Structural Estimation of a Gravity Model of Trade with the Constant-Difference-of-Elasticities

Anton Yang*

This paper presents a general equilibrium gravity model of trade based on the constant difference of elasticities of substitution preferences. Hanoch (1975) illustrates these preferences' advantages in terms of parsimony and flexibility. This paper introduces a parsimonious, non-homothetic and globally well-behaved demand model into the gravity model that both separates substitution effects from income effects and has non-constant substitution elasticities. These features of the demand model—together with the structural estimation procedure devised in this paper—allow nesting several prominent theoretical motivations for the gravity model, and exploring the merits of this more general model. They also allow identification of the elasticity of trade costs with respect to distance and asymmetric border coefficients from the elasticity of trade flows with respect to trade costs, that are not easily identified in most previous studies. This paper develops a new method of computing counterfactual equilibrium which works best for large and complex general equilibrium models, and particularly those models that use implicit demand models.

I. Introduction

The gravity model of trade is a well-known quantitative model that analyzes the impacts of the size of bilateral partners and their bilateral distances on trade patterns, which has long been a successful model in explaining multilateral trade

* Center for Global Trade Analysis, Purdue University & Department of Economics and Center for Agricultural and Rural Development, Iowa State University. I am grateful to my co-chair Russell Hillberry for his guidance and advice on this paper, and to my co-chair Thomas Hertel and committee member Farid Farrokhi for their invaluable comments, suggestions and support. I specially thank Costas Arkolakis and Yves Surry for their encouragement and substantial suggestions and fruitful discussions. I am also further indebted to Thibault Fally, Ahmad Lashkaripour, Martí Mestieri, William Novshek, Paul Preckel, Anson Soderbery, Sebastian Sotelo, Dominique van der Mensbrugghe for their helpful discussions and constructive feedback at various stages of developing this paper. I also thank seminar participants at Purdue University, Indiana University, University of Southern Denmark as well as several virtual meetings. I gratefully acknowledge the generous financial support through the Center for Global Trade Analysis, Purdue University. All errors are my own. Correspondence: yang1069@purdue.edu.

patterns and welfare gains (Anderson, 2011). The model predicts that bilateral trade flows are jointly explained by some amplifying effects of national income and the attenuating influence of average effective distance. In almost any standard gravity model the implied endogenous system of trade can be generalized to a factor demand system.[1] Thus, studying demand systems can help us enrich the gravity model.

Many existing standard gravity approaches are built on Anderson (1979)'s framework which integrated the *nearly ubiquitous* Constant Elasticity of Substitution (CES) consumer preferences (i.e., top-level utility functions) into the gravity model of trade (Head and Mayer, 2014). These models work well in cross-sectional studies, and the principal of parameter parsimony put CES on the forefront of mainstream gravity trade literature. But those preferences come with a cost; they entail homothetic preferences and unitary income elasticities implying that income shocks in a region will generate proportional changes in national factor demands. Integrating the CES assumption into global factor demand and supply (in aggregate trade flows), means that a 10% increase in relative income will also raise the relative demand for national factor supplies by 10%, with factor expenditure shares being constant. As there have been substantial changes in the relative income in the world, it is important to explore other available demand models which provide better insights to understand the cross-sectional data of international trade.

If the traditional CES-Gravity models cannot help us better understand the implications of significant changes in income, what other demand models are most consistent with the data? In the past, the CES-gravity model was constructed in such a way that it conforms with known empirical relationships. Testing the factor demand implications in the CES-gravity model requires us to bring the theory of non-homotheticity, in which income shocks affect the relative demand for imported factors.

One way to test the model is to compare it against less restrictive (i.e., non-homothetic) models of bilateral trade. Another way is to explore out-of-sample predictions, including welfare and trade responses of bilateral trade to income

[1] For example, the home demand for foreign goods can be perceived as demand for foreign factor services (which are used to produce those goods demanded by home). Adao, Costinot and Donaldson (2017) state the conceptual equivalence between neoclassical economies and "reduced exchange economies" in which countries, instead of trade goods with each other, trade global factor services. One can also see a detailed elaboration in Costinot and Rodríguez-Clare (2018), who shows that the factor service concept provides an intuitive perspective on the ACR (Arkolakis, Costinot and Rodríguez-Clare, 2012) formula.

growth. In this paper, I will test different non-homothetic demand systems under the gravity model of trade. In so doing, this paper presents a new demand component for the gravity model. The model is both flexible and parsimonious, incorporates both non-homotheticity and non-constant elasticities of substitution, removes the tie between substitution effects and income effects, and is globally regular. This paper shows how other prominent theoretical demand models for the gravity model can be nested within the more general demand model. The paper then devises a structural estimation procedure using the same data and cross-sectional estimation structure to determine which demand system best fits the pattern of trade.

If traditional CES-gravity models cannot help us better understand the implications of significant changes in income, what other demand models are most consistent with the data? In the past, the CES-gravity model was constructed in such a way that it conforms with known empirical relationships. Testing the factor demand implications in the CES-gravity model requires to highlight the theory of non-homotheticity, in which income shocks affect the relative demand for imported factors. One way to test the model is to compare it against less restrictive (i.e., non-homothetic) models of bilateral trade. Another way is to explore out-of-sample predictions, including welfare and trade responses of bilateral trade to income growth. In this paper, I contribute to the literature by testing different non-homothetic demand systems under the gravity model of trade. In so doing, this paper presents a new demand component for the gravity model. The model is both flexible and parsimonious, incorporates both non-homotheticity and non-constant elasticities of substitution, removes the tie between substitution effects and income effects, and is globally regular. This paper shows how other prominent theoretical demand models for the gravity model can be nested within a more general demand model. The paper also devises a structural estimation procedure using the same data and cross-sectional estimation structure to determine which demand system best fits the pattern of trade.

Another important contribution of this paper is to show how the proposed estimation procedure can identify the elasticity of trade costs with respect to distance and asymmetric border coefficients from the elasticity of trade flows with respect to trade costs. This paper shows that population data, along with a structural procedure permitting evaluation of cardinal values of utility, are the key to identification of the structural parameters in the flexible demand model.

Finally, the test statistic (likelihood ratio) in the log-likelihood ratio tests of CDE model against other more restricted demand models in the gravity model rejects the null hypothesis that the less general (more restricted) model is more consistent with the data.[2] This suggests that the CDE model is a better model in the gravity model of aggregate trade flows in terms of fitting to the same data used in this paper. Furthermore, the out-of-sample responses to the Chinese income and population shocks suggest that the CDE produces entirely different results as compared to the CES model which generates no variation in the trade flows across exporting countries. A test of predictiveness of the CDE model using actual changes in import shares and those predicted by the CDE model shows that there is a positive coefficient on the CDE shares which suggest that there is a sign that the CDE model is predictive.

The structure of the following sections is as follows. Section II analyses the past and recent demand models employed in the gravity models. Section III discusses the demand theory and introduces an alternative perspective in which flexible (and nested) import demand equations are adopted in gravity models. Section IV constructs a general equilibrium gravity model based on the CDE model and a nested gravity estimation framework. Section V discusses the structural estimation procedure. Section VI and Section VII discuss estimation results and likelihood ratio tests, respectively. Section VIII discusses out-of-sample predictions using different techniques. IX proposes a new counterfactual method which is applied in this paper and discusses counterfactual results. Section X concludes.

II. Demand Models in Gravity

Hanoch (1975) discusses the advantages and disadvantages of multiple classes of flexible demand models. In general, these models can be categorized as: (i) n^{th} order approximations (NOA) to the utility or cost functions;[3] (ii) explicitly direct or indirect additivity; and (iii) implicitly direct or indirect additivity. The most common types of demand models introduced in the gravity literature belong to either type (i) or (ii), while the additive preferences models in (iii) which—together with (ii) both generalized from the Bergson family (Bergson, 1936)—are

[2]CDE refers to the Constant-Difference-of-Elasticities examined in this paper (see Section III).

[3]They may be regarded as Taylor's series approximations. See Diewert (1971), Lau (1974) and Christensen, Jorgenson and Lau (1973, 1975) for different definitions. Also, Barnett (1983) proves that the terminology NOA used in economic literature is equivalent to that defined in mathematics.

still relatively unknown in the gravity literature.

Some of the NOA systems or type (i) models have been introduced into the gravity models, such as the transcendental-logarithmic (translog) demand system adopted in Novy (2013) and Tan (2013) and studied in several international trade papers such as Feenstra and Weinstein (2017) and Arkolakis, Costinot and Rodríguez-clare (2010), which is fairly rich in its substitution matrix while exhibiting homothetic preferences, and the Almost-Ideal Demand System (AIDS) developed by Deaton and Muellbauer (1980)—a combination of homothetic and non-homothetic systems of equations, which is analogous to the translog but adds non-homothetic behavior; AIDS is well-known in a variety of economic literature and has been introduced into the gravity models such as Fajgelbaum and Khandelwal (2016). However, as is commonly known, AIDS, and generally most other systems derived from the NOA collapse (precisely, the expenditure shares can fall out of regular "zero to one" range) when unbounded real income changes become significantly large (Rimmer and Powell, 1996). This is due to the fact that the parametric conditions alone constrained on the NOA systems do not typically ensure global regularity conditions, but often additionally require appropriate range of the vector of exogenous variables, e.g., see also Lewbel and Pendakur (2009)'s EASI demand system. Furthermore, these models work well (and are less restricted) for small numbers of goods, but for large number (n) of goods NOA implies that the number of estimated parameters becomes square of n, which will become much less tractable (Hanoch, 1975).

For obvious reasons, the parameter parsimony and the imposition of global regularity conditions are often desirable features for specifications of the gravity model of trade. One way to reduce the number of parameters while ensuring global regularity is to impose the separability assumption (Hanoch, 1975; Lau, 1969) or equivalently impose—explicit or implicit additivity. In general, the utility function is implicitly additive if the cardinal utility cannot be algebraically solved in terms of the model's exogenous variables and parameters, whereas the utility expressed in explicit additivity—which is a special case of implicit additivity—may be explicitly derived.[4]

The standard homothetic CES demand in gravity is a particular class of explicitly additive models in Hanoch (1975) or type (ii), which is parsimonious and

[4] The special case can be established in any specialized implicit additivity by allowing the model to be parameterized in such a way that the utility variable can be isolated alone.

has obviously achieved the parameter reduction. These demand models, however, either homothetic or non-homothetic, directly or indirectly additive, impose restrictions that tie substitution effects to income effects. In the gravity models, this restrictive assumption implies that the substitutabilities among all factors are heavily and proportionally dependent on income changes, and invariant to what can be observed from the data.

Some studies are based on non-homothetic explicitly additive models applied in gravity. For example, Caron, Fally and Markusen (2014) adopt Fieler (2011)'s nested non-homothetic demand model which they name Constant Relative Income Elasticity (CRIE). It is a version of explicitly directly additive model. They use CRIE to explain missing trade flows, which depends on demand patterns that are influenced by income growth and non-homothetic preferences through import penetration.[5] Bertoletti, Etro and Simonovska (2018) adopt an explicitly indirectly additive or a generalized version of indirect addilog model that allows additivity in unit-cost prices in their utility formula, which also yields variable substitution elasticities. Analogous to CRIE, the consequences of indirect additivity include constant differences of income elasticities (or can be referred to as CDIE) and that the ratios of elasticities of substitution equal the ratios of the affected quantities of consumer goods.

However, like the case of the CES, these explicit models do not separate substitution effects from income effects because the proportional relationship between cross-price derivatives and income or Engel derivatives are not eliminated (Houthakker, 1960; Hanoch, 1975; Rimmer and Powell, 1996). The constrained relationship between the two was first put forward by Pigou (1910) and has been extensively studied mostly in the 1970s in connection with its unfavorable implications for empirical demand analysis (see, e.g., Deaton (1974); Barten (1977)). The provision of restrictions suggests that high-income countries, which ordinarily have lower income elasticities with respect to *normal* factors, are also restricted to have lower substitutabilities among these factors—regardless of whether the data reveal that the destination prefers imported factors from one country far more than another. This dependence on income effects also implies that the relative substitution elasticities of imported factors between countries A and B with

[5]The model allows non-constancy of substitution elasticities and has constant ratios of income elasticities. It is a special case of Mukerji (1963)'s Constant Ratio of Elasticities of Substitution (CRES) and is closely related to Houthakker (1960)'s direct addilog model where the income elasticity ratios are paired with ratios of elasticities of substitution.

respect to factors imported from country C are only determined by income elasticities with respect to factors imported from A and B; they are invariant to the vector of third country's factors.

One way to remove this restriction (i.e., to allow a more flexible substitution matrix) is to replace explicit additivity with implicit additivity, or type (iii). Comin, Lashkari and Mestieri (2021)'s Non-Homothetic CES (NHCES) and Yilmazkuday (2019)'s specialized NHCES are the only gravity literature that adopt implicit additivity.[6] Nevertheless, these models have constant elasticities of substitution, as with the standard CES models. While they separate the factor substitutabilities from income effects, they imply constant substitutabilities across all factor demands in the world. Imagine a north-south trade scenario where observed trade frictions are typically expected to be lower than south-north trade, estimates calculated from this possibly misspecified gravitational relationship may not be consistent with the real world where the constant substitution elasticity is identified using the bilateral matrix of distances or other types of trade frictions. In other words, in the standard CES model, lower trade costs are typically associated with higher substitutability that is likely occurred in the north-south bilateral link, which may be inconsistent with the peculiar restriction imposed on the link between income and cross-price derivatives of north-south trade in the CES model.

This paper enriches the gravity literature by introducing a new class of non-homothetic consumer preferences. Despite its parsimony, the model relaxes both restrictions of the CES and imposition of substitution effects that are dependent on income effects (see Table 1). The preferences that will be discussed in Section (III) are the constant differences of elasticities of substitution (CDE) introduced in Hanoch (1975), which is non-homothetic, implicitly indirectly additive and non-CES.

III. Theory

I first present a general consumer preference structure which belongs to a class of *implicitly indirect additivity* models, and is a generalization of the Bergson family (Bergson, 1936). I show how the more general demand model can nest a

[6] See also Matsuyama (2019) who adopts implicitly additive model to study patterns of structural change, innovation and trade driven by demand. The supplement to his paper provides other discussions of implicit and explicit additive (direct and indirect) models.

Table 1—Demand Systems in Gravity Models

Gravity Models	**Demand Systems**	**Parsimonious**	**Sub-Inc-Effect Separation**	**Non-CES**
Standard	CES	X		
Fajgelbaum *et al.* (2016)	AIDS			X
Fieler (2011)	Explicitly Direct	X		X
Bertoletti *et al.* (2018)	Explicitly Indirect	X		X
Comin *et al.* (2015)	Implicitly Direct NHCES	X	X	
This paper	Implicitly Indirect CDE	X	X	X

"Parsimonious" ≈ number of parameters is proportional to n goods.

standard CES model as well as an NHCES model. Next, I show how these nested demand models can be generalized to the nested theoretical Hicksian demand of imported factors, in which I define that each $i \in \mathcal{I}$ is a potential source of factor supplies to any destination countries (i.e., with no trade costs adjusting nominal prices). Finally, I fully integrate the more general model into the general equilibrium (GE) gravity model, while taking into account the bilateral matrix of frictions attributed to geographic distances and asymmetric border barriers across regions.

A. *An Implicitly Indirect Additive Demand Model*

The demand system adopted in this paper is the CDE proposed by Hanoch (1975). It is defined by the form of the following identity:

$$G(\frac{\mathbf{p}}{E}, u) = \sum_i \beta_i u^{e_i(1-\alpha_i)} (\frac{p_i}{E})^{1-\alpha_i} \equiv 1, \tag{1}$$

with $\log[u^{e_i}(p_i/E)]$ replacing $u^{e_i(1-\alpha_i)}(p_i/E)^{1-\alpha_i}$ for $\alpha_i = 1$ in the limiting case, where $i \in \{1, ..., \mathcal{I}\}$ indexes consumption goods with the price vector $\mathbf{p} = \{p_i\}_{i=1}^{\mathcal{I}}$; u is the per capita utility, and E represents the per capita income; $\mathbf{p}/E$ may be referred to as the unit-cost or the normalized price of i; β_i, e_i, α_i are distribution, expansion and substitution parameters, respectively. The parametric restrictions for the demand function to be globally valid (monotonic and quasi-concave) are that (i) $\beta_i, e_i > 0 \; \forall \; i \in \mathcal{I}$, and (ii) either $\alpha_i \geq 1$ or $0 < \alpha_i < 1 \; \forall \; i \in \mathcal{I}$.

Utility is *implicitly* defined, as the indirect utility in this model. It cannot be explicitly or algebraically solved as a function of the model's exogenous variables—the price vector and income. The model is indirect (rather than direct) due to the fact that the function is additive in the $\mathcal{I}$ unit-cost prices along consumer's indifference surfaces, whereas directly separable models are additive in $\mathcal{I}$ quantities of consumer goods (Hanoch, 1975). The *quasi* Marshallian correspondence can be derived using Roy's Identity by applying the chain rule to Equation (1):

$$q_i(\mathbf{p}, E, u) = \frac{\beta_i u^{e_i(1-\alpha_i)}(1-\alpha_i)(\frac{p_i}{E})^{-\alpha_i}}{\sum_j \beta_j u^{e_j(1-\alpha_j)}(1-\alpha_j)(\frac{p_j}{E})^{1-\alpha_j}}. \tag{2}$$

Note that U remains in this demand function $q_i(\mathbf{p}, E, u) \equiv h_i(\mathbf{p}, E, u)$, where $h_i(\mathbf{p}, E, u)$ is the *Hicksian* demand correspondence for consumer goods i.[7] The evaluation of u is essential to parameter identification. Since u is not observable, we cannot directly estimate the demand model in a reduced form.[8] However, u can be estimated structurally, which will be discussed later.

Following estimation of the structural parameters α_i, β_i, e_i and u:

$$\sigma_{ij} = \alpha_i + \alpha_j - \sum_k \omega_k \alpha_k - \frac{\Delta_{ij}\alpha_i}{\omega_i}, \tag{3}$$

where $\Delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$ is the Kronecker delta, and the optimal expenditure share:

$$\omega_i \equiv \frac{p_i q_i}{E} = \frac{\beta_i(1-\alpha_i)u^{e_i(1-\alpha_i)}(\frac{p_i}{E})^{1-\alpha_i}}{\sum_j \beta_j(1-\alpha_j)u^{e_j(1-\alpha_j)}(\frac{p_j}{E})^{1-\alpha_j}}. \tag{4}$$

Two important substitution characteristics without formal proof can be stated immediately upon derived Equation (3): (i) the model has constant differences of ES: $\sigma_{ij} - \sigma_{ik} = \alpha_j - \alpha_k$ is always constant, which is independent of any goods i (for $i \neq j \neq k$). The differences of ES may be regarded as one class of *two-factor-*

[7] Also note that the demand function (2) share a common structure with other gravity models. The denominator plays a similar role as the multilateral resistance (MLR) in Anderson and van Wincoop (2003).

[8] It can be shown that u can be eliminated using the method of double log-differencing. However, with this technique, only the $N-2$ substitution parameters α's can be identified, not the e_i's and β_i's; see, e.g., Hanoch (1975) and Surry (1993).

one-price elasticities; (ii) σ_{ij} can be negative if the ES is dominated by large substitution parameters α_k, weighted by expenditure shares of k (or likewise, if the summation of α_i and α_j is sufficiently small)—implying *possibilities of factor complementarity*—which is a cogent yet missing theoretical possibility in standard gravity.[9]

It can be also shown that the income elasticities η_i is given by

$$\eta_i = \frac{e_i(1-\alpha_i) + \sum_k \omega_k e_k \alpha_k}{\sum_k \omega_k e_k} + \alpha_i - \sum_k \omega_k \alpha_k. \tag{5}$$

The income elasticity can be negative if it is dominated by a large weighted sum of substitution parameters α_k, while the expansion parameters e_i control for the elasticity of aggregate expenditure. The direction and magnitude of responses to income shocks depend on the integrated preferences of good i relative to the weighted sum of the preferences given to all other goods—implying *possibilities of inferior goods.* This feature allows destination regions to prefer to import factors from a specific region when their income is constricted, and less so as income rises.

Breaking Linkages between Substitution and Income Effects

Another key property of CDE—and other demand models under the implicit additivity assumption—is that their substitution effects (represented by Allen-Uzawa ES) are separated from income effects (represented by income elasticities). Since it is not necessarily well-known among many international trade studies, it is worth revisiting these details introduced in seminal demand papers such as Houthakker (1960) and Hanoch (1975).

In general, there are two types of linkages between substitution effects and income effects. The first one is that the derived substitution elasticities are functions of income elasticities, distinguished by whether it is multiplicative form of income elasticities in the explicitly direct case (e.g., standard CES) or additive form in the explicitly indirect case (Hanoch, 1975):

[9]This is an important advantage of implicitly indirect models over their direct dual. The indirect models that satisfy implicit additivity and *global regularity conditions* allows goods to be both complements and substitutes. In the direct cases, goods are allowed to be net complements only when the models are locally valid, but all pairs of goods are substitutes under global regularity restrictions. Moreover, factor complementarity may be especially relevant in an era of global value chains, where CDE preferences seem to be reasonable candidates for modelling factor demand.

$$\begin{cases} \sigma_{ij} = \eta_i \eta_j (\sum_k \alpha_k \omega_k) & \text{(Explicitly Direct Case)} \\ \sigma_{ij} = \eta_i + \eta_j (\sum_k \alpha_k \omega_k - 2) & \text{(Explicitly Indirect Case).} \end{cases}$$

Comparing with Equation (3), while the substitution matrix of the CDE is still restricted by its substitution parameters, they do not depend on income elasticities. The second linkage (in explicit demand models) is that there is always either a strict ratio or additive relationship between income and substitution elasticities, specifically,

$$\begin{cases} \frac{\eta_i}{\eta_j} = \frac{\sigma_{ik}}{\sigma_{jk}} & \text{(Explicitly Direct Case)} \\ \eta_i - \eta_j = \sigma_{ik} - \sigma_{jk} & \text{(Explicitly Indirect Case),} \end{cases}$$

whereas these strong peculiar linkages are eliminated in the CDE as shown in Equation (5). In other words, CDE preferences break the link between cross-price derivatives and income derivatives.

In the standard CES-gravity models (i.e., explicitly direct case), $\eta_i = \eta_j = 1$ and $\sigma = \alpha$, where σ is identified using bilateral distances or other measures of trade costs. Given a scenario of south-south trade, the model could be misspecified, e.g., a high-income region buys goods from another high-income region, where the observable trade frictions among them are typically expected to be low. We expect rich countries to have relatively lower income elasticities, but also expect the south-south substitutabilities to be higher. This is impossible in a standard CES specification. In the case of explicit non-homothetic models, the stated restriction implies that the substitutabilities among factors in the vector of countries are fixed by income effects, regardless of whether that region strictly prefers to import factors from another region relative to all other countries, or regardless of whether the elasticity of substitution among two imported factors is zero—as can be both revealed by the data.

True-Cost-of-Living Index

PROPOSITION 1: *Given an implicit indirect additive CDE model $G(\boldsymbol{\xi}, u)$, where $\boldsymbol{\xi} = (1/E)\boldsymbol{p}$ and $E = c(\boldsymbol{p}, u) \equiv \{min\ p'q : f(\boldsymbol{p}, u) \equiv 1\}$ is the expenditure function that satisfies concavity, linear homogeneity and monotone conditions, and $u(\boldsymbol{q})$ is quasi-concave, continuously differentiable, there exists an equivalent*

Lagrangian multiplier $\Lambda^* \equiv PU \equiv \lambda^{*-1}$, *where* λ^* *corresponds to the shadow price in the optimization problem as solved in the explicit direct utility function:* $max\{u(\boldsymbol{q}) : p'q \leq E\}$ *with its gradient vector evaluated at* $\boldsymbol{q}$ *at an interior optimum.*

PROOF: Let us denote by PU the *true-cost-of-living index* in the sense of money metric utility equalling the cost of per unit of utility, then following Chen (2017), the following price index can be derived by finding the total differential of the implicitly defined $G[\boldsymbol{\xi}, u]$ function in Equation (1), with respect to u and E at any arbitrary $\boldsymbol{p}$:

$$PU = \frac{\sum_i \beta_i u^{e_i - e_i\alpha_i - 1}(1-\alpha_i)p_i^{1-\alpha_i}E^{\alpha_i - 1}e_i}{\sum_j \beta_j u^{e_j(1-\alpha_j)}(1-\alpha_j)p_j^{1-\alpha_j}E^{\alpha_j-2}}, \tag{6}$$

where, for any $E = c(\boldsymbol{p}, u)$, $PU \equiv dE/du \equiv \Lambda^*$ is the marginal income (or total expenditure) of utility, which is equivalent to the inverse of the Lagrange multiplier $\lambda^* \equiv du/dE$ resulting from the Jacobian matrix of the corresponding optimization problem: $\max\{u(\mathbf{q}) : p'q \leq E\}$ at interior quantity $\boldsymbol{q}$. □

Note that Equation (6) represents a *Samuelsonian price index* that gives us information about how many dollar units one needs to possess in order to purchase an additional unit of utility. That is, for any fixed $\mathbf{p}$ there exists a cost function $E \equiv E(\mathbf{p}, \mathbf{q})$, satisfying quasi-concavity, homogeneity and monotonicity, that is increasing in the amount of utility consumers attain. Basic microfoundations provide that, for any explicit directly additive demand (e.g., the standard CES system adopted in gravity), it is the Lagrangian multiplier (its inverse) derived from the expenditure minimization (utility maximization) problem. In the implicit CDE demand system, it can be theory-consistently derived by applying total differentiation in (1) with respect to income and utility at any fixed level of prices.

PROPOSITION 2: *If the true-cost-of-living index (PU) of the CDE model is defined as the marginal expenditure required to purchase an additional unit of cardinal utility within a specific economy, then there exists a measure of the aggregate expenditure elasticity* $\theta > 0$, *which controls* PU, *and is operated through the expansion parameter* $e_i > 0 \ \forall \ i$, *where* $PU = \theta(E/u)$.

PROOF: The RHS in Equation (6) can be generalized to $\frac{G_i(\xi_i,u)}{\sum_j \xi_j G_j(\xi_j,u)}$, where $G_i(\xi_i,u) = \frac{\partial G_i(\xi_i,u)}{\partial \xi_i}$, with $\xi_i = \frac{p_i}{E}$ (Hanoch, 1975). Thus, in the pure goods-exchange economy that has one representative consumer, PU takes the following reduced form:

$$PU = \frac{G_i e_i (1-\alpha_i)}{G_j (1-\alpha_j)} \frac{E}{u} \quad \Big(\equiv \vartheta \frac{E}{u}\Big). \tag{7}$$

The first term in the RHS of Equation (7) converges to $\vartheta = \sum_i e_i \omega_i$, which is equivalent to the aggregate expenditure elasticity with respect to $G(\frac{\mathbf{p}}{E}, u)$ introduced in Hanoch (1975). □

θ can be calculated from the estimated structural demand parameters. The second term is the ratio of wealth and utility, where u cannot be substituted out via function operations due to the implicit properties. The aggregate expenditure elasticity $\vartheta(\boldsymbol{p}, e; u, \alpha, \beta,\ e)$ can be interpreted as a conditional aggregate behavioral parameter associated with the consumption level, which governs the non-homotheticity. Holding ϑ constant, higher wealth attained by consumers is associated with higher cost-of-living index. When $\vartheta = 1$, the price index will converge to the same fashion under the standard CES demand: i.e., $P \equiv E/u \equiv PU$, where P is the average price index of goods as defined in Dixit and Stiglitz (1977).

B. Generalization to the Standard CES Demand

The CDE model can be formally parameterized to arrive at a standard CES model. The parameterization provides a convenient way to test the CDE model against the CES under the gravity using the same data and estimation procedure. I show that the CES is a special version of the CDE.

PROPOSITION 3: *Let $G(\boldsymbol{\xi}, u)$ be an implicitly indirect additive utility function of Constant Difference of Elasticities (CDE), then $G(\boldsymbol{\xi}, u)$ can be parameterized to achieve an explicitly indirect Constant Elasticity of Substitution (CES) function, which is identical to its explicitly direct case; it can be further parameterized to yield the standard CES average price index of goods consumption, while satisfying the CES real wealth assumption in a standard general equilibrium framework.*

Definition 1: *A CES real wealth assumption is that price indices of aggregate* *goods consumed by a representative consumer, or cost of per capita utility, equates the per capita income adjusted by the per capita utility of the representative consumer*

(8) $$P \equiv \frac{E}{u} \equiv PU.$$

If $e = 1$ and $\alpha_i = \alpha \ \forall \ i$, then the CDE demand will converge to a standard CES demand.
Proof. See Appendix A.A1.

C. Generalization to Non-Homothetic CES (NHCES)

In this section, I show that the CDE demand nests an implicit indirect NHCES system (that is identical to its direct form). It can be mapped to the version of Comin, Lashkari and Mestieri (2021), provided that their specific NHCES aggregator C corresponds to the standard real consumption index $Q \equiv U$. The linkage to Comin, Lashkari and Mestieri (2021) shows that, while a Marshallian demand estimation procedure is applicable in the NHCES-gravity, it does not, however, work in a CDE-gravity model.

Let $\alpha_i = \alpha \ \forall \ i$, then the CDE demand becomes an implicit indirect NHCES demand [10]

(9) $$G\Big(\frac{\mathbf{p}}{E}, u\Big) = \sum_i \beta_i u^{e_i(1-\alpha)} \Big(\frac{p_i}{E}\Big)^{1-\alpha} \equiv 1,$$

which is identical to the implicit direct NHCES

(10) $$F(\mathbf{q}, u) = \sum_i k_i u^{e_i(g-1)} q_i^{1-g} \equiv 1,$$

if $\alpha = 1/g$ and $\beta_i = k_i^{\alpha}$. Both systems arrive at the same Hicksian demand

[10] The parametric restrictions are similar to the CDE function (1), except $0 < \alpha < 1$ or $\alpha > 1$. Also note that we may isolate E in (9) so that $E(u, \mathbf{p}) \equiv [\sum_i \beta_i u^{e_i(1-\alpha)} p_i^{1-\alpha}]^{\frac{1}{1-\alpha}}$, which is the expenditure function defined as Equation (3) in Comin, Lashkari and Mestieri (2021).

function

$$h_i = \frac{\beta_i u^{e_i(1-\alpha)} p_i^{-\alpha} E}{\sum_j \beta_j u^{e_j(1-\alpha)} p_j^{1-\alpha}} = \beta_i u^{e_i(1-\alpha)} \left(\frac{p_i}{E}\right)^{-\alpha}. \tag{11}$$

Proof. See Appendices A.A2 and A.A3.

Demand Mapping to Comin *et al.*

I adopt Comin, Lashkari and Mestieri (2021)'s notation to show that the demand system in their paper can be mapped from the implicitly direct NHCES written in the structural form as equation (10). I impose the parametric equalization in equation (10) as follows: 1) $g = \frac{1}{\sigma}$, so $1-g = \frac{\sigma-1}{\sigma}$ and $g-1 = \frac{1-\sigma}{\sigma}$; this also implies that $\alpha = \sigma$; 2) $\beta_i = \Omega_i$, so $k_i = \beta_i^{\frac{1}{\sigma}} = \Omega_i^{\frac{1}{\sigma}}$; and 3) $\epsilon_i = e_i \frac{g-1}{g}$=$e_i(1-\alpha)$, so $e_i = \epsilon_i \frac{g}{g-1} = \frac{\epsilon_i}{1-\alpha}$.

Following Dixit-Stiglitz-Melitz (Dixit and Stiglitz, 1977; Melitz, 2003), consumer preferences can be considered as an aggregate good index such that $Q \equiv U$. Here, I use their definitions to denote the aggregator index $Q \equiv C$ with the demand choice vector $q_i = C_i$, then equation (10) will converge to

$$\sum_i (\Omega_i C^{\epsilon_i})^{\frac{1}{\sigma}} C_i^{\frac{\sigma-1}{\sigma}} \equiv 1. \tag{12}$$

The parametric restrictions are the same as in Comin, Lashkari and Mestieri (2021) and they are fully consistent with the identity (9) in Hanoch (1975): i) $0 < \sigma = \alpha \neq 1$; ii) $\Omega_i = \beta_i > 0$; and iii) $\sigma < 1$ ($\sigma > 1$) implies $\epsilon_i > 0$ ($\epsilon_i < 0$) $\forall$ i. It is intuitive to see that, since $e_i = \epsilon_i \frac{g}{g-1} > 0 \ \forall \ i$, we must have $\epsilon_i > 0$ ($\epsilon_i < 0$) whenever $g > 1$ ($g < 1$), which implies that $\sigma = \frac{1}{g} < 1$ ($\sigma = \frac{1}{g} > 1$). Furthermore, $\epsilon_i = e_i(1-\alpha)$ indirectly controls the income elasticity of good i, through the expansion parameter e_i and the partial substitution elasticity α. The derived identity of income elasticity of NHCES demand is

$$\eta_i = \alpha + \frac{e_i(1-\alpha)}{\sum_k e_k \omega_k}. \tag{13}$$

The income elasticity of good i is increasing in ϵ_i through the direct channel $\epsilon_i = e_i(1-\alpha)$, and is decreasing in ϵ_i through the expansion preferences $e_k = \frac{\epsilon_k}{1-\alpha}$ as an indirect channel, weighted by expenditure shares of good k. In addition, the income elasticity of demand is directly increasing in α $(= \sigma)$, but is indirectly decreasing in α through the expanded substitution preferences $e_i\alpha$. Equation (13) can be rewritten in an identical form

$$\eta_i = \sigma + \frac{\epsilon_i(1-\sigma)}{\sum_k \epsilon_k \omega_k}, \tag{14}$$

where ω_k is the expenditure share of good k; since the numerator and denominator share $(1-\alpha)^{-1}$ as a common factor, such that e_i and ϵ_i can be transformed from one to another, it turns out that ϵ and σ play equivalent roles in Equation (14) as e and α in (13).

D. A Hicksian Import Demand

Here I introduce the *Hicksian approach* to CDE-gravity estimation, where an explicit average price index of goods P cannot be derived. Furthermore, in the CDE demand Equation (2) where $q_i(\mathbf{p}, E, u) = h_i(\mathbf{p}, E, u)$ is a function of both income and utility, neither the function of E and u can be explicitly derived. In this special demand model, the standard form of either Marshallian (uncompensated) or Hicksian (compensated) demand correspondence does not exist. Instead of estimating the Marshallian demand in the gravity model, the Hicksian approach targets a *Hicksian demand* for imported factors. It allows evaluating the cardinal values of per capita utility using the Hicksian price index for welfare PU.

This approach can also be applied to the standard CES-gravity as well as to the NHCES-gravity.[11] It is shown that (in III.B) for the standard CES P is equivalent to the price of utility PU. In this case, estimation procedures for Marshallian and Hicksian demand are equivalent. For the non-homothetic demand systems, the expenditure function can be algebraically solved in terms of observables, where the utility index can be linked with the average price of goods index $P = E/u$ in the Marshallian demand estimation, as implemented in Comin, Lashkari and Mestieri (2021). I first show that we can estimate a Hicksian import demand in

[11] See `www.gams.com/solvers/mpsge/markusen.htm` (James Markusen and Thomas Rutherford).

the NHCES-gravity model—which is relatively well-known while beling closely related to the CDE—using a linkage between the two price indices P and PU.

I begin by showing this linkage in the NHCES demand adopted in Comin, Lashkari and Mestieri (2021), where the average price index of goods P can be explicitly derived. I show how the Hicksian price index PU can be integrated into the Hicksian demand of the gravity model, referred to as the Hicksian approach here. Next I use this approach and integrate PU in the more general Hicksian import demand of a CDE, where P cannot be derived.

Application in the NHCES

Comin, Lashkari and Mestieri (2021) constructs an average price index of NHCES that equals real income, that is, the per capita income adjusted by real consumption index C, i.e., $P \equiv \frac{E}{C}$. The implicit demand yields an expenditure function (see Footnote 10) that can be replaced by an identity for the real wealth, while the aggregator consumption index is substituted by the function of expenditure shares and observables. It then gives rise to the real-wealth conforming non-homothetic price index of goods, equivalently, as follows

$$P = \Big[\sum_i (\beta_i p_i^{1-\alpha})^{\frac{1}{e_i}} (\omega_i E^{1-\alpha})^{\frac{e_i - 1}{e_i}}\Big]^{\frac{1}{1-\alpha}}, \tag{15}$$

with ω_i defined as the expenditure share of i, and $\frac{1}{e_i} = \frac{1-\sigma_i}{\epsilon_i}$ equaling χ_i specified in Comin, Lashkari and Mestieri (2021)'s paper.

Note that the two price indices P and PU are systematically different. Let us define the price index in the sense of money metric utility (equaling cost per unit of cardinal utility) specialized to the standard implicit indirect NHCES in (9) as

$$PU \equiv \sum_i e_i \omega_i \frac{E}{u} \Big(\equiv \sum_i \frac{\epsilon_i}{1-\sigma} \omega_i \frac{E}{C} \Big), \tag{16}$$

with $\omega_i \equiv \frac{p_i q_i}{E}$ satisfying

$$\omega_i = \beta_i u^{e_i(1-\alpha)} \left(\frac{p_i}{E}\right)^{1-\alpha}. \tag{17}$$

The expenditure share in Comin, Lashkari and Mestieri (2021) is equivalent to

$$\omega_i = \beta_i \left(\frac{E}{P}\right)^{e_i(1-\alpha)} \left(\frac{p_i}{E}\right)^{1-\alpha}. \tag{18}$$

If $PU \equiv P$, then Equations (17) and (18) cannot be both satisfied as long as $\frac{E}{PU} \not\equiv C \equiv u$, whereas $\frac{E}{PU} \equiv C$ would imply that Equation (16) is violated so long as $\epsilon_i \neq 1 - \sigma$.[12] When $\epsilon_i = 1 - \sigma$, we are back to the standard CES preferences. Hence, $PU \not\equiv P$. As can be readily seen, the per capita utility u in Hanoch (1975) corresponds to the non-homothetic CES aggregator index C for any identically calibrated per capita income. Thus we may obtain an interpretable relationship between the two price indices:

$$\frac{PU}{P} = \vartheta, \tag{19}$$

where, similar to the CDE demand, $\vartheta = \sum_i e_i \omega_i$ is the expenditure-share weighted average of expansion parameters. The price of utility equals the average price of goods consumption weighted by the non-homothetic aggregate expenditure elasticity ϑ.

If a choice bundle is dominated by sufficiently many luxury (subsistence) goods i's, indicated by higher (lower) expansion elasticity e_i relative to their weighted average in Equation (13), then the average cost of utility is higher (lower) than the average price of goods. Therefore, it seems appropriate to name ϑ the elasticity of the price of goods with respect to the cost of utility. For the general types of homothetic CES preferences, it intuitively suggests that the cost of utility is linear in the price of goods. In the standard form, it implies that the average costs of utility and goods consumption are equivalent, $PU \equiv P$ and $\vartheta = 1$, provided that $e_i = 1$ or $\epsilon_i = 1 - \sigma \; \forall \; i$.

[12] Hanoch (1975)'s demand theory suggests that, in the NHCES functions, for e_i's satisfying $\min_i e_i < 1 < \max_i e_i$, constant returns to scale (in terms of income to utility) is possible when $\sum_i e_i \omega_i = 1$.

Now rearranging the price index definition in (16) and substituting $u \equiv \sum_i e_i \omega_i \frac{E}{PU} = \vartheta \frac{E}{PU}$ into (11), yielding the Hicksian demand as a function of PU:

$$h_i = \beta_i \vartheta^{e_i(1-\alpha)} \left(\frac{E}{PU}\right)^{e_i(1-\alpha)} \left(\frac{p_i}{E}\right)^{-\alpha}, \tag{20}$$

where the derived theory-consistent price index PU, equalling unit expenditure function for utility, is jointly and analytically defined by (16) and (17). In this way, we can construct a gravity equation based on (20), given that PU can be simultaneously structurally estimated.

Application in the CDE

In a similar way, using (2) and (6) we may define the Hicksian demand of CDE as follows:

$$h_i = \frac{\beta_i(1-\alpha_i)\vartheta^{e_i(1-\alpha_i)}(\frac{E}{PU})^{e_i(1-\alpha_i)}(\frac{p_i}{E})^{-\alpha_i}}{\sum_j \beta_j(1-\alpha_j)\vartheta^{e_j(1-\alpha_j)}(\frac{E}{PU})^{e_i(1-\alpha_j)}(\frac{p_j}{E})^{1-\alpha_j}}, \tag{21}$$

except that $\omega_i[PU]$ takes the form of Equation (4) [(6)].

The Hicksian approach is suitable for the CDE framework in the gravity estimation, since i) E is implicitly defined, we may not *bypass* U (the non-homothetic real consumption) to directly evaluate P; (ii) instead, it allows evaluation of the utility index via PU, which is important in the estimation methodology to be discussed in Section (V); and iii) PU is economically defined (in terms of the definition of shadow prices and the market equilibrium condition for the real consumption), thus can be categorized in a mixed complementarity problem (MCP) of a GE framework (see Section V.B). This formulation allows the algorithm used in the estimation—Mathematical programming with equilibrium constraints (MPEC)—to tackle a numerical solution by evaluating the cost of utility as another feasibility constraint (in addition to other complementarities) in the estimation equation.

Nesting Hicksian Import Demand In the nested estimation procedure, if we allow $\alpha_i = \alpha \ \forall \ i$, then Equation (21) simplifies to (20). In this case, the

Hicksian import demand of a CDE will transform to the Hicksian import demand of an NHCES. Further, if $e_i = 1 \; \forall \; i$, then Equation (20) simplifies to (A8), and the Hicksian import demand of an NHCES will transform to the *Hicksian* import demand of a standard CES. In this case, because the aggregate expenditure elasticity $\vartheta = 1$ and $PU \equiv P$, the estimation framework is equivalent to the Marshallian demand estimation in the standard CES-gravity models.

IV. Gravity Model

I construct a GE gravity model with the non-homothetic CDE demand system. My approach is conceptually similar to that of Anderson and van Wincoop (2003) who develop and estimate a theoretic gravity model of aggregate trade flows based on a homothetic CES demand system, except that I introduce a more flexible demand system by applying an MPEC program similar to that which Balistreri and Hillberry (2007) apply to the Anderson and van Wincoop (2003) model. I combine the gravity theory and estimation methodologies of the two papers, while applying what is referred to earlier as the Hicksian approach developed by Markusen and Rutherford (see also Footnote 11).

The Hicksian approach proposed in this gravity framework is theory-consistent with the specification of the Marshallian-demand gravity estimations such as Comin, Lashkari and Mestieri (2021). The key difference between the standard Marshallian approaches and mine is that I evaluate the value of utility u using the non-homothetic shadow prices PU. This implementation gives an additional advantage in that it allows population data (in addition to income data) and the MPEC program to be critically useful in identifying the parameters of import demand and the trade cost elasticity (as well as asymmetric border frictions).

The gravity model in this paper is specialized to global factor demand. The starting point of the model setup is a transition of the per capita household demand into imported factor demand under gravity in a GE framework. Following Adao, Costinot and Donaldson (2017) and Costinot and Rodríguez-Clare (2018), I assume that the value of imported goods is realized as foreign factor payments so that each i represents a national factor. Each region trades aggregate "factors" with other countries, with the value of factors equaling the value of trade flows in aggregate terms. Henceforth, the terms "goods" and "factors" are used interchangeably. The value of trade flows under gravity are in U.S. dollars determined by aggregate quantity of goods exchanged and bilateral prices. Each region observes goods to be potentially exported from the origin associated with an import demand in the destination, including goods produced at home.

A. *Preliminaries*

I denote goods by $i = \{1, ..., \mathcal{I}\}$ and $l = \{1, ..., i-1, i+1, ..., \mathcal{I}+1\}$ as region of origin and destination when $i \neq l$, respectively, where $\mathcal{I}+1$ is the additional element denoting the rest of the world. When $i = l$, all activities and variables are domestic-specific. Let d_{il} be the distance between any pairs of countries and $\{\tau_{il} = d_{il}^{\rho}\}_{i,l \in \mathcal{I}} \in \overline{\mathbb{R}}_{++}$: $\mathbb{R}_{++} \cup \{\infty\}$ be the iceberg trade costs between them, where ρ is the trade cost elasticity with respect to distance. Following Eaton and Kortum (2002)'s Samuelsonian iceberg assumption, $\tau_{il} > 0$ $(= 1)$ for any foreign (domestic) supply and the triangle inequality is satisfied for any trilateral relationships.[13] Let L_l be destination's population and $Y_l = w_l L_l$ be the total nominal national income. Here, $p_{il} = FOB_i \tau_{il}$ is the bilateral total prices (FOB price at origin inflated by bilateral trade costs) for the output units supplied from origin i. In Section (IV.G), I allow border costs to impact τ_{il} so the trade costs are composed of both distance and border effects.

B. *Model Formulation*

I translate the CDE per capita household demand into gravitational import demand. Per capita income is calculated with national population and the aggregate income, i.e., $E_l = Y_l/L_l$. The inclusion of population terms introduces additional cross-region data variation that helps to identify structural parameters in the empirical gravity model. The CDE-gravity demand variable (i.e., $q_{il} \equiv h_{il}$) has a partial elasticity with respect to population of $-\alpha_i$:[14]

$$q_{il} = \frac{\beta_i u_l^{e_i(1-\alpha_i)} (1-\alpha_i)(FOB_i \tau_{il})^{-\alpha_i} Y_l^{\alpha_i} L_l^{-\alpha_i}}{\sum_j \beta_j u_l^{e_j(1-\alpha_j)} (1-\alpha_j)(FOB_j \tau_{jl})^{1-\alpha_j} Y_l^{\alpha_j - 1} L_l^{1-\alpha_j}}. \tag{22}$$

[13]The triangle inequality relationship, i.e., $\tau_{ij}\tau_{jl} \geq \tau_{il}$ $\forall\, i, j, l \in \mathcal{I}$, ensures that no shipments through any intermediate hub is less expensive than direct transportation.

[14]The MPEC essentially unpacks the equilibrium binding constraint defining PU and converts (21) to Equation (22) in terms of the utility index and observables.

The per capita demand of region i's output units is determined by consumer preferences, nominal incomes, populations and bilateral prices. With total quantity shipped from i to l (taking importer l's population into account, i.e., $Q_{il} = q_{il}L_l$), the corresponding national aggregate bilateral trade flows in dollar values, by multiplying Equation (22) by $FOB_i\tau_{il}$ and L_l, is

$$X_{il} = \frac{\beta_i u_l^{e_i(1-\alpha_i)}(1-\alpha_i)(FOB_i\tau_{il})^{1-\alpha_i}Y_l^{\alpha_i}L_l^{1-\alpha_i}}{\sum_j \beta_j u_l^{e_j(1-\alpha_j)}(1-\alpha_j)(FOB_j\tau_{jl})^{1-\alpha_j}Y_l^{\alpha_j-1}L_l^{1-\alpha_j}}. \tag{23}$$

Equation (23) is the gravity condition, where the value of bilateral trade flows are satisfied for any locations of i and l in the world. In a structural gravity model, this equation is used both as a crucial part of the objective estimation equation and in the counterfactual analysis of predicted bilateral trade flows.

C. *General Equilibrium Environment*

The bilateral trade flows expressed in Equation (23) equals the value of total production in i for the output demanded in l, thus in equilibrium the total national income in i satisfies $Y_i = \sum_l X_{il}$. Total bilateral trade costs follow the iceberg assumption. Let K_{il}^0 denote the distributed endowment in i that is supplied to l discounted by iceberg melt, which equals the total quantity demanded in l from i, e.g., $K_{il}^0/\tau_{il} = Q_{il}$. The associated dollar value in the bilateral shipments is $FOB_i\frac{K_{il}^0}{\tau_{il}} = FOB_iQ_{il} = X_{il}/\tau_{il}$. The total quantity shipped from origin i to the world is given by $K_i^0 = \sum_l \tau_{il}Q_{il} = \sum_l \tau_{il}q_{il}L_l$, which equals the fixed total endowment at origin i:

(24)
$$K_i^0 = \sum_l \frac{\beta_i u_l^{e_i(1-\alpha_i)}(1-\alpha_i)FOB_i^{-\alpha_i}\tau_{il}^{1-\alpha_i}Y_l^{\alpha_i}L_l^{1-\alpha_i}}{\sum_j \beta_j u_l^{e_j(1-\alpha_j)}(1-\alpha_j)(FOB_i\tau_{il})^{1-\alpha_j}Y_l^{\alpha_j-1}L_l^{1-\alpha_j}} \quad \forall i \quad \text{(Goods Market Clearing)}.$$

The fixed endowment K_i^0 can also be interpreted as the total quantity of output units that needs to be produced in order to meet the total demand across the globe. The product of K_i^0 and FOB_i yields the definition of aggregate income in each region i:

$$Y_i = FOB_i K_i^0 \quad \forall i \quad \text{(Income Definition)}, \tag{25}$$

with the definition for the marginal cost of utility in destination l:

(26)
$$PU_l = \frac{\sum_i \beta_i u_l^{e_i(1-\alpha_i)-1}(1-\alpha_i)(FOB_i\tau_{il})^{1-\alpha_i} Y_l^{\alpha_i} L_l^{1-\alpha_i} e_i}{\sum_j \beta_j u_l^{e_j(1-\alpha_j)}(1-\alpha_j)(FOB_j\tau_{jl})^{1-\alpha_j} Y_l^{\alpha_j-1} L_l^{2-\alpha_j}} \quad \forall l \quad \text{(Price Index Definition)},$$

while Equation (27) implies the *non-homothetic* exporter i's benchmark utility equals *weighted* benchmark per capita income adjusted by fitted equilibrium price index for utility:

$$u_i = \vartheta_i \frac{E_i}{PU_i} = \vartheta_i \frac{Y_i}{L_i}\Big(PU_i\Big)^{-1} \quad \forall i \quad \text{(Benchmark Utility Definition)}, \tag{27}$$

There is one single labor factor L_i in each region. It is assumed to be fixed from the benchmark population pool, since the model concentrates on the distributed output of national factors around the world (rather than the operation of individual labor markets). As we move to counterfactual analysis, the proportion of workers being hired will remain, but their output will be likely reallocated across the globe.

In the labor market, I introduce a Ricardian unit labor requirement ϕ_i. The purpose of this productivity parameter is to allow the factor income (average wages) to vary with the GDPs observed from the data, so one unit of produced goods in each origin i requires ϕ_i units of labor, satisfying

$$L_i = \phi_i K_i^0 \quad \forall i \quad \text{(Labor Market Clearing)}. \tag{28}$$

Under full employment, total population must equal the total amount of labor hired for production. Division of (25) by (28) automatically implies zero-profit condition for *firms*:

$$\phi_i \frac{Y_i}{L_i} = \phi_i E_i = FOB_i \quad \forall i \quad \text{(Zero Profit Condition for Firms)}. \tag{29}$$

Equation (29) also motivates the *cross-region differences* in the gap between average wages and prices in this non-homothetic demand system, which is captured by ϕ_i.

D. *Identification of Demand Parameters and the Elasticity of Trade Costs with Respect to Distance*

The expression of the log-linear gravity regression takes the following form. It is consistent with the suggested estimation equation in Hanoch (1975), except that I introduce the iceberg trade cost factor and the population variable:

$$\begin{aligned} \log X_{il} = {} & \log \mid \beta_i(1-\alpha_i) \mid + e_i(1-\alpha_i)\log u_l + (1-\alpha_i)\log FOB_i \\ & + \rho(1-\alpha_i)\log d_{il} + \alpha_i \log Y_l + (1-\alpha_i)\log L_l \\ & - \log \Bigg| \underbrace{\Bigg[\sum_j \beta_j u_l^{e_j(1-\alpha_j)} (1-\alpha_j)(FOB_j \tau_{jl})^{1-\alpha_j} Y_l^{\alpha_j - 1} L_l^{1-\alpha_j} \Bigg]}_{\text{Multilateral Resistance}} \Bigg| + \varepsilon_{il}. \end{aligned} \tag{30}$$

Bilateral trade flows are determined by f.o.b. prices of origins, bilateral distances, national income, population and per capita utility levels at destinations, as well as the aggregated multilateral effects of these terms.[15]

The structural procedure can separately identify the elasticity of trade costs with respect to distance ρ, the elasticity of trade with respect to trade costs τ_{il}'s, and $1-\alpha_i$. Variation in the population data is key to identification; it allows identification of α. Another equally important identification strategy is that the estimation procedure allows solving for cardinal values of u. Variation in u allows e_i's to be identified, given α_i's.

The key point is that, if we can control for the MLR, the variables Y and L can help pin down α_i's. With the identification of α_i's, the distance, e.g., d_{il}'s, can pin down parameter ρ. The estimation of e_i's, however, is impossible without

[15] The expression inside the square bracket corresponds to the MLR as the average trade barrier in Anderson and van Wincoop (2003)

evaluation of u_l, which cannot be accomplished in conventional methods (e.g., reduced-form regressions). However, an MPEC procedure allows joint calculation of MLR and cardinal values of u_l's (given equilibrium constraints and the equation that implicitly defines utility), even as vectors of α_i, e_i and τ_{il} are being estimated. The MPEC will evaluate u_l's using the definition of unit expenditure function for utility PU (26) and the implicitly indirect demand function (31):

$$(31) \qquad \sum_i \beta_i u_l^{e_i(1-\alpha_i)} (\frac{FOB_i \tau_{il}}{E_l})^{1-\alpha_i} \equiv 1 \quad \forall l \qquad \text{(CDE Defining Equation)}.$$

E. Parametric Conditions

The parameter estimates $\alpha_i \geq 1 \ \forall \ i$ are expected from Equation (30) on the basis of standard gravity assumption where distances increase trade frictions and thus $\rho > 0$. It follows that α_i must be greater than 1 to reflect distance impeding trade flows, whereas $0 < \alpha_i < 1$ is inconsistent with the key stylized fact in the gravity literature since it would imply a positive coefficient on logarithmic distances. However, for estimation purposes, I do not impose any additional restrictions on α_i's except that they are constrained to be strictly greater than zero in order to satisfy the stated regularity conditions of the demand system.

Consistent with the theoretical demand system, $\alpha_i = 1$ is allowed for some extreme cases in some l (e.g., some political regime l might have high autarky level and/or import demand biases in which trade is invariant to utility level and distances).[16] All other parametric restrictions in the CDE demand system are standard following Hanoch (1975).[17]

F. Predictions of Zero Trade Flows

The demand theory in Hanoch (1975) and parameter restrictions discussed above suggest that the model can be parameterized to predict zero trade flows.

[16]There are counter cases where evidence shows that standard gravity assumption on distance does not hold. Melitz (2007) (European Economic Review, 2007) shows that under North-South bilateral pairs there are proportional increases in trade with distance. Buch, Kleinert and Toubal (2004) (Economic Letters, 2004) shows that in the extreme case where standard gravity assumption is violated, the impacts of distance could be captured by the constants.

[17]It is worth highlighting that the MLR always increases with multilateral trade barriers (e.g., τ_{jl} $\forall j, l$) as in Anderson and van Wincoop (2003), invariant to the choice of restrictions on α_i from the CDE theory.

Specifically, the following cases allow zero bilateral or multilateral trade flows:

Case (i) *Driven by small substitution and distribution parameters:*
When $lim_{\alpha_i \to 1}$ *or* $lim_{\beta_i \to 0}$*, zero trade flows may exist between* i *and* l $\forall$ l*.*

Case (ii) *Driven by large expansion parameters:*
When $lim_{e_i \to \infty}$ $\forall$ $\alpha_i \in [1, \infty]$*, zero trade flows may exist between* i *and* l $\forall$ l*. This may also occur when* $lim_{\alpha_i \to \infty}$*, while the re-scaled normalized prices* ξ_i*'s are strictly less than 1, due to the purpose of numerical stability in the estimation problem.*

Case (iii) *Driven by high distance elasticity:*[18]
When $lim_{\rho \to \infty}$ $\forall$ $\alpha_i \in [1, \infty]$*, zero trade flows may exist between* i *and* l $\forall$ l:

Case (iv) *Driven by long exogenous bilateral distances:*
When $lim_{d_{il} \to \infty}$ $\forall$ $\alpha_i \in (1, \infty] \cap \beta_i \neq 0$*, while either condition under (ii) is not satisfied and holding* ρ *constant, then zero trade flows may only exist between bilateral* i *and* l*.*

Case (i) states that if either parameter approaches to the limits specified above, then the estimation will not succeed, invariant to exogenous variables. Case (ii) holds when α_i is extremely large even if $e_i = 0$, as long as there is a trade friction between i and l, whereas if $\tau_{il} = 1$ (e.g., no trade costs), then the expansion parameter e_i must not be anywhere near zero in order for case (ii) to hold (automatically satisfied for global validity). Case (iii) is among the most intuitive ones, if ρ approaches infinity (e.g., at the extreme where policy exclusively impedes trade), then there is no trade anywhere in the world—autarky condition, invariant to consumer preferences. Similarly, in case (iv) where distance matters without taking trade policy into account, extremely far distances between i and l will yield zero bilateral trade flows, which is a standard result in gravity models. However, case (iv) will collapse if (ii) holds simultaneously.

[18] Note that high distance elasticity *per se* do not automatically provide zero trade prediction (e.g., CES-gravity models), but the parametric condition for α_i's in the CDE model allows demand to cut the y-axis as ρ approaches infinity.

G. Asymmetric Border Effects

I now allow bilateral border charges to be included in τ_{il}. Balistreri and Hillberry (2007) propose an estimation framework that measures bilateral border costs between the U.S. and Canada based on Anderson and van Wincoop (2003), while accounting for asymmetric border effects. They introduce an identity that is equivalent to the following expression:

$$\tau_{il} = d_{il}^{\rho}[\exp(\delta_{il})]^{1-\text{dummy}_{il}}, \tag{32}$$

where $\delta_{il} \equiv \ln(1+\bar{T}_{il})$ and $\bar{T}_{il}$ is the tariff equivalent of border frictions ($\delta_{il} = \delta_l$ $\forall\ i$, $\bar{T}_{il} = \bar{T}_l\ \forall\ i$); dummy_{il} is a dummy variable set of source-home consumption, with $\text{dummy}_{il} = 0$ denoting cross-border shipments and $\text{dummy}_{il} = 1$ if otherwise. Following Balistreri and Hillberry (2007), I include the bilateral border coefficient δ_{il}, which embeds in the empirical form of Equation (32) in the transformed estimation equation specialized to Equation (30) for the purpose of econometric analysis:

$$\begin{aligned}
\log X_{il} = {} & \log \mid \beta_i(1-\alpha_i) \mid + e_i(1-\alpha_i)\log u_l + (1-\alpha_i)\log FOB_i \\
& + \rho(1-\alpha_i)\log d_{il} + \alpha_i \log Y_l + (1-\alpha_i)\log L_l \\
& + (1-\text{dummy}_{il})(1-\alpha_i)\delta_{il} \\
& - \log\left|\Big[\sum_j \beta_j u_l^{e_j(1-\alpha_j)}(1-\alpha_j)(FOB_j\tau_{jl})^{1-\alpha_j} Y_l^{\alpha_j-1} L_l^{1-\alpha_j}\Big]\right| + \varepsilon_{il}.
\end{aligned} \tag{33}$$

Equation (33) involving border coefficients is consistent with Balistreri and Hillberry (2007) and Anderson and van Wincoop (2003) . Since the parametric restrictions on α_i's is piece-wise, thus it must follow that $\delta_{il} > 0$ (< 0) when $\alpha \geq 1$ ($0 < \alpha < 1$). Since I impose $\alpha_i > 0\ \forall\ i$ in the econometric exercise, this implementation then allows the signs of δ_{il}'s to be fully determined by the data and the structural import demand designed in the GE system.

V. Structural Estimation

There are various concerns over standard OLS-based estimations of log-linearised gravity. The structural estimation procedure minimizes the sum of least squares using observed trade flow values or import shares in the objective function to predict fitted flows. Some well-known concerns include heteroskedasticity, logarithmic transformation issues, etc. Another important issue is the zero-trade flows which are often in the observed data. Using a conventional log-linear estimator may involve some imperfect ways of handling zero trade flows (e.g., omission of zero pairs, adding a small value, or Heckman's two-step), if they do occur.

An alternative method is to use the Poisson Pseudo Maximum Likelihood (PPML) estimator, which has become standard in the gravity literature. The PPML estimator allows pervasive zero bilateral trade flows to enter into the estimation framework albeit with small effective weights given by conditional mean (Larch et al., 2019). Following Gourieroux, Monfort and Trognon (1984) and Santos Silva and Tenreyro (2006), the PPML estimator based on the linear expression of bilateral trade flows in Equation (16) is given by

$$\begin{aligned} X_{il} = \exp\Bigg\{ & \Big|\log \beta_i(1-\alpha_i)\Big| + e_i(1-\alpha_i)\log u_l + (1-\alpha_i)\log FOB_i \\ & + \rho(1-\alpha_i)\log d_{il} + \alpha_i \log Y_l + (1-\alpha_i)\log L_l \\ & + (1-\text{dummy}_{il})(1-\alpha_i)\delta_{il} \\ & - \log\Bigg[\Big|\sum_j \beta_j u_l^{e_j(1-\alpha_j)}(1-\alpha_j)(FOB_j\tau_{jl})^{1-\alpha_j} Y_l^{\alpha_j-1} L_l^{1-\alpha_j}\Big|\Bigg]\Bigg\} + \Upsilon_{il} \\ = {} & \exp(x_{il}b_i) + \Upsilon_{il}, \end{aligned} \tag{34}$$

where $\exp(x_{il}b_i)$ is a reduced form in which x_{il} and b_i are exogenous variables and unknown CDE parameters (plus ρ and δ_{il}), respectively; and Υ_{il} is the new error term.

The constrained optimization problem is to maximize (35)

(35)

$$L(b_i) = constant - \sum_i \sum_l exp(x_{il}b_i) + \sum_i \sum_l y_{il}x_{il}b_i \qquad \text{(Objective Function)},$$

subject to the computable general equilibrium (24)-(29) with implicit utility defining equation (31), and the following parametric restrictions: (1) $\alpha_i, \beta_i, e_i > 0$, $\forall$ i; (2) $\phi_i > 0$ $\forall$ i; (3) $u_l > 0$ $\forall$ l at all $\frac{\mathbf{p}}{E} \gg 0$; $\rho > 0$ and (5) $\delta_{il} \geq 0$.

A. *Computation Challenges and Solution*

The estimation of the CDE demand system is an empirical barrier that is difficult enough to overcome. Due to its implicit properties, it is ideal for computational algorithms to find the *implicitly defined indirect relationships* (between variables and structural parameters) by first knowing the solution's starting values that are feasible with respect to these relationships. This procedure, however, has not been successfully implemented in the past. One example is the Global Trade Analysis Project (GTAP) Model—a computable general equilibrium model used for a large volume of applied work. The GTAP Model uses the CDE functional form as preferences of private households in its model, but because of the estimation challenges, has calibrated CDE parameters from other simpler demand models, instead of direct estimation.

Coupled with the GE gravity framework, the problem introduced in this paper is undoubtedly much more complex than estimating a pure CDE household demand. The first challenge is to find a feasible region that best characterizes the implicit indirect relationships. The second challenge is that the model involves both complementarity problems and highly nonlinear system of equations as constraints. The third challenge is to deal with an unequal number of equations and unknowns. There are effectively $6\mathcal{I}$ GE equations (with f.o.b prices as the numeraire), which solve $6\mathcal{I} + 1$ parameters: $\alpha_i, \beta_i, e_i, u_i, \phi_i$, δ_{il} and ρ.

The empirical strategy is to eliminate utilities in the *ex-ante* estimation equation (34). It involves function transformations using double log-differencing. The first log-difference would eliminate the MLR term, while the second log-difference would eliminate utilities. The technique using the first log-difference has been employed by economists who apply CDE as a production function in economic models, where u is realized as observable production outputs pulled from the database (e.g., Surry (1993)). There are two major issues with this method: (i) not all structural demand parameters can be identified from the transformed equations (only α_i and $\mathcal{I} - 2$ equations are estimated); (ii) it disconnects with this gravity framework as u is eliminated—the constraints defining national income and benchmark utilities are thus non-binding. Subsequently, without evaluation of u, we cannot identify the elasticity of trade costs with respect to distance ρ and the border coefficients δ_{il}. For these primary reasons, we need a structural procedure to identify the demand parameters in the model, while calculating the cardinal measures of u. We also need an appropriate algorithm that allows the procedure to be implemented, which is the MPEC.

B. Inequality Constraints as an MCP

Before discussing an MPEC procedure, I first illustrate the mixed complementarity problem (MCP) modeled as constraints in the constrained optimization problem.

Set of GE constraints While Equation (25) indeed belongs to a formulation of equality constraints, the rest are implicitly defined as complementarity problems, thus can be characterized in an MCP. The market clearing conditions imply that the strict equalities would hold if and only if the associated goods or factors are free of charge:[19]

$$K_i^0 \geq \sum_l \frac{\beta_i u_l^{e_i(1-\alpha_i)}(1-\alpha_i) FOB_i^{-\alpha_i} \tau_{il}^{1-\alpha_i} Y_l^{\alpha_i} L_l^{1-\alpha_i}}{\sum_j \beta_j u_l^{e_j(1-\alpha_j)}(1-\alpha_j)(FOB_i \tau_{il})^{1-\alpha_j} Y_l^{\alpha_j-1} L_l^{1-\alpha_j}}, \tag{36}$$

and

$$L_i \geq \phi_i K_i^0. \tag{37}$$

Similarly, zero-profit conditions (26) and (29) imply that *production of utility* and goods must be zero whenever strict inequality constraints hold in equilibrium:

$$\frac{\sum_i \beta_i u_l^{e_i(1-\alpha_i)-1}(1-\alpha_i)(FOB_i \tau_{il})^{1-\alpha_i} Y_l^{\alpha_i} L_l^{1-\alpha_i} e_i}{\sum_j \beta_j u_l^{e_j(1-\alpha_j)}(1-\alpha_j)(FOB_j \tau_{jl})^{1-\alpha_j} Y_l^{\alpha_j-1} L_l^{2-\alpha_j}} \geq PU_l \quad \perp \quad U_l \quad \geq 0, \tag{38}$$

[19]The product of the non-zero element(s) on each side of "⊥" or "perp" is zero, or has an either or relationship.

and

$$\phi_i \frac{Y_i}{L_i} \geq FOB_i \quad \forall i \quad \perp \quad K_i \geq 0. \tag{39}$$

In view of the optimization theory, this indirect CDE demand problem specified as a Hicksian import demand is implicitly a bilevel system involving so-called "leaders" and "followers" (Dirkse and Ferris, 1999). The optimum cardinal values of utility and parameters in a given choice of objective function (e.g., PPML objective) solves the optimum of marginal cost of utilities in (38). At fixed numeraire prices, the MPEC will then solve the shadow price of utility from the second-layer problem of "followers" as a constraint to the objective "leaders". In other words, the MPEC essentially solves the marginal cost of utility such that $PU \equiv g(u)$ is a solution to a change in u with respect to variations in observables, such as the national income and population, in the objective function as specified in (35).

C. MPEC

The MPEC was originally a mathematical optimization tool used among engineers. It has recently become popular in solving complex economic problems that involve variational inequality, complementarity problems, or multilevel optimization problems (Dirkse and Ferris, 1999). The MPEC has been applied and discussed in numerous important trade and structural estimation literature (Balistreri and Hillberry, 2007; Balistreri, Hillberry and Rutherford, 2011; Su and Judd, 2012; Tintelnot, 2017; Li, 2018; Antràs and Gortari, 2020). The MPEC program allows estimated parameters and objective values to be fully theory-consistent with the GE system. It is an appropriate program to solve this GE gravity model reformulated as follows

$$
\begin{aligned}
\max_{\mathbf{b}_i=\{\alpha_i,\beta_i,e_i,\rho\},u_l,\phi_i,\delta_{il}} \quad & \text{PPML Objective} \\
s.t. \quad & \text{GE}(\mathbf{b}_i, u_i, u_l, \phi_i, \delta_{il}) \ \text{[set of GE constraints]} \\
& \text{CDE}(\mathbf{b}_i, u_l, \delta_{il}) \equiv 1 \\
& \alpha_i, \beta_i, e_i > 0 \\
& u_i, u_l, \phi_i, \rho > 0 \\
& \delta_{il} \geq 0,
\end{aligned}
$$

where $\text{GE}(\mathbf{b}_i, \phi_i)$ characterizes the nonlinear GE system of equality and inequality constraints, and $\text{CDE}(\mathbf{b}_i)$ is the implicit utility defining equation, while the rest of inequalities are parametric restrictions specified in the model.

The problem is burdensome as the trade cost elasticity ρ enters the implicit additivity defining equation. The final prices at home, which are the values of goods adjusted by trade costs, are endogenous because of the unknown elasticity with respect to distance ρ. Moreover, as we allow border effects to impact τ_{il}, the program then also needs to determine the asymmetric border coefficients that are represented by δ_{il}. On the positive side, there is also data on population to help evaluate α_i, as well as a GE system of constraints—Equations (27) and (36)-(39)—to jointly evaluate u_l. It is also worth noting that the benchmark utility defining Equation (27), which helps pin down the value of utility, is consistent with both the GE model and the implicitly additive demand.

D. A Solution to Feasible Initial Values

The estimation procedure allows estimating several demand systems using the same data within the same GE gravity framework. This feature is not only convenient but also an important step of finding starting values that are feasible. Without letting the MPEC start with an appropriate initialization, we may not obtain a solution. The nested program eases the high computation burden of finding the optimal solution to the CDE-gravity, which contains nonlinear equality and inequality constraints. For this reason, I formulate the problem in the following way by adding two conditional equality constraints:

$$
\begin{aligned}
&\text{CDE2NHCES}: \text{if } \text{ord(i)} > 1 \quad \text{then} \quad \alpha(\mathbf{1}) = \alpha_i \\
&\text{CDE2HCDE}: \ \text{if } \text{ord(i)} > 1 \quad \text{then} \quad e(\mathbf{1}) = e_i
\end{aligned}
$$

where CDE2NHCES is a transformation equation that generalizes the CDE to an indirect version of the NHCES function, and CDE2HCDE equation transforms standard non-homothetic CDE to a homogeneous CDE function (e.g., $e_i = e \ \forall \ i$); ord is an operator which returns relative position in the set i, and ord(i) is the relative position of $i \in \mathcal{I}$, and $\mathbf{1}$ is the first element in the set $i \in \mathcal{I}$.

When both constraints are enforced, the MPEC will first solve a parameterized standard CES-gravity as stated in Proposition (3). The procedure sets parameter values by benchmarking CDE functional form that is the most restricted special case of the general model. The MPEC then solves the cardinal values of u—which is implicitly indirectly defined—that makes the defining constraints feasible, while holding the parameter values of the CES-gravity constant. The next step is to release the restriction imposed on e_i's by removing *CDE2HCDE*, and to estimate the NHCES-gravity model, as specified in Section (III.D). The final step is to release all parameter values by eliminating both *CDE2NHCES* and *CDE2HCDE* as constraints, while fully estimating the CDE-gravity model.

E. Normalization Strategy

The specification of this class of separable demand model (e.g., NHCES, CDE and the standard CES) indicates that not all demand parameters are identifiable. For the system of CDE, there are multiple degrees of freedom in the parameter space. First, it is intuitive to see that any positive constant scaling of the distribution parameters β_i's will raise the RHS of Equation (1) by the same scaling factor. Furthermore, without loss of generality, if we let $u = \gamma V_1$ for some scalar $\gamma > 0$, and $e_i = \mu f_i$ for some scalar $\mu > 0$, we will obtain that

$$(40) \qquad \sum_i \beta_i (\gamma V_1)^{\mu f_i (1-\alpha_i)} \left(\frac{p_i}{E}\right)^{1-\alpha_i} \equiv 1,$$

which is equivalent to

$$(41) \qquad \sum_i \beta_{i1} V_2^{f_i (1-\alpha_i)} \left(\frac{p_i}{E}\right)^{1-\alpha_i} \equiv 1,$$

where $\beta_{i1} = \beta_i \gamma^{\mu f_i (1-\alpha)} = \beta_i \gamma^{e_i (1-\alpha)}$, $V_2 = V_1^{\mu} = \left(\frac{u}{\gamma}\right)^{\mu}$. It is easy to check that both β_i and u have been absorbed twice by the arbitrary choices of the two scalars, transforming into an identical demand functional form with new sets of parameter combinations. The identity above implies that, for some $n\gamma$ and $m\mu$, there exists n x m possibilities of estimated β_i parameters and utilities values. Without removing the extra degrees of freedom in the gravity estimation, one may expect to find very large standard errors in terms of the parameter values in the stage of statistical testing.

Unlike Comin, Lashkari and Mestieri (2021) who choose to normalize individual base factors ("goods"), I normalize such that $\sum_i \beta_i \equiv \kappa = G(.)$ and $\sum_i e_i = \sum_i \delta_i = \mathcal{I}$, where κ is a normalizing constant determined by the data, and $\mathcal{I}$, again, is the number of goods. This normalization scheme tends to improve the scaling of unobserved utility levels as it does not restrict one of the parameters to be one so the rest of parameters could very likely be badly scaled.

VI. Data and Parameter Estimates

The model is estimated using the 1995 observations for 37 countries or regions (if applicable) in the world. The population and distance data are taken from the CEPII gravity database. The data on implied factor trade is taken from the World Input-Output Database (WIOD). The national GDP data for each region is aggregated based on aggregate trade flows from the WIOD. Using the datasets from WIOD, Figure (1) exhibits the empirical patterns of global demand for foreign factors in 1995, with higher aggregate national income represented by larger size of nodes and higher value of aggregate bilateral trade flows represented by greater weight of lines. For instance, the trade flows between the world's two "richest" economies in 1995 (the U.S. and Japan) are higher than their EU trade partners, despite longer geographic distance, and are lower than U.S.-Canada, despite higher aggregate income in Japan than in Canada.

Based on these empirical observations from the WIOD 1995, I estimate the structural parameters in the standard CES-gravity, NHCES-gravity and CDE-gravity models using the nested structure discussed in Section (V.D). The estimation program is the General Algebraic Modeling System (GAMS) version 34.3, which is solved using the MPEC-NLPEC (non-linear programming with equilibrium constraints) solver with the help of the preprocessor using GAMS-F tool.[20]

[20]See http://www.mpsge.org/inclib/gams-f.htm (Michael Ferris, Thomas Rutherford, and Collin

Figure 1. Network Chart: Global Demand for Foreign Factors in WIOD, 1995

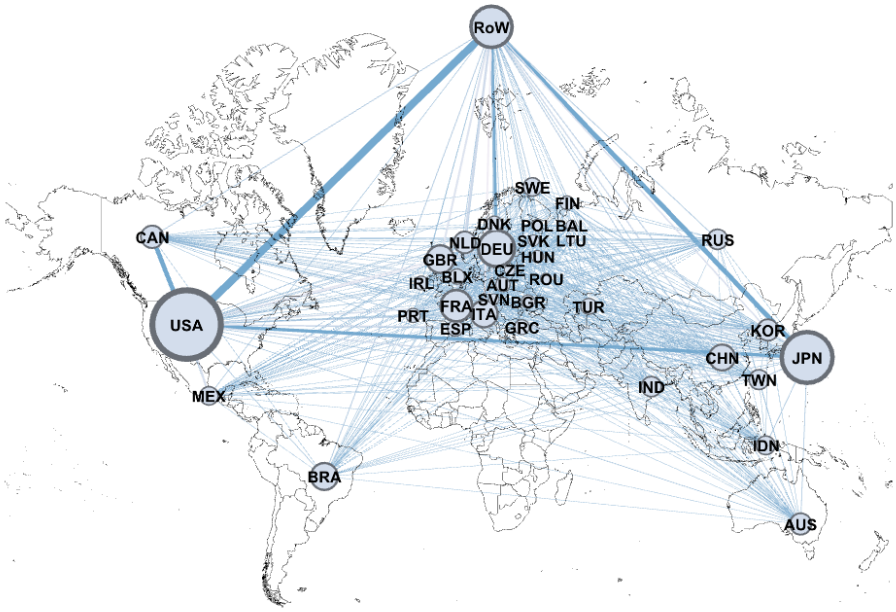


Table (2) reports estimates for 37 countries with both distance and asymmetric border effects on bilateral trade costs.

While I will discuss the likelihood ratio tests in section (VII), the results in Table (2) show that the CDE model generates higher PPML (log-likelihood value) than the NHCES and CES models (with less degrees of freedom). Both CES and NHCES models produce a σ that is in the neighborhood of 5, whereas in the CDE model some elasticities can be either as high as 9.4 (i.e., Slovakia) or as low as 2.38 (i.e., Taiwan). When comparing the NHCES model with the CES model, with both models exhibiting constant elasticities, the trade costs (i.e., distance elasticity and border coefficients) are lowered in the NHCES model. However, when we fit the CDE system to the same data, the trade costs become even larger than the CES model. This is because that the trade costs are driven by consumer preferences. In the CES and NHCES models, the constant substitutabilities (represented by the Allen-Partial elasticities) are identical to the region-generic α, whereas in the CDE model the Allen-Partial elasticities are region-specific. By calculating them, I find that there are many (home-foreign) pairs of goods that

Starkweather). The tool greatly reduces the computation complexity in calculating the trade costs.

Table 2—Parameter Estimates of the Nested Gravity Models

	CDE			NHCES			CES		
PPML (with border charges)	370.563×10^3			370.465×10^3			370.412×10^3		
ρ	0.235			0.208			0.209		
	α	e	δ	α	e	δ	α	e	δ
AUS	3.48	1.05	1.87	5.26	1.01	0.92	5.21	1.00	0.92
AUT	5.66	1.03	0.80		1.02	0.94			0.95
BAL	8.80	1.03	0.46		1.04	1.42			1.54
BGR	7.37	1.05	0.55		1.08	1.25			1.41
BLX	4.93	0.95	0.63		0.95	0.58			0.63
BRA	4.86	0.98	1.35		0.99	1.24			1.25
CAN	4.96	0.93	0.97		0.91	0.92			0.91
CHN	2.95	0.94	2.73		0.97	1.14			1.06
CZE	8.44	1.04	0.39		1.03	1.15			1.19
DEU	5.10	1.02	0.68		1.03	0.68			0.66
DNK	5.07	1.04	0.94		1.04	0.92			0.91
ESP	7.08	1.01	0.57		0.99	1.05			1.07
FIN	6.52	1.02	0.56		1.03	0.99			0.99
FRA	5.09	1.02	0.89		1.02	0.88			0.88
GBR	4.37	1.01	0.95		1.02	0.71			0.71
GRC	6.11	1.00	0.85		1.00	1.25			1.27
HUN	6.07	1.02	0.84		1.02	1.17			1.20
IDN	3.03	0.93	2.77		0.92	1.16			0.96
IND	5.23	0.92	1.52		0.93	1.60			1.31
IRL	5.06	0.95	0.73		0.95	0.74			0.75
ITA	5.03	1.01	0.89		1.01	0.86			0.87
JPN	3.73	1.01	1.20		1.00	0.68			0.68
KOR	2.96	1.05	1.91		1.03	0.66			0.66
LTU	7.80	1.04	0.54		1.06	1.35			1.48
MEX	6.39	0.97	0.66		0.94	1.13			1.02
NLD	4.86	1.00	0.61		1.00	0.54			0.55
POL	7.35	1.00	0.72		0.99	1.36			1.34
PRT	6.37	0.95	0.66		0.93	1.12			1.09
ROU	6.41	1.03	0.94		1.03	1.41			1.48
RUS	8.75	0.99	0.42		1.01	1.20			1.23
ROW	2.88	0.99	1.12		1.00	0.20			0.21
SVK	9.40	1.02	0.32		1.06	1.24			1.32
SVN	5.84	1.03	0.88		1.02	1.16			1.19
SWE	5.21	1.02	0.90		1.01	0.91			0.92
TUR	6.67	0.96	0.86		0.96	1.40			1.33
TWN	2.38	1.02	2.71		0.99	0.39			0.41
USA	5.24	0.99	0.62		0.99	0.64			0.66

have lower substitutabilities than those estimated from the CES and NHCES models. Furthermore, the lower substitutabilities between home and foreign countries are associated with higher trade costs. This suggests that the higher trade costs are driven by lower substitutabilities in the CDE model.

VII. Likelihood Ratio Tests

The estimation procedure simultaneously produces the values of the objective function, which are the PPMLs, under both CES-gravity and CDE-gravity. Given the same data and estimation procedure, and that a CDE nests a CES (where a CES is a CDE in restricted parameter spaces), it is natural to compare the two demand systems by conducting a likelihood-ratio test to examine the performance of one null demand model against the more restricted one. Formally, the test illustrated in the form of hypothesis setting will indicate how many times more likely the data fits better in one model when comparing with the other. The likelihood ratio statistics specialized to this problem is Greene (2012), Sections 14.9.3.c and 18.4.1.

$$\Lambda = -2[\log L_{CES}(\mathbf{b}_i) - \log L_{CDE}(\mathbf{b}_i)], \tag{42}$$

where Λ is the test statistic.

The hypothesis test then starts with the following assumptions:

Hypothesis $\mathbf{H_0}$: *the more restricted CES is more consistent with the data.*

Hypothesis $\mathbf{H_a}$: *the less restricted CDE is more consistent with the data.*[21]

In statistical expression:

[21]The data represents the national GDP, population, WIOD aggregate trade flows, and CEPII's distance measures.

$$\Lambda > c, \qquad \rightarrow \text{reject } H_0$$
$$\Lambda < c, \qquad \rightarrow \text{do not reject } H_0$$

for all $\beta, e \geq 0$ and for all $u, \rho > 0$, where c is a specified cutoff value that governs significance level φ, satisfying: $P\{\Lambda > c|H_0\} = \varphi$. Define the significance level ϕ to be 0.05 on the left; dimensions $\mathbf{dim}(\mathbf{b}_{CDE}) = s$ and $\mathbf{dim}(\mathbf{b}_{CES}) = m$. By Wilks' theorem, when H_0 is true (under standard regularity conditions), Λ must be asymptotically χ^2 - distributed with calculated $s - m$ degree of freedom which equals to 73 (from CDE to standard CES: 37 α's reduce to 1 alpha and 37 expansion parameters e's reduce to none, since e is one everywhere in the CES Model, i.e., 36 + 37) and the calculated cutoff value is $c^{CES} = 93.9$. The estimated Λ is 150 which is much greater than 93.9, so I reject the null hypothesis that the more restricted CES is more consistent with the data.

I test the CDE Model against the NHCES Model and find that the cutoff value is approximately $c^{NHCES} = 51$ (in this case, the calculated difference of the degree of freedom is 36, i.e., 37 α's reduce to 1 uniform α) and is less than the estimated log-likelihood ratio $\Lambda = 90$, so I can reject the null that the NHCES is more consistent with the data.

The log-likelihood tests suggest that the CDE is a better Model in this gravity model of aggregate trade flows in terms of fitting to the data, when compares to the CES and NHCES in the gravity model with the same general equilibrium constraints.

VIII. Out-of-Sample Predictions and Missing Trade

A. *Counterfactuals (*GTAP, Hertel and Dixon et al.*'s Convention)*

The analytical import demand framework provides us a convenient way to explore out-of-sample predictions. Suppose we estimated the structural model using the 1995 WIOD data, we may then forecast 2011 trade flows using different estimated models, which are then compared with actual trade flows data in the 2011 WIOD. Instead of calculating forecast of trade flows in value terms, I will implement the out-of-sample predictions in the percentage change formula. One convenience of this procedure is that we can decompose the change in trade flows into different components including the welfare response. Another advantage is that it yields a potential framework to explain possible implications of missing trade flows. Finally, it enables us to revisit the demand-system comparisons within a parsimonious analytical framework that are relatively *cleaner* than the numerical solutions from the MPEC. In what follows in this section, I follow the counterfactual conventions pioneered by Dixon and Parmenter (1996) and Hertel (1997) who provide a GEMPACK or GTAP solution to predict consequences of counterfactual experiments.[22]

Key Endogenous Model Complements in "Hat"

Using Equation (2), the total differentiation with respect to utility, the per capita income and the price vector will decompose the change in per capita quantity of trade into three major counterfactual components as follows:

$$\begin{aligned} \widehat{q}_{il} &= \left[e_i(1-\alpha_i) - \sum_j e_j(1-\alpha_j)\omega_{jl}\right]\widehat{u}_l \\ &+ \left[\alpha_i + \sum_j (1-\alpha_j)\omega_{jl}\right]\widehat{E}_l \\ &- \alpha_i \widehat{p}_{il} - \sum_j (1-\alpha_j)\omega_{jl}\widehat{p}_{jl}, \end{aligned} \tag{43}$$

Derivation. See Appendix A.A4.

where the first term in the RHS is the effect owing to changes in the non-homothetic real consumption, the second term is the per capita income effect, and the last two expressions involving $p_{il} = FOB_i \tau_{il}$ shows the effect of changes in a composite price term including components of distance and border effects.

It then follows that the counterfactual aggregate trade flows can be decomposed into the following expression with changes in population, FOB prices and trade costs:

[22] General Equilibrium Modelling, or GEMPACK, carries out counterfactual analyses using approximations and hat derivatives of large economic models, such as the GTAP Model.

$$(44) \qquad \widehat{X}_{il} = \widehat{q}_{il} + \widehat{L}_l + \widehat{FOB}_i + \widehat{\tau}_{il}$$

Combining the two hat equations above, it is intuitive to see that if we can measure the unobserved change in wealth (utilities) in the following identity, then the macro data (income, population and trade flows) will allow us to measure the uncompensated percent differences in aggregate trade flows excluding price effects (FOB prices and trade costs):

$$(45) \qquad \widehat{u}_l\Big[\widehat{w}_l, \widehat{\mathbf{FOB}}_i; \bar{\mathbf{b}}_l, \bar{\bar{\delta}}_{il}\Big]^{\bar{\mathbf{b}}_l=\{\bar{\alpha}_l,\bar{\beta}_l,\bar{e}_l,\bar{\rho}\}}_{(il)\in\mathcal{I}} = \frac{u_l^1 - u_l^0}{u_l^0},$$

which can be derived by implementing total differentiation in (1) with respect to u_l, w_l and FOB_i (with fixed border frictions):

$$(46) \qquad \widehat{u}_l = \frac{\widehat{w}_l - \sum_i \omega_{il}^0 \widehat{p}_{il}}{\sum_i e_i \omega_{il}^0} = \frac{\widehat{w}_l - \sum_i \omega_{il}^0 \widehat{p}_{il}}{\vartheta_l|_{\omega_{il}^0}},$$

Derivation. See Appendix A.A5.

where $\widehat{u}_l$ is the counterfactual change in welfare response in percentage terms; $\vartheta_l|_{\omega_{il}^0}$ is the aggregate elasticity with respect to utility evaluated at benchmark import shares. Using (43)-(46), we may i) forecast out-of-sample trade flows in percentage changes; while ii) evaluating a composite effect of factor prices on trade by bridging a gap between predicted and observed trade flows. Let $\widehat{X}_{il}$ defined in Equation (44) be the predicted percentage change in trade flows and $\widetilde{X}_{il}$ be the actual percent changes. Furthermore, denoting $\widehat{RF}$ the change in real values of factors , we may then obtain the following identities:

$$\begin{cases} \widehat{X}_{il}^{CDE}/\widetilde{X}_{il} \Longleftrightarrow \widehat{X}_{il}^{CES}/\widetilde{X}_{il} \Big|_{1995-2011} & \text{(Model Comparison)} \\ \\ \widehat{RF}_{il}|_{1995-2011} = \widetilde{X}_{il} - \widehat{X}_{il} & \text{(Missing Trade Flows)} \end{cases}$$

The model comparison shows the proportion of the actual percentage changes in trade flows between 1995 and 2011 that can be explained by one model compared with a different nested model. The *missing trade flows* explains that, within each demand model, how much missing trade flows (based on the counterfactual analysis) can be attributed to changes in relative factor prices that are not observed in the data (see an application and results of using this technique in the Online Appendix).

The Role of Population in Counterfactuals

One important advantage of adopting percent-change counterfactuals is to make decomposition intuitive in analytical formulae. This also makes it convenient to analyze the impact of change in the population. After incorporating the population effects into the per capita quantity of trade, it can be shown that

$$\widehat{q}_{il} - \widehat{L}_l = \sum_j \left(\sigma^m_{ijl}\widehat{p}_{jl}\right) + \eta_{il}\left(\widehat{Y}_l - \widehat{L}_l\right), \tag{47}$$

where $\widehat{L}_l$ is the percent change in population, $\sigma^m_{ijl} = \omega_{jl}(\sigma_{ijl} - \eta_{il})$ is the uncompensated price elasticities of demand for i with respect to price of j in l, $\widehat{p}_{jl}$ is change of goods price of j in l. Equation (47) is identical to the one in the GTAP counterfactuals. It is easy to see that there is no population effect with homothetic CES demand where income elasticities are unitary everywhere. In this case, the CES-Gravity models are likely misspecified in predicting trade flows as they only exploit income data at national level. With non-homothetic demand, such as the CDE, the income elasticities also interact with the change in population. In this case, the population is not negligible in the comprehensive income effect.

B. *Exact Hat Algebra (*Dekle, Eaton and Kortum*'s Convention)*

Key Endogenous Model Complements

Using the counterfactual technique *à la* Dekle, Eaton and Kortum (2007), I now denote $\overset{\circ}{X}_{il} \equiv X^1_{il}/X^0_{il}$ as the ratio of the unobserved counterfactual value X^1_{il} to the observed current or the baseline value X^0_{il}. Recall that Equation (23) is the gravity condition, which is restated here for convenience:

$$X_{il} = \frac{\beta_i u_l^{e_i(1-\alpha_i)}(1-\alpha_i)(FOB_i\tau_{il})^{1-\alpha_i}Y_l^{\alpha_i}L_l^{1-\alpha_i}}{\sum_j \beta_j u_l^{e_j(1-\alpha_j)}(1-\alpha_j)(FOB_j\tau_{jl})^{1-\alpha_j}Y_l^{\alpha_j-1}L_l^{1-\alpha_j}}. \tag{48}$$

Using Equation (24) and the income condition, i.e., Equation (25), goods-market must clear in each of the exporter location i:

$$Y_i = \sum_l \frac{\beta_i u_l^{e_i(1-\alpha_i)}(1-\alpha_i)(FOB_i\tau_{il})^{1-\alpha_i}Y_l^{\alpha_i}L_l^{1-\alpha_i}}{\sum_j \beta_j u_l^{e_j(1-\alpha_j)}(1-\alpha_j)(FOB_j\tau_{jl})^{1-\alpha_j}Y_l^{\alpha_j-1}L_l^{1-\alpha_j}} \quad \forall\, i. \tag{49}$$

In each of the importer location, trade is balanced (i.e., income equals expenditure; no deficit):

$$Y_l = \sum_i X_{il} \quad \forall\, l. \tag{50}$$

The labor-market clearing condition is now generalized as

$$\begin{aligned} L_i &= \phi_i K_i^0 \\ &= \phi_i \sum_l \frac{\beta_i u_l^{e_i(1-\alpha_i)}(1-\alpha_i)FOB_i^{-\alpha_i}\tau_{il}^{1-\alpha_i}Y_l^{\alpha_i}L_l^{1-\alpha_i}}{\sum_j \beta_j u_l^{e_j(1-\alpha_j)}(1-\alpha_j)(FOB_i\tau_{il})^{1-\alpha_j}Y_l^{\alpha_j-1}L_l^{1-\alpha_j}} \quad \forall\, i. \end{aligned} \tag{51}$$

In all locations, the cost of living index with respect to utility is defined as:

$$PU_l \equiv \sum_i e_i \omega_{ij} \frac{Y_l}{u_l L_l} \quad \forall\, l. \tag{52}$$

The complementarity condition $\sum_i e_i\omega_{ij}(Y_l/L_l)/u_l \geq PU_l \ \perp \ U_l \geq 0$ pins down the level of utility in Equation (52), while the utility defining equation (31) pins down the cost of living index PU_l. The counterfactual change of cost of living is defined as

$$\overset{\circ}{PU}_l \equiv \frac{PU_l^1}{PU_l^0}. \tag{53}$$

Define the ratio of the counterfactual value of welfare to the initial values of welfare as

$$\overset{\circ}{W}_l \equiv \frac{u_l^1}{u_l^0}. \tag{54}$$

Due to non-homotheticity (governed by e_i's) and that counterfactual import shares are endogenous (governed by endogenous counterfactual levels of utility), the welfare and cost of living index are affected through different channels. For the cost of living index, there is one *diminishing* channel: higher cost of living index implies lower utility level (via Equation (52)), and one (possibly) *magnifying* channel: higher utility in importer location l implies higher income in l and thus may buy more from every location (operated by expansion parameters) via Equation (4).

By proposition 1 in Allen and Arkolakis (2015) and the Perron-Frobenius theorem, there exists a unique set to income Y_i up to normalized scale (given observed income shares ω_{il} in the gravity condition—that is, the PPML objective function in the structural model) and, while scaling factors specified in Section (V.E) (pinning down the cardinal value of utility and expansion parameters), there exists a unique set of bilateral matrix of trade frictions $\mathscr{F} \equiv \{\mathscr{F}_{il}\}$, given some positive "shifters" $\mathscr{B} \equiv \{\mathscr{B}_{il}\}$, some $\boldsymbol{\alpha} \equiv \{\alpha_i\}$ and $\boldsymbol{e} \equiv \{e_i\}$ (exogenous *gravity constant* or *behavioral parameters*) (see Section (IX)).

IX. Iterative vs. Levels Solution: A General Counterfactual Procedure

Pearson (1991), Dixon and Parmenter (1996) and Hertel (1997) discuss *GEMPACK* solutions to errors with linearization which is to iterate the linearization over and over until a new solution in levels is achieved. This method is especially quite useful in large and complex models wherein a computer algorithm that *crawls* towards the solutions with a series of linearized approximations could be more robust.[23] Another trend is the levels solution along with increases in computation capacities. Without having a priori knowledge of whether solutions to levels in a particular model can be robust or successful, it is difficult to decide which convention to follow. Furthermore, the implicit demand models in GE models can create a puzzling situation where the CES income balance condition (i.e., $U.PU = Y$) is potentially redundant and it might be difficult to determine the complementarity variables. In what follows, I provide a general procedure to solve for counterfactual equilibrium.

Step 1: Constructing a full GE model with standard equilibrium conditions (IV.B) and estimate the model parameters based on historical data (V):

The identification is achieved up to a scale and an arbitrary normalization scheme, under which a presumption is made upon one's belief of so-called "deep" structural parameters in the counterfactual experiment. In a typical general equilibrium framework, what to expect to influence trade flows are some "shifters" $\mathscr{B} \equiv \{\mathscr{B}_{il}\}$, some exogenous trade frictions $\mathscr{F} \equiv \{\mathscr{F}_{il}\}$, and some *gravity constant* or *behavioral fixed parameters* $\boldsymbol{\alpha} \equiv \{\alpha_i\}$ and $\boldsymbol{e} \equiv \{e_i\}$.

One additional procedure is to calibrate the "shifters" following Balistreri and Hillberry (2008) using the same data set used in the estimation procedure. The scale and numeric of these parameters are established up to normalization. Thus, from the structural point of view, these model parameters are not fully identified from the $\{i \text{ x } l\}$ bilateral trade flows matrix. However, these shifters can be calculated from the data, while holding $\boldsymbol{\alpha}$, $\boldsymbol{e}$ and some $\rho > 0$ (exogenously) affecting $\mathscr{F}$ fixed, following the first step.

Step 2: Testing the existence of general equilibrium:

Once fixed parameters are determined, implement the test of existence of general equilibrium. This is critical for large and complex models or models that involves implicit additivity such as the CDE Model. Implement a *homogeneity test* by setting and shifting one price numeraire in the *counterfactual* experiment.

[23] Pearson (1991) introduces a *Gragg* method which is considerably fast in computations.

This can be implemented using a *mixed complementarity problem* or an MPEC that embeds an MCP). The MCP does not include a likelihood function as an optimizing objective. Therefore, when choosing an MPEC to replicate the initial equilibrium, one must also fix the value of likelihood that fits exactly the data and predicted exogenous scalars. Then, during the shift of numeraire, free up the value of likelihood to allow differentials in residuals between the model initialization and counterfactuals (code of applying this technique in a standard CES general equilibrium model is available upon request).

Step 3: Estimating the consequences of counterfactual experiments:

Here I propose an MPEC approach. The logic is similar to implementing an MCP after estimation of the model parameters using a non-linear programming (NLP) approach, but it solves the problems by treating a parametric derived from the NLP problem fixed, while setting initial benchmarks by constraining the likelihood of the objective function, such that $\overline{\mathscr{J}}(\overline{\mathscr{F}}, \overline{\boldsymbol{\alpha}}, \overline{\boldsymbol{e}}) \equiv \mathscr{J}^0(\overline{\mathscr{F}}, \overline{\boldsymbol{\alpha}}, \overline{\boldsymbol{e}})$ with the choice of numeraire, where $\overline{\mathscr{J}}$ is the estimated value of objective function (likelihood or residuals). In the second solve, release this restriction and compute the counterfactual values directly after the exogenous shocks.

The benefit of implementing an MPEC in counterfactuals is that it is computationally more "natural" than MCP (in GAMS) that is more "manual". One justification of this claim is a complementarity test of the standard CES gravity model by switching complementarity variables U and PU in the model statement of an MCP and the choice does not affect benchmark equilibrium. But in problems of implicitly additive demand models and larger model space it is more difficult for a modeler to do this manually in an MCP alone with added complexities of required numeraires used in previous NLP for identification, and determination of NLP variables that need to be freed up. This method can be used in the some subsequent papers after Anderson and van Wincoop while replicating exactly the same results, and is a better way of doing counterfactuals for larger and more complex models.

This approach is not theoretically new, but computationally more convenient. I take advantage of that 1) MPEC embeds an MCP, and 2) MPEC seeks for fitted counterfactual equilibria with fixed model parameters, and while restricting the likelihood to be the same as solved from the initial non-linear system of equations, it re-generates the benchmark equilibrium before the counterfactuals, and predicts counterfactual consequences while freeing up endogenous variables. In the standard CES general equilibrium model, as shown, the complementary variables are associated with income balance condition and unit expenditure function, which pin down each of the cost of living index and utility index. However, the complementarity slackness assumption in the implicitly additive model (and complex and large models) is not as tractable as in the CES model. The MPEC counterfactual solution helps overcome this challenge.

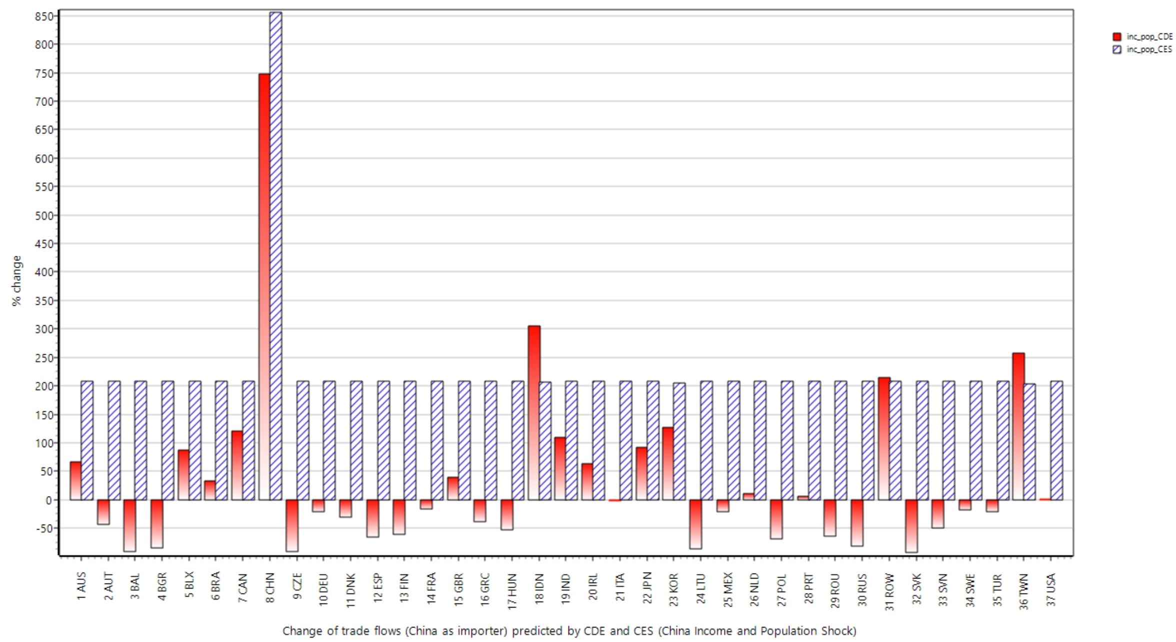
A. Counterfactual Results

Before testing the predictiveness of the CDE model, it is useful to see whether the CDE model yields largely different results than the CES in a same counterfactual scenario. Following the MPEC procedure, I use the estimated parameters from the CDE and the CES models to implement counterfactual analysis using the procedures introduced in Section (IX). Since there has been substantial income growth in China since 1995, I choose to shock the income and population in China to the year 2011 using both the CDE and the CES predictions. As shown in Table (2), the CDE model produces very different out-of-sample responses to the Chinese income and population shock.

It is clear to see that under the CES assumption the model produces the same percentage change of trade flows in China across all other export markets, whereas the CDE model under this environment predicts both positive and negative changes across different exporter countries. Indeed, this peculiar outcome alone does not automatically suggest that the CDE model is more predictive than the CES model; however, the fact that the CES model generates no variation in trade flows across all exporter countries is indeed unrealistic.

For this reason, if we regress the changes in actual import shares on CDE-predicted changes in import shares and obtain a (in fact, any) positive coefficient, then this would be a clear sign that the CDE model is predictive. To do this, I implement a counterfactual exercise which allows to shock the income and population in the world to the year 2011. The regression results show a positive coefficient of 0.309 (0.004) on the fitted shares of import with $R^2 = 0.832$, which suggests that the CDE model is predictive.

FIGURE 2. CDE VS. CES: COUNTERFACTUAL COMPARISONS



X. Conclusion

This paper contributes to the international trade literature by introducing a new preference structure that is implicitly indirectly additive, and by introducing a complete nested structural estimation procedure that is both tractable and practical in solving a sophisticated general equilibrium system, and by testing the more general model against standard CES model and other globally regular additive preference models used in gravity for the first time. The estimation methodology devised in this paper can be applied to almost any structural gravity model that has the need to test whether any new demand models can be more consistent with the data compared to existing gravity models. By devising strategies for model comparisons using fundamental demand and GE theories and standard optimization software with advanced mathematical algorithm but with reasonable computational expense, this paper significantly advances trade literature. Furthermore, I provide a solution that is computationally tractable in estimating counterfactual consequences of complex and large models, and those involving implicit additive demand models.

The gravity model of trade has been overwhelmingly successful in empirical trade for a very long time, during which many trade economists have taken it for granted without questioning the theoretical justifications in economics until several leading international trade economists, such as Helpman, Krugman and Deardorff, started to challenge the linkages between economic theories and the empirical evidence of the gravitational relationships (Deardorff, 2011). This cultivated theoretical motivations of other notable trade economists, such as Anderson and Bergstrand, who bridge theories with empirical applications of the gravity model which, for decades, is followed by many who construct gravity equations, which are mostly based on the CES preferences but are heavily grounded in the production theories to conform with known empirical relationships.

Moreover, the world's rapid but uneven economic growth has long led a substantial increase in relative incomes in the world. Since integration of the CES assumption into the world of aggregate factor demand and supply implies proportional increases of relative demand for national factor supplies with relative incomes, I ask whether the standard CES assumption is still suitable in bridging gravity theories with empirical tests. Testing gravity theories is exceedingly difficult absent more general demand models that can nest the CES model using the same data and identical estimation structure of econometric models. I revisit theoretical demand literature from the 1960s and 1970s and devise a nested structural optimization approach which allows us to overcome this challenge. I find that the more general CDE-gravity model is most consistent with the data in terms of the value of the objective log-likelihood function whether border costs are present or not, as opposed to results obtained from the NHCES and CES gravity models.

I construct an empirical gravity equation that allows (1) explicitly estimating the cardinal utility with the help of the mathematical programming; and (2) utilizing the population data that greatly aids identification. These implementations along with the structural identification strategy devised for the CDE-gravity model allows separating the elasticity of trade costs with respect to distance and asymmetric border coefficients from the elasticity of trade flows with respect to trade costs, which has been found to be challenging in most previous literature.

The test statistic in the log-likelihood ratio tests of CDE against CES and

NHCES models in the gravity model is able to reject both null hypotheses that CES and NHCES are more consistent with the data at a significance level of 0.05. This suggests that CDE is a better model in this gravity model of aggregate trade flows in terms of fitting to the data.

I develop a counterfactual method which does not need to explicitly point to any specific complementarity variable which could be a puzzle in a complex model or when utility is implicitly defined as in this paper. The counterfactual results show that the CDE model produces entirely different out-of-sample responses (from the CES model) to the Chinese income and population shock. The results also show that the CES model produces the same percentage change across all other export markets, which is unrealistic. I regress changes in actual import shares on the changes in import shares predicted by the CDE and obtain a positive coefficient of 0.309, which shows a sign that the CDE model is predictive.

REFERENCES

Adao, Rodrigo, Arnaud Costinot, and Dave Donaldson. 2017. "Non-parametric Counterfactual Predictions in Neoclassical Models of International Trade." *American Economic Review*, 107(3): 633–689.

Allen, Treb, and Costas Arkolakis. 2015. "Calibration and Counterfactuals in Gravity Trade Models." *mimeo*, 10.

Anderson, James E. 1979. "A Theoretical Foundation for the Gravity Equation." *The American Economic Review*, 69(1): 106–116.

Anderson, James E. 2011. "The Gravity Model." *Annual Review of Economics*, 3(1): 133–160.

Anderson, James E., and Eric van Wincoop. 2003. "Gravity with Gravitas: A Solution to the Border Puzzle." *American Economic Review*, 93(1): 170–192.

Antràs, Pol, and Alonso de Gortari. 2020. "On the Geography of Global Value Chains." *Econometrica*, 88(4): 1553–1598.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-clare. 2010. "Gains From Trade under Monopolistic Competition: A Simple Example with Translog Expenditure Functions and Pareto Distributions of Firm-Level Productivity." *mimeo*.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare. 2012. "New Trade Models, Same Old Gains?" *American Economic Review*, 102(1): 94–130.

Balistreri, Edward J., and Russell H. Hillberry. 2007. "Structural estimation and the border puzzle." *Journal of International Economics*, 72(2): 451–463.

Balistreri, Edward J., and Russell H. Hillberry. 2008. "The Gravity Model: An Illustration of Structural Estimation as Calibration." *Economic Inquiry*, 46(4): 511–527.

Balistreri, Edward J., Russell H. Hillberry, and Thomas F. Rutherford. 2011. "Structural estimation and solution of international trade models with heterogeneous firms." *Journal of International Economics*, 83(2): 95–108.

Barnett, William A. 1983. "Definitions of 'second order approximation' and of 'flexible functional form'." *Economics Letters*, 12(1): 31–35.

Barten, Anton P. 1977. "The Systems of Consumer Demand Functions Approach: A Review." *Econometrica*, 45(1): 23–51.

Bergson, A.(Burk). 1936. "Real Income, Expenditure Proportionality, and Frisch's New Methods." *Review of Economic Studies*, 4: 33–52.

Bertoletti, Paolo, Federico Etro, and Ina Simonovska. 2018. "International Trade with Indirect Additivity." *American Economic Journal: Microeconomics*, 10(2): 1–57.

Buch, Claudia M., Jörn Kleinert, and Farid Toubal. 2004. "The distance puzzle: on the interpretation of the distance coefficient in gravity equations." *Economics Letters*, 83(3): 293–298.

Caron, Justin, Thibault Fally, and James R. Markusen. 2014. "International Trade Puzzles: A Solution Linking Production and Preferences." *The Quarterly Journal of Economics*, 129(3): 1501–1552.

Chen, Y.-H. Henry. 2017. "The Calibration and Performance of a Nonhomothetic CDE Demand System for CGE Models." *Journal of Global Economic Analysis*, 2(1): 166–214.

Christensen, Laurits R., Dale W. Jorgenson, and Lawrence J. Lau. 1973. "Transcendental Logarithmic Production Frontiers." *Review of Economics and Statistics*, 55: 28–45.

Christensen, Laurits R., Dale W. Jorgenson, and Lawrence J. Lau. 1975. "Transcendental Logarithmic Utility Functions." *The American Economic Review*, 65(3): 367–383.

Comin, Diego, Danial Lashkari, and Martí Mestieri. 2021. "Structural Change With Long-Run Income and Price Effects." *Econometrica*, 89(1): 311–374.

Costinot, Arnaud, and Andrés Rodríguez-Clare. 2018. "The US Gains from Trade: Valuation Using the Demand for Foreign Factor Services." *Journal of Economic Perspectives*, 32(2): 3–24.

Deardorff, Alan V. 2011. "Determinants of bilateral trade: does gravity work in a neoclassical world?" In *Comparative advantage, growth, and the gains from trade and globalization: a festschrift in honor of alan v deardorff*. 267–293. WORLD SCIENTIFIC.

Deaton, Angus. 1974. "A Reconsideration of the Empirical Implications of Additive Preferences." *The Economic Journal*, 84(334): 338–348.

Deaton, Angus, and John Muellbauer. 1980. "An Almost Ideal Demand System." *The American Economic Review*, 70(3): 312–326.

Dekle, Robert, Jonathan Eaton, and Samuel Kortum. 2007. "Unbalanced Trade." *American Economic Review*, 97(2): 351–355.

Diewert, W. E. 1971. "An Application of the Shephard Duality Theorem: A Generalized Leontief Production Function." *Journal of Political Economy*, 79(3): 481–507.

Dirkse, Steven P., and Michael C. Ferris. 1999. "Modeling and Solution Environments for MPEC: GAMS & MATLAB." In *Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods.* , ed. Masao Fukushima and Liqun Qi, 127–147. Springer US.

Dixit, Avinash K., and Joseph E. Stiglitz. 1977. "Monopolistic Competition and Optimum Product Diversity." *The American Economic Review*, 67(3): 297–308.

Dixon, Peter B., and B.R. Parmenter. 1996. "Chapter 1 Computable general equilibrium modelling for policy analysis and forecasting." In *Handbook of Computational Economics.* Vol. 1, 3–85. Elsevier.

Eaton, Jonathan, and Samuel Kortum. 2002. "Technology, Geography, and Trade." *Econometrica*, 70(5): 1741–1779.

Fajgelbaum, Pablo D., and Amit K. Khandelwal. 2016. "Measuring the Unequal Gains from Trade *." *The Quarterly Journal of Economics*, 131(3): 1113–1180.

Feenstra, Robert C., and David E. Weinstein. 2017. "Globalization, Markups, and US Welfare." *Journal of Political Economy*, 125(4): 1040–1074.

Fieler, Ana Cecília. 2011. "Nonhomotheticity and Bilateral Trade: Evidence and a Quantitative Explanation." *Econometrica*, 79(4): 1069–1101.

Gourieroux, C., A. Monfort, and A. Trognon. 1984. "Pseudo Maximum Likelihood Methods: Theory." *Econometrica*, 52(3): 681–700.

Greene, William H. 2012. *Econometric analysis.* . 7th ed ed., Prentice Hall.

Green, H. A. John. 1964. *Aggregation in Economic Analysis.* Princeton University Press.

Hanoch, Giora. 1975. "Production and Demand Models with Direct or Indirect Implicit Additivity." *Econometrica*, 43(3): 395–419. Publisher: [Wiley, Econometric Society].

Head, Keith, and Thierry Mayer. 2014. "Chapter 3 - Gravity Equations: Workhorse,Toolkit, and Cookbook." In *Handbook of International Economics.* Vol. 4 of *Handbook of International Economics*, , ed. Gita Gopinath, Elhanan Helpman and Kenneth Rogoff, 131–195. Elsevier.

Hertel, Thomas Warren. 1997. *Global Trade Analysis: Modeling and Applications.* Cambridge University Press.

Houthakker, H. S. 1960. "Additive Preferences." *Econometrica*, 28(2): 244–257.

Larch, Mario, Joschka Wanner, Yoto V. Yotov, and Thomas Zylkin. 2019. "Currency Unions and Trade: A PPML Re-assessment with High-dimensional Fixed Effects." *Oxford Bulletin of Economics and Statistics*, 81(3): 487–510.

Lau, Lawrence. 1969. "Duality and the structure of utility functions." *Journal of Economic Theory*, 1: 374–396.

Lau, Lawrence J. 1974. "Application of Duality Theory: A Comment." In *Frontiers in Quantitative Economics,*. Vol. 2 of *Frontiers in Quantitative Economics*, , ed. M.D. Intriligator and D.A. Kendrick, 176–199. North-Holland Publishing Company, Amsterdam.

Lewbel, Arthur, and Krishna Pendakur. 2009. "Tricks with Hicks: The EASI Demand System." *American Economic Review*, 99(3): 827–863.

Li, Shengyu. 2018. "A structural model of productivity, uncertain demand, and export dynamics." *Journal of International Economics*, 115: 1–15.

Matsuyama, Kiminori. 2019. "Engel's Law in the Global Economy: Demand-Induced Patterns of Structural Change, Innovation, and Trade." *Econometrica*, 87(2): 497–528.

Melitz, Jacques. 2007. "North, South and distance in the gravity model." *European Economic Review*, 51(4): 971–991.

Melitz, Marc J. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica*, 71(6): 1695–1725.

Mukerji, V. 1963. "A Generalized S.M.A.C. Function with Constant Ratios of Elasticity of Substitution." *The Review of Economic Studies*, 30(3): 233–236.

Novy, Dennis. 2013. "International trade without CES: Estimating translog gravity." *Journal of International Economics*, 89(2): 271–282.

Pearson, K. 1991. "Solving Nonlinear Economic Models Accurately Via a Linear Representation." *COPS Working Paper No. IP-55.*

Pigou, A. C. 1910. "A Method of Determining the Numerical Value of Elasticities of Demand." *The Economic Journal*, 20(80): 636–640.

Rimmer, Maureen T., and Alan A. Powell. 1996. "An implicitly additive demand system." *Applied Economics*, 28(12): 1613–1622.

Santos Silva, J. M. C., and Silvana Tenreyro. 2006. "The Log of Gravity." *The Review of Economics and Statistics*, 88(4): 641–658.

Su, Che-Lin, and Kenneth L. Judd. 2012. "Constrained Optimization Approaches to Estimation of Structural Models." *Econometrica*, 80(5): 2213–2230.

Surry, Y. 1993. "The Constant Difference of Elasticities Function with Applications to the Ec Animal Feed Sector." *Journal of Agricultural Economics*, 44(1): 110–125.

Tan, Shawn W. 2013. "Structural Estimation of a Flexible Translog Gravity Model." *University of Melbourne, Department of Economics Working Paper, 1164*, 44.

Tintelnot, Felix. 2017. "Global Production with Export Platforms*." *The Quarterly Journal of Economics*, 132(1): 157–209.

Yilmazkuday, Hakan. 2019. "Welfare Gains from Trade in Multi-Sector Models: The Role of Aggregation and Income Elasticities." *mimeo.*

Mathematical Appendix

A1. Proof of Proposition 3

PROOF: Let $e_i = e$ and $\alpha_i = \alpha\ \forall i$, then Equation (1) generalizes to

$$G(\frac{\mathbf{p}}{E}, u) = \sum_i \beta_i u^{e(1-\alpha)} (\frac{p_i}{E})^{1-\alpha} \equiv 1. \tag{A1}$$

(1) *Utility function is homogeneous* The income elasticity of generalized CDE function

$$\eta_i = \frac{e_i(1-\alpha_i) + \sum_k e_k \omega_k \alpha_k}{\sum_k e_k \omega_k} + \alpha_i - \sum_k \omega_k \alpha_k, \tag{A2}$$

equals 1 $\forall i$ if $e_i = e$ and $\alpha_i = \alpha\ \forall i$.

(2) *Constant elasticity of substitution* The Allen-Uzawa ES

$$\sigma_{ij} = \alpha_i + \alpha_j - \sum_k \omega_k \alpha_k - \frac{\Delta_{ij}\alpha_i}{\omega_i}, \tag{A3}$$

equals $\alpha\ \forall\ i \neq j$ if $e_i = e$ and $\alpha_i = \alpha\ \forall i$.

(3) *Identical to Explicitly Direct CES* Rearranging (A1) by factoring common terms, denoting V as indirect utility, then it leads to the following explicitly indirect expression:

$$V = \Big[\sum_i \beta_i \Big(\frac{p_i}{E}\Big)^{1-\alpha}\Big]^{\frac{1}{e(\alpha-1)}} \quad \text{(Explicitly Indirect Homothetic CES)}, \tag{A4}$$

which is dual to the following explicitly direct CES, with U being the direct utility:

$$U = \Big[\sum_{i=1}^{N} \beta_i^{\frac{1}{\alpha}} q_i^{\frac{\alpha-1}{\alpha}}\Big]^{\frac{\alpha}{e(\alpha-1)}} \quad \text{(Explicitly Direct Homothetic CES)}. \tag{A5}$$

(A5 is the standard Armington representation in the gravity model.)

(4) Standard *CES real wealth assumption can be satisfied*

With $e = 1$ being a special case, the utility functions lead to the following:

$$V = \Big[\sum_i \beta_i \Big(\frac{p_i}{E}\Big)^{1-\alpha}\Big]^{\frac{1}{\alpha-1}} \quad \text{(special case of A4)}, \tag{A6}$$

and

$$U = \Big[\sum_{i=1}^{N} \beta_i^{\frac{1}{\alpha}} q_i^{\frac{\alpha-1}{\alpha}}\Big]^{\frac{\alpha}{\alpha-1}} \quad \text{(special case of A5)}. \tag{A7}$$

The ordinary demand from utility maximization to (A7) yields the same result as derived from (A5), whether or not $e = 1$ is additionally imposed. The demand as a function of $PU \equiv P \equiv E/u$ (as a result from A5) in this special case can then be written as follows:

$$q_i = \frac{\beta_i p_i^{-\alpha}}{PU^{1-\alpha}} E \equiv h_i. \tag{A8}$$

Hence, a standard CDE can be parameterized to achieve (A4) as an explicitly indirect homothetic CES $\iff$ $e_i = e$, $\alpha_i = \alpha \; \forall \; i$, and $\alpha > 1$ or $0 < \alpha < 1$, which is identical to explicitly direct homothetic CES in (A5). It further (i) yields standard CES price index in GE; and (ii) satisfies the CES real wealth assumption $\iff$ $e = 1$, $\alpha_i = \alpha \; \forall \; i$. $\square$

A2. Implicit Indirect NHCES to Implicitly Direct NHCES

Proof of the Parameterization using Hanoch (1975)'s Log-Linear Approach

PROOF: To be complete, let us start from the least restricted CDE demand function:

$$q_i(\mathbf{p}, E) = \frac{\left[\beta_i v^{e_i(1-\alpha_i)}(1-\alpha_i)(\frac{p_i}{E})^{-\alpha_i}\right]}{\sum_j \beta_j v^{e_j(1-\alpha_j)}(1-\alpha_j)(\frac{p_j}{E})^{1-\alpha_j}}. \tag{A9}$$

Note that if $\alpha_i = \alpha \; \forall \; i$, Equation (A9) will converge to the Marshallian demand function of an implicitly indirect NHCES system (Hanoch, 1975). This result can also obtained by applying Roy's Identity to the NHCES defining equation.

Taking the natural logarithm of both sides of (A9) (with $\xi_i = p_i/E$):

(A10)

$$\ln q_i = \ln[\beta_i(1-\alpha_i)] + e_i(1-\alpha_i)\ln u - \alpha_i \ln \xi_i - \ln \Big[\sum_j \beta_j u^{e_j(1-\alpha_j)}(1-\alpha_j)\xi_i^{1-\alpha_j} \Big].$$

Eliminating the last term in equation (A10) by using logarithmic ratio:

(A11)

$$\ln\frac{q_i}{q_1} = \ln\frac{\beta_i(1-\alpha_i)}{\beta_1(1-\alpha_1)} + [e_i(1-\alpha_i) - e_1(1-\alpha_1)]\ln u - \alpha_i \ln\xi_i + \alpha_1 \ln\xi_1$$

$$= A_i + Z_i \ln u - \alpha_i \ln\xi_i + \alpha_1 \ln\xi_1 \quad \forall i \in [2,\infty) \quad \text{(CDE)}$$

$$= \widetilde{A}_i + \widetilde{Z}_i \ln u - \alpha \ln(\frac{p_i}{p_1}) \quad \forall i \in [2,\infty) \qquad \Longleftrightarrow \quad \alpha_i = \alpha \quad \forall i, \quad \text{(Implicitly Indirect NHCES)},$$

where $A_i = \ln\frac{\beta_i(1-\alpha_i)}{\beta_1(1-\alpha_1)}$, $Z_i = e_i(1-\alpha_i) - e_1(1-\alpha_1)$; $\widetilde{A}_i = \ln\frac{\beta_i}{\beta_1}$, $\widetilde{Z}_i = (e_i - e_1)(1-\alpha)$.

Considering the following implicitly direct NHCES function in Hanoch (1975):

$$F(\mathbf{q}, u) = \sum_i k_i u^{-e_i(1-g)} q_i^{1-g} \equiv 1 \quad \text{(Implicitly Direct NHCES)}, \tag{A12}$$

which is parameterized from Mukerji (1963)'s Constant Ratios of Elasticity of Substitution (CRES) model (with $g_i = g\ \forall i$ in A13):

$$F(\mathbf{q}, u) = \sum_i k_i u^{-e_i(1-g_i)} q_i^{1-g_i} \equiv 1 \quad \text{(Implicitly Direct CRES)}, \tag{A13}$$

where the parametric restrictions are

$$\begin{cases} k_i > 0 & \text{(i)} \\ e_i > 0 & \text{(ii)} \\ g_i > 0 & \text{(iii)} \\ g_i \geq 1 \quad \text{or} \quad 0 < g_i \leq 1 & \text{(iv)} \end{cases}$$

$\forall i$ for $u = f(\mathbf{q})$ in (A13) to be globally valid (monotonic and quasi-concave).

For completeness, from the expenditure minimization problem to (A13) (CRES): $\min\{\sum_i p_i q_i : \bar{F} \leq F(\mathbf{q}, u)\}$, the first-order conditions with respect to q_i give rise to

$$p_i = \lambda k_i (1-g_i) u^{-e_i(1-g_i)} q_i^{-g_i} \qquad \forall i, \tag{A14}$$

where

(A15) $$p_1 = \lambda k_1(1-g_1)u^{-e_1(1-g_1)}q_1^{-g_1}.$$

Dividing (A14) by (A15) eliminates $\lambda = \frac{\partial F(\mathbf{q},u)}{\partial q_i} \neq \frac{\partial f(\mathbf{q})}{\partial q_i} = \frac{\partial u}{\partial q_i}$, while yielding

(A16) $$q_1^{-g_1}\frac{p_i}{p_1} = \frac{k_i(1-g_i)}{k_1(1-g_1)}u^{e_1(1-g_1)-e_i(1-g_i)}q_i^{-g_i}.$$

Solving for q_i

(A17) $$q_i = (\frac{p_i}{p_1})^{-\frac{1}{g_i}}\left[\frac{k_i(1-g_i)}{k_1(1-g_1)}\right]^{\frac{1}{g_i}}u^{\frac{e_1(1-g_1)-e_i(1-g_i)}{g_i}}q_1^{\frac{g_1}{g_i}}.$$

Taking the natural logarithm of both sides of (A17):

(A18)
$$\begin{aligned}\ln q_i &= \frac{1}{g_i}\ln\left[\frac{k_i(1-g_i)}{k_1(1-g_1)}\right] - \frac{1}{g_i}\ln(\frac{p_i}{p_1}) + \frac{e_1(1-g_1)-e_i(1-g_i)}{g_i}\ln u + \frac{g_1}{g_i}\ln q_1 \\ &= M_i - s_i\ln(\frac{p_i}{p_1}) + R_i\ln u + \frac{s_i}{s_1}\ln q_1 \quad \forall i \in [2,\infty),\end{aligned}$$

where $s_i = \frac{1}{g_i}$, $M_i = s_i \ln\left[\frac{k_i(1-g_i)}{k_1(1-g_1)}\right]$, and $R_i = s_i[e_1(1-g_1) - e_i(1-g_i)]$. Since the transformation of (A13) (CRES) to (A12) (Implicitly Direct NHCES) has arisen by restricting $g_i = g\ \forall i \implies s_i = s\ \forall i$, then (A18) converges to

(A19)
$$\ln\frac{q_i}{q_1} = \widetilde{M}_i + \widetilde{R}_i \ln u - s\ln(\frac{p_i}{p_1}) \quad \forall i \in [2,\infty) \quad \text{(Implicitly Direct NHCES)},$$

where $\widetilde{M}_i = s\ln\frac{k_i}{k_1}$, and $\widetilde{R}_i = (e_1 - e_i)(s-1)$, which is identical to (A11) $\iff$ $\widetilde{M}_i = \widetilde{A}_i$, $\widetilde{R}_i = \widetilde{Z}_i$, and $s = \alpha \implies \beta_i = k_i^{\alpha}\ \forall i$ and $\alpha = 1/g$. □

A3. Direct Proof of (A2)

Proof of the Parameterization Directly (without using the Log-linear Approach)

PROOF: Multiplying both sides of (A17) by p_i:

$$\text{(A20)} \qquad p_i q_i = p_i^{\frac{g_i-1}{g_i}} p_1^{\frac{1}{g_i}} \left[\frac{k_i(1-g_i)}{k_1(1-g_1)}\right]^{\frac{1}{g_i}} u^{\frac{e_1(1-g_1)-e_i(1-g_i)}{g_i}} q_1^{\frac{g_1}{g_i}}.$$

Then the total expenditure can be expressed as

$$\text{(A21)} \qquad \sum_i p_i q_i = E = \sum_i p_i^{\frac{g_i-1}{g_i}} p_1^{\frac{1}{g_i}} \left[\frac{k_i(1-g_i)}{k_1(1-g_1)}\right]^{\frac{1}{g_i}} u^{\frac{e_1(1-g_1)-e_i(1-g_i)}{g_i}} q_1^{\frac{g_1}{g_i}}.$$

Solving for q_1:

$$\text{(A22)} \qquad q_1 = \frac{E^{\frac{g_i}{g_1}}}{\left[\sum_i p_i^{\frac{g_i-1}{g_i}} p_1^{\frac{1}{g_i}} \left(\frac{k_i(1-g_i)}{k_1(1-g_1)}\right)^{\frac{1}{g_i}} u^{\frac{e_1(1-g_1)-e_i(1-g_i)}{g_i}}\right]^{\frac{g_i}{g_1}}}.$$

Substituting back to the first-order expression for q_i, while eliminating q_1:

(A23)

$$q_i = \left(\frac{p_i}{p_1}\right)^{-\frac{1}{g_i}} \left[\frac{k_i(1-g_i)}{k_1(1-g_1)}\right]^{\frac{1}{g_i}} u^{\frac{e_1(1-g_1)-e_i(1-g_i)}{g_i}} \frac{E}{\sum_i p_i^{\frac{g_i-1}{g_i}} p_1^{\frac{1}{g_i}} \left(\frac{k_i(1-g_i)}{k_1(1-g_1)}\right)^{\frac{1}{g_i}} u^{\frac{e_1(1-g_1)-e_i(1-g_i)}{g_i}}}$$

$$= \frac{\left(\frac{p_i}{p_1}\right)^{-\frac{1}{g_i}} \left[\frac{k_i(1-g_i)}{k_1(1-g_1)}\right]^{\frac{1}{g_i}} u^{\frac{e_1(1-g_1)-e_i(1-g_i)}{g_i}} E}{\sum_i p_i^{\frac{g_i-1}{g_i}} p_1^{\frac{1}{g_i}} \left(\frac{k_i(1-g_i)}{k_1(1-g_1)}\right)^{\frac{1}{g_i}} u^{\frac{e_1(1-g_1)-e_i(1-g_i)}{g_i}}}.$$

If $g_i = g\ \forall i$, then the implicitly Direct CRES generalizes to implicitly direct NHCES. Then we may directly apply this parameterization to (A23) and solves for q_i:

$$\text{(A24)} \quad \begin{aligned} q_i &= \frac{\left(\frac{p_i}{p_1}\right)^{-\frac{1}{g_i}} \left[\frac{k_i(1-g_i)}{k_1(1-g_1)}\right]^{\frac{1}{g_i}} u^{\frac{e_1(1-g_1)-e_i(1-g_i)}{g_i}} E}{\sum_i p_i^{\frac{g_i-1}{g_i}} p_1^{\frac{1}{g_i}} \left(\frac{k_i(1-g_i)}{k_1(1-g_1)}\right)^{\frac{1}{g_i}} u^{\frac{e_1(1-g_1)-e_i(1-g_i)}{g_i}}} \\ &= \frac{p_i^{-\frac{1}{g}} \left(\frac{k_i}{k_1}\right)^{\frac{1}{g}} u^{\frac{(e_1-e_i)(1-g)}{g}} E}{\sum_i p_i^{\frac{g-1}{g}} \left(\frac{k_i}{k_1}\right)^{\frac{1}{g}} u^{\frac{(e_1-e_i)(1-g)}{g}}} \\ &= \frac{p_i^{-\frac{1}{g}} \left(\frac{k_i}{k_1}\right)^{\frac{1}{g}} u^{\frac{(e_1-e_i)(1-g)}{g}} E}{\sum_i p_i^{\frac{g-1}{g}} \left(\frac{k_i}{k_1}\right)^{\frac{1}{g}} u^{\frac{(e_1-e_i)(1-g)}{g}}} \\ &= \frac{k_i^{\frac{1}{g}} u^{\frac{e_i(g-1)}{g}} p_i^{-\frac{1}{g}} E}{\sum_i k_i^{\frac{1}{g}} u^{\frac{e_i(g-1)}{g}} p_i^{\frac{g-1}{g}}} \end{aligned}$$

Let $g = 1/\alpha$, then we may express (A24) as follows:

$$\text{(A25)} \qquad q_i = \frac{k_i^{\alpha} u^{e_i(1-\alpha)} p_i^{-\alpha} E}{\sum_i k_i^{\alpha} u^{e_i(1-\alpha)} p_i^{1-\alpha}}.$$

Let $k_i = \beta_i^{1/\alpha} = \beta_i^{g}$, then implicitly direct NHCES is identical to implicitly indirect NHCES (which can be generalized by letting $\alpha_i = \alpha \ \forall i$ in the standard CDE), and leads q_i to the following expression:

$$\text{(A26)} \qquad q_i = \frac{\beta_i u^{e_i(1-\alpha)} p_i^{-\alpha} E}{\sum_i \beta_i u^{e_i(1-\alpha)} p_i^{1-\alpha}},$$

which is derived under the same parameterization using the linear approach in Appendix A1. □

A4. Derivation and Extension of Counterfactual Equations

Considering a closed-economy (autarky). Taking total derivatives with respect to implicit utility, the price vector, and per capita income in Equation (2) leads to the following expression:

(A27)

$$\hat{q}_i = e_i(1-\alpha_i)\hat{u} - \alpha_i\hat{p}_i + \alpha_i\hat{E} - \sum_j e_j(1-\alpha_j)\omega_j\hat{u} - \sum_j(1-\alpha_j)\omega_j\hat{p}_j + \sum_j(1-\alpha_j)\omega_j\hat{E}$$

$$= \left[e_i(1-\alpha_i) - \sum_j e_j(1-\alpha_j)\omega_j\right]\hat{u}$$

$$+ \left[\alpha_i + \sum_j(1-\alpha_j)\omega_j\right]\hat{E}$$

$$- \alpha_i\hat{p}_i - \sum_j(1-\alpha_j)\omega_j\hat{p}_j.$$

Equation (A27) can be considered as a special case with zero trade costs (e.g., $\tau_{il} = 1$). The counterfactual result of quantity consumption depends on utility and income changes, as well as price changes of own-goods and all other goods bundle. Meanwhile changes in utility, income and prices are interacted with expansion and substitution parameters, e_i and α_i, respectively with respect to goods i, as well as with the share-weighted parameter values of the two with respect to all goods $i \in I$.

Substituting income and price elasticities into (A27), the change of quantity consumption can also be expressed as a function of income and cross-price elasticities:

$$\hat{q}_i = \eta_i\hat{E} + \sum_j \sigma_{i,j}\hat{p}_j, \tag{A28}$$

where $\eta_i = \frac{e_i(1-\alpha_i)+\sum_k e\omega_k\alpha_k}{\sum_k e\omega_k} + \alpha_i - \sum_k \omega_k\alpha_k$ and $\sigma_{i,j} = \alpha_i + \alpha_j - \sum_k \omega_k\alpha_k - \frac{\Delta_{ij}\alpha_i}{\omega_i}$, using Equations (3) and (5).

Note that in the case of a general CES, the effect of the change in utility is removed:

$$\hat{q}_i = \hat{E} - \alpha\hat{p}_i - (1-\alpha)\sum_j \hat{p}_j, \tag{A29}$$

which is identical to the counterfactual result of the standard CES.

A5. *Derivation of Equation (46)*

Implementing total differentiation in (1) with respect to utility, wealth and the price vector:

$$\begin{aligned}
&\sum_i \beta_i e_i (1-\alpha_i) u^{e_i(1-\alpha_i)-1} (\frac{p_i}{E})^{1-\alpha_i} du \\
&+ \sum_i (\alpha_i - 1)\beta_i u^{e_i(1-\alpha_i)} p_i^{1-\alpha_i} E^{\alpha_i - 2} dE \\
&+ \sum_i \beta_i (1-\alpha_i) u^{e_i(1-\alpha_i)} p_i^{-\alpha_i} E^{\alpha_i - 1} dp_i \\
\equiv \quad & 0.
\end{aligned} \tag{A30}$$

By rearranging terms above, we have

$$\begin{aligned}
&\sum_i (1-\alpha_i)\beta_i u^{e_i(1-\alpha_i)} (\frac{p_i}{E})^{1-\alpha_i} \qquad \overbrace{\frac{dw}{E}}^{\text{Change of Wealth}} \\
\equiv &\sum_i \beta_i e_i (1-\alpha_i) u^{e_i(1-\alpha_i)} (\frac{p_i}{E})^{1-\alpha_i} \qquad \underbrace{\overbrace{\frac{du}{u}}}_{\text{Change of Utility}} \\
+ &\sum_i \beta_i (1-\alpha_i) u^{e_i(1-\alpha_i)} (\frac{p_i}{E})^{1-\alpha_i} \qquad \underbrace{\overbrace{\frac{dp_i}{p_i}}}_{\text{Change of Price}} .
\end{aligned} \tag{A31}$$

Rewriting the CDE expenditure share expression using a proxy T, e.g.,

$$\omega_i = \frac{\beta_i u^{e_i(1-\alpha_i)}(1-\alpha_i)(\frac{p_i}{E})^{1-\alpha_i}}{\sum_j \beta_j u^{e_j(1-\alpha_j)}(1-\alpha_j)(\frac{p_j}{E})^{1-\alpha_j}} = \frac{\beta_i u^{e_i(1-\alpha_i)}(1-\alpha_i)(\frac{p_i}{E})^{1-\alpha_i}}{T}. \tag{A32}$$

Note that if we divide both sides by T of Equation (A31), and using *hat* to denote rate of changes on corresponding terms, then the expression can be further simplified to

$$\sum_i \omega_i \hat{E} \equiv \sum_i e_i \omega_i \hat{u} + \sum_i \omega_i \hat{p}_i, \tag{A33}$$

and since $\sum_i \omega_i = 1$, and $\hat{w}$ does not depend on each i, we obtain

$$\hat{E} \equiv \sum_i e_i \omega_i \hat{u} + \sum_i \omega_i \hat{p}_i, \tag{A34}$$

and since $\hat{u}$ does not depend on each i, the change of utility can be written as

$$\hat{u} = \frac{\hat{E} - \sum_i \omega_i \hat{p}_i}{\sum_i e_i \omega_i}. \tag{A35}$$

Online Appendix and Supplementary Materials

Statement: The Online Appendix has been posted on my personal site and was previously circulated as "Universal CES Demand Systems and Counterfactuals in International Trade" The supplementary materials are developed in and supplement this chapter.

The standard CDE is a class of implicitly indirect additive demand system. It is a general case of the standard CES demand system. The nature of the use of the terminology "implicit" rather than "explicit" is that utility in the model cannot be explicitly and algebraically solved using the model's exogenous variables and model parameters. The distinction between "indirect" and "direct" is that indirect additive models are separable in the n unit-cost or normalized prices along consumers' indifference surfaces, whereas directly separable models are additive in n consumer goods (Hanoch (1975)). The standard CDE model is implicitly and indirectly defined as:

(B1) $$G(\frac{\mathbf{p}}{w}, u) = \sum_i \beta_i u^{e_i(1-\alpha_i)} (\frac{p_i}{w})^{1-\alpha_i} \equiv 1 \quad \text{(CDE)}.$$

Parametric restrictions $\forall i$ for (B1) to be globally valid (monotonic and quasi-concave), $\forall$ normalized price vectors (e.g., unit-cost prices) $\boldsymbol{\xi} = \frac{\mathbf{p}}{w} \gg \mathbf{0}$, are as follows Hanoch (1975):[24]

$$\begin{cases} \beta_i > 0 & \text{(i)} \\ e_i > 0 & \text{(ii)} \\ \alpha_i > 0 & \text{(iii)} \\ \alpha_i \geq 1 \quad \text{or} \quad 0 < \alpha_i < 1 & \text{(iv)} \end{cases}$$

Using Roy's Identity, the Marshallian or ordinary demand correspondence is given by:

(B2) $$q_i(\mathbf{p}, w) = \frac{\left[\beta_i v^{e_i(1-\alpha_i)}(1-\alpha_i)(\frac{p_i}{w})^{-\alpha_i}\right]}{\sum_j \beta_j v^{e_j(1-\alpha_j)}(1-\alpha_j)(\frac{p_j}{w})^{1-\alpha_j}}.$$

B1. Implicit Indirect Non-Homothetic CES

PROPOSITION 4: *Let $G(\boldsymbol{\xi}, u)$ be an implicit indirect utility function of Constant Difference of Elasticities (CDE), then $G(\boldsymbol{\xi}, u)$ can be parameterized to be an implicit indirect version of non-homothetic CES function, which is identical to implicit direct non-homothetic CES.*

PROOF: Let $\alpha_i = \alpha \ \forall i$, then (B1) generalizes to

(B3) $$G(\frac{\mathbf{p}}{w}, u) = \sum_i \beta_i u^{e_i(1-\alpha)} (\frac{p_i}{w})^{1-\alpha} \equiv 1 \quad \text{(Implicitly Indirect NHCES)}.$$

(1) *Constant elasticity of substitution* The cross-Allen partial elasticity is:

[24]Condition (iv) automatically satisfies (iii), but they are not equivalently meaningful for the global or local regularity condition. The model is *dampened* to be locally valid if the "same sign" condition in (iv) is not satisfied, and $\alpha_I < 0$ for some $I \in i$, which violates (iii).

$$\begin{aligned}
\sigma_{ij} &= \alpha_i + \alpha_j - \sum_k \pi_k \alpha_k - \frac{\Delta_{ij}\alpha_i}{\pi_i} \\
&= 2\alpha - \alpha \sum_k \pi_k - \frac{\Delta_{ij}\alpha}{\pi_i} \qquad \Longleftrightarrow \quad \alpha_i = \alpha \quad \forall i \\
&= \alpha\left(2 - \sum_k \pi_k - \frac{\Delta_{ij}}{\pi_i}\right) \\
&= \alpha\left(1 - \frac{\Delta_{ij}}{\pi_i}\right) \\
&= \alpha \quad \forall i \neq j.
\end{aligned} \tag{B4}$$

(2) *Identical to implicit direct NHCES* By §2.2 and §2.4 in Hanoch (1975):[25]

Taking the natural logarithm of both sides of (B2) (with $\xi_i = p_i/w$):

(B5)
$$\ln q_i = \ln[\beta_i(1-\alpha_i)] + e_i(1-\alpha_i)\ln u - \alpha_i \ln \xi_i - \ln\big[\sum_j \beta_j u^{e_j(1-\alpha_j)}(1-\alpha_j)\xi_i^{1-\alpha_j}\big].$$

Eliminating the last term in equation (B5) by using logarithmic ratio:

$$\begin{aligned}
\ln\frac{q_i}{q_1} &= \ln\frac{\beta_i(1-\alpha_i)}{\beta_1(1-\alpha_1)} + [e_i(1-\alpha_i) - e_1(1-\alpha_1)]\ln u - \alpha_i \ln \xi_i + \alpha_1 \ln \xi_1 \\
&= A_i + Z_i \ln u - \alpha_i \ln \xi_i + \alpha_1 \ln \xi_1 \quad \forall i \in [2, \infty) \\
&= \widetilde{A}_i + \widetilde{Z}_i \ln u - \alpha \ln(\frac{p_i}{p_1}) \quad \forall i \in [2, \infty) \qquad \Longleftrightarrow \quad \alpha_i = \alpha \quad \forall i,
\end{aligned} \tag{B6}$$

where $A_i = \ln\frac{\beta_i(1-\alpha_i)}{\beta_1(1-\alpha_1)}$, $Z_i = e_i(1-\alpha_i) - e_1(1-\alpha_1)$; $\widetilde{A}_i = \ln\frac{\beta_i}{\beta_1}$, $\widetilde{Z}_i = (e_i - e_1)(1-\alpha)$.

Considering the following implicitly direct NHCES function in Hanoch (1975):

$$F(\mathbf{q}, u) = \sum_i k_i u^{-e_i(1-g)} q_i^{1-g} \equiv 1 \quad \text{(Implicitly Direct NHCES)}, \tag{B7}$$

which is parameterized from Mukerji (1963)'s Constant Ratios of Elasticity of

[25]See pg.414 of Hanoch (1975).

Substitution (CRES) model (with $g_i = g\ \forall i$ in B8):

$$F(\mathbf{q}, u) = \sum_i k_i u^{-e_i(1-g_i)} q_i^{1-g_i} \equiv 1 \quad \text{(Implicitly Direct CRES)}, \tag{B8}$$

where the parametric restrictions are

$$\begin{cases} k_i > 0 & \text{(i)} \\ e_i > 0 & \text{(ii)} \\ g_i > 0 & \text{(iii)} \\ g_i \geq 1 \quad \text{or} \quad 0 < g_i \leq 1 & \text{(iv)} \end{cases}$$

$\forall i$ for $u = f(\mathbf{q})$ in (B8) to be globally valid (monotonic and quasi-concave).
For completeness, from the expenditure minimization problem to (B8) (CRES): $\min\{\sum_i p_i q_i : \bar{F} \leq F(\mathbf{q}, u)\}$, the first-order conditions with respect to q_i give rise to:

$$p_i = \lambda k_i (1 - g_i) u^{-e_i(1-g_i)} q_i^{-g_i} \qquad \forall i, \tag{B9}$$

where

$$p_1 = \lambda k_1 (1 - g_1) u^{-e_1(1-g_1)} q_1^{-g_1} \tag{B10}$$

Dividing (B9) by (B10) eliminates $\lambda = \frac{\partial F(\mathbf{q},u)}{\partial q_i} \neq \frac{\partial f(\mathbf{q})}{\partial q_i} = \frac{\partial u}{\partial q_i}$, while yielding:

$$q_1^{-g_1} \frac{p_i}{p_1} = \frac{k_i(1 - g_i)}{k_1(1 - g_1)} u^{e_1(1-g_1) - e_i(1-g_i)} q_i^{-g_i}. \tag{B11}$$

Solving for q_i:

$$q_i = (\frac{p_i}{p_1})^{-\frac{1}{g_i}} \left[\frac{k_i(1 - g_i)}{k_1(1 - g_1)}\right]^{\frac{1}{g_i}} u^{\frac{e_1(1-g_1) - e_i(1-g_i)}{g_i}} q_1^{\frac{g_1}{g_i}}. \tag{B12}$$

Taking the natural logarithm of both sides of (B12): [26]

(B13)

$$\ln q_i = \frac{1}{g_i}\ln\left[\frac{k_i(1-g_i)}{k_1(1-g_1)}\right] - \frac{1}{g_i}\ln(\frac{p_i}{p_1}) + \frac{e_1(1-g_1)-e_i(1-g_i)}{g_i}\ln u + \frac{g_1}{g_i}\ln q_1$$
$$= M_i - s_i\ln(\frac{p_i}{p_1}) + R_i\ln u + \frac{s_i}{s_1}\ln q_1 \quad \forall i \in [2,\infty),$$

where $s_i = \frac{1}{g_i}$, $M_i = s_i \ln\left[\frac{k_i(1-g_i)}{k_1(1-g_1)}\right]$, and $R_i = s_i[e_1(1-g_1) - e_i(1-g_i)]$.
Since the transformation of B8 (CRES) to B7 (Implicitly Direct NHCES) has arisen by restricting $g_i = g \ \forall i \implies s_i = s \ \forall i$, then (B13) converges to:

(B14)
$$\ln\frac{q_i}{q_1} = \widetilde{M_i} + \widetilde{R}_i \ln u - s\ln(\frac{p_i}{p_1}) \quad \forall i \in [2,\infty),$$

where $\widetilde{M_i} = s\ln\frac{k_i}{k_1}$, and $\widetilde{R}_i = (e_1 - e_i)(s-1)$;
which is identical to (B6) $\iff$ $\widetilde{M_i} = \widetilde{A}_i$, $\widetilde{R}_i = \widetilde{Z}_i$, and $s = \alpha \implies \beta_i = k_i^\alpha \ \forall i$ and $\alpha = 1/g$. $\square$
To summarize,

$$\begin{cases} \text{CDE} \longrightarrow \text{Implicitly Indirect NHCES} & \iff \alpha_i = \alpha \ \forall i \\ \text{Implicitly Indirect NHCES} \equiv \text{Implicitly Direct NHCES} & \iff \alpha = 1/g \text{ and } \beta_i = k_i^\alpha \ \forall i \\ \text{Implicit CRES} \longrightarrow \text{Implicitly Direct NHCES} & \iff g_i = g \ \forall i \text{ and } \neg\,\forall e_i = e. \end{cases}$$

B2. Explicit Indirect Homothetic CES

As shown in Hanoch (1975), the specialized CDE can be parameterized to achieve a homothetic CES demand model. Such parameterization from CDE allows testings against the standard CES in GE trade models. It is common in GE to include an additional set of equilibrium conditions that characterizes a global economy, e.g., aggregate price index (see, e.g., Dixit and Stiglitz (1977) and the *real wealth assumption*. However, it can be challenging for some general demand

[26]The following proof is established using a linear approach (using log-differencing) based on Hanoch (1975). It turns out that the resulting parametric requirements are slightly different from Hanoch. We show that, in addition to allow $\alpha = 1/g$, distribution parameter β_i in the CDE requires to be equate k_i^α $\forall i$ (instead of $\beta_i = k_i \ \forall i$ in Hanoch), in order to transform implicitly indirect NHCES (as a special case generalized from CDE) to implicitly direct NHCES.

systems to be parameterized in such a way that they always yield some desired equilibrium framework. We demonstrate a procedure under which circumstances that (1) CDE will converge to a homothetic CES demand system (as originally introduced by Hanoch); and (2) that additional parametric restrictions are required to be suitable in a GE framework, where there exists an aggregate price index and real wealth assumption is imposed.[27]

PROPOSITION 5: *Let $G(\boldsymbol{\xi}, u)$ be an implicitly indirectly additive utility function of Constant Difference of Elasticities (CDE), then $G(\boldsymbol{\xi}, u)$ can be parameterized to achieve an explicitly indirect constant Elasticity of Substitution (CES) function, which is identical to its explicitly direct case; it can be further parameterized to satisfy the standard CES price index, while satisfying the CES real wealth assumption in a general equilibrium framework.*

Definition 1: *A CES real wealth assumption is that price indices of aggregate goods consumed by a representative consumer, or cost of per capita utility, equates the per capita income adjusted by the per capita utility of the representative consumer, i.e.:*

$$P = w/u, \tag{B15}$$

where P is the aggregate or consumer price index, w is the per capita income, and u is the per capita utility.

PROOF: Let $e_i = e$ and $\alpha_i = \alpha$ $\forall i$, then (B1) generalizes to

$$G(\frac{\mathbf{p}}{w}, u) = \sum_i \beta_i u^{e(1-\alpha)} (\frac{p_i}{w})^{1-\alpha} \equiv 1. \tag{B16}$$

(1) *Utility function is homogeneous* The income elasticity of generalized CDE function is:

[27]The expression of an aggregate price index in CGE is virtually an assumption; also see Melitz (2003).

$$\eta_i = \frac{e(1-\alpha) + \sum_k e\pi_k\alpha}{\sum_k e\pi_k} + \alpha - \sum_k \pi_k\alpha = 1. \tag{B17}$$

(2) *Constant elasticity of substitution* The cross-Allen partial elasticity is:

$$\sigma_{ij} = \alpha_i + \alpha_j - \sum_k \pi_k\alpha_k - \frac{\Delta_{ij}\alpha_i}{\pi_i} = \alpha \quad \forall i \neq j. \tag{B18}$$

(3) *Identical to Explicitly Direct CES* Rearranging (B16) by factoring common terms:

$$u^{e(1-\alpha)} \sum_i \beta_i p_i^{1-\alpha} = w^{1-\alpha}, \tag{B19}$$

which can be further simplified to

$$\underbrace{u^e \left(\sum_i \beta_i p_i^{1-\alpha} \right)^{\frac{1}{1-\alpha}}}_{=P \iff e=1} = w, \tag{B20}$$

and leads to the following explicitly indirect expression by isolating u:

$$U = \Big[\sum_i \beta_i \Big(\frac{p_i}{w}\Big)^{1-\alpha} \Big]^{\frac{1}{e(\alpha-1)}} \quad \text{(Explicitly Indirect Homothetic CES)}, \tag{B21}$$

which is identical to the following explicitly direct homothetic CES:

$$U = \Big[\sum_{i=1}^{N} \beta_i^{\frac{1}{\alpha}} q_i^{\frac{\alpha-1}{\alpha}} \Big]^{\frac{\alpha}{e(\alpha-1)}} \quad \text{(Explicitly Direct Homothetic CES)}. \tag{B22}$$

Suppose the identities take the form of (B22), while global regularity conditions are satisfied as parametrically restricted in (B1) except $\alpha \neq 1$. The utility maximization solving the problem: $\max\{u(\mathbf{q}) : \sum_i p_i q_i \leq w, i = 1, ..., n\}$, leads to the following demand function:

$$q_i = \frac{\beta_i p_i^{-\alpha}}{\sum_i \beta_i p_i^{1-\alpha}} w, \tag{B23}$$

which is identical to (B2) if $e_i = e$ and $\alpha_i = \alpha \; \forall i$. Also, substitution of (B23) into first-order conditions of (B22) yields the same price index as in the implicitly indirect case.[28]

(4) *CES Real wealth assumption can be further satisfied*

With $e = 1$ as a special version in (B20), the utility function leads to the following:[29]

$$U = \Big[\sum_i \beta_i \Big(\frac{p_i}{w}\Big)^{1-\alpha}\Big]^{\frac{1}{\alpha-1}} \quad \text{(special case of B21)}. \tag{B24}$$

P in (B20) is the exact form of CES price indices defined in Dixit and Stiglitz (1977) based on Green (1964), with however the following direct homothetic CES preference:[30]

$$U = \Big[\sum_{i=1}^{N} \beta_i^{\frac{1}{\alpha}} q_i^{\frac{\alpha-1}{\alpha}}\Big]^{\frac{\alpha}{\alpha-1}} \quad \text{(special case of B22)}. \tag{B25}$$

The ordinary demand from utility maximization to (B25) yields the same result as in (B23), which is derived from (B21), invariant to whether $e = 1$ is additionally imposed. With the expression for the price index P in (B20), the demand as a function of P in this special case can then be expressed as follows, which is

[28] See Appendix C for complete mathematical derivations of Marshallian demand and price index for the explicitly direct CES. For the explicitly indirect CES, see section 4 for the price index derivation, and for the Marshallian demand it can be readily verified by parameterization to the derived CDE demand.

[29] See Appendix A for the derivation of Explicitly Indirect Homothetic CES.

[30] It turns out that the price index P solved from the explicit case is equivalent to the implicit case, which is computed from substituting U [derived from the total differentiation to (B1)] to (B22).

equivalent as derived from (B24):

$$q_i = \frac{\beta_i p_i^{-\alpha}}{P^{1-\alpha}} w. \tag{B26}$$

Hence, a standard CDE as in (B1) can be parameterized to achieve (B21) of an explicitly indirect homothetic CES $\iff$ $e_i = e$, $\alpha_i = \alpha \; \forall \; i$, and $\alpha > 1$ or $0 < \alpha < 1$, which is identical to explicitly direct homothetic CES as in (B22); and further satisfies (1) standard CES price index in GE; and (2) the real wealth assumption $\iff$ $e = 1$, $\alpha_i = \alpha \; \forall \; i$. □

Parameterization to allow a standard CDE to transform into an indirect homothetic CES system automatically makes the demand model explicit. It is due to the fact that utility can be algebraically solved in terms of the model's exogenous variables and parameters, when α and e are both uniform.

B3. Price Indices

PROPOSITION 6: *Implicitly Indirect Non-Homothetic CDE, Implicit Homogeneous CDE and Implicitly Indirect Non-Homothetic CES systems all lead to aggregate price indices that are implicitly defined with endogenized utility; Explicitly Indirect Non-Homothetic CES leads to explicit price index, where utility index can be solved in terms of exogenous variables of the system.*

PROOF: *Non-Homothetic CDE* Suppose the utility function is implicitly defined as an implicitly indirect CDE in (B1) and assume all global regularity conditions hold. Following Chen (2017), the total differentiation of (B1) with respect to u and w at a given price vector, leads to the expression as follows:

(B27)
$$\sum_i \beta_i e_i (1-\alpha_i) u^{e_i(1-\alpha_i)-1} (\frac{p_i}{w})^{1-\alpha_i} du + \sum_i (\alpha_i - 1) \beta_i u^{e_i(1-\alpha_i)} p_i^{1-\alpha_i} w^{\alpha_i - 2} dw \equiv 0.$$

Then the marginal cost of utility can be derived as

(B28)
$$\frac{dw}{du} = \left[\sum_i \beta_i u^{e_i - e_i\alpha_i - 1} (1-\alpha_i) p_i^{1-\alpha_i} w^{\alpha_i - 1} e_i \right] \left[\sum_i \beta_i u^{e_i(1-\alpha_i)} (1-\alpha_i) p_i^{1-\alpha_i} w^{\alpha_i - 2} \right]^{-1},$$

where $\frac{dw}{du} = P \equiv \lambda^{-1}$, P is the aggregate price index; and λ is the Lagrange multiplier (from the utility maximization problem as if it were solved from the explicit case), representing marginal utility of income, as can be solved from the utility maximization to any explicit direct utility functions: $\max\{u(\mathbf{q}) : p'q \leq w\}$ with its gradient vector evaluated at $\mathbf{q}$ at an interior optimum.[31]

Given (B1), (B28) can be rewritten in a reduced form using the following expression:

$$P = \eta \frac{w}{u}, \tag{B29}$$

where $\eta = \sum_i \pi_i e_i$ is the elasticity of aggregate expenditure with respect to utility, and π_i is the optimal expenditure share. Since the expenditure share is a function of an ordinal demand in (B2) where its utility is endogenized, and the irreducible summation cannot be factored out over e_i, then P is implicitly defined, depending on the utility.

Implicit Indirect NHCES By proposition 2, an implicitly indirect CDE will converge to an implicitly indirect CES if $\alpha_i = \alpha \; \forall i$. Using (B28), then the price index is expressed as

$$P = \frac{\sum_i \beta_i u^{e_i(1-\alpha)} p_i^{1-\alpha} e_i}{\sum_i \beta_i u^{e_i(1-\alpha)} p_i^{1-\alpha}} \frac{w}{u}, \tag{B30}$$

which is implicitly defined with the utility u, as in the case of non-homothetic CDE.

Implicit Homogeneous CDE Again, by Proposition 2, an implicitly indirect CDE can be parameterized to transform to an implicit homogeneous CDE by allowing $e_i = e \; \forall i$, then the price index can be expressed as

[31] The consumer price index $P \equiv \lambda^{-1}$ can be further verified in Dixit and Stiglitz (1977) and Green (1964). In Appendix B, we show a simple example of the version of Dixit-Stiglitz two goods; Appendix C is indeed a more generalized case with N goods with the expansion parameter $e \neq 1$.

$$P = e\frac{w}{u}, \tag{B31}$$

which appears to be a succinct functional form, but cannot be further simplified with elimination of u, which is endogenously determined even $e_i = e \ \forall i$.

Explicit Indirect Homothetic CES We know that the price index takes a reduced form in (B31) if $e_i = e \ \forall i$, then in the case of explicitly indirect non-homothetic CES, the price index yields the same expression when, additionally, allowing $\alpha_i = \alpha \ \forall i$, as can be also verified in (B30) if e is factored out of the summation.

From the identity in (B19), it is readily demonstrable that the price index can be algebraically solved in this explicit case. Rearranging (B31) and then substituting $u = \frac{ew}{P}$ into (B20) immediately yields the following expression:

$$\Big(\frac{ew}{P}\Big)^{e}\Big(\sum_i \beta_i p_i^{1-\alpha}\Big)^{\frac{1}{1-\alpha}} = w. \tag{B32}$$

Isolating P to derive the price index of the exact form:

$$\begin{aligned} P &= e\Big[\sum_i \beta_i p_i^{1-\alpha} w^{(e-1)(1-\alpha)}\Big]^{\frac{1}{e(1-\alpha)}} \\ &= e\Big[\sum_i \beta_i (\frac{p_i}{w^{1-e}})^{1-\alpha}\Big]^{\frac{1}{e(1-\alpha)}}. \end{aligned} \tag{B33}$$

It is easy to see from (B33), the price index leads to the same result (shown under the curly brackets in B20) as derived from the explicitly direct case in (B22). □

B4. Expenditure Shares

Implicit Indirect Non-Homothetic CDE With $\pi_i = \frac{p_i q_i}{w}$ and ordinary demand in (B2), the per capita expenditure share of i in the standard CDE can be expressed as the following:

$$\pi_i = \frac{\beta_i(1-\alpha_i)u^{e_i(1-\alpha_i)}(\frac{p_i}{w})^{1-\alpha_i}}{\sum_j \beta_j(1-\alpha_j)u^{e_j(1-\alpha_j)}(\frac{p_j}{w})^{1-\alpha_j}}. \tag{B34}$$

Implicit Indirect Homogeneous CDE With $e_i = e \ \forall i$, the standard CDE is homogenous. The expenditure share leads to the following expression:

$$\pi_i = \frac{\beta_i(1-\alpha_i)u^{-e\alpha_i}(\frac{p_i}{w})^{1-\alpha_i}}{\sum_j \beta_j(1-\alpha_j)u^{-e\alpha_j}(\frac{p_j}{w})^{1-\alpha_j}}. \tag{B35}$$

Implicit Indirect NHCES By proposition 1, the CDE function converges to Implicitly Indirect NHCES $\iff \alpha_i = \alpha \ \forall i$. Its expenditure share leads to the following expression:

$$\pi_i = \frac{\beta_i u^{e_i(1-\alpha)} p_i^{1-\alpha}}{\sum_j \beta_j u^{e_j(1-\alpha)} p_j^{1-\alpha}}. \tag{B36}$$

(B36) shows that expenditure shares in the implicit NHCES are not directly affected by changes in income, although they are responsive to incomes where utility is endogenized and is implicitly defined as a function of incomes.

Explicit Indirect Homothetic CES

By proposition 2, convergence to homothetic CES requires that $e_i = e \ \forall i$, which leads to the following expenditure shares expression:[32]

$$\pi_i = \frac{\beta_i p_i^{1-\alpha}}{\sum_j \beta_j p_j^{1-\alpha}} \implies \beta_i \left(\frac{P}{p_i}\right)^{\alpha-1} \iff e = 1. \tag{B37}$$

[32]P is the price index with respect to (B21).

(B36) shows that the expenditure shares under implicit NHCES are only affected by changes in the price vector, but are independent of any income changes.

B5. Counterfactual Welfare Responses (in Percent Changes)

Implicit Indirect Non-Homothetic CDE

Implementing total differentiation in (B1) with respect to utility, wealth and the price vector:

$$\begin{aligned}
&\sum_i \beta_i e_i (1-\alpha_i) u^{e_i(1-\alpha_i)-1} (\frac{p_i}{w})^{1-\alpha_i} du \\
&+ \sum_i (\alpha_i - 1) \beta_i u^{e_i(1-\alpha_i)} p_i^{1-\alpha_i} w^{\alpha_i - 2} dw \\
&+ \sum_i \beta_i (1-\alpha_i) u^{e_i(1-\alpha_i)} p_i^{-\alpha_i} w^{\alpha_i - 1} dp_i \\
&\equiv \quad 0.
\end{aligned} \tag{B38}$$

Equation (B38) can be simplified as:

$$\begin{aligned}
&\sum_i (1-\alpha_i) \beta_i u^{e_i(1-\alpha_i)} (\frac{p_i}{w})^{1-\alpha_i} \overbrace{\frac{dw}{w}}^{\text{Change of Wealth}} \\
\equiv &\sum_i \beta_i e_i (1-\alpha_i) u^{e_i(1-\alpha_i)} (\frac{p_i}{w})^{1-\alpha_i} \underbrace{\overbrace{\frac{du}{u}}}_{\text{Change of Utility}} \\
+ &\sum_i \beta_i (1-\alpha_i) u^{e_i(1-\alpha_i)} (\frac{p_i}{w})^{1-\alpha_i} \underbrace{\overbrace{\frac{dp_i}{p_i}}}_{\text{Change of Price}}
\end{aligned} \tag{B39}$$

Rewriting the CDE expenditure share expression in (B34), e.g., $\frac{p_i q_i}{\sum_i p_i q_i}$:

$$\pi_i = \frac{\beta_i u^{e_i(1-\alpha_i)} (1-\alpha_i) (\frac{p_i}{w})^{1-\alpha_i}}{\sum_j \beta_j u^{e_j(1-\alpha_j)} (1-\alpha_j) (\frac{p_j}{w})^{1-\alpha_j}} = \frac{\beta_i u^{e_i(1-\alpha_i)} (1-\alpha_i) (\frac{p_i}{w})^{1-\alpha_i}}{T}. \tag{B40}$$

Note that if we divide both sides by T of equation (B39), and using *hat* to denote rate of changes on corresponding terms, then the expression can be further simplified to

$$\sum_i \pi_i \hat{w} \equiv \sum_i e_i \pi_i \hat{u} + \sum_i \pi_i \hat{p}_i, \tag{B41}$$

and since $\sum_i \pi_i = 1$, and $\hat{w}$ does not depend on each i, we obtain:

$$\hat{w} \equiv \sum_i e_i \pi_i \hat{u} + \sum_i \pi_i \hat{p}_i, \tag{B42}$$

and because $\hat{u}$ does not depend on each i, the change of utility can be written as:

$$\hat{u} = \frac{\hat{w} - \sum_i \pi_i \hat{p}_i}{\sum_i e_i \pi_i}. \tag{B43}$$

where $\pi_i = \dfrac{\beta_i(1-\alpha_i) \overbrace{u^{e_i(1-\alpha_i)}}^{\text{Effects of Baseline Utility}} \overbrace{(\frac{p_i}{w})^{1-\alpha_i}}^{\text{Effects of Baseline Normalized Prices}}}{\underbrace{\sum_j \beta_j(1-\alpha_j) u^{e_j(1-\alpha_j)} (\frac{p_j}{w})^{1-\alpha_j}}_{\text{Composite Effects of Baseline Utility, Income and Prices}}}$ following ($B34$).

In this case, change of cardinal utility does not only respond to changes of income and price vectors, but is also determined by heterogeneous commodity-specific CDE parameters as well as baseline income, utility and commodity prices.

Implicit Indirect Homogeneous CDE In the case of homogenous CDE where $e_i = e\ \forall i$, the denominator in ($B43$) converges to the uniform expansion parameter e, where $\pi_i = \frac{\beta_i(1-\alpha_i)u^{-e\alpha_i}(\frac{p_i}{w})^{1-\alpha_i}}{\sum_j \beta_j(1-\alpha_j)u^{-e\alpha_j}(\frac{p_j}{w})^{1-\alpha_j}}$ that follows (B35):

$$\hat{u} = \frac{\hat{w} - \sum_i \pi_i \hat{p}_i}{e} \tag{B44}$$

Implicit Indirect NHCES With uniform α, the expression in (B43) cannot be further simplified, with however the expenditure share $\pi_i = \frac{\beta_i u^{e_i(1-\alpha)} p_i^{1-\alpha}}{\sum_j \beta_j u^{e_j(1-\alpha)} p_j^{1-\alpha}}$ that follows (B36). Comparing with the standard CDE, the implicitly indirect NHCES eliminates the effects of baseline per capita income.

Explicit Indirect Homothetic CES The only possible way of parameterization to make the CDE utility explicitly derivable (otherwise remains to be implicitly defined) leads to an explicitly indirect constant elasticity of substitution (CES), if allowing $e_i = e$ and $\alpha_i = \alpha \ \forall i$. The same formula for change of utility (B44) can be applied, with expenditure share $\pi_i = \frac{\beta_i p_i^{1-\alpha}}{\sum_j \beta_j p_j^{1-\alpha}}$ that follows (B37). In this case, the change of utility is invariant to effects of both baseline utility and income levels.

It turns out that, by transforming an implicitly indirect CDE to an explicitly indirect homothetic CES, the consumer welfare appears to be *more responsive* to counterfactual income changes, lying in the fact that both effects of baseline utility and income levels are *eliminated*. None of the cases above, however, eliminates the effects of baseline prices on changes of utility, even with the special case of CES where $e = 1$.

B6. Counterfactual Trade Responses (in Percent Changes)

Implicit Indirect Non-Homothetic CDE Similarly, taking total derivatives with respect to implicit utility, price vectors, and per capita income will lead to the following percent change expression:

(B45)
$$\hat{q}_i = e_i(1-\alpha_i)\hat{u} - \alpha_i\hat{p}_i + \alpha_i\hat{w} - \sum_j e_j(1-\alpha_j)\pi_j\hat{u} - \sum_j (1-\alpha_j)\pi_j\hat{p}_j + \sum_j (1-\alpha_j)\pi_j\hat{w}$$
$$= \left[e_i(1-\alpha_i) - \sum_j e_j(1-\alpha_j)\pi_j\right]\hat{u}$$
$$+ \left[\alpha_i + \sum_j (1-\alpha_j)\pi_j\right]\hat{w}$$
$$- \alpha_i\hat{p}_i - \sum_j (1-\alpha_j)\pi_j\hat{p}_j.$$

The analysis of quantity consumption is non-trivial and (B45) can be considered as a special case where there are zero trade costs and f.o.b. prices are normalized to one (if we care more about the counterfactual price changes but not prices at their initial values). The counterfactual result of quantity consumption depends on utility and income changes, as well as price changes of own-goods and all other goods bundle. Meanwhile changes in utility, income and prices are interacted with expansion and substitution parameters, e_i and α_i, respectively with respect to goods i, as well as with share-weighted parameter values of the two with respect to all commodities $\forall i \in I$.

Substituting (B39) into (B45), the change of quantity consumption can also be expressed as a function of income and cross-price elasticities:

(B46) $$\hat{q}_i = \eta_i \hat{w} + \sum_j \sigma_{i,j} \hat{p}_j,$$

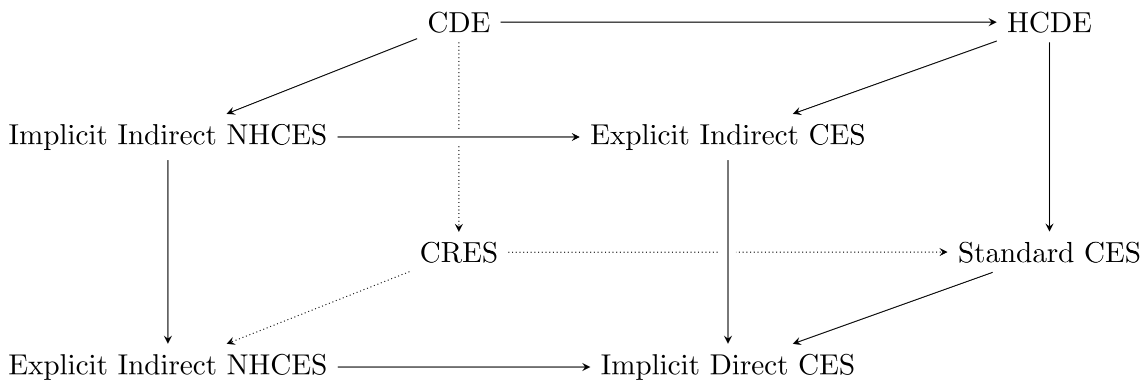
where $\eta_i = \frac{e(1-\alpha)+\sum_k e\pi_k\alpha}{\sum_k e\pi_k} + \alpha - \sum_k \pi_k\alpha$ and $\sigma_{i,j} = \alpha_i + \alpha_j - \sum_k \pi_k\alpha_k - \frac{\Delta_{ij}\alpha_i}{\pi_i}$, which follows (B17) and (B18), respectively.

Explicitly Indirect Homothetic CES Generalized explicit case removes the *effects of utility change* in (B45):

(B47) $$\hat{q}_i = \hat{w} - \alpha\hat{p}_i - (1-\alpha)\sum_j \hat{p}_j,$$

which is identical to the counterfactual result of the standard CES.

B7. Graphical Illustration of Bergson Family Transformation



CDE (Implicitly Indirect) $G(\frac{\mathbf{p}}{w}, u) = \sum_i \beta_i u^{e_i(1-\alpha_i)}(\frac{p_i}{w})^{1-\alpha_i} \equiv 1.$

Homogeneous CDE (Implicitly Indirect): $G(\frac{\mathbf{p}}{w}, u) = \sum_i \beta_i u^{e(1-\alpha_i)}(\frac{p_i}{w})^{1-\alpha_i} \equiv 1.$

NHCES (Implicitly Indirect): $G(\frac{\mathbf{p}}{w}, u) = \sum_i \beta_i u^{e_i(1-\alpha)}(\frac{p_i}{w})^{1-\alpha} \equiv 1.$

CES (Explicitly Indirect): $U = \left[\sum_i \beta_i \left(\frac{p_i}{w}\right)^{1-\alpha}\right]^{\frac{1}{e(\alpha-1)}}.$

NHCES (Implicitly Direct): $F(\mathbf{q}, u) = \sum_i k_i u^{-e_i(1-g)} q_i^{1-g}.$

CRES (Implicitly Direct): $F(\mathbf{q}, u) = \sum_i k_i u^{-e_i(1-g_i)} q_i^{1-g_i}.$

Standard CES (Explicitly Direct): $U = \left[\sum_{i=1}^{N} \beta_i^{\frac{1}{\alpha}} q_i^{\frac{\alpha-1}{\alpha}}\right]^{\frac{\alpha}{\alpha-1}}.$

B8. NHCES Counterfactual Results (in Percent Changes

Applying the counterfactual technique introduced in Section (VIII.A), Table (B1) reports the results using the estimated NHCES model of benchmark year 2011, while imputing backward the predicted percentage differences in aggregate trade flows $\widehat{X}_{il}$ to 1995 using the CEPII macro and distance data. This is used to compare with the observed "changes" in trade flows $\widetilde{X}_{il}$ using the WIOD database. Using the analytical framework described above, an $\widehat{RF}_{(AUS,USA)}$ that equals 0.220 means that, without considering the composite price effects, the uncompensated change in trade flows from Australia to the U.S.—predicted by the structural model—is 22% lower than the actual change observed in the data. The counterfactual decomposition tells us that there is about 22% missing trade that can be attributed by a decrease ("an increase") in the relative Australian factor prices in the U.S. from 1995 to 2011 (from 2011 to 1995). On the other hand, $\widehat{RF}_{(DNK,USA)} = 0.025$ means that the predicted missing trade is about 2.5% that can be explained by an increase in the relative price of Danish factor demand employed in the U.S.

Moving to the implications on Chinese factors employed in the U.S., the structural model for 2011 predicts that the bilateral trade between the U.S. and China would have been higher by a factor of 0.758 in 1995 than the 2011 benchmark; the actual data, however, reveals that the 1995 U.S.–China trade is lower by about a factor of -0.9. It suggests that the gap of trade flows can be largely explained by a surge in real wage of China realized in the U.S. during this period, and the gap is measured by a factor of -1.657 comparing the factor price of China in 2011 with 1995. The first column provides an indicator of the differences in per capita GDP. One can see that the Chinese per capita GDP in 1995 is roughly 90% lower comparing to 2011.

Table B1—Historical Validation: Selected Data and Estimation Results

2011-1995 (x 100 in %)			$\widehat{RF}_{il}$	$\widehat{X}_{il}$	$\widetilde{X}_{il}$	df_{il}	bf_{il}	$\bar{T}_{il}$
$\sigma = \alpha$	ρ		distance: $df_{il} = d_{il}^{\rho}$ border: $bf_{il} = exp(\delta_{il})^{1-dummy_{il}}$					
5.905	0.180		tariff equivalent border tax: $\bar{T}_{il} = exp(\delta_{il}) - 1$					
Origin	$\widehat{E}_l$	$\widehat{L}_l$	USA (x 100 in %)					
AUS	-0.672	-0.191	0.220	-0.951	-0.732	2.648	2.223	1.223
AUT	-0.469	-0.054	0.023	-0.750	-0.726	2.377	2.170	1.170
BAL	-0.854	0.247	-2.382	1.567	-0.815	2.368	1.685	0.685
BGR	-0.796	0.144	-1.982	1.379	-0.603	2.424	1.530	0.530
BLX	-0.488	-0.085	-0.083	-0.411	-0.494	2.337	1.636	0.636
BRA	-0.615	-0.178	-1.309	0.586	-0.723	2.375	1.445	0.445
CAN	-0.633	-0.145	1.065	-1.614	-0.548	1.860	2.167	1.167
CHN	-0.906	-0.104	-1.657	0.758	-0.899	2.518	1.085	0.085
CZE	-0.737	-0.016	0.881	-1.593	-0.712	2.371	2.844	1.844
DNK	-0.468	-0.061	-0.025	-0.574	-0.599	2.338	2.040	1.040
ESP	-0.536	-0.157	0.541	-1.391	-0.850	2.339	2.751	1.751
FIN	-0.543	-0.052	0.149	-0.742	-0.593	2.352	2.439	1.439
FRA	-0.417	-0.089	0.413	-0.774	-0.360	2.340	2.314	1.314
GBR	-0.487	-0.083	0.345	-0.809	-0.464	2.307	1.938	0.938
GRC	-0.511	-0.044	0.490	-1.070	-0.579	2.436	4.055	3.055
HUN	-0.700	0.036	0.687	-1.411	-0.724	2.391	3.930	2.930
IND	-0.745	-0.217	-0.700	-0.195	-0.896	2.592	1.073	0.073
IRL	-0.626	-0.211	-0.495	-0.392	-0.887	2.284	1.490	0.490
ITA	-0.498	-0.043	0.583	-0.979	-0.395	2.385	2.457	1.457
JPN	-0.121	-0.019	0.248	-0.268	-0.020	2.480	1.876	0.876
KOR	-0.569	-0.094	-0.939	0.448	-0.491	2.497	1.115	0.115
LTU	-0.846	0.198	-2.857	1.957	-0.899	2.374	1.354	0.354
MEX	-0.646	-0.201	-1.237	0.472	-0.764	1.918	1.002	0.002
NLD	-0.482	-0.074	-0.034	-0.399	-0.434	2.330	1.611	0.611
POL	-0.751	0.002	-2.246	1.415	-0.830	2.375	1.257	0.257
PRT	-0.492	-0.050	2.002	-2.428	-0.426	2.314	4.906	3.906
ROU	-0.806	0.126	-2.299	1.538	-0.761	2.415	1.498	0.498
RUS	-0.817	0.036	-2.138	1.253	-0.885	2.417	1.482	0.482
ROW	-0.615	-0.249	0.485	-1.163	-0.678	2.440	2.101	1.101
SVK	-0.787	-0.007	-1.358	0.521	-0.836	2.386	1.403	0.403
SVN	-0.568	-0.031	0.470	0.031	-0.439	2.382	1.661	0.661
SWE	-0.527	-0.066	0.328	-0.813	-0.485	2.340	2.117	1.117
TUR	-0.674	-0.199	-1.290	0.721	-0.569	2.454	1.344	0.344
USA	-0.426	-0.145	0.120	-0.613	-0.493	1.822	1.000	0.000