Implicit Utility and the Canonical Gravity Model☆

Russell H. Hillberry [a], Anton C. Yang [a,☆☆]

[a] *Purdue University, 403 West State Street, West Lafayette, IN 47907 USA*

Abstract

The primary advantage of structural approaches to estimating the gravity model of trade is that they allow a transparent mapping of regression coefficients to structural parameters. Unfortunately, existing structural estimation methods are unable to separately identify trade costs and the trade elasticity without incorporating external data. We demonstrate that theoretical structure is alone sufficient for identifying all of the structural parameters of the canonical constant elasticity of substitution (CES) gravity model. We accomplish this by adopting an implicitly indirect representation of utility and estimating structurally using a mathematical program with equilibrium constraints. Our estimate of the elasticity of substitution is much smaller than in much of the rest of the literature, an outcome that we attribute to Pigou's Law, which ties income and substitution elasticities together in demand systems that assume additive preferences. This restriction is undesirable in demand systems, generally, and is a critical weakness for the canonical gravity model, a model that is commonly used to interpret the geographic trade pattern and to infer the welfare gains from trade.

Keywords: Structural estimation; Gravity model; Implicit utility; Substitution elasticity; Trade costs; Border effect

JEL classification: C51; D11; F14

I. Introduction

The existence of a border effect in U.S.-Canada trade has been an important spur to the development and estimation of structural gravity models of trade.[1] In the canonical structural gravity model, Anderson (1979), the relative shortage of cross-border trade between the two countries is attributed to two

☆This research was partially financially supported by the Center for Global Trade Analysis (CGTA), Department of Agricultural Economics, Purdue University. The authors thank Farid Farrokhi and Anson Soderbery for their insightful suggestions. Anton Yang thanks Thomas Hertel and Dominique van der Mensbrugghe for their full supports. This is a preliminary draft. All errors are our own. Please check updates and link to our latest version.

☆☆Contact author of this draft: yang1069@purdue.edu.

[1] The first estimate of the border effect comes from a reduced form regression estimated by McCallum (1995). Those estimates imply that 1988 trade flows within Canada are more than twenty times larger than gravity adjusted trade flows within Canada.

parameters: an implicit trade cost imposed by the border, and an elasticity of substitution parameter (σ) that defines the response of bilateral trade to trade costs.[2] Separate identification of the two parameters is critical to a quantitative understanding of the welfare effects of the U.S.-Canada border (and of other geographic trade frictions). Unfortunately, existing approaches to structural estimation of the canonical model are under-identified; they cannot separately identify these two parameters without the use of external data.

In this paper we show that it is possible to identify the two parameters separately, *with* the canonical structure and *without* the use of additional data. We transform the representation of utility in the model from an explicit to an implicit form. This transformation eliminates *multilateral resistance* terms from the problem, but requires us to estimate directly a cardinal value for the utility of the representative agent in each region. We use a mathematical program with equilibrium constraints (MPEC) to estimate utility jointly with the model's parameters. This approach accomplishes the identification of the σ parameter, which allows, in turn, the identification of the other trade cost parameters. We are also able to estimate separately two border cost parameters—a U.S.-Canada border cost and Canada-U.S. border cost.

The theory that we adopt is the gravity model of Anderson (1979). The model was first applied to estimation of the U.S.-Canada border effect by Anderson and van Wincoop (2003), who address the identification problem we study by assuming different values of σ. Our paper is most similar to Balistreri and Hillberry (2007), who use an MPEC to estimate the parameters of the model, and conduct general-equilibrium-consistent counterfactual analysis of changes in border costs conditional on an assumed value of σ.[3] Like Balistreri and Hillberry (2007) we apply an MPEC, but we estimate a version of the model with an implicit rather than an explicit representation of utility. The implicit representation allows us to separately identify σ and several measures of trade costs, including an asymmetry between U.S.-Canada and Canada-U.S. border costs.[4]

[2]Other structural gravity models motivate trade differently, and interpret this elasticity differently, but have the same identification problem we consider here.

[3]Balistreri and Hillberry (2007) also investigate the consequences of asymmetries between U.S-Canada and Canada-US border costs. These asymmetries were also assumed in Balistreri and Hillberry (2007) because they could not be estimated.

[4]Our paper is also related to Tan (2012), who uses an MPEC to estimate a structural gravity model with translog preferences.

Our paper is part of an emerging literature that links implicit representations of utility to the pattern of trade.[5] These papers adopt demand systems with implicit utility because they are less restrictive than the CES model, and yet are sufficiently parsimonious that they allow econometric identification of key parameters. In this paper we use the estimation approach developed for that literature, in order to demonstrate its advantages for econometric identification in a system that is dual to a direct utility function. Our results are consistent with this literature's critique of restricted demand systems like the CES. We use the model estimates to show how the structure imposed by the CES form generates an estimated value of σ that is well below its conventionally understood value.

A large literature has addressed the problem of econometric identification of σ by a) incorporating external data into gravity model estimation or b) assessing joint variation in quantities and in unit values rather than in trade value. Hummels (1999), Simonovska and Waugh (2014), and Heid, Larch and Yotov (2017), among many others, employ the gravity framework but incorporate additional data on trade costs or retail prices. Feenstra (1994), Broda and Weinstein (2006), and Soderbery (2015) develop an approach that uses data on imported quantities and unit values to estimate σ at the product-level. Relative to these literatures we innovate by structurally estimating a theory-consistent aggregate σ without relying on external data. Our systemic approach to estimating the aggregate gravity model also highlights a key shortcoming of the CES structure. These shortcomings are less apparent in papers that elide an important restriction imposed by the CES form.

In the paper we first replicate the estimation of Balistreri and Hillberry (2007) under an assumption that $\sigma = 5$. Next we show that structural estimation with our implicit representation of utility produces identical results when we restrict σ to equal 5. We then demonstrate that the implicit utility representation allows σ to be estimated separately from trade costs. We are also able to separately identify asymmetric border costs. Our estimate of $\sigma = 1.624$ is quite low, relative to the values that are generally accepted in the literature. We hypothesize that imposing the full structure of the model subjects the estimates to Pigou's Law: the income elasticity of demand and the substitution elasticity are tied together by assumption. Numerical estimates from our model

[5] See for example, Comin, Lashkari and Mestieri (2015), Matsuyama (2019), Yang (2019) and Yang and Preckel (2020).

give substantial credence to this hypothesis. We also show that an even more flexible CES system can be fit with our methods, a system that allows the representative consumer in each region to have her own value of σ. This produces somewhat higher values of σ than in the version with a common elasticity, and lower estimated trade costs.

The structure of the paper is as follows. In section II we describe the explicit utility representation of the model used in Balistreri and Hillberry (2007). In section III we develop a dual representation of the model which has a representation similar to implicit utility functions. Section IV describes our approach to estimation. Section V reports results. Section VI concludes.

II. Explicit Utility Representation in Trade

Our theoretic setup on the demand side follows Anderson and van Wincoop (2003) and Balistreri and Hillberry (2007). There are N locations in the model, defined by set $S \equiv \{1, ..., N\}$. The representative consumer's preferences in region $j \in S$ are modeled by the following direct CES function:

$$U_j \equiv \left[\sum_i \alpha_i^{(1-\sigma)/\sigma} \left(\frac{T_{ij}}{t_{ij}} \right)^{(\sigma-1)/\sigma} \right]^{\sigma/(\sigma-1)}, \tag{1}$$

where T_{ij} is the quantity of shipment from region $i \in S$ to region $j \in S$, which is melted by the *iceberg trade cost* variable $t_{ij} \geq 1$, and $\{t_{ij}\}_{i,j \in S} \in \overline{\mathbb{R}}_{++}$: $\mathbb{R}_{++} \cup \{\infty\}$, satisfying triangle inequality relationship, i.e., $\tau_{ij}\tau_{jk} \geq \tau_{ik}\ \forall\ i, j, k \in S$. U_j's are the representative agent's utility in region $j \in S$. α_i's > 0 are taste parameters governing the distribution of consumer preferences over the goods from different source-region $i \in S$. $\sigma > 0$ is the constant elasticity of substitution.

Under region j's national budget constraint, i.e., $Y_j = \sum_i FOB_i T_{ij}$, the nominal trade flow equation is shown as follows:

$$FOB_i T_{ij} = Y_j \left(\frac{\alpha_i FOB_i t_{ij}}{P_j} \right)^{1-\sigma}, \tag{2}$$

where FOB_i is the domestic (i.e., the f.o.b.) price of output units and in region $i \in S$; Y_j and P_j are the national income and consumer price index in region $j \in S$, respectively.

It is easy to show that if we define P_j as the aggregate price of goods index (as in Anderson, 1979) equaling the inverse of shadow price resulted from the utility maximization, then it must be true that for any CES demand function:

$$U_j = \frac{Y_j}{P_j}. \tag{3}$$

Using Eqs. (2) and (3), Balistreri and Hillberry (2007) constructs a *4n* system of equations that is an operational general equilibrium model. They structurally estimate the gravity equation derived from Eq. (2) and bring the model to the data to minimize the sum of squared residuals between observed and fitted shares of trade flows of income. They solve the general equilibrium gravity model that satisfies the following equilibrium constraints:

Condition 1. The *income definition* is specified as region i's total income equals the product of output price and the fixed aggregate endowment E_i^0:

$$Y_i = FOB_i E_i^0 \quad \forall\, i,$$

Condition 2. In each region $i \in S$, the *goods-market clears*:

$$E_i^0 = \sum_j \left[\frac{Y_j}{FOB_i} \left(\frac{\alpha_i FOB_i t_{ij}}{P_j} \right)^{1-\sigma} \right] \quad \forall\, i,$$

Condition 3. The CES aggregate price of goods index is a measure of *unit expenditure of utility.* It is the marginal expenditure necessary to purchase an additional unit of cardinal utility in all locations:

$$P_j = \left[\sum_i (\alpha_i FOB_i t_{ij})^{1-\sigma} \right]^{1/(1-\sigma)} \quad \forall\, j,$$

which can been viewed as a zero-profit condition for the *supply of utility* that satisfies an aggregate demand for goods consumption.

Condition 4. The *income is balanced* in all locations, *CES-ly*, such that:

$$U_i P_i = Y_i \quad \forall\, i.$$

III. Implicit Representation of Utility in Trade

Balistreri and Hillberry (2007)'s approach allows them to fit a computable general equilibrium model into the bilateral trade data to study asymmetric border effects and conduct counterfactual analysis that is theory-consistent with Anderson and van Wincoop (2003)'s model. However, their estimation approach does not fully identify all structural parameters in the model, i.e., they set CES substitution elasticity $\sigma = 5$ to recover other structural parameters (namely the distance elasticity of trade costs and the ipmlicit trade cost associated with trading across a national border.

III.I. Constructing a 3n *General Equilibrium System of Equations*

We are interested in separately identifying the elasticity of substitution and the trade cost measures. We derive a general equilibrium model that is consistent with Balistreri and Hillberry (2007)'s CES-gravity Model. Unlike Balistreri and Hillberry (2007), we are explicit about the cardinal value of utility which is critical for our structural identification strategy. In so doing, we derive a *3n* general equilibrium system of equations as constraints on the econometric objective function. Using Eq. (3), we rewrite the bilateral trade prediction equation (Eq. 2) to include a cardinal value of utility U_j:

$$FOB_i T_{ij} = U_j^{1-\sigma}(\alpha_i t_{ij})^{1-\sigma} FOB_i^{1-\sigma} Y_j^{\sigma}. \tag{4}$$

This representation is also notable in that it has removed the price index term, P_j, a term that Anderson and van Wincoop label multilateral resistance.

Eq. (4) has an important motivation for the implied counterfactual relationship among quantity shipment (from origin), T_{ij}, FOB_i(or t_{ij}) and Y_j, controlled by the substitution elasticity σ. To see this, the shipment from each origin i to destination $j \in S$ in Eq. (4) can be simply isolated as a function of utilities, incomes and prices, i.e., $T_{ij} = (\alpha_i U_j)^{1-\sigma}(FOB_i t_{ij})^{-\sigma} Y_j^{\sigma}$. If utility, U_j, is being held fixed, an increase of either the prices of domestic output units (i.e., FOB_i) or the iceberg trade costs (t_{ij}) by 1% will lead T_{ij} to go down by σ%. Moreover, due to the generic σ, in order to hold U_j and T_{ij} fixed while prices increase by 1%, the region-j consumer must be compensated by exactly a 1% increase in nominal income. While holding fixed FOB_i's and U_j's, the econometric exercise exploits variations in T_{ij} and Y_j.

Same as in Balistreri and Hillberry (2007), we characterize that the aggregate income in each region $i \in S$ must satisfy:

$$Y_i = FOB_i E_i^0 \quad \forall\, i \quad \textbf{(First GE Equation).} \tag{5}$$

In equilibrium, each region i's aggregate income must equal the total value of goods purchased by region j's consumer (from all destinations $j \in S$) at the

CIF prices (i.e., taking into account the cost, insurance, and freight):

$$Y_i = \sum_j FOB_i T_{ij} \quad \forall\, i. \tag{6}$$

Thus combining with the result of the CIF demand in Eq. (4), we show that the fixed total endowment can be written as a function of utility, prices, nominal income, trade costs as well as model parameters:

$$E_i^0 = \sum_i U_j^{1-\sigma} (\alpha_i t_{ij})^{1-\sigma} FOB_i^{-\sigma} Y_j^{\sigma} \quad \forall\, i \quad \textbf{(Second GE Equation)}. \tag{7}$$

Using the income definition Eq. (5) and goods' market-clearing condition Eq. (7), we can derive an implicit representation of the indirect utility function G à la Hanoch (1975) associated with Anderson and van Wincoop (2003)'s gravity model. Specifically, it implies that each region j has the following indirect preferences of imported goods as a function of utilities, national incomes and prices of goods:

$$G(\frac{\mathbf{FOB}}{\mathbf{Y}}, \mathbf{U}) = \sum_i \alpha_i^{1-\sigma} U_j^{1-\sigma} \left(\frac{FOB_i t_{ij}}{Y_j} \right)^{1-\sigma} \equiv 1 \quad \textbf{(Third GE Equation)}. \tag{8}$$

Eq. (8) has an isomorphic structure of demand preferences as in Hanoch (1975).[6] However, it does not become a specific class of implicitly additive models, as it is obviously still strongly separable as an explicit function. We use it as an equilibrium constraint as in Yang (2019), in which the relationships among U_j, income, prices and model parameters are implicitly and indirectly defined as additive preferences.

Our first general equilibrium equation is the same as Balistreri and Hillberry (2007)'s condition 1, and our second general equilibrium equation is too iden-

[6] Yang (2019) shows that the implicitly indirect non-homothetic demand function in Hanoch (1975) can be parameterized to achieve the following additivity constraint:

$$\sum_i \beta_i U_j^{1-\sigma} \left(\frac{\mathrm{FOB}_i t_{ij}}{Y_j} \right)^{1-\sigma} \equiv 1.$$

We adopt the particular CES preferences Eq. (1) widely used in the international trade literature, while noting that the equation above is a parameter transformation of Eq. (1), i.e., $\alpha_i^{1-\sigma} = \beta_i > 0 \;\forall\, i$.

tical to Balistreri and Hillberry (2007)'s condition 2, except that we eliminate *multilateral resistance* terms by substituting region-j's price index P_j with U_j. For our estimation purpose, the unit expenditure definition is irrelevant, thus we only need an additional equation as specified in Hanoch (1975) and Yang (2019), which characterize a specific implicit relationship between model parameters and exogenous variables, to pin down U_j along with Eq. (7), using the MPEC program.

Following Balistreri and Hillberry (2007), trade costs is modeled as the following to account for both distance and asymmetric border effects on bilateral trade flows:

$$t_{ij} = d_{ij}^{\rho}\left[\exp\left(\frac{\mathbf{A}}{1-\sigma}\right)\right]^{1-\delta_{ij}}, \tag{9}$$

where d_{ij} is the distance between i and j observed from the data, ρ is the elasticity of trade costs with respect to distance; consistent with Balistreri and Hillberry (2007), $\mathbf{A} = (1-\sigma)\ln b_{ij}$ are the border coefficients, where b_{ij}'s equals one plus tariff equivalent of border costs.[7] δ_{ij}'s are the dummy variables equaling zeros if shipments cross international border and equaling ones if the shipments are taken place domestically in locations $i, j \in S$.

The nominal bilateral trade flow between i and j, $X_{ij} = FOB_i T_{ij}$, is thus given by

$$X_{ij} = \alpha_i^{1-\sigma} U_j^{1-\sigma} \left(\mathrm{FOB}_i t_{ij}\right)^{1-\sigma} (Y_j)^{\sigma}. \tag{10}$$

Given Eq. (9), the log-linearized form of Eq. (10) is

$$\begin{aligned} \log(X_{ij}) = {} & (1-\sigma)\log\alpha_i + (1-\sigma)\log U_j + (1-\sigma)\log \mathrm{FOB}_i \\ & + \rho(1-\sigma)\log d_{ij} + (1-\delta_{ij})(1-\sigma)b_{ij} + \sigma \log Y_j. \end{aligned} \tag{11}$$

III.II. Identification of Structural Parameters

A common approach to estimating a stochastic version of Eq. (11) is to use origin- and destination- fixed effects to sweep out unobserved or difficult-to-calculate terms associated with the origin i and the destination j. Our representation of the model does nothing to change the inferences that would be drawn from a fixed effect regression of this kind. Such a regression would return reduced form coefficients on the bilateral variables (d_{ij} and b_{ij}) but these

[7] $A = a2/(1-\sigma)$ if shipping from the U.S. to Canada, and $A = a3/(1-\sigma)$ if shipping from Canada to the U.S.

regression coefficients would still be unidentified in terms of the model's structural parameters. Without an estimate of σ, it is impossible to quantify the geographic trade costs.

Notably, in Eq. (11), if we know the cardinal value of utility, U_j, then we can identify σ. With σ pinned down, we can obtain the estimates b_{ij} and ρ given the information on bilateral distances. The α_i terms can be computed conditional on estimates of trade costs and σ. In any reduced-form approach to estimation, U_j can not be evaluated. With U_j's being unobserved, we would never achieve identification of other structural parameters from Eq. (10). For this reason, we follow Yang (2019)'s approach to evaluate the implicit utility using MPEC to help pinning down other structural parameters. Unlike Balistreri and Hillberry (2007), we treat Eq. (8) as one additional constraint to categorize the implicit representation of preferences among model parameters and variables, and our general equilibrium functions of value of trade flows (satisfying gravity equation) and goods market-clearing conditions are incorporated with the cardinal utilities which we estimate explicitly.

IV. Estimation Methods

We estimate the two-country model of the U.S. and Canada as in Balistreri and Hillberry (2007) and part of the exercise illustrated in Anderson and van Wincoop (2003). Our non-linear estimation procedure is conceptually similar to Anderson and van Wincoop (2003) and Balistreri and Hillberry (2007), except that Balistreri and Hillberry (2007)'s approach and ours both explicitly define a fully operational general equilibrium system as the problem constraints. We define the fitted value $\hat{z}_{ij}$ as in Balistreri and Hillberry (2007), which is obtained directly from the original trade-flows Eq. (4), except that our approach allows us structurally evaluate the U_j simultaneously using Eqs. (7) and (8):

$$\hat{z}_{ij} = (1-\sigma)\big[\log\alpha_i + \log(U_j) + \log(FOB_i) + \log t_{ij} - \log(Y_j)\big] \tag{12}$$

We minimize the sum of squared residuals between the observed shares of income z_{ij} and the fitted $(\hat{z}_{ij})$, i.e., min $\sum_i \sum_j [z_{ij} - \hat{z}_{ij}]^2$, subject to Eqs. (5), (7) and the modified (8) or Eq. (13) below:

$$G(\frac{\mathbf{FOB}}{\mathbf{Y}}, \mathbf{U}) = \sum_i \alpha_i^{1-\sigma} U_j^{1-\sigma} \left(\frac{FOB_i t_{ij}}{Y_j}\right)^{1-\sigma} \equiv \kappa. \tag{13}$$

We allow the RHS to equal some $\kappa > 0$ instead of unity, while letting it be endogenously determined from the system. This is because that the scaling factor of U_j on the LHS is not identifiable from the data. This concept is somewhat similar to the κ discussed in Allen, Arkolakis and Takahashi (2020), where they point out that $\kappa = 1$ is implicitly imposed in most trade models, whereas k is endogenously calculated in most geography models. As in Allen, Arkolakis and Takahashi (2020), we also argue that the value of κ must depend on a specific model and data characteristics. For example, Yang and Preckel (2020) estimate a specialized implicitly indirect demand system using the World Bank and GTAP databases and find that κ tends to equal the sum of distribution parameters. Most published work of econometric estimation that estimates cardinal utility also estimates κ explicitly (see, e.g., Rimmer and Powell, 1996; Cranfield *et al.*, 2002; Preckel, Cranfield and Hertel, 2010; Gouel and Guimbard, 2019). Albeit the equilibrium outcome is invariant to any specific values of κ, constraining the RHS to be one is indeed an arbitrary normalization and tends to over-restrict the solution to the constrained optimization problem.

Balistreri and Hillberry (2007) adopts Anderson and van Wincoop (2003)'s notation of the regression coefficient $a1$ which measures the product of the distance elasticity ρ and $1 - \sigma$ appeared in Eq. (11). Again, same as in Anderson and van Wincoop (2003), Balistreri and Hillberry (2007) set $\sigma = 5$, and calculate $a1$ in their *true* structural estimation, constrained by their *4n* general equilibrium system of equations, to be -1.44. We would like to empirically verify that our theoretical approach using the *3n* general equilibrium system of equations is numerically consistent with their model. That is, if we relax one degree of freedom by setting $\sigma = 5$, the theoretical structure with the U_j implicitly defined in Eqs. (7) and (8), which fit into Eq. (12), must also yield that $\rho = 0.36$, given that Balistreri and Hillberry (2007) and us use the same data and the econometric specification (i.e., least-squares).

We then solve the system of equations and inequality constraints with a formulated mathematical program using the General Algebraic Modeling System (GAMS) version 31.1.1 and the MPEC-NLPEC (non-linear programming with equilibrium constraints) solver (Ferris, Dirkse and Meeraus, 2002). This exercise is carried out on a standard Windows 64-bit operating system.

Using this approach we first replicate Balistreri and Hillberry (2007)'s model using their dual general equilibrium model. We then estimate the implicit version of the model conditional on the same choice of $\sigma = 5$. After this we estimate the model with a free σ. We then estimate a version of the model with region-specific values of the elasticity of substitution, σ_j.

IV.I. Asymmetric Border Costs as Constraints

We characterize locations of the U.S. states and Canadian borders as subsets such that the set of the U.S. States is defined as $S^{US} \equiv \{1, ..., M\}$ and the set of Canadian Provinces is defined as $S^{CAN} \equiv \{M+1, ..., N\}$, where all regions $S \equiv \{1, ..., N\}$. Thus, the constraint with respect to asymmetric border cost is given by, $\forall$ $\{i, j\} \equiv \{S^{US}, S^{CAN}\}$ or $\forall$ $\{i, j\} \equiv \{S^{CAN}, S^{US}\}$:

$$\begin{aligned} b_{ij} &= b(1, M+1) \longleftarrow \{i, j\} \equiv \{S^{US}, S^{CAN}\} \\ &\quad + b(M+1, 1) \longleftarrow \{i, j\} \equiv \{S^{CAN}, S^{US}\}. \end{aligned} \tag{14}$$

With cross-border dummy variables in Eq. (9) embedded into Eq. (11) which yield zero border costs for interstate and inter-provincial trade, Eq. (13) controls that the border costs from all states to all provinces as well as from all provinces to states are the same.

IV.II. Excess Degree of Freedom

Yang and Preckel (2020) show that in a similar additive preferences model as in Eq. (8), utility can only be identified up to a strictly increasing transformation. It can be shown that the scaling factor on U_j is not identifiable from the data. Since the quality of the estimated gravitational relationship is invariant to such transformations due to unknown scalars, the model parameters are not fully identified without specifying the scale of utility. For this reason, we follow a similar method as in Anderson and van Wincoop (2003) and Balistreri and Hillberry (2007) to establish the scale of utility in the estimation:

$$\left(\frac{Y_{\text{Alabama}}}{U_{\text{Alabama}}}\right)^{1-\sigma} = \sum_i \left[\frac{U_i t_{\text{Alabama},j}}{\sum_j Y_j}\right]^{1-\sigma}. \tag{15}$$

Once the excess degree of freedom is removed due to the unknown scaling factor of utility, we ask whether there are additional degrees of freedom in the parameter space that can be removed and whether additional normalization is needed as we are moving to free σ in our estimation. We follow Yang and Preckel (2020)'s strategy to determine the latitude for changing our estimated model parameters. We construct a new problem by maximizing and minimizing σ, subject to the original problem constraints in the estimation, Eqs. (5), (7), (8), (13) and (14), while requiring that the sum of squared residual is at least as small as computed from the estimation problem. If the change to the new objective is essentially zero, then we conclude that the parameter values cannot be changed locally without increasing the residuals, and hence, no additional normalization is needed for this estimation problem.

V. Estimation Results

We present the results of these estimation exercises in Table 1. Before turning to the results we discuss the estimates we report. The first three parameter estimates a_1, a_2 and a_3 are regression coefficients similar to those estimated in Anderson and van Wincoop (2003), though we follow Balistreri and Hillberry (2007) in allowing a_2 and a_3 to capture asymmetries in US and Canadian border frictions. From a structural point of view, these parameters are underidentified. In each table we also report structural estimates for σ, ρ and the border cost variables $\ln b_{US-CA}$ and $\ln b_{CA-US}$. We also report the sum of squared residuals in each column, which is the value of our objective function.[8] The standard errors we report are calculated with the boostrap procedure outlined above.

In the first column we replicate the estimation of Balistreri and Hillberry (2007). All of our results match theirs. Assuming a value of $\sigma = 5$, as Balistreri and Hillberry (2007) do, allows the trade cost variables to be estimated, conditionally, along with conditional standard errors. In column 2, we report results for our system of equations, holding σ fixed at 5. The results match exactly those of Balistreri and Hillberry (2007). This should be expected, as ours is simply a dual representation of their system of equations.

In column 3 we allow σ to be freely estimated, along with the trade cost parameters. We achieve identification through the structure of the demand system, and the bootstrapped standard errors are quite small. Our estimate of $\sigma = 1.62$ is much lower than those extant in the literature on the aggregate trade elasticity. A much lower value of σ in the CES structure automatically implies much larger geographic trade frictions ρ and $\ln b$.

V.I. Implications of Pigou's Law

Before turning to an even more flexible estimation model in column 4 we stop to consider how identification is achieved in column 3. Our estimated

[8]We leave for future work estimation using the PPML likelihood function. The results from these exercises is not very different from our least squares estimates.

Table 1: Structural estimation with implicit and explicit representation of utility

	BH replication with explicit representation	BH replication with implicit representation	Structural estimation with implicit representation	Implicit representation with regional specific $\sigma's$
	(1)	(2)	(3)	(4)
$a1 = (1-\sigma)\rho$	-1.44	-1.44	-1.44	
$a2 = (1-\sigma)\ln b_{US-CA}$	-1.85	-1.85	-1.85	
$a3 = (1-\sigma)\ln b_{CA-US}$	-1.85	-1.85	-1.85	
$\bar{a1} = (1-\bar{\sigma})\rho$				-1.47
$\bar{a2} = (1-\bar{\sigma})\ln b_{US-CA}$				-1.39
$\bar{a3} = (1-\bar{\sigma})\ln b_{CA-US}$				-1.39
σ	5 (assigned)	5 (assigned)	1.62	
			(0.005)	
$\bar{\sigma}$				1.81
				(0.02)
ρ	0.36	0.36	2.31	1.82
	(0.005)	(0.005)	(0.03)	(0.06)
$\ln b_{US-CA}$	0.46	0.46	2.96	1.72
	(0.02)	(0.02)	(0.14)	(0.12)
$\ln b_{CA-US}$	0.46	0.46	2.96	1.72
	(0.02)	(0.02)	(0.14)	(0.12)
N	1511	1511	1511	1511
Sum of squared residuals	2262.84	2262.84	2262.84	1286.03

Standard errors across columns (1) - (4) in "()" obtained from 2,000 bootstrap resamples.

value of $\sigma = 1.624$ lies well below the conventional views about the magnitude of this parameter. Anderson and van Wincoop (2003) do back-of-the-envelope calculations assuming σ's of 5, 10 and 20. Balistreri and Hillberry (2007) assume $\sigma = 5$, a value that is commonly used as a default when taking this parameter 'from the literature.' Simonovska and Waugh (2014) use a somewhat different theoretic structure (with very similar implications) and estimate an aggregate trade elasticity of 4. Romalis (2007) studies trade responses after the North American Free Trade Agreement and estimates a value of 6.9.[9] Clearly our low estimate of σ is an outlier. Why is that so?

Because we exploit the full structure of the CES model in our estimation, our estimate of σ is driven by Pigou's Law, a relationship developed in Pigou (1910) and named as such in Deaton (1974). Pigou's Law says that, in additively separable demand systems, the income elasticity of demand is approximately proportional to the uncompensated own price elasticity of demand. The coefficient of proportionality in this relationship is (approximately) ϕ a measure of inverse of the elasticity of the marginal utility of income to income (Barten,

[9] Soderbery (2015) provides the most recent and thorough estimates of σ using estimators of the kind suggested by Feenstra (1994). The central tendencies of these estimates for individual countries are consistently in the 3-4 range.

1977). Formally Pigou's Law takes the form:

$$\epsilon_{ij} = \delta_{ij}\phi\eta_i + \eta_i w_j(1 - \phi\eta_j), \tag{16}$$

where ϵ_{ij} is the cross-price elasticity of demand, δ_{ij} is the Kronecker delta, ϕ is defined by $1/\phi = dln\lambda/dlnM$ with λ the marginal utility of income and M is the consumer's incomeη_i is the income elasticity of demand for good x, and w_j is the average budget share. Since our demand system is homothetic, $\eta_i = 1$. In the CES system, $\epsilon_{ij} = \sigma$. Since we have 41 goods, w_j is small, and the second term is dominated by the first. The implication of Pigou's law is that $\sigma \approx \phi$.

Since ϕ is an elasticity associated with a Lagrange multiplier, it is not commonly observed in econometric studies. We, however are able to recover the estimates of λ from our structural procedure.[10] In order to estimate ϕ, we collect values of λ_i from the numerical solution to each region's demand system, and estimate the following regression:

$$\begin{aligned} \ln \lambda_i = &- \underset{(1.17)}{7.11^{***}} \; + \; \underset{(0.103)}{0.658^{***}} \; \ln GDP_i + u_i \\ & \qquad\qquad\qquad\qquad R^2 = 0.512, \end{aligned}$$

where λ_i is the marginal utility of income recovered from the structural estimation, GDP_i are the measures of regional GDP used in the structural estimation, and u_i is a normally distributed error term. The coefficient of 0.658 on the GDP variable is our (average) estimate of elasticity of the marginal utility of income to income across the sample. Since $1/0.658 \approx 1.52$, Pigou's law implies that σ should be in the neighborhood of 1.52. Our structural estimate of 1.62 is sufficiently close to this figure to suggest that Pigou's Law is driving our estimate of σ.

V.II. Region-specific σ's

Our final exercise is an estimation of the model that allows each region to have its own value of σ. The results of this exercise are reported in column 4 of table 1. We estimate 41 different values of $\sigma(j)$, along with associated standard

[10]The value of λ is not directly calculated in our estimation. In the dual representation of the CES model, marginal utility of income is represented as the inverse of the unit expenditure function. We calculate the value of the unit expenditure function for each region i as the ratio of income Y_i to Utility, U_i, and treat this ratio as the region's value of λ_i^{-1}.

errors, but report only the mean of the σ estimates as well as their standard deviation. The mean (standard deviation) of our estimated region-specific σ's is 1.81 (0.02). We estimate common values of the trade cost parameters, which we report in column 4 along with associated standard errors. The added flexibility of the CES parameter pushes the trade cost estimates downward. Note that the addition of 40 more substitution parameters substantially improves the overall fit our of estimator. The sum of least squares falls from 2262.84 to 1286.03.

VI. Conclusion

Many papers in the quantitative trade literature develop a structural model and calibrate with estimated trade elasticities 'from the literature'. While there are exceptions (e.g., Simonovska and Waugh, 2014), this approach is pervasive. A key reason for the pervasiveness of this approach is the econometric identification problems that we solve in this paper, gravity regressions exploiting the geographic trade pattern are unable to separately identify the structural parameters defining trade costs from the trade elasticity. We show how this problem can be solved by representing the utility function in an implicit form, and applying the structural estimation tools that have been used to estimate implicit demand systems.

We solve the identification problem, but run into another problem with the CES demand system. Under the CES demand system's strong assumption of homotheticity, the identification we achieve operates through Pigou's Law, which imposes a relationship between σ and the income elasticity of demand. Although we believe it is useful to show that the CES system can be fully identified, we view identification through Pigou's Law is unsatisfactory. More flexible implicit additive systems, such as those estimated by Comin, Lashkari and Mestieri (2015), Yang (2019) and Yang and Preckel (2020) are likely to be more useful than the CES system.

One simple way to introduce flexibility to the model is to allow the σ parameters to vary across regions. We show that allowing each of the region's representative consumers to have its own value of σ substantially improves the fit of the model, and reduces the implied magnitude of the geographic trade frictions.

References

[1] Allen, Treb, Costas Arkolakis, and Yuta Takahashi. Universal gravity., *Journal of Political Economy*, 128.2, 000–000, 2020.

[2] Anderson, James E. A Theoretical Foundation for the Gravity Equation., *American Economic Review*, 69, 106–116, 1979.

[3] Anderson, James E., and Eric Van Wincoop. Gravity with gravitas: a solution to the border puzzle., *American Economic Review* , 93.1: 170-192, 2003.

[4] Balistreri, Edward J. and Russell H. Hillberry. Structural Estimation and the Border Puzzle., *Journal of International Economics*, 72: 451–463, 2007.

[5] Barten, Anton P. The systems of consumer demand functions approach: a review., *Econometrica: Journal of the Econometric Society*, 23–51, 1977.

[6] Broda, Christian, and David E. Weinstein. Globalization and the Gains from Variety., *The Quarterly journal of economics*, 121.2: 541–585, 2006.

[7] Comin, Diego A., Danial Lashkari, and Martí Mestieri. Structural change with long-run income and price effects., *National Bureau of Economic Research*, No. w21595, 2015.

[8] Cranfield, J. A. L., Paul V. Preckel, James S. Eales, and Thomas W. Hertel. Estimating consumer demands across the development spectrum: maximum likelihood estimates of an implicit direct additivity model., *Journal of Development Economics*, 68.2: 289–307, 2002.

[9] Deaton, Angus. A reconsideration of the empirical implications of additive preferences., *The Economic Journal*, 84.334: 338-348, 1974.

[10] Feenstra, Robert C. New product varieties and the measurement of international prices., *The American Economic Review*, 157-177, 1994.

[11] Ferris, Michael C., Steven P. Dirkse, and Alexander Meeraus. Mathematical programs with equilibrium constraints: Automatic reformulation and solution via constrained optimization. In *Technical Report NA-02/11*. Oxford University Computing Laboratory, 2002.

[12] Gouel, Christophe, and Houssein Guimbard. Nutrition transition and the structure of global food demand., *American Journal of Agricultural Economics*, 101.2: 383-403, 2019.

[13] Hanoch, Giora. Production and demand models with direct or indirect implicit additivity., *Econometrica: Journal of the Econometric Society*, 395-419, 1975.

[14] Heid, Benedikt, Mario Larch, and Yoto Yotov. Estimating the Effects of Non-discriminatory Trade Policies within Structural Gravity Models., 2017.

[15] Hummels, D. Towards a Geography of Trade Costs., Mimeo, Krannert School of Management, Purdue University, 1999.

[16] McCallum, John. National borders matter: Canada-US regional trade patterns., *The American Economic Review*, 85.3: 615-623, 1995.

[17] Pigou, Arthur Cecil. A method of determining the numerical value of elasticities of demand., *The Economic Journal*, 20.80: 636-640, 1910.

[18] Preckel, Paul V., J.A.L. Cranfield, and Thomas W. Hertel. A Modified, Implicit, Directly Additive Demand System., *Applied Economics*, 42(2), 143–155, 2010.

[19] Rimmer, Maureen T., and Alan A. Powell. An implicitly additive demand system., *Applied Economics*, 28.12: 1613-1622, 1996.

[20] Romalis, John. NAFTA's and CUSFTA's Impact on International Trade., *The review of Economics and Statistics*, 89.3: 416-435, 2007.

[21] Matsuyama, Kiminori. Engel's law in the global economy: Demand-induced patterns of structural change, innovation, and trade., *Econometrica*, 87.2: 497-528, 2019.

[22] Simonovska, Ina, and Michael E. Waugh. The elasticity of trade: Estimates and evidence., *Journal of international Economics*, 92.1: 34-50, 2014.

[23] Soderbery, Anson. Estimating import supply and demand elasticities: Analysis and implications., *Journal of International Economics*, 96.1: 1-17, 2015.

[24] Tan, Shawn. Structural Estimation of a Flexible Translog Gravity Model., university of Melbourne, Department of Economics Working Paper, 1164 (2012).

[25] Yang, Anton. Structural Estimation of a Gravity Model of Trade with the Constant-Difference-of-Elasticities Preferences., *Midwest Economic Theory and International Trade Meetings*, 2019.

[26] Yang, Anton C., and Paul V. Preckel. Estimation of an Implicit Additive Indirect Demand System., *23rd Annual Conference on Global Economic Analysis (Virtual Conference)*, 2020.