Understanding Interstate Trade Patterns*

Hakan Yilmazkuday†

June 2008

Abstract

This paper attempts to find the motivation behind interstate trade patterns within the U.S. by using a partial equilibrium trade model. In particular, why does a region import more goods from a particular region while importing fewer from others? Why does a region import more of a particular good, while importing less of another one? The answers emerge from the simple model set forth by this paper that considers the geographical and technological differences across regions and goods. By construction, the model, together with the estimation methodology, controls for possible local (i.e., wholesale and retail) distribution costs, insurance costs, local taxes, markup differences in production, and intermediate input trade as well as zero trade observations, which are ignored in most gravity type studies. Moreover, the elasticity of substitution across goods, the elasticity of substitution across varieties of each good, and the good specific elasticity of distance measures are all identified in the empirical analysis, which is also not the case in most gravity type studies. Compared to empirical international trade literature, the elasticity of substitution is estimated to be lower, while the elasticity of distance is estimated to be higher intranationally.

JEL Classification: R12, R13, R32

Key Words: Trade Ratios; Transportation; The United States

*I thank Eric Bond, Mario Crucini, Kevin Huang, David Hummels, Meredith Crowley, Brown Bag Seminar participants at Vanderbilt University, Midwest International Economics Meeting participants in University of Illinois at Urbana-Champaign, SRSA Annual Meeting participants in Arlington, VA-Washington, DC for their helpful comments and suggestions. All errors are my own responsibility.

†Deparment of Economics, Vanderbilt University, Nashville, TN, 37235, USA; Tel: +1-615-343-2472; fax: +1-615-343-8495; e-mail: hakan.yilmazkuday@vanderbilt.edu

1. Introduction

What is the main motivation behind intranational trade? Compared to *relatively* complex models in the literature, this paper contributes by introducing a simple partial equilibrium model to analyze the motivations behind bilateral trade patterns of regions at the *disaggregate* level. We attempt to find why regions do import more goods from some regions while importing fewer from others. We also investigate why a region imports more of a good while importing less of another one.

In particular, we introduce a type of monopolistic competition model consisting of a finite number of regions. There are two types of goods, namely traded and nontraded. Each region produces and consumes a unique nontraded good. Each region may also consume all varieties of all traded goods, while it can produce only one variety of each traded good. While the traded goods are produced by a perfectly mobile unique factor, the only nontraded good in each region is produced by the same mobile factor together with traded intermediate inputs.

According to this setup, as is standard in the literature, we show that the trade of a variety of a particular traded good across any two regions depend on the relative price of the variety and the total demand (final consumption demand plus intermediate input demand) of the good in the destination (importer) region. Our contribution comes into the picture when we take the ratio between imports of varieties from different sources (exporters). We find that a region imports more goods (measured in values) from the lower price regions and fewer goods from the more distant regions.

We show that our model together with our estimation methodology has several empirical and analytical benefits compared to gravity models in the following senses:

(i) In our model, there is no identification problem in terms of estimating the elasticity of substitution and the elasticity of distance at the same time. However, even Anderson and van Wincoop's (2003) most popular gravity model suffers from this problem (also see Wei 1996; Hummels 1999, 2001). By distinguishing between aggregate level and disaggregate level trade data, together with considering the production side of our model, we can also estimate the elasticity of substitution across goods.

(ii) Our methodology controls for a possible issue of overstating the distance measures (due to using calculated distances, such as great circle distances) mentioned by Hillberry and Hummels

(2001).

(iii) By construction, the model is capable of controlling for the effects of local (i.e., wholesale and retail) distribution costs, insurance costs, local taxes, markup differences in production, international trade (under reasonable assumptions), and intermediate input trade, each of which are possible topics for separate debates in the literature (see Anderson and van Wincoop 2004).

(iv) There is an exogenous solution for the estimated trade expression, and thus, there is no need for any income data for estimation, *given* the technological levels.

We estimate the model using bilateral trade data belonging to the states of the U.S. The estimated parameters correspond to: a) elasticity of substitution across varieties of a good, each produced in a different region; b) elasticity of substitution across goods, each consisting of different varieties; c) elasticity of distance, which governs good specific trade costs; and d) heterogeneity of individual tastes, governing geographic barriers and the so-called *home-bias*. We pursue several strategies to estimate these parameters and support our results with different sensitivity analyses. Overall, our model is capable of explaining the interstate trade data up to 84% at the disaggregate level, and up to 77% at the aggregate level.

Our estimated parameters give insights about a number of issues related to interstate trade patterns within the U.S.: a) compared to empirical international studies, elasticity of substitution is lower intranationally; b) compared to empirical international studies, elasticity of distance is higher intranationally; c) there is evidence of home-bias even at the intranational level; d) trade costs are mostly good specific even at the intranational level; e) source-specific fixed effects are important for bilateral trade patterns, effects usually ignored in the literature; f) production technologies are both good and region specific rather than country specific; g) elasticity of substitution across varieties is good specific.

Related Literature

In this subsection, we briefly describe how this paper relates to its closest antecedents, especially gravity type studies. The gravity models are popular mostly due to their empirical success.[1] When we look at the theoretical background of gravity type studies, Anderson (1979) is the first one to model gravity equations. The main motivation behind Anderson's (1979)

[1] Deardorff (1984) reviews the earlier gravity literature. For recent applications, see Wei (1996), Jensen (2000), Rauch (1999), Helpman (1987), Hummels and Levinsohn (1995), and Evenett and Keller (2002).

gravity model is the assumption that each region is specialized in the production of only one good.[2] Despite its empirical success, as Anderson and van Wincoop (2003) point out, the specialization assumption suppresses finer classifications of goods, and thus makes the model useless in explaining the trade data at the disaggregate level. Another deficiency of Anderson's (1979) gravity model is the lack of a production side. Bergstrand (1985) bridges this gap by introducing a one-factor, one-industry, N-country general equilibrium model in which the production side is considered. In his following study, Bergstrand (1989) extends his earlier gravity model to a two-factor, two-industry, N-country gravity model.[3]

The main deficiency of the gravity models is that they cannot control for good specific transportation costs, good specific local (i.e., wholesale and retail) distribution costs, good specific insurance costs, good specific local taxes, region specific markup differences in production, good specific intermediate input trade or international trade. Moreover, as we have mentioned above, one cannot estimate the elasticity of substitution and the elasticity of distance at the same time by using gravity equations. However, our paper controls for all of these situations.

None of the papers mentioned above empirically deal with the trade patterns within a country. Recently, Wolf (2000), Hillberry and Hummels (2001), and Millimet and Thomas (2007) bridged this gap by analyzing the interstate trade patterns within the U.S. However, these studies have a deficiency, because they use the aggregate level (i.e., total bilateral) trade data, while our paper uses disaggregate level bilateral data that give more insight related to good specific analyses. Another deficiency of these studies is that they use gravity type models which suffer from the same issues mentioned above. Also, these studies cannot distinguish between the elasticity of substitution across varieties of a good, elasticity of substitution across goods, and the elasticity of distance at the same. By taking the ratio between imports of varieties from different origins (exporters), by taking the ratio between imports of different goods, and by including intermediate input trade into the model, our paper takes care of all of these issues by construction.

[2] In the Appendix of his paper, Anderson (1979) extends his basic model to a model in which multiple goods are produced in each region.

[3] Also see Suga (2007) for a monopolistic-competition model of international trade with external economies of scale, Lopez et al. (2006) for an analysis on home-bias on U.S. imports of processed food products, and Gallaway et al. (2003) for an empirical study to estimate short-run and long-run industry-level U.S. Armington elasticities.

Nevertheless, our paper is not the first one that considers trade ratios. For instance, studies such as Head and Ries (2001), Head and Mayer (2002), Eaton and Kortum (2002), and Romalis (2007), among others, have somehow also considered trade ratios in their gravity type models. Most of these studies have attempted to eliminate price measures from the gravity equation since they see them as nuisances. In order to get rid of those price measures, one cannot simply take the ratio among imports of varieties from different origins; they also have to consider the ratio among imports of varieties within regions. This process results with having an index for freeness of trade that helps us determine the impacts of borders (mostly related to international trade literature) rather than explaining the intranational trade patterns. Although this approach seems fine up to a point, it has the deficiency of not considering the production side at all and not having a structure to analyze the disaggregate level trade. The closest study to this paper is by Romalis (2007). However, by eliminating the source specific marginal costs (i.e., the production side), Romalis (2007) cannot identify the elasticity of substitution and the elasticity of distance at the same time; instead, he can only estimate the elasticity of substitution. By considering the production side, this paper can estimate the elasticity of substitution and the elasticity of distance at the same time. Moreover, all of these studies also don't take into account zero trade observations that have a high share in overall observations.[4] This paper contributes to the literature by controlling for all of these issues, therefore by having more accurate empirical results.

Plan of the Paper

The rest of the paper is organized as follows. Section 2 introduces our regional trade model. Section 3 provides insights and depicts our estimation methodology. Section 4 gives the empirical results, while Section 5 concludes. The data are described in the Appendix.

2. The Model

We model an economy consisting of a finite number of regions. In each region, there are two types of goods, namely traded and nontraded. While a unique nontraded good is produced and consumed within all regions (thus, the nontraded goods market is in equilibrium in each region

[4] Helpman et al. (2008) show that almost 50% of the observations are zero trade observations in international trade.

separately), each region may consume all varieties of all traded goods and can produce only one variety of each traded good. Since we only care about the partial equilibrium bilateral trade implications of our model, in many instances, we skip the irrelevant details of the model in order to keep it as simple as possible.[5]

Each traded good is denoted by $j = 1, ..., J$. Each variety is denoted by i that is also the notation for the region producing that variety. We make our analysis for a typical region, r. In the model, generally speaking, $H_{a,b}(j)$ stands for the variable H, where a is related to the region of consumption, b is related to the variety (and thus, the region of production), and j is related to the good.

2.1. Individuals

The representative agent in region r maximizes utility $U\left(C_r^T, C_r^{NT}\right)$ where C_r^T is a composite index of traded goods and C_r^{NT} is a unique nontraded good. The composite index of traded goods, C_r^T, is given by:

$$C_r^T \equiv \left(\sum_j (\gamma_j)^{\frac{1}{\varepsilon}} \left(C_r^T(j)\right)^{\frac{\varepsilon-1}{\varepsilon}}\right)^{\frac{\varepsilon}{\varepsilon-1}}$$

where $C_r^T(j)$ is given by:

$$C_r^T(j) \equiv \left(\sum_i (\beta_r \theta_i)^{\frac{1}{\eta(j)}} \left(C_{r,i}^T(j)\right)^{\frac{\eta(j)-1}{\eta(j)}}\right)^{\frac{\eta(j)}{\eta(j)-1}}$$

where $C_{r,i}^T(j)$ is the variety i of traded good j imported from region i; $\varepsilon > 0$ is the elasticity of substitution across goods; $\eta(j) > 1$ is the elasticity of substitution across varieties of good j; γ_j is a good specific taste parameter; β_r is a destination (i.e., importer) specific taste parameter; and finally, θ_i is a source (i.e., exporter) specific taste parameter. For different varieties, while having only one bilateral taste parameter, which is both destination and source specific, is standard in the literature, decomposing it into β_r and θ_i is new in this paper.[6] In particular, both β_r and θ_i can be used as fixed effects in a regression analysis; i.e., they together represent

[5] A general equilibrium framework is not necessary in our analysis. It would only complicate our model with unnecessary details.

[6] Distinguishing between destination and source specific taste parameters has useful properties in terms of our estimation. Our reason will be clearer when we move to Section 3.

a unique bilateral taste parameter between regions r and i. Moreover, by putting restrictions on θ_i, one can easily measure home-bias implications of our model. Besides, one can also control for issues such as migration by using θ_i (see Millimet and Thomas, 2007). Our claim will be clearer when we show the bilateral trade implications of our model in Section 3. We will also test for the validity of this assumption in Section 4.

The optimal allocation of any given expenditure within each variety of goods yields the following demand functions:

$$C_{r,i}^{T}(j) = \beta_r \theta_i \left(\frac{P_{r,i}^{T}(j)}{P_r^{T}(j)} \right)^{-\eta(j)} C_r^{T}(j) \tag{2.1}$$

and

$$C_r^{T}(j) = \gamma_j \left(\frac{P_r^{T}(j)}{P_r^{T}} \right)^{-\varepsilon} C_r^{T} \tag{2.2}$$

where $P_r^T(j) \equiv \left(\sum_i \beta_r \theta_i P_{r,i}^T(j)^{1-\eta(j)} \right)^{\frac{1}{1-\eta(j)}}$ is the price index of the traded good j (which is composed of different varieties), and $P_r^T \equiv \left(\sum_i \gamma_j P_r^T(j)^{1-\varepsilon} \right)^{\frac{1}{1-\varepsilon}}$ is the cost of living index in region r. It is implied that $P_r^T(j)\, C_r^T(j) = \sum_i P_{r,i}^T(j)\, C_{r,i}^T(j)$.

2.2. Firms

Since there are two types of goods, namely traded and nontraded, there are two types of firms in each region.

2.2.1. Production of Traded Goods

Traded good j in region r (i.e., variety r of good j) is produced by the following production function:

$$Y_r^T(j) = A_r(j)\, L_r^T(j) \tag{2.3}$$

where $A_r(j)$ represents the good and region specific technology, and $L_r^T(j)$ represents a completely mobile factor of production of which hour is worth W in all regions.[7]

[7] One can easily assume $L_r(j)$ to be labor and/or capital, but our results are not affected at all by these details, because we don't employ any factor market in order to keep the model as simple as possible in our analysis.

The firm chooses $L_r^T(j)$, taking as given its price W. The cost minimization problem of the firm implies that the marginal cost of producing variety r of good j (in region r) is given by:

$$MC_r^T(j) = \frac{W}{A_r(j)} \tag{2.4}$$

Note that $MC_r^T(j)$ is good and region specific.

2.2.2. Production of Nontraded Goods

The unique nontraded good in region r is produced by a production function $Y_r^{NT}\left(L_r^{NT}, G_r^{NT}\right)$ where L_r^{NT} represents the completely mobile factor of production (i.e., the same factor used in the production of traded goods) and G_r^{NT} is the counterpart of C_r^T given by:

$$G_r^{NT} \equiv \left(\sum_j (\gamma_j)^{\frac{1}{\varepsilon}} \left(G_r^{NT}(j)\right)^{\frac{\varepsilon-1}{\varepsilon}}\right)^{\frac{\varepsilon}{\varepsilon-1}}$$

where $G_r^{NT}(j)$ is given by:[8]

$$G_r^{NT}(j) \equiv \left(\sum_i (\beta_r\theta_i)^{\frac{1}{\eta(j)}} \left(G_{r,i}^{NT}(j)\right)^{\frac{\eta(j)-1}{\eta(j)}}\right)^{\frac{\eta(j)}{\eta(j)-1}}$$

The optimal allocation of any given expenditure within each variety of intermediate inputs yields the following demand functions:

$$G_{r,i}^{NT}(j) = \beta_r\theta_i \left(\frac{P_{r,i}^T(j)}{P_r^T(j)}\right)^{-\eta(j)} G_r^{NT}(j) \tag{2.5}$$

and

$$G_r^{NT}(j) = \gamma_j \left(\frac{P_r^T(j)}{P_r^T}\right)^{-\varepsilon} G_r^{NT} \tag{2.6}$$

Note that the firms share the same taste parameters, β_r, θ_i and γ_j, with the individuals. Although this is somehow a restrictive assumption, it has very nice properties in terms of bilateral trade implications that are discussed in Section 3.

[8] Since we care about the bilateral trade implications of our model, after assuming that the nontraded goods market is in equilibrium in each region, the exact functional forms of $Y_r^{NT}\left(L_r^{NT}, G_r^{NT}\right)$, demand for L_r^{NT}, and the marginal cost implication for the nontraded goods are all irrelevant in our analysis.

2.2.3. Trade Cost

We have to define our trade cost first. Anderson and van Wincoop (2004) categorize the trade costs under two names, costs imposed by policy (tariffs, quotas, etc.) and costs imposed by the environment (transportation, wholesale and retail distribution, insurance against various hazards, etc.). Since we analyze trade within a country (i.e., the U.S.), we ignore the first category and focus on the second one. Instead of assuming an iceberg transport cost, we assume that the transportation is achieved by a transportation sector, which is not modeled here.[9] This assumption is important to distinguish between the export income received by the exporter and the transportation income received by the transporter. The implications of this assumption will be clearer below. In particular, we assume that, if there is a trade between regions, it is subject to a transportation cost:[10]

$$\begin{aligned} P_{i,r}^{T}(j) &= (1+\tau_{i,r}(j))\left(P_{r,r}^{T}(j)\right) \\ &= (D_{i,r})^{\delta(j)}\left(P_{r,r}^{T}(j)\right) \end{aligned} \tag{2.7}$$

where $P_{r,r}^{T}(j)$ is the price of the traded good at the factory gate (i.e., the source); $\tau_{i,r}(j) > 0$ is a good specific net transportation cost from region r to region i; $D_{i,r}$ is the distance between regions r and i; and $\delta(j)$ is the elasticity of distance.[11] This assumption is commonly used in the literature (see Anderson and van Wincoop 2003, 2004). The cost implications of our model in terms of wholesale distribution, retail distribution, insurance or local taxes, will be provided below.

[9] Since we consider only the partial equilibrium bilateral trade implications of our model, the actual role/model of the transportation sector is irrelevant in our analysis after assuming that transportation is achieved by using the completely mobile factor of production (i.e., the same factor used in the production of traded/nontraded goods).

[10] Needless to say, the existence of trade is determined by Equation 2.1 for all i, r and j. As we will discuss in detail in the following sections, we consider the absence of trade, besides the existence of it in our empirical analysis.

[11] For the distance within each state (i.e., the internal distance), we use the proxy developed by Wei (1996), which is one-fourth the distance of a region's capital from the nearest capital of another region.

2.2.4. Equilibrium

Since we care about the partial equilibrium bilateral trade implications of our model, we naturally assume that the nontraded good market, of which details are not shown here, is in equilibrium in each region. So, in this subsection, we depict equilibrium in the traded goods market. In particular, for each variety r of traded good j produced in region r, the market clearing condition implies:

$$Y_r^T(j) = \sum_i C_{i,r}^T(j) + \sum_i G_{i,r}^{NT}(j) \tag{2.8}$$

where $C_{i,r}^T(j)$ is the demand of region i for variety r of traded good j (produced in region r); and $G_{i,r}^{NT}(j)$ is the intermediate input demand for variety r of traded good j (produced in region r) demanded for the production of the nontraded good in region i . Equation 2.8 basically says that variety r of final good j produced in region r is either consumed locally or by other regions, either for final consumption or as an intermediate input.

2.2.5. Price Setting

Since we care about the partial equilibrium bilateral trade implications of our model, the price setting behavior of the firms producing the unique nontraded good in each region is irrelevant in our analysis. For the traded goods, in region r, we assume that a typical firm that produces variety r of traded good j faces the following profit maximization problem:

$$\max_{P_{r,r}^T(j)} Y_r^T(j)\left[P_{r,r}^T(j) - MC_r^T(j)\right]$$

subject to Equation 2.8 and the symmetric versions of Equation 2.1 and 2.5. The first order condition for this problem is as follows:[12]

$$Y_r^T(j)\left[1 - \frac{\eta(j)}{P_{r,r}^T(j)}\left(P_{r,r}^T(j) - MC_r^T(j)\right)\right] = 0$$

[12]Notice that the firm takes the composite consumption index of good j (i.e., $C_r^T(j)$), the composite index of intermediate demand for good j (i.e., $G_r^{NT}(j)$) and the composite price index of good j (i.e., $P_r^T(j)$) in each region as given in the optimization problem.

which implies that:

$$P_{r,r}^{T}(j) = \left(\frac{\eta(j)}{\eta(j)-1}\right) MC_{r}^{T}(j) \tag{2.9}$$
$$= \left(\frac{\eta(j)}{\eta(j)-1}\right) \frac{W}{A_{r}(j)}$$

where $\frac{\eta(j)}{\eta(j)-1}$ represents the gross mark-up. For the second line, we have used Equation 2.4 which implies that, for a specific good, the factory price of the product differs in each region only because of the region specific technology levels.

2.3. Bilateral Trade

We distinguish between disaggregate and aggregate level trade in our analysis. While the disaggregate level trade considers bilateral ratios of imports of a region for different varieties of a particular good, the aggregate level trade considers bilateral ratios of imports of different goods for a particular region.

2.3.1. Disaggregate Level Trade

By using Equations 2.1, 2.5 and 2.7, we obtain our key expression for the ratio of imports of region r across regions a and b, which is expressed by:

$$\frac{X_{r,a}(j)}{X_{r,b}(j)} = \frac{\theta_a}{\theta_b}\left(\frac{P_{b,b}^{T}(j)}{P_{a,a}^{T}(j)}\right)^{\eta(j)-1}\left(\frac{D_{r,b}}{D_{r,a}}\right)^{\delta(j)\eta(j)} \tag{2.10}$$

where $X_{r,k}(j) = \left(C_{r,k}^{T}(j) + G_{r,k}^{NT}(j)\right) P_{k,k}^{T}(j)$ is the value of total imports of region r from region k measured at the source for good j.[13] Equation 2.10 says that a region imports more goods (measured in values) from the lower price regions and fewer goods from the more distant regions.

[13] If we had an iceberg cost in our analysis, we would have had $X_{r,k}(j) = \left(C_{r,k}^{T}(j) + G_{r,k}^{NT}(j)\right) P_{k,k}^{T}(j)\left(1+\tau_{r,k}(j)\right)$ as the export income received by the exporter region for good j. However, this is not the case in the real world that distinguishes between the exporter sector and the transportation sector. For instance, our data set of Commodity Flow Survey includes only the export income received by the firms, not the transportation income.

Substituting Equation 2.9 into Equation 2.10 results in the general form of our estimated equations for our disaggregated level trade analysis:

$$\frac{X_{r,a}(j)}{X_{r,b}(j)} = \frac{\theta_a}{\theta_b}\left(\frac{A_a(j)}{A_b(j)}\right)^{\eta(j)-1}\left(\frac{D_{r,b}}{D_{r,a}}\right)^{\delta(j)\eta(j)} \tag{2.11}$$

Note that Equation 2.11 is an exogenous solution for our estimated disaggregate level trade expression, and thus, there is no need for any endogenous data such as income for the estimation *given* the technology levels. Moreover, it can easily be estimated in log terms.

2.3.2. Aggregate Level Trade

By using Equations 2.2 and 2.6, we obtain our key expression for the ratio of imports of region r in terms of goods j and k as follows:

$$\frac{X_r(j)}{X_r(k)} = \frac{\gamma_j}{\gamma_k}\left(\frac{P_r^T(j)}{P_r^T(k)}\right)^{1-\varepsilon} \tag{2.12}$$

where $X_r(m) = \left(C_r^T(m) + G_r^{NT}(m)\right) P_r^T(m) = \sum_i \left(\left(C_{r,i}^T(m) + G_{r,i}^{NT}(m)\right) P_{i,i}^T(j)\, D_{r,i}^{\delta(j)}\right)$ is the value of total imports of region r in terms of good m measured at the destination.[14] Equation 2.12 says that a region imports more (less) of a good which has a lower (higher) destination price.

Substituting Equations 2.7, 2.9, and $P_r^T(j) \equiv \left(\sum_i \beta_r \theta_i P_{r,i}^T(j)^{1-\eta(j)}\right)^{\frac{1}{1-\eta(j)}}$ into Equation 2.12 results in the general form of our estimated equations for our aggregate level trade analysis:

$$\frac{X_r(j)}{X_r(k)} = \frac{\gamma_j}{\gamma_k}\left(\frac{\left(\frac{\eta(j)}{\eta(j)-1}\right)\left(\sum_i \beta_r\theta_i D_{r,i}^{\delta(j)}\left(A_i(j)\right)^{\eta(j)-1}\right)^{\frac{1}{1-\eta(j)}}}{\left(\frac{\eta(k)}{\eta(k)-1}\right)\left(\sum_i \beta_r\theta_i D_{r,i}^{\delta(k)}\left(A_i(k)\right)^{\eta(k)-1}\right)^{\frac{1}{1-\eta(k)}}}\right)^{1-\varepsilon} \tag{2.13}$$

Note that Equation 2.13 is an exogenous solution for our estimated aggregate level trade expression, and thus, there is again no need for any endogenous data such as income for the estimation *given* the technology levels. Moreover, it can easily be estimated in log terms after estimating the disaggregate level expression given by Equation 2.11 (i.e., after obtaining estimates for $\eta(j)$'s and $\delta(j)$'s).

[14] Since the data set of Commodity Flow Survey includes only the export income received by the firms, we have to distinguish between the value of exports at the source and at the destination. We will control for these issues in our empirical analysis below.

3. Remarks and Estimation Methodology

We employ a two-step estimation process. First, we test the empirical power of our model at the disaggregate level and obtain estimates of elasticity of substitution across varieties of each good (i.e., $\eta(j)$'s), and good specific distance elasticities (i.e., $\delta(j)$'s) in the disaggregate level estimation, by which we obtain good specific price indices (i.e., $P_i(j)$'s). Second, we test the empirical power of our model at the aggregate level and obtain the elasticity of substitution across goods (i.e., ε).

3.1. Disaggregate Level Trade Estimation

In this subsection, we provide the implications of Equation 2.11, and we empirically test different log versions of it by using the Commodity Flow Survey (CFS) that covers bilateral interstate trade data within the U.S. The details of data are described in the Appendix.

Although Equation 2.11 holds on average, it doesn't hold for each bilateral trade ratio. In empirical terms, following Santos Silva and Tenreyro (2006), and Henderson and Millimet (2008), to address the unobservable nature of bilateral trade ratios, we assume that there is an error term associated with each ratio, which implies that:

$$\frac{X_{r,a}(j)}{X_{r,b}(j)} = \frac{\theta_a}{\theta_b}\left(\frac{A_a(j)}{A_b(j)}\right)^{\eta(j)-1}\left(\frac{D_{r,b}}{D_{r,a}}\right)^{\delta(j)\eta(j)} + \mu_{r,a,b,j}$$

where $E\left[\mu_{r,a,b,j}\left|\frac{\theta_a}{\theta_b}, \frac{A_a(j)}{A_b(j)}, \frac{D_{r,b}}{D_{r,a}}\right.\right] = 0$. This can be rewritten as:

$$\frac{X_{r,a}(j)}{X_{r,b}(j)} = \frac{\theta_a}{\theta_b}\left(\frac{A_a(j)}{A_b(j)}\right)^{\eta(j)-1}\left(\frac{D_{r,b}}{D_{r,a}}\right)^{\delta(j)\eta(j)} \upsilon_{r,a,b,j} \tag{3.1}$$

where

$$\upsilon_{r,a,b,j} = 1 + \frac{\mu_{r,a,b,j}}{\frac{\theta_a}{\theta_b}\left(\frac{A_a(j)}{A_b(j)}\right)^{\eta(j)-1}\left(\frac{D_{r,b}}{D_{r,a}}\right)^{\delta(j)\eta(j)}} \tag{3.2}$$

and $E\left[\upsilon_{r,a,b,j}\left|\frac{\theta_a}{\theta_b}, \frac{A_a(j)}{A_b(j)}, \frac{D_{r,b}}{D_{r,a}}\right.\right] = 1$. Taking the log of both sides in Equation 3.1 results in the following log-linear expression for the bilateral disaggregate level trade *ratios*:

$$\log\left(\frac{X_{r,a}(j)}{X_{r,b}(j)}\right) = \log\left(\frac{\theta_a}{\theta_b}\right) + (\eta(j)-1)\log\left(\frac{A_a(j)}{A_b(j)}\right) + \delta(j)\eta(j)\log\left(\frac{D_{r,b}}{D_{r,a}}\right) + \log(\upsilon_{r,a,b,j}) \tag{3.3}$$

To obtain a consistent estimator of the slope parameters by the Ordinary Least Squares (OLS), we assume that $E\left[\log\left(\upsilon_{r,a,b,j}\right)\middle|\frac{\theta_a}{\theta_b},\frac{A_a(j)}{A_b(j)},\frac{D_{r,b}}{D_{r,a}}\right]$ does not depend on the regressors.[15] Because of Equation 3.2, this condition is met only if $\mu_{r,a,b,j}$ can be written as follows:

$$\mu_{r,a,b,j}=\frac{\theta_a}{\theta_b}\left(\frac{A_a(j)}{A_b(j)}\right)^{\eta(j)-1}\left(\frac{D_{r,b}}{D_{r,a}}\right)^{\delta(j)\eta(j)}\xi_{r,a,b,j}$$

where $\xi_{r,a,b,j}$ is a random variable statistically independent of the regressors. In such a case, $\upsilon_{r,a,b,j}=1+\xi_{r,a,b,j}$ and therefore is statistically independent of the regressors, implying that $E\left[\log\left(\upsilon_{r,a,b,j}\right)\middle|\frac{\theta_a}{\theta_b},\frac{A_a(j)}{A_b(j)},\frac{D_{r,b}}{D_{r,a}}\right]$ is a constant. Following Santos Silva and Tenreyro (2006), we relax the assumption of $E\left[\log\left(\upsilon_{r,a,b,j}\right)\middle|\frac{\theta_a}{\theta_b},\frac{A_a(j)}{A_b(j)},\frac{D_{r,b}}{D_{r,a}}\right]$ not depending on the regressors, in our Sensitivity Analysis #4, below, by considering the Poisson Pseudo-Maximum Likelihood (PPML) estimator.

In our estimation, we run *only one* OLS (or PPML) regression for the pooled sample by including relevant dummy variables for each $\frac{\theta_a}{\theta_b}$ and $\delta(j)$ in Equation 3.3. Although $A_i(j)$'s are region and good specific technology levels in Equation 3.3, they don't necessarily capture all the source specific fixed effects. This is why source specific taste parameters (i.e., θ_i's) may play an important role in our estimation. For instance, in additional to the technology levels, source specific fixed effects *may* capture possible differences in source specific production markups, source specific production taxes, and so on. We test the validity of having both these fixed effects and technology levels at the same time in Version B and Version G of our empirical estimation, below.[16]

According to Equation 3.3, the following propositions are implied:

[15] It is well known that modeling zero interregional flows using a normal error process leads to problems. If the dependent variable cannot take a value below zero, then a normal error process is a poor approximation. Nevertheless, we don't have such a concern, because our log-linearized equation does have values below zero, by considering the (log) ratio of bilateral trade values.

[16] Multicollinearity is less of a problem in a cross-sectional analysis like ours that has a high sample size. The reasoning is that we run only one regression instead of good specific regressions; if we were running good specific regressions, then θ_i's and $A_i(j)$'s would have been perfectly correlated, because, in such a case, we would have good specific θ_i's. Moreover, the individual effects of technology and source specific taste parameters can both be assessed when there are sufficient number of observations of high technology regions with low fixed effects and low technology regions with high fixed effects. Besides, the theoretical consequences of multicollinearity is still a debate, because even if the multicollinearity is very high, the OLS estimators still remain to be the best linear unbiased estimators. The only possible problem arises due to having wide confidence intervals in the presence of

Proposition 1. *Both $\delta(j)$ and $\eta(j)$ can be identified in Equation 3.3 which is not the case in most gravity models (see Anderson and van Wincoop 2003; Hummels 1999, 2001; Wei 1996).*

P roof. The identification is realized via the technology levels which are usually ignored in gravity models. In particular, since both $(\eta(j) - 1)$ and $\delta(j)\eta(j)$ can be estimated by Equation 3.3, one can identify both $\eta(j)$ and $\delta(j)$ while also calculating their standard errors by employing the Delta method. ■

Proposition 2. *All the variables in Equation 3.3 are exogenous, which leaves an applied researcher free from a possible endogeneity problem. Moreover, there is no need for income data given the exogenous technology levels.*

P roof. The proof follows through Equation 2.11.[17] ■

Proposition 3. *Assuming that overstatement of a distance is proportional to the distance itself, the model controls for such an issue (because of the use of calculated distance measures such as great circle distances) as mentioned by Hillberry and Hummels (2001).*

P roof. Assuming that overstatement of a distance is proportional to the distance itself, the distance ratio in Equation 3.3 is not affected at all. See sensitivity analysis #3 in Section 4 for details. ■

Proposition 4. *By construction, the model is capable of controlling for the effects of local (i.e., wholesale and retail) distribution costs, insurance costs or local taxes, each of which are possible topics for separate debates in the literature (see Anderson and van Wincoop 2004).*

P roof. To see this, consider Equation 2.1 by including such possible good specific proportional costs. For instance, say that there is a proportional (net) cost of $\varphi(j)$ for good j in region

multicollinearity. However, by having very low confidence intervals, our estimation results below are robust to a possible multicollinearity problem. See Achen (1982) and Gujarati (1995) for more details.

[17]If trade leads to technology transfer, than technology may be correlated with past trade levels. And, if there are unobservables omitted that are serially correlated, then technology will be endogenous. Nevertheless, these are not issues in our case, because we have a static rather than a dynamic analysis. Moreover, we have unobservable (source specific, destination specific and bilateral specific) fixed effects in our analysis.

r. Then, it follows that:

$$C_{r,a}(j) = \theta_a \left(\frac{P_{r,a}(j)(1+\varphi(j))}{P_r(j)(1+\varphi(j))} \right)^{-\eta(j)} C_r(j)$$

and

$$C_{r,b}(j) = \theta_b \left(\frac{P_{r,b}(j)(1+\varphi(j))}{P_r(j)(1+\varphi(j))} \right)^{-\eta(j)} C_r(j)$$

The same logic applies for Equation 2.5, which together with the expressions above, implies exactly the same expression as in Equation 2.10. ■

Proposition 5. *By construction, the model is capable of controlling for the effects of intermediate input trade.*

P roof. Proof follows through the definition of $X_{r,k}(j) = \left(C^T_{r,k}(j) + G^T_{r,k}(j) \right) P^T_{k,k}(j)$ in Equation 2.10. ■

Under certain assumptions, the model is also capable of controlling for the effects of international trade. In particular, we assume that the international trade partners of the U.S. share similar tastes with the states in which the customs are located. Our justification comes from the fact that, in CFS, international export (import) shipments are included, with the domestic destination (source) defined as the U.S. port, airport, or border crossing of exit from the U.S. After this reasonable assumption, it follows that our estimated trade ratio given by Equation 3.3 is not affected at all by international trade, since the inclusion of international trade will be proportional in such a case.

By using the general form in Equation 3.3, we test several restricted versions of it along with its unrestricted version. These restrictions are not only important for econometric significance tests, but they are also important for economic intuition in terms of the contribution of each variable in Equation 3.3 to explain the interstate trade patterns. In particular, we test for the following versions of Equation 3.3:

Version A) Unrestricted version of Equation 3.3 in which we estimate $\boldsymbol{\eta}^{\mathbf{A}}$ (the vector consisting of $\eta(j)$'s), $\boldsymbol{\theta}^{\mathbf{A}}$ (the vector consisting of $\frac{\theta_a}{\theta_b}$'s), and $\boldsymbol{\delta}^{\mathbf{A}}$ (the vector consisting of $\delta(j)$'s) for all r, j, a and b. This is our benchmark equation by which we use $\frac{\theta_a}{\theta_b}$ values as fixed effects in our regression, by which we can estimate $\eta(j)$'s, by which we can estimate $\delta(j)\,\eta(j)$'s, and thus, by which we can obtain estimates of $\delta(j)$ and relative standard errors through the use of the Delta method.

Version B) Restricted version of Equation 3.3 in which $\theta_i = \theta$ for all i, thus, in which we estimate $\boldsymbol{\eta}^{\mathbf{B}}$ (the vector consisting of $\eta(j)$'s), and $\boldsymbol{\delta}^{\mathbf{B}}$ (the vector consisting of $\delta(j)$'s) for all j. Recall that in the unrestricted version of Equation 3.3, θ_i values serve as source specific fixed effects in the regression analysis. When $\theta_a = \theta_b$, it follows that $\log\left(\frac{\theta_a}{\theta_b}\right) = 0$. Thus, the purpose of this restricted version is that it helps us evaluate whether or not there are source specific fixed effects. This is also important in terms of testing our assumption of source specific taste parameters in our CES consumption/intermediate input functions. We can also see the contribution of these fixed effects in explaining the interstate trade patterns by comparing the results of this version with the results of version A through an additional restriction test.

Version C) Restricted version of Equation 3.3 in which $\delta(j) = \delta$ and $\eta(j) = \eta$ for all j, and thus, in which we estimate η, δ and $\boldsymbol{\theta}^{\mathbf{C}}$ (the vector consisting of $\frac{\theta_a}{\theta_b}$'s) for all r, a and b. The purpose of this restriction is that it helps us decide whether or not the trade costs and elasticities of substitution across varieties are good specific. This restriction is important, because most of the gravity type studies ignore good specific variations that affect the accuracy of the estimation results. Together with Version H, this restriction is also used to figure out whether or not the trade costs are good specific.

Version D) Restricted version of Equation 3.3 in which $\theta_i = \theta$ for all i; and in which $\delta(j) = \delta$ and $\eta(j) = \eta$ for all j; thus, in which we estimate η and δ. This restriction is used to test whether or not there are source specific taste parameters when there are common trade costs and common elasticity of substitution across varieties for different goods.

Version E) Restricted version of Equation 3.3 in which $\frac{\theta_r}{\theta_b} = \theta_H$ and $\frac{\theta_a}{\theta_b} = 1$ for all r, $a(\neq r), b(\neq r)$; and in which $\delta(j) = \delta$ and $\eta(j) = \eta$ for all j, thus, in which we estimate θ_H, η and δ. Since we make our analysis for a typical region r, $\frac{\theta_r}{\theta_b} = \theta_H$ and $\frac{\theta_a}{\theta_b} = 1$ together means that the goods purchased within a region are different from the goods imported from other regions, i.e., the so-called *home-bias*. Together with $\delta(j) = \delta$ and $\eta(j) = \eta$, the main purpose of this restriction is to find whether or not there is any home-bias, even at the intranational level, when trade costs and elasticities of substitution across varieties are the same across goods.

Version F) Restricted version of Equation 3.3 in which $\frac{\theta_r}{\theta_b} = \theta_H$ and $\frac{\theta_a}{\theta_b} = 1$ for all r, $a(\neq r), b(\neq r)$; thus, in which we estimate $\theta_H, \boldsymbol{\eta}^{\mathbf{F}}$ (the vector consisting of $\eta(j)$'s) and $\boldsymbol{\delta}^{\mathbf{F}}$ (the vector consisting of $\delta(j)$'s) for all j. This is the same as version E except that the trade costs are now good specific. Thus, the main purpose of this restriction is to find whether or not there is any home-bias, even at intranational level, when elasticities of substitution across varieties, and trade costs are good specific.

Version G) Restricted version of Equation 3.3 in which $A_a(j) = A_b(j)$ for all j (which is equivalent, since we talk about the ratios, saying that $A_i(j) = A$ for all i and j); thus, in which we estimate $\boldsymbol{\eta}^{\mathbf{G}}\boldsymbol{\delta}^{\mathbf{G}}$ (the vector consisting of $\eta(j)\,\delta(j)$'s) and $\boldsymbol{\theta}^{\mathbf{G}}$ (the vector consisting of $\frac{\theta_a}{\theta_b}$'s). The purpose of this restriction is to evaluate whether the technology levels are region specific or country specific.

Version H) Restricted version of Equation 3.3 in which $\eta(j) = \eta$ for all j, thus, in which we estimateη, $\boldsymbol{\theta}^{\mathbf{A}}$ (the vector consisting of $\frac{\theta_a}{\theta_b}$'s), and $\boldsymbol{\delta}^{\mathbf{A}}$ (the vector consisting of $\delta(j)$'s) for all r, j, a and b. The purpose of this restriction is that it helps us decide whether or not the elasticity of substitution across varieties is good specific. This restriction is important, because most of the gravity type studies ignore good specific $\eta(j)$'s which affect the accuracy of the estimation results.

3.2. Aggregate Level Trade Estimation

In this section, we introduce our methodology to estimate Equation 2.13. We empirically test it by using CFS data set and the estimation results of the disaggregate level trade estimation. Analogous to the disaggregate level trade equation, although Equation 2.13 holds on average, it doesn't hold for each bilateral trade ratio. Therefore, we assume that there is an error term associated with each ratio, which implies that:

$$\frac{X_r(j)}{X_r(k)} = \frac{\gamma_j}{\gamma_k}\left(\frac{\left(\frac{\eta(j)}{\eta(j)-1}\right)\left(\sum_i \beta_r\theta_i D_{r,i}^{\delta(j)}\left(A_i(j)\right)^{\eta(j)-1}\right)^{\frac{1}{1-\eta(j)}}}{\left(\frac{\eta(k)}{\eta(k)-1}\right)\left(\sum_i \beta_r\theta_i D_{r,i}^{\delta(k)}\left(A_i(k)\right)^{\eta(k)-1}\right)^{\frac{1}{1-\eta(k)}}}\right)^{1-\varepsilon} + \mu_{r,j,k}$$

where $E\left[\mu_{r,j,k}\left|\frac{\gamma_j}{\gamma_k},\frac{\left(\frac{\eta(j)}{\eta(j)-1}\right)\left(\sum_i\beta_r\theta_i D_{r,i}^{\delta(j)}(A_i(j))^{\eta(j)-1}\right)^{\frac{1}{1-\eta(j)}}}{\left(\frac{\eta(k)}{\eta(k)-1}\right)\left(\sum_i\beta_r\theta_i D_{r,i}^{\delta(k)}(A_i(k))^{\eta(k)-1}\right)^{\frac{1}{1-\eta(k)}}}\right.\right]=0$. This can be rewritten as:

$$\frac{X_r\left(j\right)}{X_r\left(k\right)}=\frac{\gamma_j}{\gamma_k}\left(\frac{\left(\frac{\eta(j)}{\eta(j)-1}\right)\left(\sum_i\beta_r\theta_i D_{r,i}^{\delta(j)}\left(A_i\left(j\right)\right)^{\eta(j)-1}\right)^{\frac{1}{1-\eta(j)}}}{\left(\frac{\eta(k)}{\eta(k)-1}\right)\left(\sum_i\beta_r\theta_i D_{r,i}^{\delta(k)}\left(A_i\left(k\right)\right)^{\eta(k)-1}\right)^{\frac{1}{1-\eta(k)}}}\right)^{1-\varepsilon}\upsilon_{r,j,k} \tag{3.4}$$

where

$$\upsilon_{r,j,k}=1+\frac{\mu_{r,j,k}}{\frac{\gamma_j}{\gamma_k}\left(\frac{\left(\frac{\eta(j)}{\eta(j)-1}\right)\left(\sum_i\beta_r\theta_i D_{r,i}^{\delta(j)}(A_i(j))^{\eta(j)-1}\right)^{\frac{1}{1-\eta(j)}}}{\left(\frac{\eta(k)}{\eta(k)-1}\right)\left(\sum_i\beta_r\theta_i D_{r,i}^{\delta(k)}(A_i(k))^{\eta(k)-1}\right)^{\frac{1}{1-\eta(k)}}}\right)^{1-\varepsilon}} \tag{3.5}$$

and $E\left[\upsilon_{r,j,k}\left|\frac{\gamma_j}{\gamma_k},\frac{\left(\frac{\eta(j)}{\eta(j)-1}\right)\left(\sum_i\beta_r\theta_i D_{r,i}^{\delta(j)}(A_i(j))^{\eta(j)-1}\right)^{\frac{1}{1-\eta(j)}}}{\left(\frac{\eta(k)}{\eta(k)-1}\right)\left(\sum_i\beta_r\theta_i D_{r,i}^{\delta(k)}(A_i(k))^{\eta(k)-1}\right)^{\frac{1}{1-\eta(k)}}}\right.\right]=1$. Taking the log of both sides in Equation 3.4 results in the following log-linear expression for the bilateral disaggregate level trade *ratios*:

$$\log\left(\frac{X_r\left(j\right)}{X_r\left(k\right)}\right)=\log\left(\frac{\gamma_j}{\gamma_k}\right)+\left(1-\varepsilon\right)\log\left(\frac{\left(\frac{\eta(j)}{\eta(j)-1}\right)\left(\sum_i\beta_r\theta_i D_{r,i}^{\delta(j)}\left(A_i\left(j\right)\right)^{\eta(j)-1}\right)^{\frac{1}{1-\eta(j)}}}{\left(\frac{\eta(k)}{\eta(k)-1}\right)\left(\sum_i\beta_r\theta_i D_{r,i}^{\delta(k)}\left(A_i\left(k\right)\right)^{\eta(k)-1}\right)^{\frac{1}{1-\eta(k)}}}\right)+\log\left(\upsilon_{r,j,k}\right) \tag{3.6}$$

To obtain a consistent estimator of the slope parameters by the OLS, we assume that $E\left[\log\left(\upsilon_{r,j,k}\right)\left|\frac{\gamma_j}{\gamma_k},\frac{\left(\frac{\eta(j)}{\eta(j)-1}\right)\left(\sum_i\beta_r\theta_i D_{r,i}^{\delta(j)}(A_i(j))^{\eta(j)-1}\right)^{\frac{1}{1-\eta(j)}}}{\left(\frac{\eta(k)}{\eta(k)-1}\right)\left(\sum_i\beta_r\theta_i D_{r,i}^{\delta(k)}(A_i(k))^{\eta(k)-1}\right)^{\frac{1}{1-\eta(k)}}}\right.\right]$ does not depend on the regressors. Because of Equation 3.5, this condition is met only if $\mu_{r,j,k}$ can be written as follows:

$$\mu_{r,j,k}=\frac{\gamma_j}{\gamma_k}\left(\frac{\left(\frac{\eta(j)}{\eta(j)-1}\right)\left(\sum_i\beta_r\theta_i D_{r,i}^{\delta(j)}\left(A_i\left(j\right)\right)^{\eta(j)-1}\right)^{\frac{1}{1-\eta(j)}}}{\left(\frac{\eta(k)}{\eta(k)-1}\right)\left(\sum_i\beta_r\theta_i D_{r,i}^{\delta(k)}\left(A_i\left(k\right)\right)^{\eta(k)-1}\right)^{\frac{1}{1-\eta(k)}}}\right)^{1-\varepsilon}\xi_{r,j,k}$$

where $\xi_{r,j,k}$ is a random variable statistically independent of the regressors. In such a case, $\upsilon_{r,j,k}=1+\xi_{r,j,k}$ and therefore is statistically independent of the regressors, implying that $E\left[\log\left(\upsilon_{r,j,k}\right)\left|\frac{\gamma_j}{\gamma_k},\frac{\left(\frac{\eta(j)}{\eta(j)-1}\right)\left(\sum_i\beta_r\theta_i D_{r,i}^{\delta(j)}(A_i(j))^{\eta(j)-1}\right)^{\frac{1}{1-\eta(j)}}}{\left(\frac{\eta(k)}{\eta(k)-1}\right)\left(\sum_i\beta_r\theta_i D_{r,i}^{\delta(k)}(A_i(k))^{\eta(k)-1}\right)^{\frac{1}{1-\eta(k)}}}\right.\right]$ is a constant. As in the disaggregate level analysis, for robustness, in addition to the OLS regression, we relax the assumption of $E\left[\log\left(\upsilon_{r,a,b,j}\right)\left|\frac{\gamma_j}{\gamma_k},\frac{\left(\frac{\eta(j)}{\eta(j)-1}\right)\left(\sum_i\beta_r\theta_i D_{r,i}^{\delta(j)}(A_i(j))^{\eta(j)-1}\right)^{\frac{1}{1-\eta(j)}}}{\left(\frac{\eta(k)}{\eta(k)-1}\right)\left(\sum_i\beta_r\theta_i D_{r,i}^{\delta(k)}(A_i(k))^{\eta(k)-1}\right)^{\frac{1}{1-\eta(k)}}}\right.\right]$ not depending on the regressors, by considering a PPML regression.

In our estimation, we run *only one* OLS (or PPML) regression for the pooled sample by including relevant dummy variables for each $\frac{\gamma_j}{\gamma_k}$ in Equation 3.6. After having the estimates for $\eta(j)$'s and $\delta(j)$'s coming from the disaggregate level estimation, we have data and parameters for everything in Equation 3.6 except for β_i's and θ_i's. In particular, β_i's cannot be estimated by our disaggregate level analysis, because they are cancelled out after considering trade ratios. Moreover, we cannot uniquely identify each and every θ_i in our disaggregate level analysis due to overidentification issues. Hence, we restrict ourselves to a special case in which $\beta_i = \theta_i = 1$ for all i in our aggregate level analysis.

Although calculated $P_i(j)$'s are region and good specific price levels in Equation 3.6, they don't necessarily capture all the good specific fixed effects, especially the actual preferences of the individuals for specific goods. This is why good specific taste parameters (i.e., γ_i's) may play an important role in our estimation. Below, we test the validity of having both these fixed effects and price levels at the same time.

4. Empirical Results

The empirical results for disaggregate and aggregate level trade estimations are given in the following subsections. Before we continue, we have to take care of one more issue: How should we include zero trade observations in our log-linear estimated equation? For the sensitivity of our analysis, we follow three different approaches: 1) Assume that zero (trade) observations are equal to one U.S. dollar's worth; 2) assume that zero (trade) observations are equal to one U.S. cent's worth; 3) ignore the zero (trade) observations.[18] Although the last one will be biased toward low elasticities of substitution compared to the other two, it is worth presenting it for the sake of sensitivity. Moreover, we also use that third approach to compare the effects of using great circle distances and actual CFS distances, which is mentioned by Hillberry and Hummels (2001). The estimation based on the first approach will be presented as the Benchmark Case,

[18] Unfortunately, we cannot employ a tobit specificition to account for the zeros, because we consider the trade ratios instead of the trade itself. In particular, when we have a zero trade observation, then either the numerator or the denominator of the left hand side of Equation 3.3 (or both) is equal to zero. This would make the trade ratio equal to either zero or infinity (or indeterminate), and thus, employing a tobit specification would not be plausible in log terms.

and the estimation based on the others will be presented as the Sensitivity Analyses.

4.1. Disaggregate Level Trade Estimation Results

The disaggregate level trade estimation results for the benchmark case (i.e., the first approach in which zero trade observations are set equal to one U.S. dollar's worth) are given in Table 1. Table 1 distinguishes between different versions of the estimated equation. Note that versions B,C,D,E, F, and H are all restricted versions of version A, and version G is a special case of version A. Thus, we can test for those restrictions and decide whether or not they are valid. The test results for these restrictions are given in Table 4. As is evident, all the restrictions are rejected according to our F-test results. This suggests that Version A, which is obtained through our model, is selected among all of our versions. This implies that:

- Source specific fixed effects are found to be significant in version B, which supports our assumption of source specific taste parameters in the utility function.
- Trade costs are found to be good specific in version C, which supports our assumption of good specific trade costs.
- Production technology for each good is found to be region specific in version G, which further supports our model.
- Elasticity of substitution across varieties is found to be good specific in version H, which supports our disaggregate level model.

[Tables 1-4 are about here]

As is evident by Version A in Table 1, the elasticity of substitution across regions is estimated as 5.24 on average.[19] The disaggregate level estimates are given in Table 2. Since the intranational studies within the U.S. such as Wolf (2000), Hillberry and Hummels (2001), and Millimet and Thomas (2007) use gravity equations, they cannot estimate for the elasticity of

[19] Also note that our estimates are highly significant. For robustness, we have also considered Moulton standard errors, and our (t-test) results are almost the same. These results are available upon request. See Moulton (1986) and Donald and Lang (2007) for the details of Moulton standard errors.

substitution and the elasticity of distance at the same time. So, we compare our results with the ones in empirical international trade literature and see that our estimates for the elasticity of substitution are lower on average. In particular, Hummel's (2001) estimates range between 4.79 and 8.26; the estimates of Head and Ries (2001) range between 7.9 and 11.4; the estimate of Baier and Bergstrand (2001) is about 6.4; Harrigan's (1993) estimates range from 5 to 10; Feenstra's (1994) estimates range from 3 to 8.4; the estimate by Eaton and Kortum (2002) is about 9.28; the estimates by Romalis (2007) range between 6.2 and 10.9; the (mean) estimates of Broda and Weinstein (2006) range between 4 and 17.3. This difference may be due to the distinction between intranational and international data sets as well as the ignored factors in the literature such as local distribution costs, insurance costs, local taxes and intermediate input trade. Since our model controls for all of these factors, we claim that we have more accurate results intranationally. Someone may claim that the difference between our estimates and the estimates in the literature may also be due to our inclusion of zero trade observations; however, as we will show in our sensitivity analyses, the difference between our estimates and the ones in the literature gets higher when we ignore zero trade observations which is what the studies mentioned above actually do.

According to Version A in Table 1, the distance elasticity is estimated as 0.60 on average. The disaggregate level estimates are given in Table 3. These numbers are higher than the distance elasticity estimates found by the literature, which are about 0.3 (see Hummels, 2001; Limao and Venables, 2001; Anderson and van Wincoop, 2004). This difference is most probably again due to using different frameworks or data sets, as well as due to our inclusion of zero (trade) observations into our analysis. We are going to check for the latter possibility in our sensitivity analysis. Another possible explanation for the difference between our distance elasticity estimates and the ones in the literature may be the mode of transportation for interstate trade. In particular, it may well be the case that the interstate trade is done by air through couriers like UPS, FedEX, and so forth, while the international trade is done in transportation modes different from those. We will also check for this possibility by considering different distance measures in our sensitivity analysis. Another reason may be the usual assumption of iceberg transport costs in the literature. As can be shown, if we had used that assumption instead of having a transportation sector, our distance elasticities would have had lower estimates.[20] However, since our data set of CFS provides only the income received by the exporter firms (and excludes transportation income), we prefer to distinguish between the exporter income and the transporter income, which is against the iceberg cost assumption.

Although version A (implied by our model) is selected among all estimated versions by our restriction tests, we can still have inference from other versions. Note that versions E and F represent the cases by which we can analyze whether or not there is a home-bias. Again according to Table 1, the values for θ_H are positive and significant, which according to our definitions for versions E and F, suggest that there is a home-bias across the states of the U.S. This bias is estimated as 5.73 by Equation E and 2.25 by Equation F. However, since Equation E is a restricted version of Equation F, we can test this restriction. We find that the restriction is rejected, which means that a home-bias of 2.25 is more plausible compared to 5.73. In particular, a typical state has a taste parameter θ for locally produced goods about 2.25 times more than imported goods. This number is very close to the intranational home-bias estimated by Hillberry and Hummels (2003) which is $exp(0.99) = 2.69$. Therefore, although the literature overestimates the elasticity of substitution measures and underestimates the elasticity of distance measures with respect to our results, the measures of home-bias seem to be similar. One explanation is due to the interaction between the two elasticity measures in Equation 3.3. In particular, if two elasticity measures operate in opposite signs (i.e., if one is overestimated and the other is underestimated), then the results for the fixed effects captured by θ values are not affected too much since two estimation errors cancel each other out to some degree.

Finally, the high adjusted R^2 value of 0.42 for Equation A also supports our model. Although version A (implied by our model) is selected among all estimated versions by our significance tests, we can still compare the contribution of each variable in Equation 3.3 in explaining interstate trade patterns by considering the adjusted R^2 values of each version. In particular, we see that the highest difference of adjusted R^2 values takes place between versions A and D&E, which means that source specific fixed effects and good specific trade costs together play an important role in our estimations. The second highest difference of adjusted R^2 values takes

[20] It can be shown easily that the average $\delta(j)$ estimate given in Table 1 (i.e.,0.59) would be replaced by 0.43 under the iceberg cost assumption.

place between versions A and B&F, which means that source specific fixed effects are significant individually. The third highest difference of adjusted R^2 values takes place between versions A and C, which means that good specific trade costs are also significant individually. Finally, the lowest difference of adjusted R^2 values takes place between versions A, G and H, which means that good and region specific technology parameters and elasticities of substitution across goods, besides the source specific fixed effects, play a lesser role compared to other parameters, which makes sense since we are making our analysis within a highly integrated economy, the U.S.

4.1.1. Sensitivity Analyses

In order to support our empirical results, in this section, we employ four sensitivity analyses. The first two are related to zero (trade) observations, the third one is related to distance measures, and the last one is related to a possible biasedness of the OLS estimator in log-linearized models.

Sensitivity Analysis #1 We start our sensitivity analysis by setting zero (trade) observations equal to one U.S. cent's worth. In such a case, the estimation results in Table 1 are replaced by the ones in Table 5. Note that we can again test for the restrictions of versions B, C, D, E, F, G, and H with respect to version A. The test results for these restrictions are given in Table 6. As is evident, all the restrictions are again rejected according to our F-test results. This suggests that version A is again selected among all of our equations. The high adjusted R^2 value of 0.40 for Equation A again supports our model.

[Tables 5-6 are about here]

As is evident by Version A in Table 5, the elasticity of substitution is estimated as 6.27 on average. The disaggregate level estimates are given again in Table 2. Although these values are slightly higher than the ones in our benchmark case, they are still lower than the estimates in the literature on average.[21]

The distance elasticity is estimated as 0.61 on average, which is very close to our initial estimate in Table 1, yet higher than the ones in the literature. Moreover, the disaggregated

[21] Even if we set zero trade observations equal to 0.01 U.S. cent worth, the elasticity of substitution is estimated as 6.50 on average.

distance elasticities given in Table 3 are very close to the ones that we estimated initially.

Again according to Table 5, the values for θ_H are positive and significant, which according to our definitions for versions E and F, suggest that there is a home-bias across the states of the U.S. After a restriction analysis between versions E and F, the restriction in E is again rejected. Thus, a typical state has a taste parameter θ for locally produced goods about 1.95 times more than imported goods. This number is close to our initial estimate of 2.25.

Sensitivity Analysis #2 For our second sensitivity analysis, we ignore the zero (trade) observations. In such a case, the estimation results in Table 1 are replaced by the ones in Table 7. We again test for the restrictions of versions B, C, D, E, F, G, and H with respect to version A. The test results for these restrictions are given in Table 8. As is evident, all the restrictions are again rejected according to our F-test results. This suggests that version A is again selected among all of our equations. The high adjusted R^2 value of 0.60 for Equation A again supports our model.

[Tables 7-8 are about here]

This time, according to Version A in Table 7, the elasticity of substitution is estimated as 2.70 on average. The disaggregate level estimates are given again in Table 2. These estimates are very low compared to the studies mentioned above even though they also ignore zero trade observations (as in this subsection). We had given possible explanations for this difference above, so we won't repeat them here.

The distance elasticity is estimated as 0.37 on average. The disaggregate level estimates are again given in Table 3. Although these numbers are closer to the distance elasticity estimates in the literature (that we mentioned above, which are about 0.3), they are still higher. Thus, the difference between our first two estimates of distance elasticities (i.e., our initial estimate and our first sensitivity analysis) and the estimates in the literature can, to some degree, be explained by the fact that we have included zero (trade) observations in our first two estimations. Nevertheless, the difference doesn't disappear completely.

According to Table 7, the values for θ_H are again positive and significant, which according to our definitions for Equations E and F, suggest that there is a home-bias across the states of the U.S. In particular, a typical state has a taste parameter θ for locally produced goods about

1.93 times more than imported goods after testing for the restriction between Equations E and F and rejecting it. This number is lower compared to our initial estimates and the estimates of Hillberry and Hummels (2003).

Sensitivity Analysis #3 As we detail in the Appendix, until now, we have used great circle distances instead of actual CFS distances, because average distance measures are not provided by CFS for zero (trade) observations. However, as is shown by Hillberry and Hummels (2001), using great circle distances, instead of actual distances provided by CFS, may overstate the distance measure as in Wolf (2000). In one of our propositions, we had claimed that we already control for this issue by taking the ratio of imports as our dependent variable. Moreover, the coefficient of correlation between the great circle distances and actual distances provided by CFS is calculated as 0.98, after ignoring zero trade observations. Nevertheless, as our third sensitivity analysis, we repeat our sensitivity analysis #2, this time by using the average distance measure provided by CFS instead of the great circle distance measure that we have used until now. In this way, we can compare the effects of great circle distances and the CFS distances on our empirical results.

When we use the CFS distances, the estimation results of sensitivity analysis #2 given in Table 7 are replaced by the ones in Table 9. We again test for the restrictions of versions B, C, D, E, F, G, and H with respect to version A. The test results for these restrictions are given in Table 10. As is evident, all the restrictions are again rejected according to our F-test results. This suggests that version A is again selected among all of our equations. The high adjusted R^2 value of 0.60 for Equation A again supports our model.

[Tables 9-10 are about here]

As is evident by Version A in Table 9, the elasticity of substitution is estimated as 2.74 on average. The disaggregate level estimates are given in Table 2. The distance elasticity is estimated as 0.38 on average. The disaggregate level estimates are given in Table 3. All of these estimates are very close to the ones presented for Sensitivity Analysis #2.

According to Table 9, the values for θ_H are again positive and significant, which according to our definitions for Equations E and F, suggest that there is a home-bias across the states of the U.S. In particular, a typical state has a taste parameter θ for locally produced goods about

2.03 times more than imported goods after testing for the restriction between Equations E and F, and rejecting it. Although this number is close to our initial estimates, it is slightly higher compared to Table 7.

Overall, if we compare the numbers in Table 7 and Table 9, we see that they don't change significantly. This result supports our claim that we already control for overstating distances mentioned by Hillberry and Hummels (2001).

Sensitivity Analysis #4 For our last sensitivity analysis, we repeat our analysis for the benchmark case and the first three sensitivity analyses by using the Poisson Pseudo-Maximum Likelihood (PPML) estimator. As Santos Silva and Tenreyro (2006), and Henderson and Millimet (2008) suggest, under heteroskedasticity, the parameters of log-linearized models estimated by OLS may lead to biased estimates; thus, PPML should be used. To show this, Equation 3.3 can be written as follows:

$$\frac{X_{r,a}(j)}{X_{r,b}(j)} = \exp\left(\log\left(\frac{\theta_a}{\theta_b}\right) + (\eta(j) - 1)\log\left(\frac{A_a(j)}{A_b(j)}\right) + \delta(j)\,\eta(j)\log\left(\frac{D_{r,b}}{D_{r,a}}\right)\right)\upsilon_{r,a,b,j} \quad (4.1)$$

Assuming $E[\upsilon_{r,a,b,j}|\frac{\theta_a}{\theta_b}, \frac{A_a(j)}{A_b(j)}, \frac{D_{r,b}}{D_{r,a}}] = 1$, then Equation 4.1 may be estimated consistently using the Poisson Pseudo-Maximum Likelihood (PPML) estimator (see Santos Silva and Tenreyro, 2006). Since Version A has been chosen among all versions earlier, we repeat our analysis only for Version A here. The results are given in Table 11.

[Table 11 is about here]

As is evident, the (average) elasticity of substitution across varieties η ranges between 2.17 and 2.63, and the (average) elasticity of distance δ ranges between 0.30 and 0.59, which are both consistent with our earlier claim that our η estimates are lower than and our δ estimates are higher than the ones in the international trade literature.

4.2. Aggregate Level Trade Estimation Results

The aggregate level trade estimation results are given in Table 12. As we have done for the disaggregate level analysis, we consider four different approaches with two different estimation methods, OLS and PPML. As is evident, the elasticity of substitution across goods is estimated as 1.38 by OLS (2.19 by PPML) in our benchmark case, which is the one that sets zero trade observations equal to one U.S. dollar's worth. When zero trade observations are set equal to one U.S. cent's worth, the elasticity of substitution across goods is estimated as 1.27 by OLS (2.19 by PPML). When zero trade observations are ignored, it is estimated as 1.95 by OLS (2.44 by PPML). Finally, when CFS distance measures are used instead of great circle distances, it is estimated as 1.92 by OLS (2.97 by PPML).

[Table 12 is about here]

Although the estimates of ε are lower than the elasticity of substitution across varieties estimates (i.e., $\eta(j)$'s), as expected, according to OLS estimator, they are very close to each other according to PPML estimator. This result is consistent with the view that when goods are aggregated, the elasticity of substitution across them decreases. Nevertheless, these numbers are significantly lower than the estimates in the literature that we have discussed above. As in our disaggregate level analysis, we claim that this difference may be due to distinction between intranational and international data sets as well as the ignored factors in the literature such as local distribution costs, insurance costs, local taxes and intermediate input trade. Since our model controls for all of these factors, we claim that we have more accurate results intranationally. Our results are supported by several sensitivity analyses with high explanatory powers.[22]

5. Conclusions

We have written a partial equilibrium model to find motivations for bilateral trade ratios across regions. In particular, we have shown that a region imports more goods from the higher technology regions and fewer goods from the more distant regions, subject to an elasticity of substitution across varieties. Moreover, a region imports more of a good, of which price is lower, subject to an elasticity of substitution across goods. As we have explained in detail in the text, our model has several empirical and analytical benefits compared to the gravity models. Thanks

[22] We have also tested different restricted versions of Equation 3.6, such as common γ's or common $P_i(j)$'s, in our aggregate level analysis. We find that none of the restrictions are valid, and therefore Equation 3.6 is selected among all versions. These restriction test results are available upon request.

to the disaggregate (state) level data set combined from the Commodity Flow Survey and the U.S. Census Bureau, we are also able to show that our simple model is capable of explaining the interstate trade patterns within the U.S. In particular, we show that the elasticity of substitution measures are overestimated in the literature, while the elasticity of distance measures (thus, trade costs) are underestimated in the literature relative to our estimates.

We have shown that source specific fixed effects and good specific taste parameters are important for bilateral trade patterns, which are usually ignored in the literature. We have also shown that elasticities of substitution across varieties, and trade costs are good specific, which is not a considered fact in most of the aggregate level gravity type studies. Moreover, production technology for each good is found to be region specific rather than country specific. Our sensitivity analyses support our results.

The best strategy for possible future research would be to extend the model of this paper toward explaining international trade patterns. Such an analysis would be more convenient with a general equilibrium framework, although a partial equilibrium framework was good enough for this paper after assuming factor mobility for the production of traded goods.

References

[1] Achen, C.H., (1982), "Interpreting and Using Regression", Sage Publications, Beverly Hills, California.

[2] Alvarez, Fernando and Lucas, Robert Jr., (2007), "General equilibrium analysis of the Eaton-Kortum model of international trade," *Journal of Monetary Economics*, 54(6): 1726-1768.

[3] Anderson, J.E., (1979), "A Theoretical Foundation for the Gravity Equation", *American Economic Review*, 69(1): 106-116.

[4] Anderson, J.E., and Wincoop, E.V., (2003), "Gravity with Gravitas: A Solution to the Border Puzzle", *American Economic Review*, 93(1): 170-192.

[5] Anderson, J.E., and Wincoop, E.V., (2004), "Trade Costs", *Journal of Economic Literature*, 42: 691-751.

[6] Bergstrand, J.H., (1985), "The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence", *The Review of Economics and Statistics*, 67(3): 474-481.

[7] Bergstrand, J.H., (1989), "The Generalized Gravity Equation, Monopolistic Competition, and the Factor-Proportions Theory in International Trade", *The Review of Economics and Statistics*, 71(1): 143-153.

[8] Berry, W.D., Fording, R.C., and Hanson, R.L., (2003), "Cost of Living Index for the American States, 1960-2003", Inter-University Consortium for Political and Social Research (ICPSR) Study No: 1275.

[9] Broda, C., and Weinstein, D.E., (2006), "Globalization and the Gains from Variety", *Quarterly Journal of Economics*, 121(2): 541-585.

[10] Corsetti, G., and Pesenti, P., (2005), "International Dimensions of Optimal Monetary Policy", *Journal of Monetary Economics*, 52: 281-305.

[11] Deardorff, A. V. (1984): "Testing Trade Theories and Predicting Trade Flows," in Handbook of International Economics, Volume I, ed. by R. Jones and P. Kenen. Amsterdam: North-Holland.

[12] Dixit, Avinash K. and Stiglitz, Joseph E., (1977), "Monopolistic Competition and Optimal Product Diversity", *American Economic Review*, 67(3): 297-308.

[13] Donald, S.G. and Lang, K., (2007), "Inference with Difference in Differences and Other Panel Data", 89(2): 221-233

[14] Dornbusch, R., Fischer, S., and Samuelson, P.A., (1977), "Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods", *The American Economic Review*, 67(5): 823-839. .

[15] Dornbusch, R., Fischer, S., and Samuelson, P.A., (1980), "Heckscher-Ohlin Trade Theory with a Continuum of Goods", *Quarterly Journal of Economics*, 95(2): 203-224.

[16] Eaton, J., and Kortum, S., (2002), "Technology, Geography and Trade", *Econometrica*, 70(5): 1741-1779.

[17] Erceg, C.J., Henderson, D.W., and Levin A.T., (2000), "Optimal Monetary Policy with Staggered Wage and Price Contracts", *Journal of Monetary Economics*, 46: 281-313.

[18] Evenett, S. J., and W. Keller (2002): "On Theories Explaining the Success of the Gravity Equation," *Journal of Political Economy*, 110, 281–316.

[19] Gali, J. and Monacelli, T., (2005), "Monetary Policy and Exchange Rate Volatility in a Small Open Economy", *Review of Economic Studies*, 72: 707-734.

[20] Gallaway, Michael P. , McDaniel, Christine A., and Rivera, Sandra A., (2003), "Short-run and long-run industry-level estimates of U.S. Armington elasticities", *North American Journal of Economics and Finance* 14: 49–68

[21] Grossman, G. M., and E. Helpman (1995): "Technology and Trade," in The Handbook of International Economics, Volume III, ed. by G. M. Grossman and K. Rogoff. Amsterdam: North-Holland.

[22] Grossman, G., (1998), "Comment on Deardorff", in The Regionalization of the World Economy. J.A. Frankel, ed. Chicago: U. Chicago Press, pp. 33–57.

[23] Gujarati, D.N., (1995), "Basic Econometrics", McGraw-Hill, Singapore.

[24] Harrigan, J., (1996), "Openness to Trade in Manufactures in the OECD", *Journal of International Economics*, 40: 23-39.

[25] Head, K. and Ries, J., (2001), "Increasing Returns versus National Product Differentiation as an Explanation for the Pattern of U.S.-Canada Trade", *American Economic Review*, 91(4): 858-876.

[26] Head, K. and Mayer, T., (2001), "Illusory Border Effects: Distance Mismeasurement Inflates Estimates of Home Bias in Trade", CEPII Working Papers 2002-01.

[27] Helpman, E. (1987): "Imperfect Competition and International Trade: Evidence from Fourteen Industrial Countries," *Journal of Japanese and International Economics*, 1: 62–81.

[28] Helpman, E., Melitz, M. and Rubinstein, Y., (2008), "Estimating Trade Flows: Trading Partners and Trading Volumes", *Quarterly Journal of Economics,* 123(2): 441–487.

[29] Henderson, D.J. and Millimet, D.L., (2008), "Is gravity linear?," *Journal of Applied Econometrics*, 23(2): 137-172.

[30] Hillberry, R., and Hummels D., (2003), "Intranational Home Bias: Some Explanations", *The Review of Economic and Statistics*, 85(4):1089-1092.

[31] Hummels, D., (1999), "Have International Transportation Costs Declined?", *mimeo.*

[32] Hummels, D., (2001), "Toward a Geography of Trade Costs", *mimeo.*

[33] Hummels, D., and J. Levinsohn (1995):"Monopolistic Competition and International Trade: Reconsidering the Evidence," *Quarterly Journal of Economics*, 110: 799–836.

[34] Krugman, P., (1980), "Scale Economies, Product Differentiation, and the Pattern of Trade", *American Economic Review,* 70(5): 950-959.

[35] Krugman, P., (1991), "Increasing Returns and Economic Geography", *Journal of Political Economy*, 99(3): 483-499.

[36] Krugman, P. R. (1995): "Increasing Returns, Imperfect Competition, and the Positive Theory of International Trade," in The Handbook of International Economics, Volume III, ed. by G. M. Grossman and K. Rogoff. Amsterdam: North-Holland.

[37] Krugman, P. R., and A. J. Venables (1995): "Globalization and the Inequality of Nations," *Quarterly Journal of Economics*, 110: 857–880.

[38] Lopez, Rigoberto A., Pagoulatos, Emilio, Gonzalez, Maria A., (2006), "Home bias and U.S. imports of processed food products", *North American Journal of Economics and Finance* 17: 363–373

[39] Matsuyama, K., (2000), "A Ricardian Model with a Continuum of Goods under Nonhomothetic Preferences: Demand Complementarities, Income Distribution, and North-South Trade", *Journal of Political Economy,* 108(6): 1093-1120.

[40] Millimet, D., Osang, T., (2007), "Do state borders matter for U.S. intranational trade? The role of history and internal migration", *Canadian Journal of Economics*, 40(1): 93-126(34).

[41] Moulton, Brent R., (1986), "Random group effects and the precision of regression estimates," *Journal of Econometrics*, 32(3): 385-397.

[42] Romalis, J., (2007), "NAFTA's and CUSFTA's Impact on International Trade", *The Review of Economics and Statistics*, 89(3): 416-435

[43] Rossi-Hansberg, E., (2005), "A Spatial Theory of Trade", *American Economic Review*, 95(5): 1464-1491.

[44] Santos Silva, J. M. C. and Tenreyro, S., (2006). "The Log of Gravity," *The Review of Economics and Statistics*, 88(4): 641-658.

[45] Wei, SJ, (1996), "Intranational versus International Trade: How Stubborn are Nations in Global Integration?", NBER Working Paper No: 5531.

[46] Wolf, H., (2000), "Intra-National Home Bias in Trade", *Review of Economics and Statistics*, 82(4): 555-563.

Appendix A - The Data and Descriptive Analysis

For the bilateral trade analysis, we use the state-level Commodity Flow Survey (CFS) data obtained from the Bureau of Transportation Statistics for the United States for the year 2002. In particular, we use bilateral interstate trade data for the 2-digit Standard Classification of Transported Goods (SCTG) commodities of which codes are given in Table A.1 and of which names are given in the first column of Table 2 in the respective order.

Table A.1 - Crosswalk Between NAICS and SCTG

SCTG	NAICS	SCTG	NAICS	SCTG	NAICS
2	311 – 312*	17	324	31	327
3	311 – 312*	18	324	32	331 – 324*
4	311 – 312*	19	324 – 325*	33	332
5	311 – 312*	20	325	34	333
6	311 – 312*	21	325	35	334 – 335*
7	311 – 312*	22	325	36	336
8	311 – 312*	23	325	37	336
9	311 – 312*	24	326	38	334
11	212**	26	321	39	337
12	212**	27	322	40	339
13	212**	28	322	41	313 – 331*
14	212**	29	323	43	MIX OF ALL
15	212**	30	313		

Notes: The source is National Transportation Library of the Bureau of Transportation Statistics. * means that an average of the relevant NAICS industries has been used to obtain technology levels. ** means that there is no corresponding production data for that specific NAICS industry in the U.S. Census Bureau data set; thus, we assume that the technology levels are the same across states for those industries. Finally, SCTG 43 corresponds to mixed freight for which an average of all other NAICS industries in the table are used to obtain technology levels.

The CFS captures data on shipments originating from select types of business establishments located in all states of the U.S. However, because of data availability, we exclude Alaska, District of Columbia and Hawaii from our analysis. In CFS, shipments traversing the U.S. from a foreign location to another foreign location (e.g., from Canada to Mexico) are not included, nor are shipments from a foreign location to a U.S. location. Shipments that are shipped through a foreign territory with both the origin and destination in the U.S. are included in the CFS data. The mileages calculated for these shipments exclude the international segments (e.g., shipments from New York to Michigan through Canada do not include any mileages for Canada). International export (import) shipments are also included in CFS, with the domestic destination (source) defined as the U.S. port, airport, or border crossing of exit from the U.S.

In order to obtain the technology levels, we first use an approximate crosswalk between 3-digit North American Industry Classification System (NAICS) and 2-digit SCTG obtained from the National Transportation Library of the Bureau of Transportation Statistics. This crosswalk is given in Table A.1. After that, we use $A_i(j) = \log\left(\frac{V_i(j)}{P_i L_i(j)}\right)$ as our measure for the technology levels, where $V_i(j)$ is the industry/region specific value added; P_i is the cost of living index for state i borrowed from Berry et al. (2003); and $L_i(j)$ is the industry/region specific hours of labor supplied by the production workers. For the value added of each NAICS industry in each state, we use the state level U.S. Census Bureau data for the relevant industries in 2002.[23]

For distance measures, we calculate great circle distance between states by using latitudes and longitudes of capital cities of each state published by U.S. Census Bureau. Note that we don't use the average distance measures given by CFS in our initial analysis, because those measures are available only for realized trade observations. Since we consider zero (trade)observations in our analysis, we use the great circle distance measures that are not included in CFS. Moreover, because we use the ratio of imports of a region (and thus, the ratio of distances), we already control for a possible issue of overstating the distance measures mentioned by Hillberry and Hummels (2001). Nevertheless, we compare our estimation result obtained by great circle

[23] Although we use value added for each industry to calculate technology levels, this should not be necessary the case if we already had a better measure of technology. In other words, our claim in the text saying "We don't need any income data *given* the technology levels" still holds. Although the state-level production functions typically include public and private capital, to be consistent with the model, technology is defined on the basis of value added by labor.

distances and with the one obtained by CFS distances in our sensitivity analysis #3 in the text.

Table 1 - OLS Estimation Results

	Equation							
	A	B	C	D	E	F	G	H
η	[5.24]	[1.16]	5.31	1.12	1.08	[1.13]	–	5.24
	[(0.08)]	[(0.16)]	(0.06)	(0.03)	(0.03)	[(0.17)]	–	(0.03)
δ	[0.60]	[2.76]	0.59	2.85	2.87	[2.78]	–	[0.60]
	[(0.03)]	[(0.13)]	(0.01)	(0.02)	(0.03)	[(0.14)]	–	[(0.29)]
θ	$\boldsymbol{\theta}^{\mathbf{A}}$	–	$\boldsymbol{\theta}^{\mathbf{C}}$	–	5.73	2.25	$\boldsymbol{\theta}^{\mathbf{G}}$	$\boldsymbol{\theta}^{\mathbf{H}}$
	$(\cdot)$	–	$(\cdot)$	–	(0.29)	(0.22)	$(\cdot)$	$(\cdot)$
$\delta\eta$	–	–	–	–	–	–	$\eta\boldsymbol{\delta}^{\mathbf{G}}$	–
	–	–	–	–	–	–	$(\cdot)$	–
R-bar sqd.	0.42	0.31	0.37	0.26	0.26	0.31	0.41	0.41

Notes: The standard errors calculated by Delta method are in parenthesis. The sample size for all estimations is 47,819 which is found after considering the independent observations (ratios) and after ignoring the missing observations. The average of the estimated vectors of $\eta^{\mathbf{A}}$ and $\delta^{\mathbf{A}}$are given in brackets of which full vectors are given in Table 2. The estimated vectors of $\theta^{\mathbf{A}}$, $\theta^{\mathbf{C}}$, $\theta^{\mathbf{G}}$ and $\theta^{\mathbf{H}}$ (all having a size of 505) are omitted to save space. For equations E and F, θ corresponds to θ_H. For Equations A-F, the estimates for $\delta\eta$ are omitted since the estimates for δ and η are already given separately.

Table 2 - Estimated Vectors of Elasticity of Substitution across Varieties $(\boldsymbol{\eta}^{\mathbf{A}})$

	Benchmark	SA#1	SA#2	SA#3
Cereal grains	5.01	6.02	2.22	2.32
	(0.06)	(0.06)	(0.18)	(0.19
Other agricultural products	5.05	6.04	2.57	2.45
	(0.06)	(0.07)	(0.09)	(0.10)
Animal feed and products of animal origin	5.01	5.99	2.46	2.57
	(0.06)	(0.07)	(0.10)	(0.10)
Meat, fish, seafood, and their preparations	5.54	6.65	2.60	2.64
	(0.06)	(0.06)	(0.07)	(0.07)
Milled grain products and bakery products	5.14	6.16	2.69	2.74
	(0.06)	(0.07)	(0.07)	(0.07)
Other prepared foodstuffs and fats and oils	5.23	6.27	2.61	2.62
	(0.06)	(0.07)	(0.06)	(0.06)
Alcoholic beverages	5.25	6.29	2.81	2.83
	(0.06)	(0.06)	(0.08)	(0.08)
Tobacco products	5.28	6.33	3.08	3.19
	(0.06)	(0.06)	(0.14)	(0.13)
Natural sands	5.25	6.29	2.55	2.44
	(0.06)	(0.06)	(0.12)	(0.13)
Gravel and crushed stone	5.22	6.25	2.78	2.85
	(0.06)	(0.06)	(0.10)	(0.10)
Nonmetallic minerals	5.04	6.01	2.84	2.90
	(0.06)	(0.07)	(0.10)	(0.10)
Metallic ores and concentrates	5.23	6.27	3.09	3.38
	(0.06)	(0.06)	(0.19)	(0.17)
Coal	5.07	6.06	1.65	1.90
	(0.06)	(0.06)	(0.19)	(0.18)

Table 2 - Estimated Vectors of Elasticity of Substitution across Varieties $(\boldsymbol{\eta}^{\mathbf{A}})$

	Benchmark	SA#1	SA#2	SA#3
Gasoline and aviation turbine fuel	5.13	6.14	3.09	3.12
	(0.06)	(0.06)	(0.09)	(0.09)
Fuel oils	5.10	6.10	2.69	2.87
	(0.06)	(0.06)	(0.10)	(0.11)
Coal and petroleum products	5.14	6.15	2.59	2.74
	(0.06)	(0.07)	(0.09)	(0.09)
Basic chemicals	5.32	6.38	2.54	2.55
	(0.06)	(0.07)	(0.07)	(0.07)
Pharmaceutical products	5.71	6.87	2.69	2.71
	(0.06)	(0.06)	(0.07)	(0.07)
Fertilizers	5.08	6.08	4.07	4.08
	(0.06)	(0.07)	(0.09)	(0.11)
Chemical products and preparations	5.16	6.18	2.65	2.66
	(0.07)	(0.07)	(0.07)	(0.07)
Plastics and rubber	5.16	6.17	2.72	2.73
	(0.07)	(0.07)	(0.06)	(0.06)
Wood products	4.98	5.95	2.65	2.67
	(0.07)	(0.07)	(0.06)	(0.06)
Pulp, newsprint, paper, and paperboard	5.74	6.90	2.72	2.74
	(0.06)	(0.06)	(0.06)	(0.06)
Paper or paperboard articles	5.21	6.23	2.66	2.67
	(0.06)	(0.07)	(0.07)	(0.07)
Printed products	5.19	6.22	2.63	2.64
	(0.07)	(0.07)	(0.06)	(0.06)
Textiles, leather, and articles of textiles or leather	5.21	6.23	2.70	2.72
	(0.07)	(0.07)	(0.06)	(0.06)

Table 2 - Estimated Vectors of Elasticity of Substitution across Varieties $(\boldsymbol{\eta}^{\mathbf{A}})$

	Benchmark	SA#1	SA#2	SA#3
Nonmetallic mineral products	5.35	6.42	2.55	2.56
	(0.06)	(0.07)	(0.07)	(0.07)
Base metal	5.38	6.45	2.68	2.70
	(0.06)	(0.07)	(0.06)	(0.06)
Articles of base metal	5.15	6.15	2.70	2.73
	(0.07)	(0.07)	(0.06)	(0.06)
Machinery	5.40	6.48	2.68	2.70
	(0.06)	(0.07)	(0.06)	(0.06)
Electronic, electrical and office equipment	5.37	6.41	2.76	2.78
	(0.06)	(0.07)	(0.06)	(0.06)
Motorized and other vehicles (including parts)	5.58	6.69	2.71	2.72
	(0.06)	(0.06)	(0.06)	(0.06)
Transportation equipment	5.22	6.25	2.58	2.57
	(0.07)	(0.07)	(0.08)	(0.08)
Precision instruments and apparatus	4.94	5.87	2.87	2.88
	(0.07)	(0.08)	(0.06)	(0.06)
Furniture, mattresses, lamps, lighting fittings	5.45	6.54	2.71	2.73
	(0.06)	(0.07)	(0.07)	(0.07)
Miscellaneous manufactured products	5.35	6.40	2.69	2.70
	(0.06)	(0.07)	(0.06)	(0.06)
Waste and scrap	5.17	6.19	2.69	2.69
	(0.06)	(0.07)	(0.16)	(0.16)
Mixed freight	5.21	6.24	2.68	2.69
	(0.07)	(0.07)	(0.06)	(0.06)

Notes: The standard errors calculated by Delta method are in parenthesis. For more details, see Table 1. SA#1,2,3 stand for Sensitivity Analysis #1,2,3, respectively.

Table 3 - Estimated Vectors of Elasticity of Distance $\left(\boldsymbol{\delta}^{\mathbf{A}}\right)$

	Benchmark	SA#1	SA#2	SA#3
Cereal grains	0.21	0.22	0.24	0.29
	(0.03)	(0.02)	(0.02)	(0.11)
Other agricultural products	0.46	0.47	0.27	0.29
	(0.03)	(0.03)	(0.03)	(0.04)
Animal feed and products of animal origin	0.49	0.51	0.53	0.46
	(0.03)	(0.03)	(0.03)	(0.06)
Meat, fish, seafood, and their preparations	0.95	0.98	0.33	0.32
	(0.03)	(0.03)	(0.03)	(0.02)
Milled grain products and bakery products	0.83	0.86	0.35	0.36
	(0.03)	(0.03)	(0.03)	(0.03)
Other prepared foodstuffs and fats and oils	0.85	0.86	0.41	0.41
	(0.03)	(0.03)	(0.03)	(0.02)
Alcoholic beverages	0.65	0.67	0.40	0.40
	(0.03)	(0.02)	(0.02)	(0.03)
Tobacco products	0.23	0.24	0.22	0.27
	(0.03)	(0.03)	(0.03)	(0.07)
Natural sands	0.43	0.46	0.51	0.55
	(0.03)	(0.03)	(0.03)	(0.06)
Gravel and crushed stone	0.60	0.63	0.69	0.66
	(0.03)	(0.02)	(0.02)	(0.05)
Nonmetallic minerals	0.42	0.45	0.26	0.21
	(0.04)	(0.03)	(0.03)	(0.06)
Metallic ores and concentrates	0.16	0.17	0.09	0.35
	(0.03)	(0.03)	(0.03)	(0.17)
Coal	0.26	0.27	0.33	0.29
	(0.03)	(0.02)	(0.02)	(0.08)

Table 3 - Estimated Vectors of Elasticity of Distance $\left(\boldsymbol{\delta^{A}}\right)$

	Benchmark	SA#1	SA#2	SA#3
Gasoline and aviation turbine fuel	0.67	0.69	0.47	0.47
	(0.03)	(0.02)	(0.02)	(0.05)
Fuel oils	0.63	0.65	0.55	0.54
	(0.03)	(0.02)	(0.02)	(0.05)
Coal and petroleum products	0.78	0.81	0.42	0.43
	(0.03)	(0.03)	(0.03)	(0.05)
Basic chemicals	0.77	0.80	0.26	0.27
	(0.03)	(0.03)	(0.03)	(0.03)
Pharmaceutical products	0.71	0.72	0.46	0.46
	(0.04)	(0.03)	(0.03)	(0.02)
Fertilizers	0.37	0.38	0.23	0.16
	(0.03)	(0.03)	(0.03)	(0.06)
Chemical products and preparations	0.80	0.83	0.38	0.38
	(0.03)	(0.03)	(0.03)	(0.02)
Plastics and rubber	0.50	0.49	0.38	0.38
	(0.03)	(0.03)	(0.03)	(0.01)
Wood products	0.89	0.91	0.39	0.39
	(0.03)	(0.02)	(0.02)	(0.02)
Pulp, newsprint, paper, and paperboard	0.84	0.87	0.32	0.31
	(0.03)	(0.03)	(0.03)	(0.02)
Paper or paperboard articles	0.99	1.03	0.39	0.39
	(0.03)	(0.03)	(0.03)	(0.02)
Printed products	0.62	0.63	0.43	0.42
	(0.03)	(0.03)	(0.03)	(0.01)
Textiles, leather, and articles of textiles or leather	0.34	0.33	0.29	0.29
	(0.03)	(0.03)	(0.03)	(0.01)

Table 3 - Estimated Vectors of Elasticity of Distance $\left(\boldsymbol{\delta}^{\mathbf{A}}\right)$

	Benchmark	SA#1	SA#2	SA#3
Nonmetallic mineral products	0.93	0.96	0.49	0.47
	(0.03)	(0.03)	(0.03)	(0.02)
Base metal	0.85	0.87	0.45	0.44
	(0.03)	(0.03)	(0.03)	(0.02)
Articles of base metal	0.64	0.65	0.40	0.41
	(0.03)	(0.03)	(0.03)	(0.02)
Machinery	0.42	0.41	0.36	0.35
	(0.03)	(0.03)	(0.03)	(0.01)
Electronic, electrical and office equipment	0.30	0.28	0.32	0.31
	(0.03)	(0.03)	(0.03)	(0.01)
Motorized and other vehicles (including parts)	0.76	0.78	0.37	0.37
	(0.03)	(0.03)	(0.03)	(0.02)
Transportation equipment	0.33	0.34	0.23	0.22
	(0.04)	(0.04)	(0.04)	(0.03)
Precision instruments and apparatus	0.69	0.71	0.25	0.25
	(0.04)	(0.03)	(0.03)	(0.02)
Furniture, mattresses, lamps, lighting fittings	0.73	0.75	0.36	0.35
	(0.03)	(0.03)	(0.03)	(0.02)
Miscellaneous manufactured products	0.27	0.25	0.34	0.34
	(0.03)	(0.03)	(0.03)	(0.01)
Waste and scrap	0.30	0.31	0.40	0.41
	(0.03)	(0.03)	(0.03)	(0.09)
Mixed freight	1.05	1.05	0.59	0.60
	(0.03)	(0.03)	(0.03)	(0.02)

Notes: The standard errors calculated by Delta method are in parenthesis. For more details, see Table 1. SA#1,2,3 stand for Sensitivity Analysis #1,2,3, respectively.

Table 4 - Restriction Test Results

	Equation						
	B	C	D	E	F	G	H
F-test	18.74	48.91	23.97	23.86	18.74	7.10	7.29
d.f. 1	505	74	579	578	504	38	37
d.f. 2	$47,743$	$47,312$	$47,817$	$47,816$	$47,742$	$47,276$	$47,275$
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Notes: The null hypothesis is that the restrictions are valid.

Table 5 - OLS Estimation Results for Sensitivity Analysis #1

	Equation							
	A	B	C	D	E	F	G	H
η	[6.27]	[1.19]	6.37	1.15	1.11	[1.17]	–	6.26
	[(0.07)]	[(0.18)]	(0.06)	(0.03)	(0.03)	[(0.20)]	–	(0.03)
δ	[0.61]	[3.31]	0.61	3.41	3.46	[3.34]	–	[0.62]
	[(0.03)]	[(0.17)]	(0.01)	(0.03)	(0.03)	[(0.17)]	–	[(0.30)]
θ	$\boldsymbol{\theta}^{\mathbf{A}}$	–	$\boldsymbol{\theta}^{\mathbf{C}}$	–	6.00	1.95	$\boldsymbol{\theta}^{\mathbf{G}}$	$\boldsymbol{\theta}^{\mathbf{H}}$
	$(\cdot)$	–	$(\cdot)$	–	(0.26)	(0.26)	$(\cdot)$	$(\cdot)$
$\delta\eta$	–	–	–	–	–	–	$\eta\boldsymbol{\delta}^{\mathbf{G}}$	–
	–	–	–	–	–	–	$(\cdot)$	–
R-bar sqd.	0.40	0.30	0.36	0.25	0.25	0.30	0.39	0.39

Notes: The standard errors calculated by Delta method are in parenthesis. The sample size for all estimations is 47,819 which is found after considering the independent observations (ratios) and after ignoring the missing observations. The average of the estimated vectors of $\eta^{\mathbf{A}}$ and $\delta^{\mathbf{A}}$are given in brackets of which full vectors are given in Table 2. The estimated vectors of $\theta^{\mathbf{A}}$, $\theta^{\mathbf{C}}$, $\theta^{\mathbf{G}}$ and $\theta^{\mathbf{H}}$ (all having a size of 505) are omitted to save space. For equations E and F, θ corresponds to θ_H. For Equations A-F, the estimates for $\delta\eta$ are omitted since the estimates for δ and η are already given separately.

Table 6 - Restriction Test Results for Sensitivity Analysis #1

	Equation						
	B	C	D	E	F	G	H
F-test	18.03	46.96	23.06	23.00	18.05	6.91	7.09
d.f. 1	505	74	579	578	504	38	37
d.f. 2	47,743	47,312	47,817	47,816	47,742	47,276	47,275
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Notes: The null hypothesis is that the restrictions are valid.

Table 7 - OLS Estimation Results for Sensitivity Analysis #2

	Equation							
	A	B	C	D	E	F	G	H
η	[2.70]	[1.06]	2.63	1.05	1.01	[0.98]	–	2.68
	[(0.09)]	[(0.18)]	(0.06)	(0.02)	(0.02)	[(0.20)]	–	(0.04)
δ	[0.37]	[0.92]	0.38	0.86	0.83	[0.93]	–	[0.37]
	[(0.03)]	[(0.02)]	(0.01)	(0.01)	(0.01)	[(0.12)]	–	[(0.18)]
θ	$\boldsymbol{\theta}^{\mathbf{A}}$	–	$\boldsymbol{\theta}^{\mathbf{C}}$	–	2.03	1.93	$\boldsymbol{\theta}^{\mathbf{G}}$	$\boldsymbol{\theta}^{\mathbf{H}}$
	$(\cdot)$	–	$(\cdot)$	–	(0.08)	(0.15)	$(\cdot)$	$(\cdot)$
$\delta\eta$	–	–	–	–	–	–	$\eta\boldsymbol{\delta}^{\mathbf{G}}$	–
	–	–	–	–	–	–	$(\cdot)$	–
R-bar sqd.	0.60	0.33	0.58	0.32	0.32	0.33	0.59	0.59

Notes: The standard errors calculated by Delta method are in parenthesis. The sample size for all estimations is 12,581 which is found after considering the independent observations (ratios) and after ignoring the missing observations together with zero observations. The average of the estimated vectors of $\eta^{\mathbf{A}}$ and $\delta^{\mathbf{A}}$are given in brackets of which full vectors are given in Table 2. The estimated vectors of $\theta^{\mathbf{A}}$, $\theta^{\mathbf{C}}$, $\theta^{\mathbf{G}}$ and $\theta^{\mathbf{H}}$ (all having a size of 709) are omitted to save space. For equations E and F, θ corresponds to θ_H. For Equations A-F, the estimates for $\delta\eta$ are omitted since the estimates for δ and η are already given separately.

Table 8 - Restriction Test Results for Sensitivity Analysis #2

	Equation						
	B	C	D	E	F	G	H
F-test	14.56	7.40	13.89	13.70	14.40	4.61	4.72
d.f. 1	701	74	775	774	700	38	37
d.f. 2	$12,740$	$12,113$	$12,814$	$12,813$	$12,739$	$12,077$	$12,076$
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Notes: The null hypothesis is that the restrictions are valid.

Table 9 - OLS Estimation Results for Sensitivity Analysis #3

	Equation							
	A	B	C	D	E	F	G	H
η	[2.74]	[1.08]	2.66	1.06	1.01	[1.01]	–	2.70
	[(0.09)]	[(0.18)]	(0.06)	(0.02)	(0.01)	[(0.20)]	–	0.04
δ	[0.38]	[0.92]	0.38	0.86	0.83	[0.92]	–	[0.38]
	[(0.04)]	[(0.12)]	(0.01)	(0.01)	(0.01)	[(0.13)]	–	[(0.18)]
θ	$\boldsymbol{\theta}^{\mathbf{A}}$	–	$\boldsymbol{\theta}^{\mathbf{C}}$	–	2.16	2.03	$\boldsymbol{\theta}^{\mathbf{G}}$	$\boldsymbol{\theta}^{\mathbf{H}}$
	$(\cdot)$	–	$(\cdot)$	–	(0.09)	(0.19)	$(\cdot)$	$(\cdot)$
$\delta\eta$	–	–	–	–	–	–	$\eta\boldsymbol{\delta}^{\mathbf{G}}$	–
	–	–	–	–	–	–	$(\cdot)$	–
R-bar sqd.	0.60	0.32	0.59	0.31	0.32	0.33	0.59	0.59

Notes: The standard errors calculated by Delta method are in parenthesis. The sample size for all estimations is 12,581 which is found after considering the independent observations (ratios) and after ignoring the missing observations together with zero observations. The average of the estimated vectors of $\eta^{\mathbf{A}}$ and $\delta^{\mathbf{A}}$are given in brackets of which full vectors are given in Table 2. The estimated vectors of $\theta^{\mathbf{A}}$, $\theta^{\mathbf{C}}$, $\theta^{\mathbf{G}}$ and $\theta^{\mathbf{H}}$ (all having a size of 709) are omitted to save space. For equations E and F, θ corresponds to θ_H. For Equations A-F, the estimates for $\delta\eta$ are omitted since the estimates for δ and η are already given separately.

Table 10 - Restriction Test Results for Sensitivity Analysis #3

	Equation						
	B	C	D	E	F	G	H
F-test	14.36	6.88	13.71	13.51	14.20	4.40	4.53
d.f. 1	709	74	783	782	708	38	37
d.f. 2	12,505	11,870	12,579	12,578	12,504	11,834	11,833
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Notes: The null hypothesis is that the restrictions are valid.

Table 11 - PPML Estimation Results

	$\eta^{\mathbf{A}}$	$\delta^{\mathbf{A}}$	R-bar sqd.
Benchmark Case	2.17	0.59	0.54
	(0.01)	(0.01)	
Sensitivity Analysis #1	2.17	0.59	0.54
	(0.01)	(0.01)	
Sensitivity Analysis #2	2.63	0.30	0.83
	(0.04)	(0.01)	
Sensitivity Analysis #3	2.54	0.34	0.84
	(0.04)	(0.01)	

Notes: The standard errors calculated by Delta method are in parenthesis. For each case, the sample size is the same as in the earlier tables. The average of the estimated vectors of $\eta^{\mathbf{A}}$ and $\delta^{\mathbf{A}}$are presented.

Table 12 - Estimation Results for Elasticity of Substitution Across Goods (ε)

	ε	R-bar sqd.
Benchmark Case (OLS)	1.38	0.68
	(0.14)	
Sensitivity Analysis #1 (OLS)	1.27	0.68
	(0.13)	
Sensitivity Analysis #2 (OLS)	1.95	0.77
	(0.14)	
Sensitivity Analysis #3 (OLS)	1.92	0.77
	(0.14)	
Benchmark Case (PPML)	2.19	0.13
	(0.01)	
Sensitivity Analysis #1 (PPML)	2.19	0.13
	(0.01)	
Sensitivity Analysis #2 (PPML)	2.44	0.08
	(0.01)	
Sensitivity Analysis #3 (PPML)	2.97	0.07
	(0.02)	

Notes: The standard errors are in parenthesis The sample size for all estimations is 1,319 which is found after considering the independent observations (ratios) and after ignoring the missing observations.