Comment on deardorff, *in* J. Frankel, ed., 'The Regionalization of the World Economy', U. Chicago Press, pp. 33–57.

Head, K. and Mayer, T. (2013), Gravity equations: Workhorse, toolkit, and cookbook, mimeo, university of british columbia.

Head, K. and Ries, J. (1998), 'Immigration and trade creation: Econometric evidence from Canada', *Canadian Journal of Economics* **31**(1), 47–62.

Helpman, E., Melitz, M. and Rubinstein, Y. (2008), 'Estimating trade flows: Trading partners and trading volumes', *The Quarterly Journal of Economics* **123**(2), 441–487.

Luo, Y. (2008), 'Consumption dynamics under information processing constraints', *Review of Economic Dynamics* **11**(2), 366–385.

Luo, Y. and Young, E. R. (2009), 'Rational inattention and aggregate fluctuations', *The BE Journal of Macroeconomics* **9**(1).

Mackowiak, B. and Wiederholt, M. (2009), 'Optimal sticky prices under rational inattention', *The American Economic Review* **99**(3), 769–803.

Matejka, F. (2012), Rationally inattentive seller: Sales and discrete pricing.l, mimeo.

Matejka, F. and McKay, A. (2012), Rational inattention to discrete choices: A new foundation for the multinomial logit model, CERGE-EI Working Papers 442.

McCallum, J. (1995), 'National borders matter: Canada-us regional trade patterns', *The American Economic Review* **85**(3), 615–623.

Mondria, J. (2010), ‘Portfolio choice, attention allocation, and price comovement’, *Journal of Economic Theory* **145**(5), 1837–1864.

Morales, E., Sheu, G. and Zahler, A. (2011), Gravity and extended gravity: Estimating a structural model of export entry, mimeo, princeton university.

Peng, L. and Xiong, W. (2006), ‘Investor attention, overconfidence and category learning’, *Journal of Financial Economics* **80**(3), 563–602.

Portes, R. and Rey, H. (2005), ‘The determinants of cross-border equity flows’, *Journal of International Economics* **65**(2), 269 – 296.

Rauch, J. E. (1999), ‘Networks versus markets in international trade’, *Journal of International Economics* **48**(1), 7 – 35.

Rauch, J. E. and Trindade, V. (2002), ‘Ethnic chinese networks in international trade’, *Review of Economics and Statistics* **84**(1), 116–130.

Sims, C. A. (2003), ‘Implications of rational inattention’, *Journal of Monetary Economics* **50**(3), 665 – 690.

Sims, C. A. (2006), ‘Rational inattention: Beyond the linear-quadratic case’, *American Economic Review* **96**(2), 158–163.

Tutino, A. (2013), ‘Rationally inattentive consumption choices’, *Review of Economic Dynamics* **16**(32), 421–439.

Van Nieuwerburgh, S. and Veldkamp, L. (2009), ‘Information immobility and the home bias puzzle’, *The Journal of Finance* **64**(3), 1187–1215.

Van Nieuwerburgh, S. and Veldkamp, L. (2010), ‘Information acquisition and under-diversification’, *The Review of Economic Studies* **77**(2), 779–805.

Woodford, M. (2009), ‘Information-constrained state-dependent pricing’, *Journal of Monetary Economics* **56**, S100–S124.

Appendix

TO BE ADDED