

Inattentive Importers ^{*}

Kunal Dasgupta

Jordi Mondria

University of Toronto

University of Toronto

August 2014

Abstract

Importers rarely observe the price of every good in every market because of information frictions. In this paper, we seek to explain how the presence of such frictions shape the flow of goods between countries. To this end, we introduce rationally inattentive importers in a multi-country Ricardian trade model. Under specific assumptions about preferences and technology, we provide an information-theoretical foundation of the gravity equation that links bilateral trade flows with the cost of processing information faced by importers. A distinguishable feature of our model is that importers buy the same good from several countries. In a more general setting, we analyze how small reductions in observable trade costs may have large effects on trade flows as importers endogenously process different amounts of information across countries. We also show that, unlike traditional trade costs, information costs have non-monotonic and asymmetric effects on bilateral trade flows. Finally, we contribute to the rational inattention literature by providing a closed-form solution to a discrete choice problem with asymmetric prior beliefs.

^{*}We would like to thank Victor Aguirregabiria, Treb Allen, Martin Burda, Davin Chor, Arnaud Costinot, Jonathan Eaton, Michal Fabinger, Filip Matejka, Peter Morrow, German Pupato, Kim Ruhl, Nicolas Schmitt, Daniel Trefler, Kei-Mu Yi, and seminar participants at 2013 NASMES, 2013 CEA, 2013 SED, 2013 AMES, 2014 NAWMES, Wilfrid Laurier and Ryerson for helpful discussions. Kunal Dasgupta and Jordi Mondria gratefully acknowledge financial support from the Social Sciences and Humanities Research Council of Canada.

1 Introduction

Imperfect information plagues international commerce. Importers rarely observe the price and attributes of every good in every market. This absence of perfect information is bound to have an impact on the flow of goods between countries. Yet, despite a widespread agreement among economists that imperfect information could create significant barriers to trade, we lack a framework that formalizes the link between information and trade. The difficulty in developing such a framework lies in the absence of a standard way of modeling information.

Borrowing tools from the rational inattention literature [[Sims, 2003, 2006](#), [Matejka and McKay, 2014](#)], in this paper we develop a model of information and trade. The central premise of this theory is that although information is freely available, agents have a limited capacity to process information.¹ Faced with a capacity constraint, agents must decide how to process information and choose the amount of information they want to process about their objects of interest. We introduce rational inattention into a multi-country Ricardian model of trade with two goods - an agricultural and a manufactured good. The productivity and hence the price of the manufactured good are stochastic. Inattentive importers would like to buy the manufactured good from the country with the lowest price adjusted for trade costs. Prices, however, are not perfectly observable. Hence, importers choose to process information about each country, and based on this information they choose the country where to buy the manufactured good. The more information importers process about a particular country, the better the precision of their information about the realization of productivity in that country. The optimal solution of a rationally inattentive importer is to choose probabilistically where to buy the manufactured good, with this probability distribution following an adjusted multinomial logit.

Our paper makes three main contributions. First, we provide an information-theoretical micro-foundation for the gravity equation. The gravity equation has become a cornerstone of the literature that tries to shed light on the factors that impede trade flows across countries.² Second, we illustrate formally how, in the presence of information processing costs, small differences in observed trade costs can generate large differences in trade flows. The large effects of observable trade costs on trade flows has led researchers to speculate that information frictions could be creating large barriers to trade.³ Third, we find closed-form solutions to a discrete choice problem

¹See [Kahneman \[1973\]](#) and [Pashler \[1998\]](#) for evidence on the human brain's limited capacity to process information.

²For alternate micro-foundations for the gravity equation, see [Anderson \[1979\]](#), [Bergstrand \[1985\]](#), [Deardorff \[1998\]](#), [Eaton and Kortum \[2002\]](#), [Chaney \[2008, 2013\]](#) and [Helpman et al. \[2008\]](#).

³In their definitive survey on trade costs, [Anderson and van Wincoop \[2004\]](#) highlight the need for more careful modeling of information frictions.

with rationally inattentive agents and asymmetric prior beliefs.

Using a specific preference structure and distribution of productivity, we derive trade shares that are observationally equivalent to [Eaton and Kortum \[2002\]](#). We then obtain a gravity equation that, for the first time in international trade literature, formalizes the link between trade flows and information costs. Our paper shares the prediction with [Eaton and Kortum \[2002\]](#) that the unconditional probability that country j imports from country i is positive for every i . Unlike [Eaton and Kortum](#) however, the corresponding posterior or conditional probability in our model is also positive for every i . Even after productivity draws are realized, importers in country j do not perfectly observe prices and hence attach a positive probability to every country i having the lowest price. The implications are twofold. First, a country can buy the same good from different source countries. Second, a country can import and export the same good at the same time. Currently, such patterns in the data are rationalized by appealing to intra-industry trade.

In a more general setting, the model is able to rationalize how small reductions in trade costs may have large effects on trade flows. The endogenous processing of information affects the response of trade flows to a change in observable trade costs between trading partners. When such a 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. Our model has an additional information effect. Faced with a cost of processing information, importers in country j choose how much information to process about every source country. A lower 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 observable trade costs could have large effects on trade flows - there is a *magnification effect*. This is a key insight of our model with inattentive importers.

The key parameter in our model is the cost of processing information. We show that, unlike traditional trade costs, changes in information cost can have non-monotonic, as well as asymmetric effects on bilateral trade flows. When there is an increase in information cost faced by an importing country, the share of imports from any other trading partner first rises but eventually declines (this happens for all but the most remote trading partner). With an increase in information costs, importers may also choose to process more information about countries that are close, resulting in an increase in imports from these countries, to the detriment of countries that are farther away - it is *as if* the importing country is imposing import tariffs that are higher for far

away countries. A uniform increase in standard trade costs can never generate such an outcome.

Our paper is related to a number of papers that have provided evidence of information frictions in international trade. Using data on trade in agricultural goods in Philippines, [Allen \[2012\]](#) demonstrated that most of the reduction in trade can be attributed to information frictions. [Steinwender \[2014\]](#) shows how the establishment of trans-Atlantic telegraph lines, that speeded up the flow of information between the U.S. and U.K., led to a convergence in prices and higher trade volumes. In a highly influential paper, [Rauch \[1999\]](#) provided evidence that proximity, common language and colonial ties are more important for trade in differentiated products, which are presumably more dependent on information, than for products traded on organized exchanges and those that have reference prices. [Gould \[1994\]](#) presented evidence 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 \[1998\]](#) found the same for Canada. Although higher imports from the home country could simply reflect greater demand from the immigrants, the same cannot be said of exports. [Rauch and Trindade \[2002\]](#) found that for differentiated goods, the presence of ethnic Chinese networks in both the trading partners increases trade.

One of the few papers that explicitly uses proxies for information costs to explain trade flows is [Portes and Rey \[2005\]](#). They run a standard gravity equation and find that information 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 an 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. The absence of perfect information about foreign markets is also a feature of the exporting models by [Eaton et al. \[2010\]](#) and [Albornoz et al. \[2012\]](#). [Chaney \[2014\]](#) 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 the sellers. Producers in [Allen \[2012\]](#) search for the most attractive market, while importers in our model process information to find 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, and is therefore silent on how the degree of inattention itself could respond to trade costs.

Our decision to model information as in the rational inattention theory is guided by two considerations. First, rational inattention has the appealing feature that information available to agents is endogenously chosen. Unlike most papers in international trade, where the information structure of agents is exogenously given, importers in our model choose to have different amounts of information about prices in different source countries in equilibrium. Second, there already exists a discrete choice model with rationally inattentive agents developed by [Matejka and McKay \[2014\]](#). We are able to contribute to this literature by finding closed-form solutions to a discrete choice problem with a priori heterogeneous options. This allows us to perform simple comparative static exercises despite facing the challenging problem of adding information acquisition to an otherwise standard trade model.

We are the first, to the best of our knowledge, to apply rational inattention to the study of international trade. Recent work in the areas of macroeconomics and finance has also 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.

The rest of the paper is organized as follows. In Section 2, we specify the production structure and introduce inattentive importers. In Section 3, we implicitly solve for the unconditional probabilities of the general model and provide closed-form solutions for the trade shares when preferences and the productivity distribution take a particular form. In the last part of this section, we numerically solve the model with a Fréchet productivity distribution and analyze its properties. Section 4 concludes. All the proofs are in the Appendix.

2 The Model

We consider a world with N countries. There are two goods - a manufactured good and an agricultural good, both of which are consumed and traded.

Agricultural good. The agricultural good is traded freely and is not plagued by information frictions. We assume that every country produces some amount of this good. The price of the agricultural good, which is equalized across countries, is set equal to one. The role of the agricultural good is to allow countries to trade; nothing substantial hinges on the assumptions made about this good.

Manufactured good. The market for the manufactured good is perfectly competitive and is plagued by information frictions. Instead of defining the production function of the manufactured good, we consider its dual, the cost function. The cost of importing one unit of the manufactured good into country j from country i is given by $1/\tilde{z}_{ij}$, where

$$\tilde{z}_{ij} = \frac{A_i \tilde{z}_i}{c_i \tau_{ij}}.$$

We call \tilde{z}_{ij} adjusted productivity (adjusted for trade costs). c_i is the cost of a standardized bundle of inputs required for producing one unit of the manufactured good in country i . The observable trade cost for the manufactured good between exporting country i and importing country j is captured by τ_{ij} , where $\tau_{jj} = 1$ and $\tau_{ij} > 1$ for $i \neq j$. We assume that τ_{ij} is an iceberg cost, i.e., country i has to ship τ_{ij} units in order to sell one unit of a good in country j . The trade cost τ_{ij} includes both policy barriers such as import tariffs and export subsidies, as well as non-policy barriers such as transportation costs, border costs and time costs. Importantly, τ_{ij} does not include information costs. The numerator in the above expression is the productivity of the inputs in country i . It has two components: a scale parameter A_i and a random component \tilde{z}_i . Separating the scale parameter from the random productivity realization simplifies our analysis. The scale parameter A_i is related to the average input productivity in country i and \tilde{z}_i is independently drawn for each country from the same distribution. Let \tilde{Z} be the vector of productivity realizations of all countries that is distributed according to the prior distribution $G(\tilde{Z})$.

There is a large number of importers in country j that want to pay the lowest price for the manufactured good. They would ideally like to import (where import is broadly defined to include purchase from home) the manufactured good from the country with the highest adjusted productivity. Importers maximize utility $u(\tilde{z}_{ij})$, where $u'(\tilde{z}_{ij}) > 0$ and $u''(\tilde{z}_{ij}) \leq 0$, when choosing the

country where to purchase the manufactured good.⁴

We introduce information frictions in the manufactured good sector by assuming that realizations of \tilde{z}_i are not perfectly observable. We also assume that the price of the manufactured good in country i becomes fully observable to importers in country j once country i has been chosen to supply the good. This assumption of perfect observability *ex-post*, combined with perfect competition in the market for the manufactured good, implies that the producers of this good in any country do not engage in strategic price setting.⁵ The price at which producers in country i are willing to sell the manufactured good to importers in country j is then given by

$$p_{ij} = \frac{1}{\tilde{z}_{ij}}, \quad (1)$$

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

Inattentive importers. Importers are rationally inattentive. They choose the information structure about productivity of each country taking into account that information is costly to acquire. 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. Despite this added complexity, however, there is no need to model the signal structure explicitly - it is enough to solve for the optimal distribution of actions conditional on the realization of the variables of interest [Sims, 2003, Matejka and McKay, 2014]. In our model, importers choose the probability that a particular country is selected for importing the manufactured good conditional on the realization of the adjusted productivity in each country.

Following Sims [2003], we use tools from information theory to model the limited information processing capabilities of importers. We define λ_j as the unit cost of information faced by country j importers about the manufactured good and κ_j as the total amount of information processed by country j about suppliers of this good.⁶ Information theory measures information as the reduction in uncertainty. By paying a cost $\lambda_j \kappa_j$, country j importers can reduce their uncer-

⁴If $u''(\tilde{z}_{ij}) < 0$, then importers are risk averse in \tilde{z}_{ij} . In words, importers may not like to buy goods from countries with highly volatile adjusted productivity.

⁵In the presence of information frictions, firms selling a homogeneous good might choose to charge a price greater than marginal cost even with free entry.

⁶Note that λ_j is a parameter while κ_j is a variable.

tainty about the realization of \tilde{Z} by κ_j . We use *entropy* as the measure of uncertainty about \tilde{Z} and *mutual information* as the measure of uncertainty reduction [Shannon, 1948].

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

$$H(X) = - \int_{x \in \mathbb{X}} p(x) \ln p(x) dx,$$

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

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

$$I(X; Y) = \int_{x \in \mathbb{X}} \int_{y \in \mathbb{Y}} p(x, y) \ln \frac{p(x, y)}{p(x)p(y)} dx dy,$$

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

Mutual information can be re-written as

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

where $H(X|Y) = - \int_x p(x|y) \ln p(x|y) dx$ is the entropy of X conditional on Y . Intuitively, mutual information measures the reduction in uncertainty of X caused by the knowledge of Y . The following property of mutual information will be useful later on:

$$\text{PROPERTY 1: } H(X) - E_y[H(X|Y)] = H(Y) - E_x[H(Y|X)].$$

In words, property 1 says that the amount of information contained in Y about X is exactly the same as the corresponding information contained in X about Y .

3 Trade Shares

3.1 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 the manufactured good. The object of interest is the likelihood of importing the good from a particular country. Let us define $f_{ij}(\tilde{Z})$ as the

probability that country j importers buy the good from country i conditional on the realization of \tilde{Z} , or the posterior probability, and π_{ij} as the unconditional probability that country j buys the good from country i . Then, we can write π_{ij} as

$$\pi_{ij} = \int_{\tilde{Z}} f_{ij}(\tilde{Z}) dG(\tilde{Z}). \quad (2)$$

Importers in country j process information about \tilde{Z} to reduce the entropy $H(\tilde{Z})$. An influential property of rational inattention is that there is no need to model explicitly the signal structure that importers receive about \tilde{Z} . The reason is that it is optimal for importers to associate one action (country selection) with at most one particular signal. Two different signals cannot lead to the same action. As actions are associated with at most one specific signal, the information processed by importers in country j can be calculated as the mutual information between adjusted productivities (variable of interest) and selected country (action). Therefore, instead of explicitly modeling the optimal signal structure, rational inattention allows us to measure uncertainty reduction as the mutual information between \tilde{Z} and the country i chosen by importers in country j :

$$\kappa_j = H(\tilde{Z}) - E[H(\tilde{Z}|ij)],$$

where $H(\tilde{Z}|ij)$ is the entropy of \tilde{Z} , conditional on country j importers purchasing the manufactured good from country i . Using Property 1, the above equation can be re-written as

$$\kappa_j = - \sum_{i=1}^N \pi_{ij} \ln \pi_{ij} + \int_{\tilde{Z}} \left(\sum_{i=1}^N f_{ij}(\tilde{Z}) \ln f_{ij}(\tilde{Z}) \right) dG(\tilde{Z}), \quad (3)$$

where the first term on the right-hand side captures the *ex-ante* uncertainty that importers in country j have about choosing to import from country i . Once the importers observe \tilde{Z} , albeit imperfectly, their uncertainty is reduced. The resulting difference is the information that country j importers have about the productivity of suppliers across the world.

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

$$\max_{\{f_{ij}(\tilde{Z})\}_{i=1}^N} \sum_{i=1}^N \int_{\tilde{Z}} u(\tilde{z}_{ij}) f_{ij}(\tilde{Z}) dG(\tilde{Z}) - \lambda_j \kappa_j, \quad (4)$$

subject to

$$f_{ij}(\tilde{Z}) \geq 0 \quad \forall i, \quad (5)$$

$$\sum_{i=1}^N f_{ij}(\tilde{Z}) = 1, \quad (6)$$

where κ_j is given by (3) and $1/\tilde{z}_{ij}$ is the cost of importing one unit of the manufactured good from country i . The first term in (4) is the expected utility of importers. Rationally inattentive importers in country j choose the probability of selecting to import from country i conditional on the realization of \tilde{Z} . The second term in (4) is the cost of processing information. As shown by [Matejka and McKay \[2014\]](#), this optimization problem is equivalent to solving the following two-stage optimization. In the first stage, importers choose to observe signals about the adjusted productivity of each country to reduce their uncertainty. In the second, stage, given the information provided by the signals, importers choose to buy the manufactured good from the country offering the lowest expected price. Following [Matejka and McKay \[2014\]](#), the next proposition derives the equilibrium conditional probabilities.

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

$$f_{ij}(\tilde{Z}) = \frac{\pi_{ij} e^{u(\tilde{z}_{ij})/\lambda_j}}{\sum_{k=1}^N \pi_{kj} e^{u(\tilde{z}_{kj})/\lambda_j}}, \quad (7)$$

where π_{ij} is given by (2).

The conditional choice probabilities have a structure similar to a multinomial logit [[McFadden, 1989](#)], except that they are adjusted by the unconditional probabilities, π_{ij} . The unconditional probabilities π_{ij} are the probabilities that importers from country j choose to buy the manufactured good from country i before any information is processed. These π_{ij} are independent of productivity realizations and only depend on the primitives of the model: prior distribution of adjusted productivities, informations costs and preferences. When the cost of information λ_j is high, conditional choice probabilities attach a high weight to unconditional probabilities as importers process small amounts of information. In this case, if a country i is seen as highly productive *ex-ante*, then it has a high probability of being chosen even if its actual productivity realization

is low. When the cost of information λ_j is low, the conditional choice probabilities attach a high weight to productivity realizations as importers process large amounts of information.

Importers in a given country have homogeneous initial beliefs about which source country has the lowest price for the manufactured good. Their actions, however, may be heterogeneous. If $f_{ij}(\tilde{Z}) > 0$ and $f_{hj}(\tilde{Z}) > 0$, then a fraction $f_{ij}(\tilde{Z})$ of importers in country j will choose to purchase the manufactured good from country i , while a fraction $f_{hj}(\tilde{Z})$ of importers will choose to import from country h .⁷ The following corollary makes an important observation:

Corollary 1. *If $\pi_{ij} > 0$, then $f_{ij}(\tilde{Z}) > 0$.*

An implication of the above corollary is that importers in one country could buy the same product from different countries. If importers have a positive unconditional probability of buying the manufactured good from a particular country, then a fraction of importers will select to purchase the good from that country after processing information about adjusted productivities. The fraction of importers purchasing the good will depend on the realization of adjusted productivities. If the realization of adjusted productivity is high, then the fraction of importers selecting this country will be large. Therefore, if there is more than one country with positive unconditional probabilities, then there will be imports of the manufactured good from more than one country. Notice that corollary 1 contrasts sharply with the result in [Eaton and Kortum \[2002\]](#). In their paper, even though the unconditional probability that country j imports a good from country i is positive for every i , the probability after information is processed drops to zero for every exporting country but the one with the highest productivity realization. Under full information, once the productivities are drawn, importers in country j purchase a good almost surely from one country, the country with the lowest cost of delivering that good to country j . In our model, this is not true any more. Importers never fully observe the productivity draws and believe that every country can have the cheapest product with some probability.

In the literature, when a narrowly defined good is imported from many countries, it is usually assumed that this good represents different varieties [[Klenow and Rodriguez-Clare, 1997](#)]. In our model, the exact same good could still be imported from multiple countries because of information frictions. Furthermore, if the unconditional probabilities that country j both imports from as well as exports to country i are positive, then so are the conditional probabilities. Hence, in equilibrium, we could observe the same good being traded in both directions by two countries. This feature, which is shared by [Allen \[2012\]](#), can never be generated in a full information model of trade.

The next proposition characterizes the properties of the unconditional probability π_{ij} .

⁷Here we are applying a Law of Large Numbers.

Proposition 2. π_{ij} has the following properties:

1. π_{ij} is increasing in the scale parameter A_i and decreasing in input costs c_i and trade costs τ_{ij} .
2. If countries are *ex-ante* identical, $A_i = A$; $c_i = c$; $\tau_{ij} = \tau$ for all i , then $\pi_{ij} = 1/N$ for all i .
3. If countries are *ex-ante* identical but $\tau_{ij} > \tau_{jj}$ for all $i \neq j$, then as $\lambda_j \rightarrow \infty$, $\pi_{ij} \rightarrow 0$ for all $i \neq j$ and $\pi_{jj} \rightarrow 1$.

The first property of Proposition 2 states that *ex-ante*, importers in country j are less likely to purchase goods from countries with a low expected adjusted productivity. A decrease in A_i leads to a decrease in the average productivity of country i and reduces the probability that the price of the manufactured good in country i is the lowest price. In a full information model, this results in a lower probability of purchasing the good 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, π_{ij} drops further - the presence of information costs creates a *magnification effect*.

The second property of Proposition 2 establishes that in a world with no trade costs and where countries are *a priori* identical, all countries have the same *ex-ante* probability of being selected to import the manufactured good. In this case, the choice probabilities $f_{ij}(\tilde{Z})$ in equation (7) follow a standard multinomial logit.

The third property of Proposition 2 demonstrates that all else equal, if the information processing cost becomes extremely large, consumers stop importing from other countries. Intuitively, when information processing costs are high, importers incorporate less information into their decision making and attach a greater weight to the primitives, A_i , c_i and τ_{ij} . Because in the home country trade costs are the lowest and expected adjusted productivity is the highest among all countries, an increase in importance of the primitives raises the likelihood of buying the good from the home country.

The intuition of the model can be explained by Figure 1. Imagine for simplicity that there are two identical countries in the world (h : Home, f : Foreign) with the same productivity distribution and input costs. The only difference between countries is that, for a given productivity realization, foreign producers charge higher prices in the home market because of trade costs. Figure 1 plots the utility of the adjusted productivity distribution for the two countries. The adjusted productivity

distribution for the home country lies to the right 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 buy from the country with the higher productivity realization. For example, if the home country draws $u(z_{hh}) = a$, while the foreign country draws $u(z_{fh}) = b$, then home importers should buy the good from the foreign country.

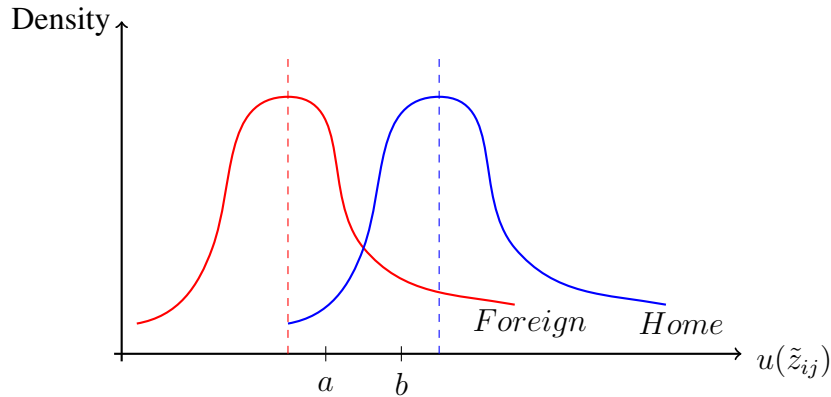


Figure 1: Distribution of price

If information is imperfect, however, importers choose to purchase the good from the country they *believe* has the highest adjusted productivity. In the extreme case, when the information costs are prohibitively high and no information is processed, importers choose to buy the good only from the home country. Intuitively, in the absence of any new information, importers choose to purchase the good from the country with higher *ex-ante* expected utility (of adjusted productivity). Since the two countries are identical, except that the good from the foreign country faces trade costs, it is always the case that the good produced at home has the higher expected adjusted productivity. Hence, using a continuity argument between the perfect information case and the no information case, one can see that when information processing costs increase, there is a decrease in the likelihood of purchasing the manufactured good from the foreign country.

Note that while deriving Propositions 1 and 2, we did not specify a particular functional form for the utility function $u(\tilde{z}_{ij})$ or for the distribution of productivity $G(\tilde{Z})$. In particular, these results are satisfied for any $u(\tilde{z}_{ij})$ and $G(\tilde{Z})$. At this level of generality, however, we cannot explicitly solve for the unconditional probabilities π_{ij} as there is no general solution for the integral in equation (2). In Section 3.2, we find a closed-form solution for π_{ij} using additional assumptions and use numerical integration for the general case in Section 3.3.

3.2 Closed-form solution for π_{ij}

In this section, we make additional assumptions to derive a closed-form solution for π_{ij} . First, we set $u(\tilde{z}_{ij}) = \ln(\tilde{z}_{ij})$. Second, we impose $\lambda_j = \lambda$ for all j . Third, we use a specific form for $G(\tilde{Z})$. Following Cardell [1997], we define a distribution $C(\beta)$.

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

$$g_{\beta}(z) = \frac{1}{\beta} \sum_{n=0}^{\infty} \left[\frac{(-1)^n e^{-nz}}{n! \Gamma(-\beta n)} \right]. \quad (8)$$

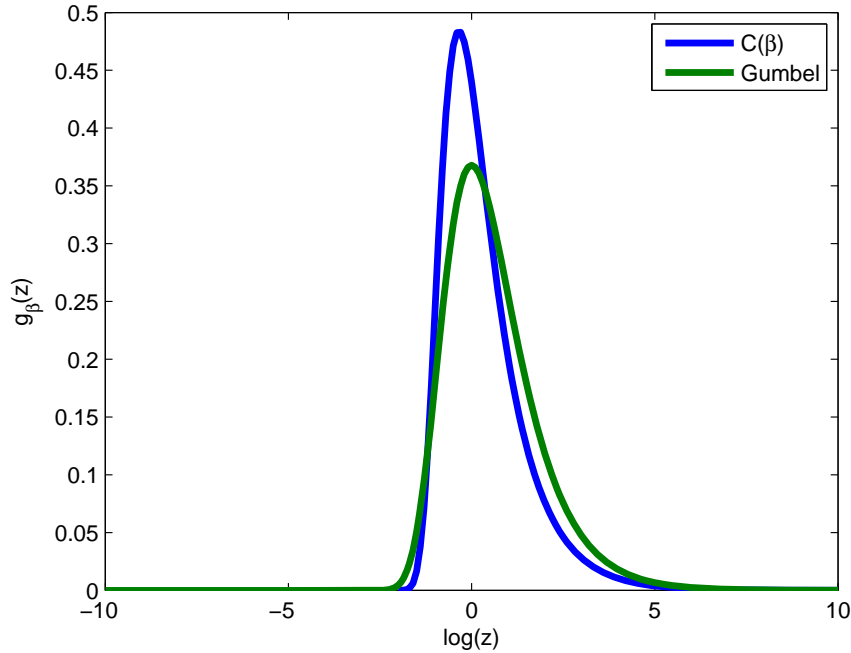


Figure 2: Distribution $C(\beta)$

The main property of the $C(\beta)$ 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 $C(\beta)$, then $\nu + \beta\epsilon$ is a random variable distributed as Type I extreme value. The relation between $C(\beta)$ and a Gumbel distribution is shown in Figure 2.⁸ It is clear that qualitatively, the two distributions are very similar. The next proposition shows that when the logarithm of the random productivities is distributed as $C(\beta)$, there exists a closed-form solution for π_{ij} .

⁸For this plot, we choose $\beta = 0.5$.

Proposition 3. *If $\beta = \lambda$ and $\ln(\tilde{z}_i)$ is drawn independently from a cumulative distribution $C(\lambda)$ where $0 < \lambda < 1$, then π_{ij} is given by*

$$\pi_{ij} = \frac{\bar{A}_i(c_i\tau_{ij})^{\frac{-1}{1-\lambda}}}{\sum_{h=1}^N \bar{A}_h(c_h\tau_{hj})^{\frac{-1}{1-\lambda}}}, \quad (9)$$

where $\bar{A}_i = A_i^{\frac{1}{1-\lambda}}$.

The *unconditional* probability that country j buys the manufactured good from country i has a form that also bears resemblance to a multinomial logit.⁹ The expression for π_{ij} is observationally equivalent to Eaton and Kortum [2002]. There are, however, a few key differences. First, as stated in Corollary 1, importers from the same country choose to import the same good from several countries, *ex-post*. Second, in Eaton and Kortum [2002], π_{ij} is also the probability that country i offers the lowest price for a good to country j . It is derived entirely from the primitive productivity distributions in each country - it is a feature of technology.¹⁰ In contrast, the unconditional probability π_{ij} in our model is derived from the conditional probabilities $f_{ij}(\tilde{Z})$ that are *chosen* by inattentive importers.¹¹ Third, the expression for π_{ij} instead of depending on a parameter that captures the dispersion in the productivity distribution as in Eaton and Kortum [2002], depends on a parameter that captures both information frictions and productivity dispersion. The effect of these parameters on trade flows is the same: a lower dispersion, as well as higher information costs make trade flows more sensitive to trade costs.¹² The underlying mechanisms are quite different, however. A low productivity dispersion makes comparative advantage forces weak, whereby small trade costs cause large differences in trade flows. Instead, when processing information is costly, importers tend to place a greater weight on their *ex-ante* expectations about adjusted productivity. Consequently, countries with smaller trade costs capture a disproportionately large share of the market.

The unconditional probabilities are not observed in the data. They are, however, equal to a variable that we do observe - the share of country j 's expenditure on imports from country i . Let X_{ij} be the value of imports by country j from country i , E_j be the total expenditure by country j

⁹Setting $u(\tilde{z}_{ij}) = \ln(\tilde{z}_{ij})$ gets rid of the exponential terms in equation (9).

¹⁰A Fréchet distribution for the random productivity draws generates a Weibull distribution for the price distributions, resulting in a closed-form expression for π_{ij} .

¹¹For this derivation, first, we set $u(\tilde{z}_{ij}) = \ln(\tilde{z}_{ij}) = \ln(A_i/c_i\tau_{ij}) + \ln(\tilde{z}_i)$ so that we can separate deterministic components from random components. Second, we use a specific distribution for \tilde{Z} (the $C(\beta)$ distribution) and set $\beta = \lambda_j$ to find a closed-form solution of the integral in (2). There is no general solution for this integral. Third, we set $\lambda_j = \lambda$ to ensure that all countries draw from a distribution that has the same shape and differs only in terms of the mean.

¹²The variance of $\ln \tilde{z}_i$ is given by $(1 - \lambda^2)\pi^2/6$.

on the manufactured good, and $X_i = \sum_j X_{ij}$ be the total income of country i from the production of the same good. Proposition 3 and the above observation imply that

$$X_{ij} = \frac{(\tau_{ij})^{\frac{-1}{1-\lambda}}}{\Psi_i \Omega_j} X_i E_j, \quad (10)$$

where Ψ_i and Ω_j measure the average remoteness of countries i and j from the rest of the world.¹³ The gravity equation (10) connects bilateral trade flows to trade costs, production costs and information processing costs. Therefore, our model provides a micro-foundation for the assumption in the literature that bilateral trade flows depend on information frictions.

The next proposition describes some properties of π_{ij} . Without loss of generality, we assume that for a given country j , the exporting countries are ordered with respect to $\frac{A_i}{c_i \tau_{ij}}$ with $\frac{A_1}{c_1 \tau_{1j}}$ being the largest and $\frac{A_N}{c_N \tau_{Nj}}$ being the smallest.

Proposition 4. *The closed-form solution of π_{ij} has the following additional properties:*

1. $\frac{\partial \ln \pi_{ij}}{\partial \ln \tau_{ij}}$ (the trade elasticity) is increasing in λ .
2. π_{1j} is monotone increasing while π_{Nj} is monotone decreasing in λ . For any other i , π_{ij} , as a function of λ , has a hump-shape.
3. Define λ_{ij} as the value of λ that satisfies $\frac{\partial \pi_{ij}}{\partial \lambda} = 0$. λ_{ij} is decreasing in i .

The parameter λ has a dual role in the closed-form solution of π_{ij} as it governs both the cost of processing information and the shape of the productivity distribution in each country. Therefore, one needs to be careful in interpreting the comparative static exercises involving λ as the results are driven by both changes in the shape of the productivity distribution and changes in information costs.

The first property of Proposition 4 discusses the characteristics of the trade elasticity. The trade elasticity, when our model permits a closed-form solution, is simply $\frac{1}{1-\lambda}$. If our model had different types of manufactured goods with good-specific information processing costs, then goods with high λ would have a higher trade elasticity than goods with low λ . If we assume that differentiated goods have higher λ than reference-priced goods,¹⁴ then our model is consistent with the findings in the seminal paper by Rauch [1999], where he showed that the elasticity

¹³The parameters Ψ_i and Ω_j are given by $\Psi_i = \sum_{j=1}^N \tau_{ij}^{\frac{1}{1-\lambda}} \frac{E_j}{\Omega_j}$ and $\Omega_j = \sum_{i=1}^N \tau_{ij}^{\frac{1}{1-\lambda}} \frac{X_i}{\Psi_i}$.

¹⁴Reference-priced goods are those that are not transacted in centralized exchanges, but whose prices are published in trade journals.

of trade with respect to distance is higher for differentiated goods relative to reference-priced goods. Rauch conjectured that the cost of learning about differentiated goods is higher relative to reference-priced goods as differentiated goods have multiple attributes and might require search and matching.

The second property of Proposition 4 sheds light on a property of the model that highlights a novel insight from rational inattention theory. As the cost of information rises, the share of imports 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 standard trade costs, as stated in Proposition 2, where increases in input costs c_i and trade costs τ_{ij} have monotonic effects.

The third property of Proposition 4 suggests that unlike traditional trade costs, a change in information cost also has an *asymmetric effect* on trade shares. To see this, consider the following simple example. Let there be three identical countries, indexed by 1, 2 and 3, with the same productivity distribution and input costs, but different trade costs $\tau_{11} < \tau_{21} < \tau_{31}$. We are interested in the properties of the probability of purchasing the manufactured good by importers in country 1. Suppose country 1 imposes a uniform tariff on imports from both countries 2 and 3. It is easy to show that this will cause π_{11} to rise and both π_{21} and π_{31} to fall as a higher cost of trade causes country 1 to import relatively less from all the trading partners. If λ rises however, the import shares from countries 2 and 3 do not necessarily decline. Rather, when λ is small, an increase in λ could actually lead to an increase in π_{21} . When information costs increase, importers in country 1 reallocate attention to countries with higher *ex-ante* adjusted productivity,¹⁵ to the detriment of other countries. Thus, an increase in λ may lead to an increase in the attention allocated to countries 1 and 2, but a reduction of attention to country 3, resulting in an increase of π_{11} and π_{21} , and a decrease of π_{31} . It is *as if* country 1 is imposing differential import tariffs on goods imported from countries 2 and 3, with the tariff being higher for the country that is farther away. This result highlights that information costs differ from more traditional trade costs in interesting ways.

3.3 When productivity draws follow a Fréchet

In this section, we solve the model using numerical methods. To facilitate comparison with Eaton and Kortum [2002], we make the following assumptions. First, we set $u(\tilde{z}_{ij}) = \tilde{z}_{ij}$. Second, we assume that $G(\tilde{Z})$ follows a Fréchet distribution, with a shape parameter θ . This assumption allows us to study the effect of information processing costs λ_j on trade flows, holding the

¹⁵In this example, higher *ex-ante* adjusted productivity is equivalent to lower trade costs.

shape of the productivity distribution unchanged. And third, we allow countries to have different information costs λ_j , a flexibility permitted by our general model.

The following proposition shows the limiting properties of our model with rationally inattentive importers when the cost of information processing converges to zero.

Proposition 5. *When $\lambda_j \rightarrow 0$ for all j , our model is ex-ante equivalent to [Eaton and Kortum \[2002\]](#).*

With zero information processing costs, importers have unlimited capacity to process information and purchase the good from the country that offers the lowest price. In this case, the only difference with [Eaton and Kortum](#)'s model is that our model has only one manufactured good. Consequently, after productivities have been realized, a country imports this manufactured good from only one country. In [Eaton and Kortum \[2002\]](#), a country imports from every other country even ex-post because there are many manufactured goods. Ex-ante, however, the two models generate identical predictions regarding trade flows, i.e., when $\lambda_j = 0$ for all j , π_{ij} is the same in both models.

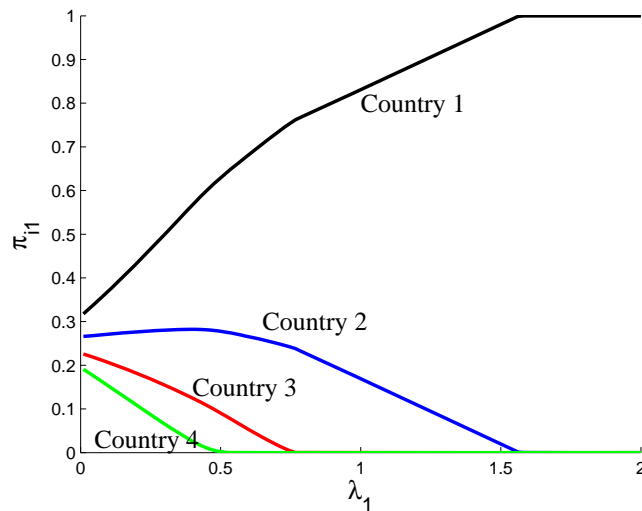


Figure 3: π_{i1} as a function of information cost

In this general setting with positive information processing costs, there is no closed-form solution to the unconditional probabilities π_{ij} as there is no general solution to the integral in equation (2). As a result, we obtain π_{ij} by computing the integrals defined in (2) using Monte-Carlo methods. For the next numerical exercise, we assume there are four identical countries, indexed by 1, 2, 3 and 4, with the same productivity distribution and input costs, but different

trade costs. We ordered countries by their “remoteness” from country 1 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 trade cost of selling to country 1 that is one percent higher than its neighboring country. We assume that productivity follows a Fréchet distribution so that $G(\tilde{z}_i) = e^{-\tilde{z}_i^{-\theta}}$ with $\theta = 8.28$ as in [Eaton and Kortum \[2002\]](#). Our main object of interest is π_{i1} , the probability that country 1 importers buy the manufactured good from country i .

Trade Elasticity. Figure 3 shows the unconditional probabilities π_{i1} ($i = 1, \dots, 4$) for different levels of information processing costs λ_1 . For positive information processing costs, this figure shows how small differences in trade costs may have large effects on trade flows. For example, for $\lambda_1 = 0.45$, the equilibrium trade shares are the following: $\pi_{11} = 60\%$; $\pi_{21} = 28\%$; $\pi_{31} = 11\%$; $\pi_{41} = 1\%$. In this model, small differences in trade costs may lead to large differences in endogenous attention. Importers optimally allocate more attention to countries with higher *ex-ante* expected adjusted productivity, which is equivalent in this example to lower trade costs, resulting in more trade with those countries (in expectation).

Our model provides a possible explanation for the puzzle that the elasticity of trade costs with respect to distance is much smaller than what is needed to explain trade data using traditional models. Let us define the elasticity of trade with respect to distance δ as the product of two elasticities: the elasticity of trade with respect to trade costs ϵ and the elasticity of trade costs with respect to distance ρ . [Disdier and Head \[2008\]](#) review a number of papers and find that δ takes a value, on average, close to -1 . In traditional trade models that generate a gravity equation, ϵ is usually a structural parameter with an average value of around -4 . This implies that $\rho = 0.25$ to be consistent with $\delta = -1$. Most studies that estimate ρ using measures of freight costs have found a value of $\rho = 0.02$, which is an order of magnitude less than what is required to explain the data [[Limao and Venables, 2001](#), [Anderson and van Wincoop, 2003](#), [Hummels, 2007](#)]. If one believes that distance affects trade only through freight costs, then a 1 percent fall in distance should increase trade flows by about 0.1 percent instead of 1 percent as the data suggests.

Turning to Figure 3 we see that when $\lambda_1 = 0.35$, for example, a 1 percent difference in trade costs between country 2 and country 3 generates bilateral trade flows from country 2 to country 1 that are about 50 percent higher than the corresponding trade flows from country 3 to country 1. If we believe that a 1 percent decline in distance causes trade costs to fall by 0.02 percent, then for $\lambda_1 = 0.35$, a 0.02 percent decline in trade costs will cause trade to rise by about 1 percent, thereby resolving the puzzle. [Grossman \[1998\]](#) was one of the first researchers to point out that freight costs are not enough to account for the effect of distance on trade. In fact, [Grossman](#) suggested that distance could be a proxy for other barriers such as information frictions.

Non-monotonicity Figure 3 shows that changes in information processing costs generate non-monotonic changes in trade flows. Specifically, the unconditional probability of importing from country 2 displays a hump-shaped behaviour with respect to information costs. For a small λ_1 , an increase in λ_1 leads to an increase in the imports from country 2. Intuitively, when there is an increase in information costs, importers decrease the total amount of information processed and substitute their attention from countries with high trade costs to countries with low trade costs. This attention reallocation results in an increase of trade flows with countries that are close, at the expense of countries that are farther away.

Zero trade flows. Figure 3 displays that for a large λ_1 , an increase in λ_1 leads to a decrease towards zero of the unconditional probability of importing from countries 2, 3 and 4. These probabilities are never equal to zero, a property that our model shares with Eaton and Kortum [2002], but they get very close to zero for large enough λ_1 . When information costs are high, importers in country 1 choose to process only a limited amount of information. Hence, they process information about only a few countries and trade almost exclusively with those countries. If trade flows are truncated below, our model provides a possible explanation for observing zero trade flows when information costs are high.¹⁶

4 Conclusion

In this paper, we make three main contributions. First, we provide an information-theoretical micro-foundation for the gravity equation. Second, we illustrate formally how, in the presence of information processing costs, small differences in observed trade costs can generate large differences in trade flows. We also show that, unlike traditional trade costs, information costs have non-monotonic and asymmetric effects on bilateral trade flows. Third, we find closed-form solutions to a discrete choice problem with rationally inattentive agents and asymmetric prior beliefs.

In a recent survey on globalization, Head and Mayer [2013] point out that at most 30 percent of the variation in trade flows can be explained by observable freight costs - the remaining 70 percent of the variation is a “dark” trade cost. We believe that a full understanding of these “dark” costs will require more knowledge about the role of information in facilitating trade. By providing a new framework to study the impact of information frictions on international trade, we have taken a step towards unraveling these “dark costs”. Much needs to be done.

¹⁶Haveman and Hummels [2004] and Helpman et al. [2008] provide evidence that there is zero trade flows between a large number of country pairs.

References

- F. Albornoz, H. F. C. Pardo, G. Corcos, and E. Ornelas. Sequential exporting. *Journal of International Economics*, 88(1):17 – 31, 2012.
- T. Allen. Information frictions in trade. mimeo, Northwestern University, 2012.
- J. E. Anderson. A theoretical foundation for the gravity equation. *The American Economic Review*, 69(1):106–116, 1979.
- J. E. Anderson and E. van Wincoop. Gravity with gravitas: A solution to the border puzzle. *The American Economic Review*, 93(1):170–192, 2003.
- J. E. Anderson and E. van Wincoop. Trade costs. *Journal of Economic Literature*, 42:691–751, 2004.
- C. Arkolakis, J. Eaton, and S. Kortum. Staggered adjustment and trade dynamics. mimeo, Yale University, 2012.
- J. H. Bergstrand. The gravity equation in international trade: some microeconomic foundations and empirical evidence. *The review of economics and statistics*, pages 474–481, 1985.
- N. S. Cardell. Variance components structures for the extreme-value and logistic distributions with application to models of heterogeneity. *Econometric Theory*, 13(02):185–213, 1997.
- T. Chaney. Distorted gravity: the intensive and extensive margins of international trade. *The American Economic Review*, 98(4):1707–1721, 2008.
- T. Chaney. The gravity equation in international trade: An explanation. mimeo, toulouse school of economics, 2013.
- T. Chaney. A network structure of international trade. *The american economic review*, forthcoming, 2014.
- A. Deardorff. Determinants of bilateral trade: does gravity work in a neoclassical world? In *The regionalization of the world economy*, pages 7–32. University of Chicago Press, 1998.
- A.-C. Disdier and K. Head. The puzzling persistence of the distance effect on bilateral trade. *The Review of Economics and Statistics*, 90(1):37–48, 2008.

- J. Eaton and S. Kortum. Technology, geography, and trade. *Econometrica*, 70(5):1741–1779, 2002.
- J. Eaton, M. Eslava, M. Kugler, and J. Tybout. A search and learning model of export dynamics. mimeo, pennsylvania state university, 2010.
- D. M. Gould. Immigrant links to the home country: Empirical implications for U.S. bilateral trade flows. *The Review of Economics and Statistics*, 76(2):302–16, 1994.
- G. Grossman. Comment on deardorff. In J. Frankel, editor, *The Regionalization of the World Economy*, pages 33–57. U. Chicago Press, 1998.
- J. Haveman and D. Hummels. Alternative hypotheses and the volume of trade: the gravity equation and the extent of specialization. *Canadian Journal of Economics*, 37(1):199–218, 2004.
- K. Head and T. Mayer. What separates us? sources of resistance to globalization. *Canadian Journal of Economics*, 46(4):1196–1231, 2013.
- K. Head and J. Ries. Immigration and trade creation: Econometric evidence from Canada. *Canadian Journal of Economics*, 31(1):47–62, 1998.
- E. Helpman, M. Melitz, and Y. Rubinstein. Estimating trade flows: Trading partners and trading volumes. *The Quarterly Journal of Economics*, 123(2):441–487, 2008.
- D. Hummels. Transportation costs and international trade in the second era of globalization. *The Journal of Economic Perspectives*, 21(3):131–154, 2007.
- D. Kahneman. *Attention and effort*. Prentice-Hall series in experimental psychology. Prentice-Hall, 1973. ISBN 9780130505187. URL <http://books.google.co.in/books?id=7kvuAAAAMAAJ>.
- P. Klenow and A. Rodriguez-Clare. Quantifying variety gains from trade liberalization. mimeo, university of chicago, 1997.
- N. Limao and A. J. Venables. Infrastructure, geographical disadvantage, transport costs, and trade. *The World Bank Economic Review*, 15(3):pp. 451–479, 2001.
- Y. Luo. Consumption dynamics under information processing constraints. *Review of Economic Dynamics*, 11(2):366–385, 2008.

- Y. Luo and E. R. Young. Rational inattention and aggregate fluctuations. *The BE Journal of Macroeconomics*, 9(1), 2009.
- B. Mackowiak and M. Wiederholt. Optimal sticky prices under rational inattention. *The American Economic Review*, 99(3):769–803, 2009.
- F. Matejka. Rationally inattentive seller: Sales and discrete pricing.I. mimeo, 2012.
- F. Matejka and A. McKay. Rational inattention to discrete choices: A new foundation for the multinomial logit model. CERGE-EI Working Papers 442, 2014.
- D. McFadden. A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica*, 57(5):995–1026, 1989.
- J. Mondria. Portfolio choice, attention allocation, and price comovement. *Journal of Economic Theory*, 145(5):1837–1864, 2010.
- E. Morales, G. Sheu, and A. Zahler. Gravity and extended gravity: Estimating a structural model of export entry. mimeo, princeton university, 2011.
- H. Pashler. *Attention*. Studies in cognition. Psychology Press, 1998. ISBN 9780863778124. URL <http://books.google.ca/books?id=ol8yMiB57RUC>.
- L. Peng and W. Xiong. Investor attention, overconfidence and category learning. *Journal of Financial Economics*, 80(3):563–602, 2006.
- R. Portes and H. Rey. The determinants of cross-border equity flows. *Journal of International Economics*, 65(2):269 – 296, 2005.
- J. E. Rauch. Networks versus markets in international trade. *Journal of International Economics*, 48(1):7 – 35, 1999.
- J. E. Rauch and V. Trindade. Ethnic chinese networks in international trade. *Review of Economics and Statistics*, 84(1):116–130, 2002.
- C. E. Shannon. A mathematical theory of communication. *Bell Technical Journal*, 3(27):379–423, 1948.
- C. A. Sims. Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665 – 690, 2003.

- C. A. Sims. Rational inattention: Beyond the linear-quadratic case. *American Economic Review*, 96(2):158–163, 2006.
- C. Steinwender. Information frictions and the law of one price: ”when the states and the kingdom became united”,. mimeo, london school of economics, 2014.
- A. Tutino. Rationally inattentive consumption choices. *Review of Economic Dynamics*, 16(32): 421–439, 2013.
- S. Van Nieuwerburgh and L. Veldkamp. Information immobility and the home bias puzzle. *The Journal of Finance*, 64(3):1187–1215, 2009.
- S. Van Nieuwerburgh and L. Veldkamp. Information acquisition and under-diversification. *The Review of Economic Studies*, 77(2):779–805, 2010.
- M. Woodford. Information-constrained state-dependent pricing. *Journal of Monetary Economics*, 56:S100–S124, 2009.

Appendix

A Proof of Proposition 1

This proof follows closely the steps in the proof of Theorem 1 in [Matejka and McKay \[2014\]](#). If $\lambda_j > 0$, then the Lagrangian of importers in country j is given by

$$\begin{aligned} \mathcal{L} = & \sum_{i=1}^N \int_{\tilde{Z}} u(\tilde{z}_{ij}) f_{ij}(\tilde{Z}) dG(\tilde{Z}) + \\ & - \lambda_j \left[- \sum_{i=1}^N \pi_{ij} \ln \pi_{ij} + \int_{\tilde{Z}} \left(\sum_{i=1}^N f_{ij}(\tilde{Z}) \ln f_{ij}(\tilde{Z}) \right) dG(\tilde{Z}) \right] + \\ & + \int_{\tilde{Z}} \xi_{ij}(\tilde{Z}) f_{ij}(\tilde{Z}) dG(\tilde{Z}) - \int_{\tilde{Z}} \mu(\tilde{Z}) \left(\sum_{i=1}^N f_{ij}(\tilde{Z}) - 1 \right) dG(\tilde{Z}) \end{aligned}$$

where $\xi_{ij}(\tilde{Z}) \geq 0$ and $\mu(\tilde{Z}) \geq 0$ are the Lagrange multipliers of equations (5) and (6) respectively. If $\pi_{ij} > 0$, the first order condition with respect to $f_{ij}(\tilde{Z})$ is given by

$$u(\tilde{z}_{ij}) + \xi_{ij}(\tilde{Z}) + \mu(\tilde{Z}) + \lambda_j \left(\ln \pi_{ij} + 1 - \ln f_{ij}(\tilde{Z}) - 1 \right) \quad (11)$$

As (5) does not bind, then the first order condition can be rearranged to

$$f_{ij}(\tilde{Z}) = \pi_{ij} e^{(u(\tilde{z}_{ij}) - \mu(\tilde{Z})) / \lambda_j} \quad (12)$$

Plugging (12) into (6), we obtain

$$e^{\mu(\tilde{Z}) / \lambda_j} = \sum_{i=1}^N \pi_{ij} e^{u(\tilde{z}_{ij}) / \lambda_j}$$

If we plug this expression back into (12), we get (7). Equation (7) holds even for $\pi_{ij} = 0$, as otherwise equation (2) would not hold.

B Proof of Proposition 2

This proof follows closely the steps in the proof of Lemma 1 and Proposition 3 in [Matejka and McKay \[2014\]](#). The following lemma will be helpful to prove Proposition 4.

Lemma 1. *The optimization problem of consumers in country j can be equivalently formulated as a maximization over the unconditional probabilities, $\{\pi_{ij}\}_{i=1}^N$:*

$$\max_{\{\pi_{ij}\}_{i=1}^N} \int_{\tilde{Z}} \lambda_j \ln \left(\sum_{i=1}^N \pi_{ij} e^{u(\tilde{z}_{ij})/\lambda_j} \right) dG(\tilde{Z}) \quad (13)$$

subject to (5) and (6).

Proof. Substituting equation (3) into the objective function, we get

$$\sum_{i=1}^N \int_{\tilde{Z}} u(\tilde{z}_{ij}) f_{ij}(\tilde{Z}) dG(\tilde{Z}) + \lambda_j \left[\sum_{i=1}^N \pi_{ij} \ln \pi_{ij} - \int_{\tilde{Z}} \left(\sum_{i=1}^N f_{ij}(\tilde{Z}) \ln f_{ij}(\tilde{Z}) \right) dG(\tilde{Z}) \right]$$

Rearranging this expression and using (7), we obtain

$$\begin{aligned} &= \int_{\tilde{Z}} \sum_{i=1}^N f_{ij}(\tilde{Z}) \left[u(\tilde{z}_{ij}) - \lambda_j \ln \left(\frac{\pi_{ij} e^{u(\tilde{z}_{ij})/\lambda_j}}{\sum_{k=1}^N \pi_{kj} e^{u(\tilde{z}_{kj})/\lambda_j}} \right) \right] dG(\tilde{Z}) + \lambda_j \sum_{i=1}^N \pi_{ij} \ln \pi_{ij} \\ &= \int_{\tilde{Z}} \sum_{i=1}^N f_{ij}(\tilde{Z}) \lambda_j \left[-\ln \pi_{ij} + \ln \left(\sum_{k=1}^N \pi_{kj} e^{u(\tilde{z}_{kj})/\lambda_j} \right) \right] dG(\tilde{Z}) + \lambda_j \sum_{i=1}^N \pi_{ij} \ln \pi_{ij} \\ &= \int_{\tilde{Z}} \sum_{i=1}^N f_{ij}(\tilde{Z}) \lambda_j \ln \left(\sum_{k=1}^N \pi_{kj} e^{u(\tilde{z}_{kj})/\lambda_j} \right) dG(\tilde{Z}) \\ &= \int_{\tilde{Z}} \lambda_j \ln \left(\sum_{k=1}^N \pi_{kj} e^{u(\tilde{z}_{kj})/\lambda_j} \right) dG(\tilde{Z}) \end{aligned}$$

□

B.1 Property 1

When we include the constraint (6) into the objective function (13), the optimization problem of consumers in country j described in Lemma 1 can be rewritten as

$$\max_{\{\pi_{ij}\}_{i=1}^N} \int_{\tilde{Z}} \lambda_j \ln \left[\sum_{i=1}^{N-1} \pi_{ij} e^{u(\tilde{z}_{ij})/\lambda_j} + \left(1 - \sum_{i=1}^{N-1} \pi_{ij} \right) e^{u(z_{Nj})/\lambda_j} \right] dG(\tilde{Z})$$

subject to (6). Let us focus on the case where the constraint (5) is not binding as the case where (5) binds is trivial. The gradient of the objective function with respect to π_{1j} is given by

$$\Delta_1 \equiv \lambda_j \int_{\tilde{Z}} \frac{e^{u(z_{1j})/\lambda_j} - e^{u(z_{Nj})/\lambda_j}}{\sum_{i=1}^N \pi_{ij} e^{u(\tilde{z}_{ij})/\lambda_j}} dG(\tilde{Z}). \quad (14)$$

where $\pi_{Nj} = 1 - \sum_{i=1}^{N-1} \pi_{ij}$. Differentiating with respect to either c_1 or τ_{1j} leads to $\frac{\partial \Delta_1}{\partial c_1} < 0$ or $\frac{\partial \Delta_1}{\partial \tau_{1j}} < 0$ respectively. This establishes that at the original optimum, an increase in either c_1 or τ_{1j} leads to a decrease of the gradient of the objective function with respect to the probability of the first option. Thus, consumers in country j will decrease π_{1j} .

B.2 Property 2

If countries are *ex-ante* identical, $A_i = A$; $c_i = c$; $\tau_{ij} = \tau$ for all i , then $G(\tilde{Z})$ is invariant to permutations of its arguments. Therefore, as showed by [Matejka and McKay \[2014\]](#), the solution for unconditional probabilities is unique and given by $\pi_{ij} = 1/N$ for all i .

B.3 Property 3

When $\lambda_j \rightarrow \infty$, importers process no information and decisions are based on *ex-ante* expectations. Given that all countries are identical except that home has lower trade costs, then home has the highest *ex-ante* expected $u(\tilde{z}_{ij})$ and $\pi_{jj} \rightarrow 1$.

C Proof of Corollary 1

If $\pi_{ij} > 0$, then the numerator $\pi_{ij} e^{u(\tilde{z}_{ij}/\lambda_j)} > 0$ and $f_{ij}(\tilde{Z}) > 0$.

D Proof of Proposition 3

The unconditional probability π_{ij} of importers in country j choosing to purchase the manufactured good from country i is given by (2), where the conditional probability $f_{ij}(\tilde{Z})$ in equation (7) can be rewritten as

$$f_{ij}(\tilde{Z}) = \frac{e^{(u(\tilde{z}_{ij}) + \lambda \ln \pi_{ij})/\lambda}}{\sum_{k=1}^N e^{(u(\tilde{z}_{kj}) + \lambda \ln \pi_{kj})/\lambda}}. \quad (15)$$

This conditional probability $f_{ij}(\tilde{Z})$ is a multinomial logit expression. There is no general solution to an integral such as in equation (2). However, an observationally equivalent model will allow us to find closed-form solutions of π_{ij} for a particular set of assumptions using well-known results in discrete choice theory. The logit formula (15) can be identically derived from a standard random utility model, where the importer in country j evaluates the utility of choosing country i with some noise ϵ_{ij} . Specifically, the two models are observationally equivalent if importers maximize the following random utility $\max_i u(\tilde{z}_{ij}) + \lambda \ln \pi_{ij} + \lambda \epsilon_{ij}$. When the noise ϵ_{ij} follows a Type I extreme value distribution and it is independently distributed, then the probability of choosing country i conditional on the realization of \tilde{z}_{ij} is given by (15) and the unconditional probability is given by (2).

If we set $u(\tilde{z}_{ij}) = \ln \tilde{z}_{ij} = \ln(A_i/c_i \tau_{ij}) + \ln \tilde{z}_i$, then the random utility maximized by importers is given by

$$\max_i \ln(A_i/c_i \tau_{ij}) + \lambda \ln \pi_{ij} + \ln \tilde{z}_i + \lambda \epsilon_{ij}.$$

The deterministic components of the random utility model are given by $\ln(A_i/c_i \tau_{ij}) + \lambda \ln \pi_{ij}$. The random utility model has two random components: first, ϵ_{ij} 's are independent and identically distributed according to a Type I extreme value distribution and second, $\ln \tilde{z}_i$'s are independent and identically distributed according to a $C(\lambda)$ distribution introduced in equation 8. According to Theorem 2.1 in Cardell [1997], if a random variable ε follows a Type I extreme value distribution, a random variable ν follows a $C(\lambda)$ distribution and both random variables are independent, then $\nu + \lambda \varepsilon$ is distributed as Type I extreme value for $0 < \lambda < 1$. Therefore, following Cardell's theorem, the addition of the two random components in the random utility model $\ln \tilde{z}_i + \lambda \epsilon_{ij}$ is distributed as a Type I extreme value. Hence, the unconditional choice probabilities of the random utility model can be written as

$$\pi_{ij} = \frac{e^{\ln(A_i/c_i \tau_{ij}) + \lambda \ln \pi_{ij}}}{\sum_{k=1}^N e^{\ln(A_k/c_k \tau_{kj}) + \lambda \ln \pi_{kj}}} \quad (16)$$

After solving the system of equations, the unconditional choice probabilities are given by (9).

E Proof of Proposition 4

E.1 Property 1

$|\frac{\partial \ln \pi_{ij}}{\partial \ln \tau_{ij}}| = \frac{1}{1-\lambda}$. The right-hand side is increasing in λ .

E.2 Property 2

Re-write π_{ij} as

$$\pi_{ij} = \frac{1}{1 + \sum_{h \neq i} \left(\frac{\alpha_{hj}}{\alpha_{ij}} \right)^{\frac{1}{1-\lambda}}}.$$

where $\alpha_{ij} = A_i / (c_i \tau_{ij})$. Differentiating with respect to λ ,

$$\frac{\partial \pi_{ij}}{\partial \lambda} = -\frac{\pi_{ij}^2}{(1-\lambda)^2} \sum_{h \neq i} \left(\frac{\alpha_{hj}}{\alpha_{ij}} \right)^{\frac{1}{1-\lambda}} \ln \left(\frac{\alpha_{hj}}{\alpha_{ij}} \right).$$

Re-call that the countries are ordered such that α_{1j} is the maximum and α_{Nj} is the minimum. When $i = 1$, the last term on the right-hand side in the above equation (involving the logs) is negative for all h . Accordingly, $\frac{\partial \pi_{1j}}{\partial \lambda} > 0$. Similarly, when $i = N$, the same term is positive for all h and $\frac{\partial \pi_{Nj}}{\partial \lambda} < 0$.

Let λ_{ij} be the value of λ such that $\frac{\partial \pi_{ij}}{\partial \lambda} = 0$. Then λ_{ij} must satisfy

$$\sum_{h=1}^{i-1} \left(\frac{\alpha_{hj}}{\alpha_{ij}} \right)^{\frac{1}{1-\lambda_{ij}}} \ln \left(\frac{\alpha_{hj}}{\alpha_{ij}} \right) + \sum_{h=i+1}^N \left(\frac{\alpha_{hj}}{\alpha_{ij}} \right)^{\frac{1}{1-\lambda_{ij}}} \ln \left(\frac{\alpha_{hj}}{\alpha_{ij}} \right) = 0.$$

When $\lambda < \lambda_{ij}$, $\left(\frac{\alpha_{hj}}{\alpha_{ij}} \right)^{\frac{1}{1-\lambda}}$ goes up for $h < i$. But these are the weights on the positive terms $\ln \left(\frac{\alpha_{hj}}{\alpha_{ij}} \right)$. At the same time, $\left(\frac{\alpha_{hj}}{\alpha_{ij}} \right)^{\frac{1}{1-\lambda}}$ goes down for $h > i$. But these are the weights on the negative terms $\ln \left(\frac{\alpha_{hj}}{\alpha_{ij}} \right)$. Accordingly, $\frac{\partial \pi_{ij}}{\partial \lambda}$ becomes positive. Similarly, it can be shown that $\frac{\partial \pi_{ij}}{\partial \lambda}$ becomes negative for $\lambda > \lambda_{ij}$. Therefore, π_{ij} is hump-shaped for all i except $i = 1$ and $i = N$.

E.3 Property 3

Consider the above equation. A little bit of algebra is required to show that when this equation is satisfied, $\frac{\partial \lambda_{ij}}{\partial \alpha_{ij}} > 0$. Because α_{ij} is decreasing in i , λ_{ij} must be decreasing in i .

F Proof of Proposition 5

The case for $\lambda_j = 0$ is trivial. When $\lambda_j = 0$, importers have perfect information and choose to import from the country with the lowest price with probability 1.