University of Toronto
Department of Economics

Working Paper 566

Inattentive Importers

By Kunal Dasgupta and Jordi Mondria

October 17, 2016

Inattentive Importers *

Kunal Dasgupta — University of Toronto

Jordi Mondria — University of Toronto

SEPTEMBER 2016

Abstract

Information frictions prevent importers from observing the price of a good in every market. 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. The amount of information importers process about each country is endogenous and reacts to changes in observable trade costs. Unlike traditional trade costs, changes in information processing costs have non-monotonic and asymmetric effects on bilateral trade flows. We go on to show quantitatively how small differences in distance generate large differences in trade flows, thereby shedding light on the distance elasticity puzzle. The model also generates a novel prediction regarding the relationship between information processing costs and the concentration of import distributions that finds support in the data.

KEYWORDS : Rational inattention, incomplete information, distance elasticity.

JEL Classification : D83, F10, F19, L15.

*We would like to thank participants at various conferences and workshops, and especially Victor Aguirregabiria, Treb Allen, Thomas Chaney and Peter Morrow for helpful discussions. Kunal Dasgupta and Jordi Mondria gratefully acknowledge financial support from the Social Sciences and Humanities Research Council of Canada (SSHRC).

1 Introduction

Incomplete information plagues international commerce. Importers rarely observe the price and attributes of a good in every market. These informational barriers are bound to have an impact on the flow of goods between countries. Yet, despite a widespread agreement among economists that incomplete information could create significant barriers to trade, we lack a framework that formalizes the link between information and trade.[1]

In this paper, we attempt to develop such a framework. Our paper makes three main contributions. First, we formally derive a relationship between import shares and the costs of processing information. A characteristic feature of our framework is that import shares not only respond directly to any change in model parameters, but also indirectly through a change in information processed. Second, we show quantitatively how, in the presence of information processing costs, small increases in distance get translated into large increases in trade flows. In the process, we provide one possible explanation for the distance elasticity puzzle. Third, we provide evidence that the import distribution for a product is, on average, more concentrated in countries that seem to have lower costs of processing information. We argue that this finding is consistent with our model of inattention, but is not predicted by standard full information models of trade.

We introduce rational inattention [Sims, 2003, 2006] into a multi-country, Ricardian model of trade. Every period, producers draw productivity stochastically. Importers would like to import a product from the country that has the lowest price. But importers have a limited capacity to process information about prices. Faced with a capacity constraint, importers must decide how much information to process about prices in each country. More information increases the precision of the noisy signals received by the importers, but comes at a higher cost. The rational importer weighs the marginal benefit of an extra unit of information against the marginal cost.

A key insight of our model is that the endogenous processing of information affects the response of trade flows to a change in observable trade costs between trading partners. When a trade cost, such as transport cost, between importing country j and exporting country i declines, country j importers start to purchase more from country i because the expected 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

[1] In their survey on trade costs, Anderson and van Wincoop [2004] highlight the need for more careful modelling of information frictions.

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 can have large effects on trade flows – there is a *magnification effect*.

Following Matejka and McKay [2014], we show that the optimal solution of a rationally inattentive importer is to choose probabilistically the country from where to buy a given product, with this probability distribution following an adjusted multinomial logit. Our paper shares the prediction with the full information model of Eaton and Kortum [2002] that the prior probability that country j imports a product from country i is positive for every i. Unlike Eaton and Kortum however, the corresponding posterior 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.

The key parameter in our model is the cost of processing information. We show that, unlike traditional trade costs, information cost may have non-monotonic effects on bilateral trade flows as the share of imports first rises but eventually declines when information cost increases. We also show that, unlike traditional trade costs, information cost may have asymmetric effects on bilateral trade flows. An increase in information cost may lead importers to choose to process more information about countries that have lower expected price,[2] 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 has imposed import tariffs that are higher for countries that have higher expected price. A uniform increase in standard trade costs can not generate such an outcome.

A long-standing puzzle in the trade literature is the distance elasticity puzzle – the finding that the measured elasticity of trade with respect to distance cannot be reconciled with estimates of structural parameters that usually govern this elasticity. In particular, estimates of the responsiveness of trade costs to distance suggest that a 1 percent increase in distance should lead to a 0.1 percent decline in trade, which is an order of magnitude smaller than the 1 percent decline in trade that is typically observed in the data [Disdier and Head, 2008]. In the second part of the paper, we show how our model of inattention may quantitatively resolve the distance elasticity puzzle.[3]

We solve a multi-country model with rationally inattentive importers. The unknown parameter in the model is the cost of processing information, which we calibrate to match the average share

[2] A lower expected price could arise either due to lower bilateral trade cost or higher average productivity.
[3] An alternative explanation for why the distance elasticity is large is provided by Krautheim [2012].

of own imports. To quantitatively evaluate the ability of our model to rationalize observed trade flows, we then run gravity regressions on the simulated data. In the benchmark scenario, we choose parameter values such that with zero information costs, the distance elasticity is only 0.1. With positive information cost, the distance elasticity in our preferred specification is an order of magnitude larger. Our results suggest that information frictions could be one possible reason why distance seems to have such a large effect on trade.[4]

Our model generates a novel prediction linking the distribution of imports with the cost of processing information. In the absence of information cost, importers purchase from one country only – the country offering the lowest price. Accordingly, the import distribution is degenerate. On the other hand, when information cost is infinitely high, importers again purchase from one country only – the country with the lowest *expected* price. For intermediate values of the information cost, importers diversify. In fact, we show that the concentration of the import distribution for a given product and importing country exhibits a U shape with respect to information cost. In the final part of the paper, we show that countries that have lower cost of processing information, proxied by greater internet usage, also tend to have more concentrated import distributions. This relationship is robust to a number of controls that could potentially explain our results. We also note that none of the standard models of trade generate a systematic relation between the concentration of imports on the one hand, and importing country characteristics on the other. Of course, one could always argue that the simultaneous variation in import concentration and information cost proxies in the data are being driven by a third omitted factor. To address this concern, we exploit the variation in goods in terms of their information intensity. In particular, we show that the relation between import concentration and information costs is stronger for products that are more information-intensive.

We are the first, to the best of our knowledge, to apply rational inattention to the study of international trade. Our decision to model information as a theory of attention allocation is guided by the following consideration: attention is a major area of investigation in education, psychology and neuroscience, and its influence is growing in economics and finance. As suggested by Kahneman [1973], the human mind is bounded by cognitive limits and even if individuals had access to full information, their mind would be unable to process all the available information. Individuals would then have to choose where to allocate their limited cognitive attention resources to process information when making decisions. Hence, selectively focusing more cognitive resources to one option would result in a decrease of cognitive attention to alternative options. In the context of

[4] That distance proxies for information frictions has also been suggested by work, among others, on FDI [Head and Ries, 2008], asset trade [Portes et al., 2001, Portes and Rey, 2005], e-bay transactions [Hortaçsu et al., 2009, Lendle et al., 2012] and patent citations [Jaffe et al., 1993, Peri, 2005].

international trade, a consequence of rational inattention is that unlike most papers that deal with information frictions, importers in our model choose to process different amounts of information about prices in different source countries.

In one of the first papers to highlight the role of information frictions in shaping international trade flows, 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. 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 – a fact suggestive of the presence of informational barriers. Drawing an analogy with astrophysics, Head and Mayer [2013] point out that at most 30 percent of the variation in trade flows can be explained by observable freight costs, while the remaining 70 percent of the variation is due to a "dark" trade cost. The authors argue that one significant component of these dark costs must be information costs.

Two recent papers have provided further evidence of informational barriers in goods trade. Looking at the market for agricultural goods in Philippines, Allen [2014] demonstrated that a number of features of the data can be explained by a model with information frictions, but are not consistent with a full information model. 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 for cotton. [5]

The paper that is closest in spirit to our paper is Allen [2014]. Unlike our static model though, he considers a model where producers sequentially search for the highest price across markets. In Allen's paper, information frictions manifest in the form of (i) a fixed cost that producers have to pay to learn about prices in each market, and (ii) an exogenous probability of searching each market. Allen [2014] goes on to show that the probability that producers in market j will search market i depends on a number of bilateral variables, the most important being distance. Our model suggests why this might be the case. *Ceteris paribus*, rationally inattentive importers process more information about markets that are close, or in other words, markets with low expected prices. This, in turn, makes it more likely that there will be a transaction between two markets, over and above what can be explained by pure transport costs.

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

[5] Other papers that provide evidence that is consistent with the presence of information frictions in trade include Gould [1994], Head and Ries [1998], Rauch and Trindade [2002], Freund and Weinhold [2004], Fink et al. [2005], Combes et al. [2005] and Chan [2016].

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, a feature that generates many of the novel results in our paper.

The rest of the paper is organized as follows. In Section 2, we specify the production structure and introduce inattentive importers into a Ricardian trade model. In Section 3, we implicitly solve for the equilibrium trade shares and discuss some of their novel properties. In Section 4, we quantitatively evaluate the role of information processing cost. In Section 5, we examine some empirical implications of our model. Section 6 concludes. All the proofs are in the Appendix.

2 The Model

We consider a world with N countries. Each country is populated with a positive measure of workers who also consume. Countries produce a continuum of symmetric products using labour and intermediate inputs.

Preference. The indirect utility function V is given by:

$$V = \int_{\omega} \log \left[\frac{1}{p(\omega)}\right] d\omega + \log Y, \tag{1}$$

where Y and $p(\omega)$ refer to total income and the price of product ω respectively. Without loss of generality, we assume that there is a (fixed) measure one of products.

A convenient feature of the indirect utility function in (1) is that the share of expenditure on each product ω is exogenously fixed. Accordingly, a consumer simply spends a fixed fraction of her income on a product, and the price of that product determines utility. In a full information world, this amounts to the consumer purchasing each product from only one country, the country that offers the lowest price for that product. That is our point of departure in this paper.

Technology. The markets for the different products are perfectly competitive. Instead of defining the production function of a product, we consider its dual, the cost function. The cost of importing one unit of product ω into country j from country i is given by $1/\tilde{z}_{ij}(\omega)$, with

$$\tilde{z}_{ij}(\omega) = \frac{z_i(\omega)}{c_i \tau_{ij}}, \tag{2}$$

where c_i is the average cost of a standardized bundle of inputs required for producing one unit of any product in country i. For now, we take c_i as given, but endogenize it in Section 4. The observable trade cost between exporting country i and importing country j is captured by the iceberg cost τ_{ij}, 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. Finally, $z_i(\omega)$ is a random productivity draw for product ω in country i. We assume that $\ln[z_i(\omega)]$ is drawn jointly for each country i and good ω from a mean zero distribution. Any deterministic determinant of productivity is included in c_i.

Information frictions. We introduce information frictions by assuming that the productivity realizations, $z_i(\omega)$, are not perfectly observable at the decision stage. We also assume that the cost of product ω produced in country i is fully revealed to consumers (henceforth importers) in country j once country i has been chosen to supply the product. This assumption of perfect observability *ex-post*, combined with perfect competition in the market for each variety, implies that the producers in any country do not engage in strategic price setting.[6] The price at which producers in country i are willing to sell product ω to importers in country j is then given by

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

i.e., producers choose to sell their goods at marginal cost. It must be emphasized that $p_{ij}(\omega)$ 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 $p_{ij}(\omega)$ may not be the lowest price for product ω faced by the importers in country j.

It is worth pointing out two observations. First, once an importer in country j chooses to purchase from country i, the transaction always takes place. This is because the preferences in (1) imply that importers always want to purchase a positive quantity.[7] Second, we do not make an ad-hoc assumption that importers have more information about the productivity draws in their own country relative to foreign countries. Rather, as we shall see below, this scenario may arise as an equilibrium outcome.

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

[7] We implicitly rule out the possibility that within a given period, the importer can choose a country other than i in the event that the price in i is revealed to be too high. This assumption is not as restrictive as it seems. The final price paid by importers could involve rounds of negotiations. If these negotiations take time and the importer has to purchase within a given period, the importer may not want to switch suppliers even if the realized price turns out to be much higher than expected.

Rational inattention. Importers are rationally inattentive. They choose to learn about the random productivity draws for each product in every country, taking into account that information is costly to acquire. The innovation provided by rational inattention theory is that importers are not constrained to learn about these productivity draws with a particular signal structure; rather, they 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 in country j choose the probability that a product ω is purchased from country i, conditional on the productivity realizations.

Following Sims [2003], we use tools from information theory to model the limited information processing capabilities of importers. At this point, we define two mathematical objects that form an integral part of our analysis.

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

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

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

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

where $H(X|Y) = -\sum_x p(x|y) \ln p(x|y)$ is the entropy of X conditional on Y.

Intuitively, mutual information measures the reduction in the entropy of X caused by the knowledge of Y. We use entropy as the measure of uncertainty about the productivity draws and mutual information as the measure of uncertainty reduction or information [Shannon, 1948]. The following property of mutual information will be useful later on:

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

Importer's problem. For each product ω, a positive measure of importers in country j choose the source country for the product. Because the products are symmetric, we can simply focus on a representative product. Let us define $f_{ij}(Z)$ as the posterior probability that country j importers purchase a variety from country i conditional on the productivity draws for that product across countries, Z. Defining π_{ij} as the expected probability that country j importers buy the product from country i, we have

$$\pi_{ij} = \int_Z f_{ij}(Z) dG(Z), \tag{3}$$

where $G(Z)$ is the distribution of Z across products. Note that because there is a continuum of products, π_{ij} is also the prior or *unconditional* probability that country j importers purchase any product from country i.

Importers in country j process information about Z to reduce the entropy $H(Z)$. As already discussed, an influential property of rational inattention is that there is no need to explicitly model the signal structure that importers choose to learn about Z. Rather, the information processed by importers in country j can be calculated as the mutual information between productivities Z (variable of interest) and the country i chosen by importers in country j (action):

$$\kappa_j = H(Z) - E\big[H\big(Z|j(i)\big)\big],$$

where $H\big(Z|j(i)\big)$ is the entropy of Z, conditional on country j importers purchasing a product from country i. When an importer purchases a product from a particular country, that action reduces their uncertainty about the productivity draws. The resulting difference in uncertainty is the information that country j importers have about productivity draws across the world.

If information could be processed freely, an importer would find out the true realization of Z. There are, however, a multitude of costs involved in processing information about the true productivity of a supplier, which are incurred by the destination country. We denote all these costs by λ_j. Hence, by paying a cost $\lambda_j \kappa_j$, country j importers can reduce their uncertainty about the realization of Z by κ_j.[8]

Given the additive preference structure, importers will maximize the expected utility of each variety, taking into account the information processed about the productivity draws. That is, importers in country j solve the following optimization problem:

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

[8] Note that λ_j is a parameter while κ_j is a variable.

subject to

$$\kappa_j = -\sum_{i=1}^{N} \pi_{ij} \ln \pi_{ij} + \int_Z \Big(\sum_{i=1}^{N} f_{ij}(Z) \ln f_{ij}(Z) \Big) dG(Z), \tag{4}$$

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

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

where $\tilde{z}_{ij}$ is given by (2). The first term in the objective function is the expected utility of importers from purchasing a product, while the second term is the cost of processing information. Rationally inattentive importers in country j choose the probability of importing from country i conditional on the realization of Z. In deriving the amount of information processed in (4), we have used Property 1. Equations (5) and (6) simply say that $f_{ij}(Z)$ must be a probability mass function.

3 Equilibrium.

As shown by Matejka and McKay [2014], the above optimization problem is equivalent to solving the following two-stage optimization: In the first stage, importers choose to observe signals to reduce their uncertainty about the productivity draws in each country. In the second stage, given the information provided by the signals, importers choose to buy the variety from the country offering the lowest expected price. Following Matejka and McKay, the next proposition derives the equilibrium posterior probability of purchasing a given variety:

Proposition 1. *If $\lambda_j > 0$, then conditional on the realization of Z, the probability that importers in country j choose to purchase a product from country i is given by*

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

where π_{ij} is given by (3).

The posterior choice probabilities have a structure similar to a multinomial logit [McFadden, 1989], except that they are adjusted by the prior probabilities, π_{ij}. These π_{ij}-s are independent of productivity realizations of individual products and only depend on exogenous objects such as the productivity distribution, informations costs, preferences, and input costs. When the cost of information is high, posterior choice probabilities attach a high weight to prior 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 as the source for a product even if its actual productivity in that product is low. When the cost of information is low, the posterior choice probabilities attach a high weight to the actual productivity realizations, Z, as importers process large amounts of information and receive signals about Z that are much more precise. The following proposition discusses an important property of $f_{ij}(Z)$.

Proposition 2. *Suppose* $\pi_{ij} > 0$.
(i) If $\lambda_j > 0$*, then* $f_{ij}(Z) > 0$.
(ii) If $\lambda_j \to 0$*, then* $f_{ij}(Z) \to 1$ *if* $\tilde{z}_{ij} = max \tilde{z}_{kj} \forall k$ *and* $f_{ij}(Z) \to 0$ *otherwise.*

An implication of the above proposition is that importers in one country could buy the same product from different countries.[9] Notice that Proposition 1 contrasts sharply with the result in Eaton and Kortum [2002]. In that paper, even though *a priori* importers in country j can buy a given product from any country, after the productivity draws are realized, this probability drops to zero for every exporting country but the one with the lowest price. In fact, as Proposition 1 shows, as the cost of information becomes negligible and our model converges to a full information model, the conditional probabilities converge to either zero or one. But as long as there are positive information costs, this is not true any more. Importers never observe the true productivity draws and believe that every country can have the cheapest product with some probability.

In the literature, when a narrowly defined product is imported from many countries, it is usually assumed that different countries produce different varieties of the same product [Klenow and Rodriguez-Clare, 1997]. In our model, the exact same product could still be imported from multiple countries because of information frictions. Furthermore, if the prior probabilities that country j both imports from as well as exports to country i are positive, then so are the posterior probabilities. Hence, in equilibrium, we could observe the same product being traded in both directions by two countries. This feature, which is shared by Allen [2014], can never be generated in a full information model of trade. The next proposition discusses some properties of the prior probability π_{ij}.

Proposition 3. π_{ij} *has the following properties:*

[9] All importers in a given country have the same initial beliefs about which source country has the lowest price for a product. Their actions, however, may be heterogeneous. If $f_{ij}(Z) > 0$ and $f_{hj}(Z) > 0$, then a fraction $f_{ij}(Z)$ of importers in country j will choose to purchase the product from country i, while a fraction $f_{hj}(Z)$ of importers will choose to import from country h. Intuitively, even though all importers in a country choose the same signal structure, different importers could receive different signals about the productivity draws at a given point in time, and could end up buying from different countries based on these signals.

1. π_{ij} *is decreasing in input costs* c_i *and trade costs* τ_{ij}.

2. *If there exists country* i *such that* $c_i\tau_{ij} = min c_k\tau_{kj}\forall k$*, then as* $\lambda_j \rightarrow \infty$*,* $\pi_{kj} \rightarrow 0$ *for all* $k \neq i$ *and* $\pi_{ij} \rightarrow 1$.

3. *If trade is frictionless and countries are ex-ante identical, i.e.* $\tau_{ij} = 1, c_i = c$ *for all* i*,* $\pi_{ij} = 1/N$ *for all* i.

The first property of Proposition 3 states that *ex-ante*, importers in country j are less likely to purchase a product from countries with a high expected cost. Holding everything else constant, an increase in c_i or τ_{ij} reduces the probability that the price of that product in country i is the lowest price among all countries.

The second property of Proposition 3 demonstrates that all else equal, if the information processing cost becomes extremely large, importers tend to purchase from only one country. Intuitively, when information processing cost is high, importers incorporate less information into their decision making and attach a greater weight to the primitives, c_i and τ_{ij}. If the expected cost of importing from country i is the lowest, then an increase in importance of the primitives raises the likelihood that country j importers will buy from country i.

The third property of Proposition 3 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 as the source for a variety by importers in country j. In this case, the choice probabilities $f_{ij}(Z)$ in equation (7) follow a standard multinomial logit.

It is straightforward to show that $f_{ij}(Z)$ is also country j's share of expenditure on a particular variety imported from i. Given this observation, a novel property of our model concerns the effect of information costs on the concentration of the import distribution for a variety. To see this, consider a world with complete information ($\lambda_j = 0$). In such a world, importers in country j buy a product almost surely from just one country, as shown in Proposition 2. In this case, the distribution of $f_{ij}(Z)$ for that variety will be degenerate. But as information cost rises, the true productivity realizations are not observed any more. Accordingly, importers diversify their purchases, causing the import distribution to become non-degenerate. At the same time, part 2 of Proposition 3 shows that importers in country j buy a product almost surely from just one country if they face arbitrarily high costs of processing information ($\lambda \rightarrow \infty$). We state this result formally in the next proposition:

Proposition 4. *Starting from zero information cost, an increase in information cost causes the distribution of imports of a given product to become less concentrated. Starting from an infinitely* *high information cost, a reduction in information cost causes the distribution of imports of a given product to become less concentrated.*

Note that while deriving Propositions 1, 2, 3 and 4, we did not specify a distribution of productivity $G(Z)$. In particular, these results are satisfied for any $G(Z)$. None of the known family of distributions, however, permit analytical solutions for the π_{ij}-s and the f_{ij}-s as there is no solution for the integral in (3). Therefore, we use numerical integration to derive more comparative statics results.

Simulation. For this exercise, we assume there are four countries, indexed by 1,...,4, that have identical input costs (i.e., we set $c_i = 1$ for all i). We order countries by their cost of exporting to country 1, τ_{i1}, and assume that $\tau_{11} < \ldots\ldots < \tau_{41}$. Finally, we assume that the log productivities are drawn independently from a normal distribution with mean zero and standard deviation σ. We draw a vector of productivities for the four countries one hundred thousand times (corresponding to one hundred thousand products).[10] The next lemma establishes a useful property of π_{ij} that we shall be exploiting below.

Lemma 1. *π_{ij} equals country j's share of expenditure on goods imported from country i.*

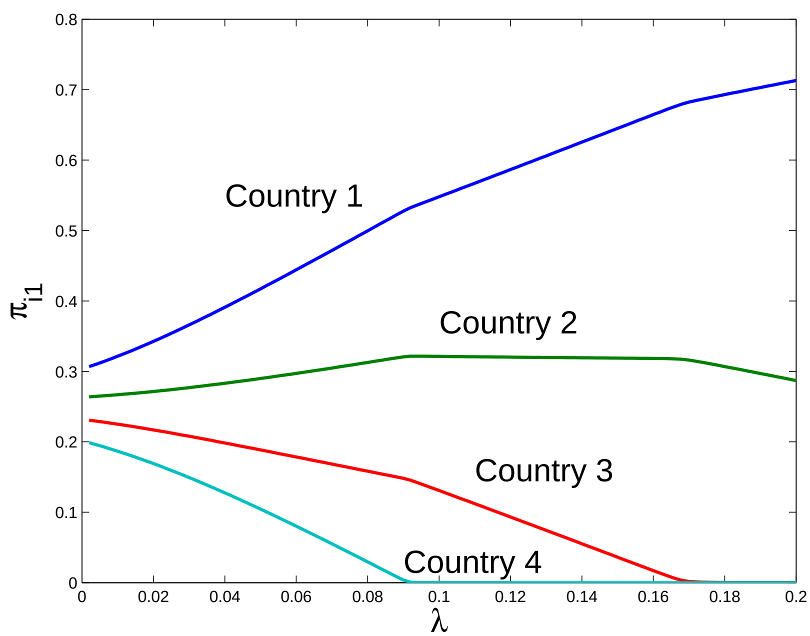


Figure 1: How π_{i1} varies with information cost

[10] In particular, we assume that $\tau_{11} = 1.000$, $\tau_{21} = 1.005$, $\tau_{31} = 1.010$, $\tau_{41} = 1.015$ and $\sigma = 0.05$.

Figure 1 shows the prior probabilities π_{i1} ($i = 1, ..., 4$), which are the same as trade shares, for different levels of information processing cost λ. As Figure 1 illustrates, in the presence of information processing costs, as traditional trade costs increase, trade declines much more than what would be predicted in a full information world. In a full information model ($\lambda \to 0$) such as Eaton and Kortum [2002], when traditional trade costs increase, the expected price of country i products decline. In our model when $\lambda > 0$, 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 the lowest price in country i declines, so does the information processed by country j importers about country i. Consequently, π_{ij} drops even more – the presence of information costs creates a *magnification effect*.

Figure 1 also sheds light on two properties of the model that highlight novel insights from rational inattention theory – *asymmetry* and *non-monotonicity* of the π_{ij}-s. Unlike traditional trade costs, a change in information cost has an asymmetric effect on trade shares. If λ increases, the import shares from countries other than 1 do not necessarily decline. Rather, when λ is small, an increase in λ actually leads to an increase in π_{21}. Intuitively, when information costs increase, importers in country 1 reallocate attention to countries with lower expected costs, to the detriment of other countries. Thus, an increase in λ leads to an increase in the attention allocated to countries 1 and 2, but a reduction of attention to countries 3 and 4, resulting in an increase in π_{11} and π_{21}, and a decrease in π_{31} and π_{41}. It is *as if* country 1 imposed differential import tariffs on goods imported from the other countries, with the tariff being higher for the country that is farther away. Hence, information cost has asymmetric effects on bilateral trade shares as they may increase the share of imports from countries with low expected costs and decrease the share of imports from countries with high expected costs.

Figure 1 also shows that the probability of country 1 importers buying a variety from country 2 displays a hump-shaped behaviour with respect to information costs. This contrasts with the response of import shares to a change in standard trade costs, as stated in Proposition 3, where increases in input costs c_i and trade costs τ_{ij} have monotonic effects. As discussed above, when there is an increase in information costs starting from low levels, importers in country 1 reduce the total amount of information processed and substitute their attention from countries 3 and 4 (countries with high trade costs) to countries 1 and 2. But for high enough information costs, country 1 importers re-allocate attention from country 2 to 1, resulting in a decline in imports from country 2. Hence, the effect of information costs on trade shares from country 2 is non-monotonic.

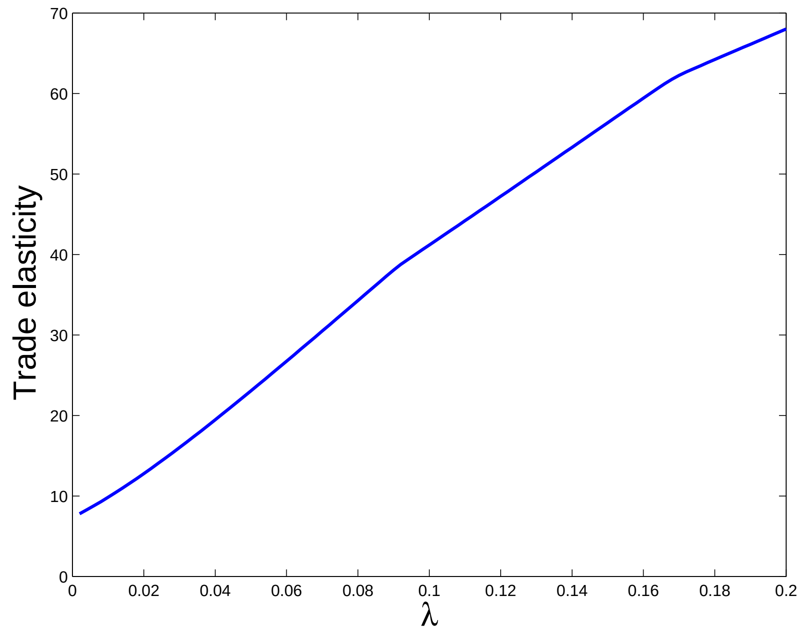


Figure 2: How trade elasticity varies with information costs

Figure 2 plots the trade elasticity against the information processing cost.[11] It shows the trade elasticity increasing in information costs, as suggested by Figure 1. When information costs are high, importers in country 1 optimally allocate more attention to countries with lower trade costs, resulting in disproportionately more trade with those countries. Small trade costs impose heavy penalties on countries that are ex-ante not very attractive sources for a product.

Figure 2 suggests that if our model had different types of products with product-specific information processing costs, then products with high λ would have a higher trade elasticity than those with low λ. If we assume that differentiated products have higher λ than reference-priced products, then our model is consistent with the findings by Rauch [1999], where he showed that the elasticity of trade with respect to distance is higher for differentiated goods relative to reference-priced goods.[12] Rauch conjectured that the cost of learning about differentiated products is higher relative to reference-priced products as the former have multiple attributes and might require search and matching. In other words, the cost of processing information about differentiated products might be higher.

Our model also provides a possible explanation for the distance elasticity puzzle. This puzzle refers to the issue that the elasticity of trade costs with respect to distance is much smaller than

[11] Trade elasticity is measured as $d\ln\pi_{ij}/d\ln\tau_{ij}$.

[12] Reference-priced products are those that are not transacted in centralized exchanges, but whose prices are published in trade journals.

what is needed to explain trade data using traditional models. 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. Although our model may qualitatively resolve the distance elasticity puzzle, the bigger question remains: how much can it actually explain? In Section 4, we attempt to provide an answer.

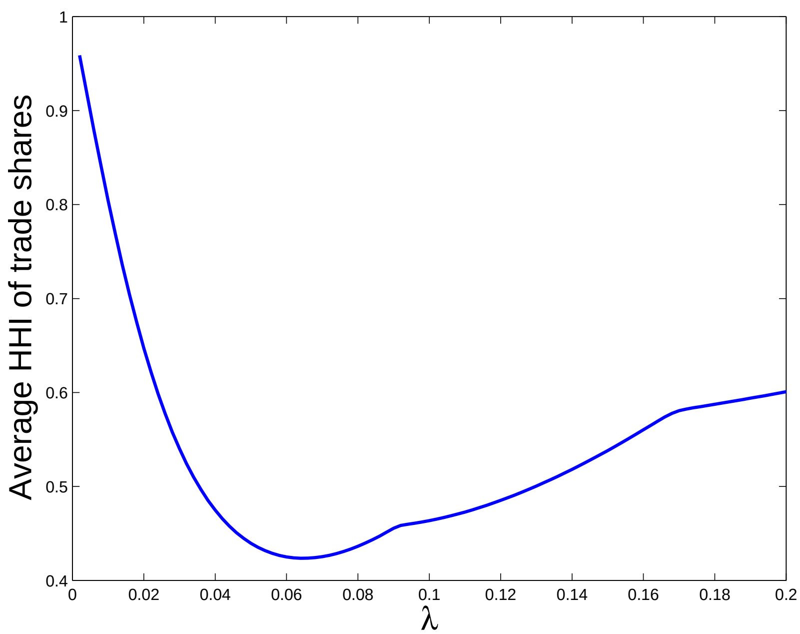


Figure 3: Import concentration as a function of information cost

Recall that in deriving a theoretical relation between information processing cost λ and the concentration of import distribution, we could characterize only the end-points ($\lambda = 0$ and $\lambda = \infty$). Numerically, we can characterize import concentration over the entire range of λ. Figure 3 shows the average import concentration for country 1, measured by the Herfindahl-Hirschman index (HHI), plotted against λ.[13] The figure shows that as λ increases, the import distribution of a product tends to become less concentrated initially, before its concentration starts to rise.

To summarize, a model with rationally inattentive importers magnifies the effect of traditional trade barriers on trade. A change in a trade barrier such as transport cost not only has a direct effect on trade volumes, but by changing how importers process information, has an indirect effect

[13]For a variable taking T distinct values with the corresponding shares being s_t, ($\sum_{t=1}^{T} s_t = 1$), the Herfindahl-Hirschman index is given by $HHI = \sum_{t=1}^{T} s_t^2$. The HHI lies between 0 and 1, with 1 corresponding to a de-generate (completely concentrated) distribution.

too. In the following section, we calibrate a multi-country model to evaluate the importance of information costs in facilitating cross-country trade.

4 Quantitative exercise

In this section, we quantitatively solve a multi-country, general equilibrium model with inattentive importers. The purpose is twofold: First, it provides us with a measure of information frictions. Second, it allows us to perform comparative static exercises with respect to information costs, thereby shedding light on how global trade is shaped by such frictions. To do so, first we need to close the model.

We assume that each differentiated product is manufactured using labour and an intermediate input that is a composite of all the available products. This leads to the following cost function for a product in country i:

$$c_i = w_i^{\alpha} P_i^{1-\alpha},$$

where w_i is the nominal wage in country i and P_i, the input price index, is given by

$$log(P_i) = \int \Big(\sum_{j=1}^{N} f_{ji}(Z) p_{ji}(Z) \Big) dZ. \tag{8}$$

Because each product is purchased from multiple countries, the term within the parenthesis in the above equation is the expected price for a product conditional on the productivity draw Z. The (logarithm of) aggregate price index is the integral of the expected prices across products. Because labour is the only factor of production, $w_i L_i$ is the aggregate income of country i.

To close the model, we impose a balanced trade condition. This requires that for each country i, the value of total imports must equal the value of total exports:

$$\sum_{j=1}^{N} \pi_{ji} w_i L_i = \sum_{j=1}^{N} \pi_{ij} w_j L_j,$$

where L_i is the size of the labour force in country i.[14] Noting that $\sum_{i=1}^{N} \pi_{ji} = 1$, the balanced trade equation can re-written as

$$w_i = \sum_{j=1}^{N} \pi_{ij} w_j L_j / L_i. \tag{9}$$

[14]Because π_{ji} is the share of income spent by country i on country j's products, $\pi_{ji} w_i L_i$ is the total expenditure by country i on country j's products. Summing over j gives us total import of i.

Given the prior probabilities and labour endowments, equation (9) is a system of equations in wages. We are finally in a position to define an equilibrium of the model:

Definition 1. *A competitive equilibrium of the model is a set of posterior probabilities $f_{ij}(Z)$, prior probabilities π_{ij}, aggregate prices P_i, and wages w_i, such that $f_{ij}(Z)$ and π_{ij} are given by (7) and (3) respectively, P_i satisfies (8) and w_i solves (9).*

The solution strategy for this model is as follows: First, for a given vector of w_i, P_i and π_i, use (7) to solve for the f_{ij}-s. Second, use (7) to check whether the resulting π_{ij} is the same as the initial guess and iterate till it is. Third, use (8) to check whether the resulting P_i is the same as the initial guess and iterate till it is. And finally, use (9) to solve for w_i and compare with the initial guess. Therefore, solving for the equilibrium involves solving a series of fixed-point problems.

Because of the high-dimensionality of the model, computational time increases exponentially with the number of countries and products. In what follows, we choose $N = 25$ and $K = 100,000$. We picked 25 countries that accounted for more than 80 percent of all imports in 1990.[15] The products are assumed to be symmetric.

4.1 Calibration

To quantify the model, we need to choose a distribution of productivities $G(Z)$ and parameter values for α (the share of value-added in production), τ_{ij} (trade costs) and λ_j (information cost). Following Eaton and Kortum [2002], we assume that $G(Z)$ follows a Fréchet distribution, with a shape parameter θ. Simonovska and Waugh [2014] recently argued that $\theta = 4$. For the benchmark model, we use their preferred value for θ. We assume that the scale parameter for the Fréchet distribution is country-specific. We use total factor productivity (TFP), obtained from the Penn-World Tables, as a proxy for these parameters.

We assume that α, the labour share of output, equals 0.3 [Waugh, 2010]. We also use an extremely parsimonious specification for observable trade costs τ_{ij}. In particular, we assume that τ_{ij} depends only on distance, with the elasticity of τ_{ij} with respect to distance being 0.025 [Limao and Venables, 2001, Hummels, 2007]. For distance, we use the great circle distance between capital cities, obtained from CEPII. Thus, we ignore all other geographic barriers to trade commonly used in the gravity literature such as borders and contiguity, cultural barriers proxied by common language and common colonial origin, as well as policy barriers such as tariffs and NTBs.

[15]The countries in the sample include 18 OECD countries – Australia, Austria, Belgium, Canada, Switzerland, Germany, Denmark, Spain, France, United Kingdom, Ireland, Italy, Japan, South Korea, Netherlands, Singapore, Sweden, United States – and 7 non-OECD countries – Brazil, China, India, Mexico, Malaysia, Russia and Thailand.

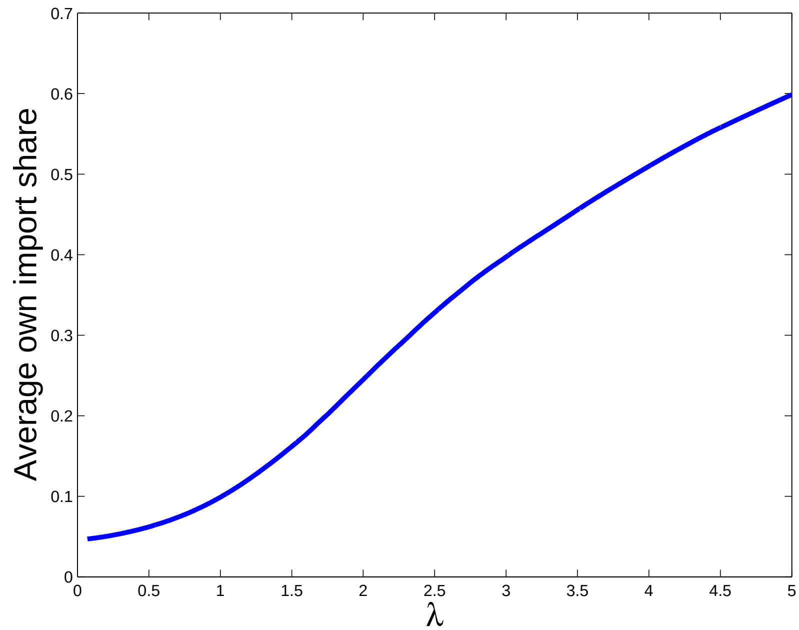
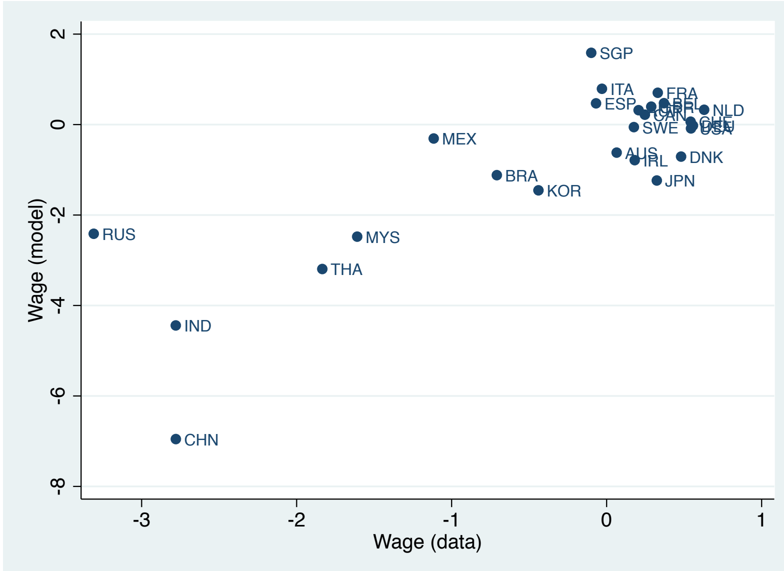


Figure 4: Average own import share as a function of λ

Finally, we assume that $\lambda_j = \lambda$, i.e., the cost of information is the same across countries. To calibrate this parameter, we exploit another property of the model. As λ rises and information processed falls, importers start receiving more noisy signals about the prices. This causes them to rely more on their prior beliefs about prices. As a result, importers tend to import more from the country that has the lowest expected price. In most cases, this happens to be the domestic country. Hence, when λ goes up, countries become less open. This is confirmed in Figure 4. The average own import share in the data for our sample of countries is 0.5, and we calibrate λ to match this number. The corresponding value of λ turns out to be 3.9.

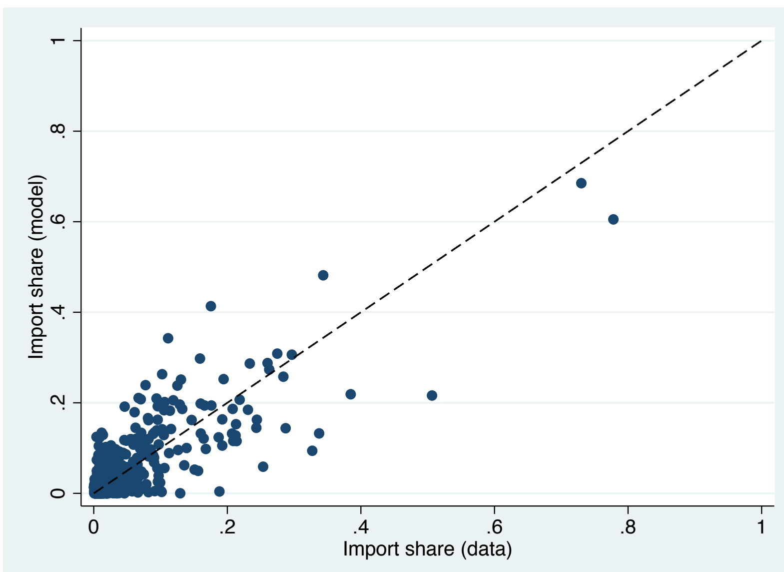
The benchmark value of λ corresponds to an average capacity (κ) of about 2.8. That is how much information importers process on average. To put this in context, the literature on rational inattention has found a number of estimates for information processing capacity, ranging between 0.1 and 3. Our benchmark value of capacity is close to the upper bound of this range and it is consistent with parameter values found by Mackowiak and Wiederholt [2009] and Pasten and Schoenle [2016].

Figure 5 plots the wage in the data against the wage generated by the model for the benchmark scenario. There is a strong positive relation between the two (raw correlation of 0.82). To further evaluate the fit of the model, Figure 6 plots the import shares generated by the model, excluding the own import shares, against their data counterpart. As the plot suggests, there is a clear corre-



Note: Wage is expressed as logarithms of deviations from mean.
Source: World Bank's World Development Indicators (WDI) and model.

Figure 5: Wage (model versus data)



Source: NBER-UN World Trade Flows and model.

Figure 6: Import shares, excluding own (model versus data)

lation between the model and the data (raw correlation of 0.79). Recall that the only component of trade cost that we use in our model is distance. It is, therefore, re-assuring that a large part of the variation in the trade shares can be explained using our parsimonious model.

A direct consequence of incomplete information is that for most products, importers in any country are no longer paying the minimum available price. Figure 7 shows the distribution of the ratio of price paid by U.S. importers over the minimum price. As the figure suggests, there is a large variation in the actual price paid for a product, relative to the minimum, with the median price being between two to three times the minimum. One implication is that cross-country price differences may not always reflect the presence of conventional trade barriers.

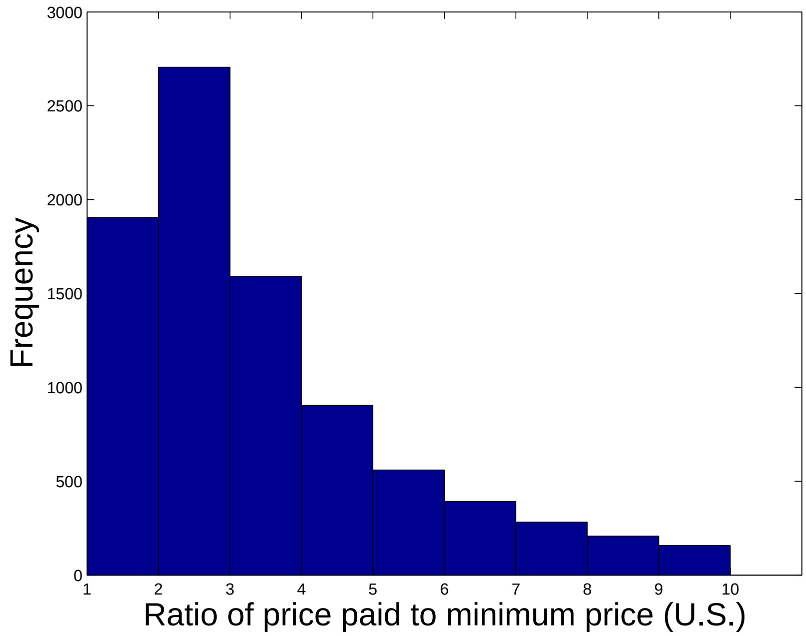


Figure 7: Distribution of prices paid by U.S. importers

4.2 Distance elasticity of trade

Next, we revisit the distance elasticity puzzle. 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. Conducting a meta-analysis, Disdier and Head [2008] find that δ takes a value, on average, of around -1. This elasticity has been surprisingly stable over time, even increasing slightly under some specifications [Coe et al., 2007, Berthelon and Freund, 2008]. In traditional trade models that generate a gravity equation, ϵ is usually a structural parameter with an average value of around -4.[16] This implies that ρ must take a value of 0.25 to be consistent with $\delta = -1$. Most studies that estimate ρ using measures of freight costs have found a value closer to $\rho = 0.025$ [Limao and Venables, 2001, Hummels, 2007]. In other words, 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, almost an order of magnitude lower than the 1 percent as suggested by the data.

To examine the relationship between distance and trade flows in our model, we run multiple specifications of the following gravity regression using the model-generated data:

$$\ln \pi_{ij} = \beta_x EX_i + \beta_m IM_j + \beta_d \ln dist_{ij} + \nu_{ij},$$

where EX_i is a set of exporting country i-specific variables, IM_j is a set of importing country j-specific variables and $dist_{ij}$ is the bilateral distance between countries i and j. The error term ν_{ij} is assumed to be orthogonal to the regressors.

It has been recognized for sometime now that a regression with only the bilateral distance between the exporting and importing countries and their GDPs produces biased estimates of the gravity coefficients – one needs to take into account how distant a country is with respect to *all* its trading partners [Anderson and van Wincoop, 2003]. Consequently, in our first specification, apart from bilateral distance and exporter GDP, we introduce two remoteness variables, one each for the exporter and another for the importer [Baldwin and Harrigan, 2011].[17] The remoteness of country i with respect to the rest of the world is given by

$$remoteness_i = \left(\sum_{j=1}^{C} gdp_j / dist_{ij} \right)^{-1}.$$

From an exporting country's perspective, remoteness is a distance-weighted measure of market access, while from an importing country's perspective, remoteness is a distance-weighted measure of supplier potential. Note that the remoteness of a country is the same, whether the country is an exporter or an importer. The results, with the remoteness variables included, are shown in Column 1 of Table 1. The coefficient on remoteness of the importing country is positive and significant. When the importing country j is located far from the large exporters, there is less

[16] In Armington type national-product-differentiation models or monopolistic competition models of trade with homogenous firms, ϵ is the elasticity of substitution. In perfectly competitive models of trade, ϵ is the Fréchet shape parameter, while in monopolistic competition models of trade with heterogenous firms, ϵ is the Pareto shape parameter.

[17] We do not include the importing country j's GDP on the right-hand side because the denominator of the ratio on the left-hand side is a multiple of country j's GDP.

Table 1: Gravity regressions

Dep. variable: Import share	$\theta = 4$		$\theta = 3$		$\theta = 5$	
	(1)	(2)	(3)	(4)	(5)	(6)
Distance	-2.97***	-2.61***	-1.94***	-1.77***	-3.88***	-3.34***
	(0.20)	(0.17)	(0.09)	(0.08)	(0.34)	(0.27)
Exporter GDP	2.10***		1.50***		2.65***	
	(0.16)		(0.07)		(0.28)	
Importer remoteness	2.16***		1.37***		2.88***	
	(0.30)		(0.13)		(0.50)	
Exporter remoteness	3.33***		1.86***		4.72***	
	(0.27)		(0.11)		(0.46)	
Importer FE	No	Yes	No	Yes	No	Yes
Exporter FE	No	Yes	No	Yes	No	Yes
Observations	625	625	625	625	625	625
R^2	0.46	0.77	0.67	0.87	0.37	0.70

Note: Robust standard errors in parantheses. *, ** and *** refer to significance at the 10%, 5% and 1% levels respectively. All variables are expressed in logarithms.

competition to sell at home. This raises imports from country i. The coefficient on remoteness of the exporting country is positive and significant as well. When the exporting country i is located far from large markets, the demand for its products is low. This raises exports to country j. The coefficient on country i's GDP is positive and significant, in line with most studies that find a coefficient on the exporting country's GDP of around 1.

The problem with the remoteness measures is that they are a-theoretic. Feenstra [2015] argues that a theory-consistent way of dealing with remoteness is to include exporter and importer country fixed effects. Such fixed effects can be interpreted as multi-lateral resistance, as opposed to bilateral resistance that is captured by bilateral trade costs. Column 2 of Table 1 reports the results of this specification. All variables specific to exporter and importer countries are now soaked up by the fixed effects. Adding the fixed effects also improves the fit of the model – R^2 rises by more than 30 percentage points.

So how does our model perform in terms of reconciling the distance elasticity puzzle? Observe that in the absence of information frictions (i.e., $\lambda = 0$), our model is equivalent to the one in Eaton and Kortum [2002]. Consequently, the distance elasticity is the product of θ and the elasticity of trade cost with respect to distance, which, we assumed was 0.025, in line with micro-evidence on trade costs. When $\theta = 4$, the distance elasticity in such a full information model (i.e., $\lambda = 0$) would be 0.1. In contrast, with the same value of parameters but a positive information cost, our model generates a distance elasticity that is an order of magnitude larger. For the benchmark scenario of $\theta = 4$, the estimated distance elasticity with exporter and importer fixed effects, -2.61, is about two and a half times larger than the median value of -1 as found by Disdier and Head [2008].

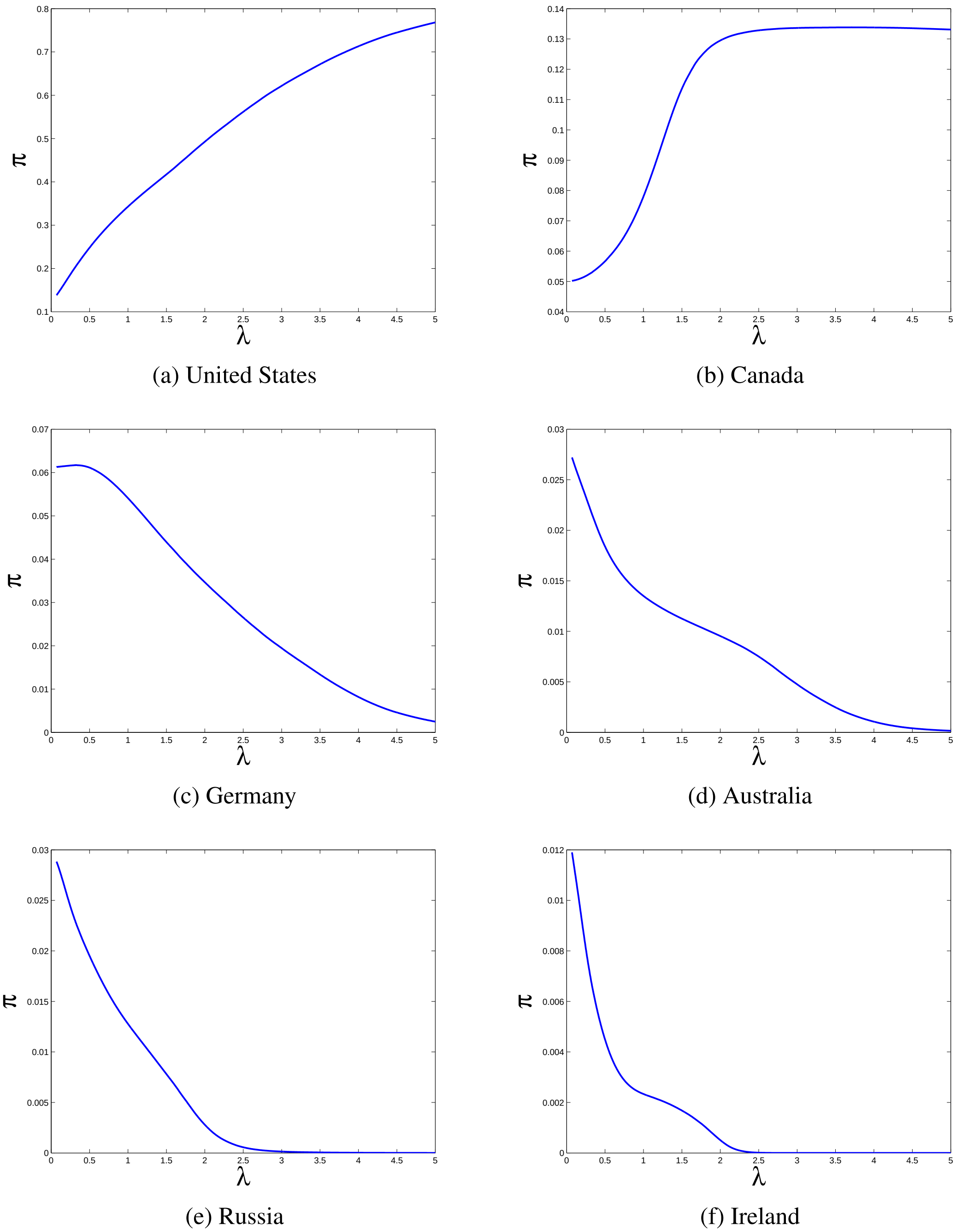
Notice that our estimate of distance elasticity depends on the value of θ, the Fréchet shape parameter used in the quantitative exercise. A lower θ implies a higher variance in productivity, which makes comparative advantage forces stronger. To examine the sensitivity of our estimates to θ, we solve the model with a smaller value ($\theta = 3$) and a larger value ($\theta = 5$), while adjusting λ to keep the average own share of imports at 0.5. Columns (3) and (5) show the results when we include the remoteness variables while columns (4) and (6) show the results with fixed effects. The distance elasticity is quite stable. In fact, when $\theta = 3$, the elasticity is even closer to -1. Intuitively, stronger comparative advantage forces tend to reduce the effect of distance on trade flows.

The reason we can generate larger effects of distance compared to other models is of course due to information processing by importers. Faced with cost of processing information, importers pay less attention to a source as it becomes more distant. This endogenous increase in information friction seems to have a much larger negative effect on trade flows than more traditional frictions such as transport costs. 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. Our model sheds light on one possible component of these dark trade costs – information friction.

4.3 Comparative statics

Having solved the model, we are now in a position to perform comparative statics with respect to the cost of information, λ. We consider how trade shares respond to a change in λ.

Figures 8 and 9 show the import and export shares of the U.S. with respect to a select group of its trade partners. In Figure 8, we plot the evolution of import shares of the largest country in terms of imports into the U.S., as well as countries in the 80th, 60th, 40th and 20th percentiles,



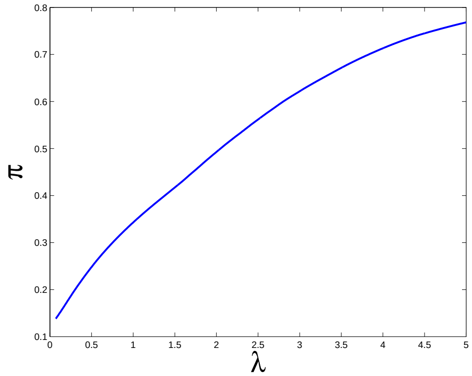
(a) United States (b) Canada

(c) Germany (d) Australia

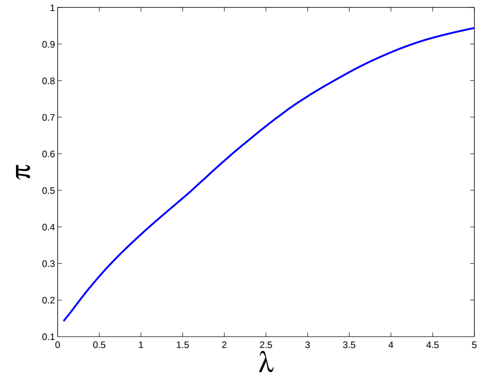
(e) Russia (f) Ireland

Figure 8: U.S. imports

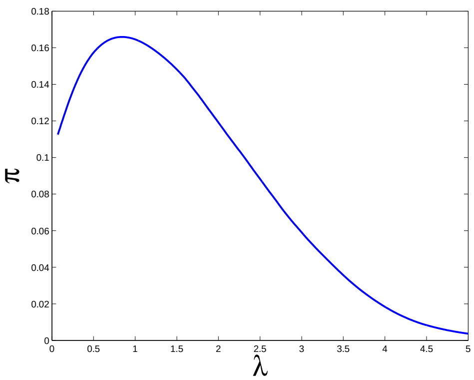
where the countries are ranked in terms of their imports into U.S. for the benchmark scenario of $\lambda = 3.9$. From the largest to the smallest, these countries are Canada, Germany, Australia,



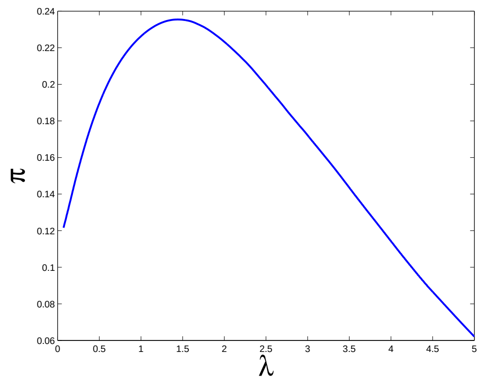
(a) United States



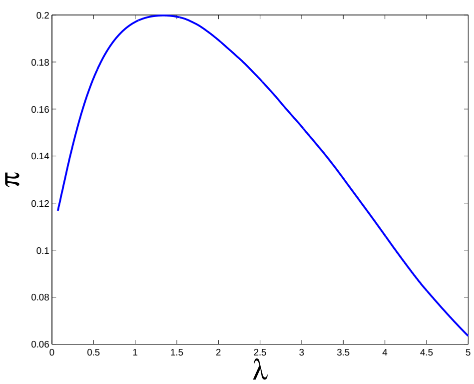
(b) Canada



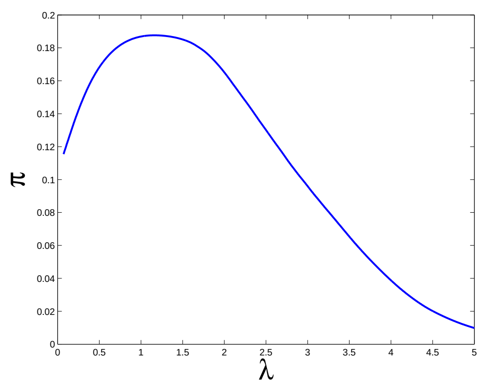
(c) Germany



(d) Australia



(e) Russia



(f) Ireland

Figure 9: U.S. exports

Russia and Ireland.[18] A decline in λ from its calibrated level of 3.9 reduces U.S. imports from itself, as well as from its largest trade partner, Canada, while causing imports from the other four

[18]The raw correlation between import shares in the model and the data for the U.S. is 0.75.

trade partners to rise. Furthermore, the import share for Germany exhibits the non-monotonicity, a unique feature of this model that we discussed before. Interestingly, for very low information costs, the U.S. is predicted to import more from Germany than Canada, suggesting how its close proximity to the U.S. gives Canada an extra edge in an incomplete information world. When we look at U.S. exports to the same set of countries in Figure 9, the patterns are similar. A drop in λ from 3.9 reduces U.S. exports to Canada, while increasing it for the other four countries.

Figures 8 and 9 confirm our assertions in Section 2 that a change in λ has asymmetric and non-monotonic effects on trade shares, even in a full-fledged general equilibrium model of trade. As λ falls, importers across the 25 countries start to process more information on average. At the same time, they start to pay relatively more attention to countries that are either (a) distant, or (b) have lower average productivity. For example, as information costs fall, U.S. importers purchase relatively more from Germany (a distant country) and Russia (a low productivity country), essentially at the expense of Canada (and the U.S. itself). Before concluding, we review some features of the data that are consistent with our model but would not necessarily be predicted by a full information model of trade.

5 Evidence

Our model of inattention generates the following prediction: for a given product and importing country, the concentration of imports is initially decreasing and then increasing in information cost. To examine this prediction, we carry out a cross-country analysis, and conjecture that information cost varies across countries. A caveat is in order. All we uncover in this section is a correlation. We do not make any attempt to infer causality. Nevertheless, the prediction relating information costs with import concentration is novel and we think of this as a preliminary attempt to examine whether the data is indeed consistent with this prediction.

From the NBER-UN database, we construct a sample of 770 4-digit SITC product categories and 84 importing countries that together accounted for more than 99 percent of all imports in 1999. Our measure of the concentration of a distribution is the Herfindahl-Hirschman index (HHI). Capturing the cost of processing information is challenging. We consider the inverse of the number of internet users (per 100 people) as a proxy for information cost. The number of users would be a function of both the fixed cost of getting an internet connection, as well as the variable cost of using that service (data usage). The latter is a better indicator of the cost of processing information in our model.[19] The data for internet usage come from the World Bank's World Development Indicators (WDI). For our sample of 84 countries, the average number of internet connections per 100 people is 10.3. Not surprisingly, there is large variation with the 90th percentile being 35.8 while the 10th percentile being smaller than 1.

To investigate the relationship between import concentration and information costs, we run the following regression at the product-importing country level:

$$\ln HHI_j^h = \beta^h + \beta_1 \ln I_j + \beta_2 X_j + \beta_3 Y_j^h + \epsilon_j^h,$$

where HHI_j^h is a measure of concentration of imports for product h in importing country j, I_j is the inverse of internet usage, a measure of information cost in country j, X_j is a vector of other country-specific regressors and Y_j^h is a set of product-importing country specific regressors. The coefficient β^h captures product fixed effects. Finally, ϵ_j^h is an error term orthogonal to the regressors. The results are displayed in Table 2.

In column 1, we regress HHI on information cost and the square of information cost. The coefficient on information cost is positive and significant while that on the square of information cost is not significantly different from zero. The result seems to suggest that as information costs increase, import distributions become more concentrated and this relation is monotone.

An issue with the correlation uncovered in column 1 of Table 2 is that an entirely different mechanism might be driving these results. To see this, observe that if country j is equally likely to import from N_j different countries, the HHI of its import distribution reduces to $1/N_j$. Consequently, the more sources a country imports from, the lower is its HHI. Now, if countries with lower information cost also import from more sources (probably because these are also the richer countries), the correlation in column 1 could be explained without using our model of rational inattention.[20] Therefore, in column 2, we include the number of source countries for a product. As hypothesized above, the import distribution seems to be less concentrated, the higher is the number of source countries. In fact, once we add this to our existing set of regressors in column 2, the sign of the coefficient on information cost flips. Controlling for the number of countries a product is sourced from, the import distribution seems to be more concentrated, the lower is cost of information.

In column 3, we add a number of other regressors. The first regressor is population. The argu-

[19] Other proxies for information frictions used in the literature include the number of web hosts located in a country [Freund and Weinhold, 2004] and bilateral telecommunication prices Fink et al. [2005].

[20] Although absent in our model, fixed costs of exporting/importing could generate systematic predictions about the extensive margin of trade.

Table 2: Country-product level regressions

Dependant variable:	HHI			
	(1)	(2)	(3)	(4)
Information cost	0.08***	-0.02***	-0.05***	-0.05***
	(0.017)	(0.005)	(0.008)	(0.009)
(Information cost)^2	-0.01	0.00	0.00	0.00
	(0.008)	(0.004)	(0.004)	(0.004)
Number of exporters per product		-0.54***	-0.60***	-0.60***
		(0.014)	(0.011)	(0.011)
Population			0.04***	0.04***
			(0.011)	(0.011)
Average exporter income			-0.03***	-0.03***
			(0.010)	(0.010)
HHI of exporter productivity			-0.05**	-0.05**
			(0.020)	(0.020)
Internet usage x Homogenous				0.01*
				(0.005)
Product FE	Yes	Yes	Yes	Yes
Observations	50660	50660	50660	46115
R^2	0.31	0.71	0.72	0.72

Note: Standard errors are clustered at the product level. *, ** and *** refer to significance at the 10%, 5% and 1% levels respectively. All variables are in logs. Dependant variable is HHI for a given importer-product. A homogenous product is defined according to the Rauch classification [Rauch, 1999]. It includes both goods that are traded in organized exchanges, as well as goods that have a reference price.
Source: NBER-UN World Trade Flows for trade, World Bank's World Development Indicators (WDI) for internet usage, Penn World Tables for per capita GDP and population.

ment behind including population is that there could be various aspects of the cost of processing information which are not directly observable (at least, for which comparable, cross-country data may not be available) but which could be correlated with a country's size. For example, if a certain information technology exhibits scale economies, the technology may be adopted only if there exists enough potential users.[21] The coefficient on population is positive and significant. If higher population indeed captures lower information cost, the result suggests once again that countries with lower cost of information have more concentrated import distributions.

It is quite possible that the relationship between information cost and import concentration is driven by selection. Suppose that poor countries import only from other poor countries and it is much harder to find out what the true prices are in those countries. At the same time, rich countries import primarily from other rich countries, with the information about producers in those countries being much easier to obtain. In this case, one could argue that importers in poor countries face a higher risk and would diversify their purchases across trading partners, resulting in less concentrated import distributions, while rich country importers would tend to purchase larger shares from countries with lower prices, resulting in more concentrated import distributions. To check this possibility, we compute the average income of source countries for each product that a country imports. Inclusion of this variable in column 3 of Table 2 does not alter the results. Furthermore, the coefficient on this variable is negative, the opposite of what we had conjectured.

Another possibility is that rich countries happen to be closer to countries with high productivity while poor countries are not. For example, Canada is much closer to a high productivity country such as the U.S. while it is far from a low productivity country such as Peru. This would cause Canadian importers to purchase much more from the U.S. relative to Peru, resulting in a highly concentrated import distribution. On the other hand, Honduras, which is roughly equidistant from both the U.S. and Peru, would tend to buy less from the U.S. and more from Peru, both relative to Canada, resulting in a less concentrated import distribution.

To examine this possibility, we define a measure of concentration of distance-adjusted productivity for country j, the HHI of exporter productivity:

$$\Phi_j = \sum_i \hat{\phi}_{ij}^2,$$

where

$$\hat{\phi}_{ij} = \frac{\phi_i/dist_{ij}}{\sum_l \phi_l/dist_{lj}}.$$

and ϕ_i is average productivity of country i. *Ceteris paribus* $\hat{\phi}_{ij}$ is higher, the more productive is country i, or the smaller the distance between i and j. Because the $\hat{\phi}_{ij}$-s are shares, a large $\hat{\phi}_{ij}$ for some i implies that Φ_j is close to one. In the example above, Φ_j for Canada would be

[21] We also included per capita GDP, but the coefficient on this variable turns out to be insignificant owing to the very high correlation with information cost.

larger than Φ_j for Peru. Inclusion of this variable in the above regression generates a negative and significant coefficient.[22] This suggests that countries which are close to a few productive sources (high Φ_j) actually have less concentrated import distributions for individual products. Hence, this observation cannot explain our findings either.

We should point out that none of the standard models of trade have any systematic prediction about the distribution of imports for a particular product on the one hand, and characteristics of the importing country on the other.[23] Of course, despite all the controls, we cannot fully rule out the possibility that the variations in import concentration and information cost are being driven by some third factor. To provide further support for our theory, we then exploit how the relation between import concentration and information cost varies across different types of products.

In particular, we divide products into those that are less information-sensitive and those that are more. Rauch [1999] argued that products falling into the former category are the homogenous products, while those falling into the latter category are the differentiated ones. If our theory is correct, we should expect that the concentration of imports should vary less with information costs for products that are not very information-intensive. When information about a product is easily obtained, information barriers are not binding. Accordingly, one should not expect significant differences in how a product is sourced across importing countries with different information costs. To test this hypothesis, we include a term that interacts the internet usage of a country with a dummy variable that takes a value of one if the product is homogenous. Our hypothesis is that the coefficient on this term should be positive. Column 4 of Table 2 confirms our hypothesis – the effect of lower information cost on import concentration is 20 percent smaller for homogenous goods relative to differentiated goods.

To summarize, the import distribution of a product is less concentrated in countries that seem to have high cost of processing information. Furthermore, our finding cannot be explained by importing countries varying systematically in terms of whether they import a product from primarily rich or primarily poor countries and whether they happen to be close to a few productive source countries. Observe that our model predicts a non-monotonic relationship between import concentration and information cost. The results in Table 2, however, indicate that the relationship is monotone decreasing, at least over the range of information costs that we consider (the quadratic term involving information cost is insignificant under all specifications). One possi-

[22]For average productivity, we use TFP measures from the Penn World Tables.

[23]A wide class of models generate the following demand in country j for goods produced in country i: $\pi_{ij} = (p_{ij}/P_j)^{-\chi}$, where P_j is an aggregate price index for goods available in country j and χ, the elasticity of trade cost, is a structural parameter whose interpretation varies across models. In this case, the relative imports from two sources, a and b, is given by $(p_{aj}/p_{bj})^{-\chi}$, a term independent of country j's characteristics.

ble interpretation of this result is that we are indeed in the region with low information costs. Of course, a monotone increasing relationship would also be consistent with our model. Either way, as argued earlier, none of these relationships are predicted by any of the standard trade models. Furthermore, we show that the monotone relationship is stronger for products that are information-intensive, a finding that is again consistent with our model.

6 Conclusion

In this paper, we make three main contributions. First, we establish a formal link between trade flows and the costs of processing information. We show how endogenous processing of information by importers generates novel comparative static results involving import shares and information costs. Second, calibrating the information cost parameter in the model, we show quantitatively how small changes in distance get translated into large changes in trade flows. In the process, we provide one possible explanation for the distant elasticity puzzle. And third, we provide evidence that the import distribution for a product is, on average, more concentrated in countries that seem to have lower costs of processing information and this relation is more pronounced for products that are information-intensive. While none of the standard full-information models predict a systematic variation in the import concentration across countries, our model does.

In their survey on the resistance to 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 in order to shed light on these "dark" costs, we need a better understanding of the role of information in facilitating trade. Borrowing from the theory of rational inattention, we have developed a framework that allows us to do just that. Much needs to be done, however.

In a recent paper, Dickstein and Morales [2015] ask a related but slightly different question: what is it that exporters know? Dickstein and Morales show that exporters typically have information on a very limited set of variables – distance to a destination, aggregate exports to that destination in the previous year and own productivity in the previous year. Accordingly, their expectations of future profits and consequently, entry decisions, are based on information sets that are far from full. In the context of our framework, the finding of Dickstein and Morales raises an interesting question: if exporters are rationally inattentive, then what is the optimal signal? Is focussing on the above-mentioned variables optimal? Or can exporters do better, given their information processing constraints? We leave the answers for future work.

References

Treb Allen. Information frictions in trade. *Econometrica*, 82(6):2041–2083, 2014.

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

James E. Anderson and Eric van Wincoop. Trade costs. *Journal of Economic Literature*, 42: 691–751, 2004.

Costas Arkolakis, Jonathan Eaton, and Samuel Kortum. Staggered adjustment and trade dynamics. mimeo, Yale University, 2012.

Richard Baldwin and James Harrigan. Zeros, quality and space: Trade theory and trade evidence. *American Economic Journal: Microeconomics*, 3(2):60–88, 2011.

Matias Berthelon and Caroline Freund. On the conservation of distance in international trade. *Journal of International Economics*, 75(2):310–320, 2008.

Jeffrey Chan. Soldiers and the structure of trade: Evidence from 50 years of U.S. military deployment. University of Toronto, mimeo, 2016.

Thomas Chaney. The network structure of international trade. *The American Economic Review*, 104(11):3600–3634, 2014.

David Coe, Arvind Subramanian, and Natalia Tamirisa. The missing globalization puzzle: Evidence of the declining importance of distance. *IMF Staff Papers*, 54:34–58, 2007.

Pierre-Philippe Combes, Miren Lafourcade, and Thierry Mayer. The trade-creating effects of business and social networks: Evidence from France. *Journal of International Economics*, 66 (1):1–29, 2005.

Michael J Dickstein and Eduardo Morales. What do exporters know? Technical report, National Bureau of Economic Research, 2015.

Anne-Célia Disdier and Keith Head. The puzzling persistence of the distance effect on bilateral trade. *The Review of Economics and Statistics*, 90(1):37–48, 2008.

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

Robert C Feenstra. *Advanced international trade: theory and evidence*. Princeton University Press, 2015.

Carsten Fink, Aaditya Mattoo, and Ileana Cristina Neagu. Assessing the impact of communication costs on international trade. *Journal of International Economics*, 67(2):428–445, 2005.

Caroline L Freund and Diana Weinhold. The effect of the internet on international trade. *Journal of International Wconomics*, 62(1):171–189, 2004.

David 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.

Gene Grossman. Comment on Deardorff. In J.A. Frankel, editor, *The Regionalization of the World Economy*, pages 33–57. U. Chicago Press, 1998.

Keith Head and Thierry Mayer. What separates us? Sources of resistance to globalization. *Canadian Journal of Economics*, 46(4):1196–1231, 2013.

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

Keith Head and John Ries. FDI as an outcome of the market for corporate control: Theory and evidence. *Journal of International Economics*, 74(1):2–20, 2008.

Ali Hortaçsu, F Martínez-Jerez, and Jason Douglas. The geography of trade in online transactions: Evidence from eBay and mercadolibre. *American Economic Journal: Microeconomics*, 1(1): 53–74, 2009.

David Hummels. Transportation costs and international trade in the second era of globalization. *The Journal of Economic Perspectives*, 21(3):131–154, 2007.

Adam B Jaffe, Manuel Trajtenberg, and Rebecca Henderson. Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*, pages 577–598, 1993.

Daniel Kahneman. *Attention and effort*. Prentice-Hall series in experimental psychology. Prentice-Hall, 1973.

P Klenow and A Rodriguez-Clare. Quantifying variety gains from trade liberalization. mimeo, University of Chicago, 1997.

Sebastian Krautheim. Heterogeneous firms, exporter networks and the effect of distance on international trade. *Journal of International Economics*, 87(1):27–35, 2012.

Andreas Lendle, Marcelo Olarreaga, Simon Schropp, and Pierre-Louis Vézina. There goes gravity: How eBay reduces trade costs. Cepr discussion paper no. dp9094, 2012.

Nuno Limao and Anthony J. Venables. Infrastructure, geographical disadvantage, transport costs, and trade. *The World Bank Economic Review*, 15(3):pp. 451–479, 2001.

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

Filip Matejka and Alisdair McKay. Rational inattention to discrete choices: A new foundation for the multinomial logit model. *The American Economic Review*, 105(1):272–298, 2014.

Daniel McFadden. A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica*, 57(5):995–1026, 1989.

Ernesto Pasten and Raphael Schoenle. Rational inattention, multi-product firms and the neutrality of money. *Journal of Monetary Economics*, 80:1 – 16, 2016.

Giovanni Peri. Determinants of knowledge flows and their effect on innovation. *Review of Economics and Statistics*, 87(2):308–322, 2005.

Richard Portes and Helene Rey. The determinants of cross-border equity flows. *Journal of International Economics*, 65(2):269 – 296, 2005.

Richard Portes, Hélène Rey, and Yonghyup Oh. Information and capital flows: The determinants of transactions in financial assets. *European Economic Review*, 45(4):783–796, 2001.

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

James E. Rauch and Vitor Trindade. Ethnic Chinese networks in international trade. *Review of Economics and Statistics*, 84(1):116–130, 2002.

Claude Elwood Shannon. A mathematical theory of communication. *Bell Technical Journal*, 3 (27):379–423, 1948.

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

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

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

Claudia Steinwender. Information frictions and the Law of One Price: When the States and the Kingdom became United. mimeo, London School of Economics, 2014.

Michael Waugh. International trade and income differences. *The American Economic Review*, 100(5):2093–2124, 2010.

Appendix

Proof of Proposition 1. This proof follows in the steps of 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

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

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

$$log(\tilde{z}_{ij}) + \xi_{ij}(Z) + \mu(Z) + \lambda_j \left(\ln \pi_{ij} + 1 - \ln f_{ij}(Z) - 1 \right) = 0.$$

As (5) does not bind, then the first order condition can be re-arranged to yield

$$f_{ij}(Z) = \pi_{ij} e^{(log(\tilde{z}_{ij}) - \mu(Z))/\lambda_j} \tag{10}$$

Plugging (10) into (6), we obtain

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

If we plug this expression back into (10), we get (7). Equation (7) holds even for $\pi_{ij} = 0$, as otherwise equation (3) would not hold. □

Proof of Proposition 2. Part (i) is trivial. For part (ii), divide (7) by $e^{log(\tilde{z}_{ij})/\lambda_j}$ to obtain

$$f_{ij}(Z) = \frac{\pi_{ij}}{\pi_{ij} + \sum_{k \neq j} \pi_{kj} e^{\frac{1}{\lambda_j}\left[log(\tilde{z}_{kj}) - log(\tilde{z}_{ij})\right]}}.$$

Suppose $\tilde{z}_{ij} = max \tilde{z}_{kj} \forall k$. Then, as $\lambda_j \to 0$, $\forall k \; \frac{1}{\lambda_j}\left[log(\tilde{z}_{kj}) - log(\tilde{z}_{ij})\right] \to -\infty$. It follows that in this case,

$$f_{ij}(Z) \to \frac{\pi_{ij}}{\pi_{ij}} = 1.$$

If, on the other hand, suppose $\exists h$ such that $\tilde{z}_{hj} > \tilde{z}_{ij}$. Then, as $\lambda_j \to 0$, $\frac{1}{\lambda_j}\left[log(\tilde{z}_{hj}) - log(\tilde{z}_{ij})\right] \to \infty$. In this case,

$$f_{ij}(Z) \to \frac{\pi_{ij}}{\infty} = 0.$$

$\square$

Proof of Lemma 1. Because there is a measure one of symmetric varieties, the average expenditure on variety ω is simply

$$E[X_{ij}(\omega)] = X_{ij}. \tag{11}$$

Let the import share for variety ω (in value terms) be denoted by $s_{ij}(\omega)$. Then we can write,

$$X_{ij}(\omega) = s_{ij}(\omega)X_j(\omega),$$

where $X_j(\omega)$ is total expenditure by j on variety ω. Now under *trade separability*, the expenditure of j on ω is independent of the allocation of this expenditure across different source countries. Hence, we can write

$$E[X_{ij}(\omega)] = E[s_{ij}(\omega)]E[X_j(\omega)]. \tag{12}$$

Now,

$$\begin{aligned} E[s_{ij}(\omega)] &= Prob(j \text{ imports } \omega \text{ from } i) \cdot E[s_{ij}(\omega)|_{j \text{ imports } \omega \text{ from } i}] \\ &= \pi_{ij} \cdot 1, \end{aligned}$$

where the second line follows from the fact that at any given point in time, a variety is purchased from only one country and hence $E[s_{ij}(\omega)|_{j \text{ imports } \omega \text{ from } i}] = 1$. Furthermore, under Cobb-Douglas preference (a form of trade separable utility function), the expenditure shares are constant, i.e.,

$$E[X_j(\omega)] = X_j(\omega) = X_j,$$

where X_j is aggregate expenditure by j.[24] Therefore (12) can be re-written as

$$E[X_{ij}(\omega)] = \pi_{ij}X_j.$$

[24] $X_j = \int X_j(\omega)d\omega = X_j(\omega)$ due to symmetry and measure one of varieties.

Replacing this in (11) and re-arranging, we have

$$\frac{X_{ij}}{X_j} = \pi_{ij}.$$

□

Proof of Proposition 3. This proof follows in the steps of the proof of Lemma 1 and Proposition 3 in Matejka and McKay [2014]. Note that the optimization problem of consumers in country j can be equivalently formulated as a maximization over the unconditional probabilities, $\{\pi_{ij}\}_{i=1}^{N}$:

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

subject to (5) and (6). To see this, substitute equation (4) into the objective function to get

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

Rearranging this expression and using (7), we obtain

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

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

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

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

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

where $\pi_{Nj} = 1 - \sum_{i=1}^{N-1} \pi_{ij}$.

1. 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}.

2. When $\lambda_j \to \infty$, importers in j process no information and decisions are based on *ex-ante* expectations. Given that country i has lowest expected price, *ex-ante* expected $u(\tilde{z}_{ij})$ is the highest and $\pi_{ij} \to 1$.

3. If countries are *ex-ante* identical, $A_i = A; c_i = c; \tau_{ij} = \tau$ for all i, then $G(Z)$ 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. $\square$