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

Anton Yang

April 2019

**Abstract**

This paper presents a general equilibrium gravity model of trade based on the preferences of the constant difference of elasticities of substitution. Hanoch (1975) illustrates its advantages in terms of parsimony and flexibility. It is the first parsimonious non-homothetic model introduced into the gravity model that both separates substitution effects from income effects and has non-constancy substitution elasticities. These features of the demand model—together with the structural estimation procedure devised in this paper—allow nesting several prominent theoretical motivations (e.g., the standard and non-homothetic CES models) for the gravity model, and exploring the merits of this more general model. They also allow identifying 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.

*Keywords*: Structural estimation; Gravity model; Non-homothetic preferences; Model comparison; Trade costs

*JEL classification*: C51; C52; F10

*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 comments, suggestions and support. This is an incomplete draft. All errors are my own. Correspondence: yang1069@purdue.edu.

# 1 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. It has long been a successful tool in empirical international economics to explain multilateral trade 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 factor demand models. Thus, studying demand systems alone can help us enrich the gravity model.

Many existing standard gravity approaches follow Anderson (1979) who integrates CES (constant elasticity of substitution) consumer preference into the gravity model. These models work well in cross-sectional studies, but meanwhile it entails homothetic preferences and unitary income elasticities. It implies that income shocks in a country will generate proportional changes in national factor demands. Integrating the CES assumption into the world of factor demand and supply (in aggregate trade flows), it means that a 10% increase in relative income will raise the relative demand for national factor supplies also by 10%, with factor expenditure shares being constant. Given that there has been substantial growth in relative incomes in the world, we need to ask whether traditional gravity mechanisms are still appropriate in the international trade literature.

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 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 (e.g., non-homothetic) models of bilateral trade. Another way is to explore out-of-sample predictions, including welfare and trade responses of bilateral trade of income 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 model to the gravity model. The model is both flexible and parsimonious, and incorporates both non-homotheticity and non-constant elasticities of substitution. 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 will be the best fit to the general gravity model.

Another potential contribution of this paper is showing 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 the variation in the population data—together with the structural procedure permitting evaluation of cardinal values of utility—are key to identification.

## 2 Literature Review

Hanoch (1975) discusses advantages and disadvantages across 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;[1] (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 structural demand models in (iii) are still relatively unknown to the readers of the gravity literature.

Some of the NOA or type (i) models have been introduced into gravity models, such as transcendental-logarithmic (translog) adopted in Tan (2013), which is rich in its substitution matrix while yet exhibits homothetic preferences, and Almost-Ideal Demand System (AIDS) developed by Deaton and Muellbauer (2016)—a combination of homothetic and non-homothetic system of equations, which is analog to translog but adds non-homothetic behaviors; AIDS is well-known in a variety of economic literature and has been introduced into gravity models such as Fajgelbaum and Khandelwal (2016). However, as is commonly known, AIDS collapses (precisely, the expenditure shares can fall out of regular "zero to one" range) when real income changes become significantly large (Rimmer and Powell, 1992). Furthermore, these models work well (and are less restricted) for small number of goods but for large $n$ goods NOA implies the number of estimated parameters becomes square of $n$, which will become less tractable (Hanoch, 1975).

One way to reduce the number of parameters is to impose separability assumption (Lau, 1969) or equivalently say—explicit or implicit additivity. In general, the utility function is implicitly additive if the cardinal values of utility cannot be algebraically solved in terms of the model's exogenous variables, whereas the utility expressed in explicit additivity—which is a special case of implicit additivity—may be explicitly derived.[2]

The standard homothetic CES demand in gravity is one class of explicitly additive models in Hanoch (1975) or type (ii), which is parsimonious and has obviously achieved the parameter reduction. These demand models, however, whether homothetic or non-homothetic, and whether directly or indirectly additive, impose restrictions that tie the substitution effects with income effects.[3] In gravity, this restrictive assumption implies that the substitutabilities

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

[2] 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.

[3] The key distinction between "direct" and "indirect" is that the indifference curve of the former (latter) is additive (separable) in consumption quantities (prices and incomes, or "unit-cost prices"). The dependence of substitution effects on income effects in the two cases of explicit models are different in their income-

among all factors are heavily dependent of income changes, invariant to data consistencies.

There have been a couple of non-homothetic explicitly additive models applied in gravity. 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.[4] 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 generates 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 affected quantities of consumer goods.

However, these explicit models do not break the tie between substitution effects and income effects (Houthakker, 1960; Hanoch, 1975). This restriction suggests that high-income countries, which typically 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 it prefers imported factors from one country far more than another. This dependence on income effects also implies that the relative substitution preferences of imported factors between countries A and B with respect to factors imported from country C are only determined by income elasticities with respect to factors imported from A and B, invariant to the vector of third country's factors.

One way to remove this restriction (to allow a more flexible substitution matrix) is to replace explicit additivity with implicit additivity assumption, or type (iii). Comin, Lashkari and Mestieri (2015)'s non-homothetic CES (NHCES) and Yilmazkuday (2018)'s specialized NHCES are the only gravity literature that adopt implicit additivity. Nevertheless, these models have constant elasticities of substitution as the standard CES models. While they separate the factor substitutabilities from income effects, they imply constant substitutabilities across all factor demands in the world.

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 imposed on substitution effects that is dependent on income effects (see Table 1). The preferences that will be discussed in Section (3) is the constant differences of elasticities of substitution (CDE) introduced in Hanoch (1975), which is non-homothetic, implicitly indirectly additive and non-CES.

substitution relationship–being multiplicative (direct) or additive (indirect) (Hanoch, 1975).

[4] 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.

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.

# 3 Theory

I first present a more general consumer demand model introduced in Hanoch (1975), which is a class of *implicitly indirect additivity*. I show how the more general demand model can nest a standard CES model as well as an NHCES model. Next, I show how these nested demand models can be linked to the 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 (with no trade costs). Finally, I fully integrate the more general model into the general equilibrium (GE) gravity model.

## 3.1 An Implicitly Indirect Additive Demand Model

The demand system adopted in this paper is Constant Difference of Elasticities (CDE) introduced in 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 $log(u^{e_i} \frac{p_i}{E})$ replacing $u^{e_i(1-\alpha_i)}(\frac{p_i}{E})^{1-\alpha_i}$ for $\alpha_i = 1$, 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 $i$; $\beta_i, e_i, \alpha_i$ 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) $\alpha_i > 0$, and 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 by the model's exogenous variables—the price vector and income. The model is indirect (rather than direct) is 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 $\mathcal{I}$ 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 $i$. Since $U$ is not observable, we cannot directly estimate the demand in a reduced form.[5] It can, however, be estimated structurally, which will be discussed later. Also note that the demand function shares a common structure with other gravity models. Here, the denominator plays a similar role as the multilateral resistance (MLR) in Anderson and van Wincoop (2003).

Subsequently, the Allen-Uzawa elasticities of substitution (ES) $\sigma_{i,j}$ can be derived

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

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 CDE 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-one-price* elasticities; (ii) $\sigma_{ij}$ can be negative if the ES is dominated by large substitution parameters $\alpha_k$, weighted by expenditure shares of $k$ (or likewise, if the summation of $\alpha_i$ and $\alpha_j$ is sufficiently small)—implying *possibilities of factor complementarity*—which is a cogent yet missing theoretical foundation in standard gravity.

It can be also shown that the income elasticities $\eta_i$ 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 $\alpha_k$, while the expansion parameters $e_i$ controls 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 weighted sum of the preferences given to all other goods—implying *possibilities of inferior goods.*

[5]It can be shown that $U$ can be eliminated using the method of double log-differencing. In this case, only the substitution parameters $\alpha$'s can be identified, while the separability constraint cannot be guaranteed.

### 3.1.1 Breaking Linkages between Substitution and Income Effects

Another key property of CDE—and other demand models under 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 is a function of income elasticities, distinguished by multiplicative form of income elasticities in the explicitly direct case (e.g., standard CES) and additive form in the explicitly indirect case (Hanoch, 1975)

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

Comparing with Equation (3), while the substitution matrix of 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 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 the standard CES-gravity models (e.g., explicitly direct case), it implies that $\eta_i = \eta_j = 1$ and $\sigma = \alpha$, where $\sigma$ 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 country buys goods from another high-income country, where the observable trade frictions among them are typically expected to be low. In the case of explicit non-homothetic models, it implies that the substitutabilities among factors in the vector of countries are fixed by income effects, regardless of whether that country strictly prefers to import factors from another country 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.

### 3.1.2 True-Cost-of-Living Index

Denoting by $PU$ the *true-cost-of-living index* in the sense of money metric utility (equalling the cost of per unit of utility). It can be derived by finding the total differential of the implicitly defined $G$ function in Equation (1) as implemented in Chen (2017)

$$PU = \Big[ \sum_i \beta_i u^{e_i - e_i\alpha_i - 1}(1-\alpha_i) p_i^{1-\alpha_i} E^{\alpha_i - 1} e_i \Big] \Big[ \sum_i \beta_i u^{e_j(1-\alpha_j)}(1-\alpha_j) p_j^{1-\alpha_j} E^{\alpha_j - 2} \Big]^{-1}. \quad (6)$$

Note that this is a Samuelsonian price index that gives us information about how much dollar units one needs to possess in order to earn 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 utilities consumers attain. Basic microfoundations provide that, for any explicitly direct 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 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.

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). In the pure goods-exchange economy that has one representative consumer, I show that $PU$ takes the following reduced form

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

The first term in 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). It 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 $\vartheta$ 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: e.g., $P \equiv E/u \equiv PU$, where $P$ is the average price index of goods as defined in Dixit and Stiglitz (1977).

### 3.2 Generalization to the standard CES Demand

It is important to highlight that 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 1** *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*

$$P \equiv E/u \equiv PU. \tag{8}$$

If $e = 1$ and $\alpha_i = \alpha \; \forall \; i$, then the CDE demand will converge to a standard CES demand.

*Proof.* See Appendix A.

## 3.3 Generalization to Non-Homothetic CES (NHCES)

In this section, I show that the CDE demand nests an implicitly indirect NHCES system (that is identical to its direct form). It can be mapped to the version of Comin, Lashkari and Mestieri (2015), provided that their specific NHCES aggregator $C$ corresponds to the standard real consumption index $Q \equiv U$. Linking to Comin, Lashkari and Mestieri (2015) is an important procedure. It 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 implicitly indirect NHCES demand [6]

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

which is identical to the implicitly direct NHCES

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

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

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

*Proof.* See Appendices B and C.

[6] 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 (2015).

#### 3.3.1 Demand Mapping to Comin *et al.*

I adopt Comin, Lashkari and Mestieri (2015)'s notations 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 Comin, Lashkari and Mestieri (2015)'s 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 (2015) 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 $\alpha$. 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 $\epsilon_i$ through the direct channel $\epsilon_i = e_i(1-\alpha)$, and is decreasing in $\epsilon_i$ 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 $\alpha$ ($= \sigma$), but is indirectly decreasing in $\alpha$ 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 $\omega_k$ 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 $\epsilon_i$ can be transformed from one to another, it turns out that $\epsilon$ and $\sigma$ play equivalent roles in Equation (14) as $e$ and $\alpha$ in (13).

## 3.4 A Hicksian Import Demand

It is crucial to introduce the *Hicksian approach* in the 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 do not exist. Instead of estimating the Marshallian demand in the gravity model, the Hicksian approach targets a *Hicksian demand* of 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.[7] It is shown that in (3.2) 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 (2015). I first show that we can estimate a Hicksian import demand in 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 (2015), 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.

### 3.4.1 Application in the NHCES

Comin, Lashkari and Mestieri (2015) constructs an average price index of NHCES that equals real income, that is, the per capita income adjusted by real consumption index $C$, e.g., $P \equiv \frac{E}{C}$. The implicit demand yields an expenditure function (see Footnote 6) 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 $\omega_i$ defined as the expenditure share of $i$, and $\frac{1}{e_i} = \frac{1-\sigma_i}{\epsilon_i}$ equaling $\chi_i$ specified in Comin, Lashkari and Mestieri (2015)'s paper.

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

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 of per unit of cardinal utility) specialized to the standard implicitly 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

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

The expenditure share in Comin, Lashkari and Mestieri (2015) is equivalent to

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

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 $\epsilon_i \neq 1-\sigma$.[8] 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 $\vartheta$.

If a choice bundle is dominated by sufficiently many luxury (subsistence) goods $i$'s, indicated by higher (lower) expansion elasticity $e_i$'s 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 $\vartheta$ the elasticity of 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 $\epsilon_i = 1-\sigma \ \forall \ i$.

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 $PU$

[8] 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 $\sum_i e_i \omega_i = 1$.

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

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.

### 3.4.2 Application in the CDE

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

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

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 (5); 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 5.2). 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) will converge 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) will converge to (A.8), 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.

# 4 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 as implemented in Balistreri and Hillberry (2007) for 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 James Markusen and Thomas Rutherford (see also Footnote 7).

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 (2015). 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 to allow the population data (in addition to the income data) and the MPEC program to be critically useful in identifying the parameters of import demand and trade cost elasticity (as well as asymmetric border frictions).

The gravity model in this paper is specialized to the global factor demand. The starting point of the model setups is a transition of the per capita household demand into imported factor demand under gravity in a GE framework. Following 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 country trades aggregate "factors" with other countries, with the value of factors equaling the value of trade flows in aggregate term. Henceforth, "goods" and "factors" may be 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 country observes goods to be potentially exported from the origin associated with an import demand in the destination, including goods produced at home.

## 4.1 Preliminaries

I denote $i = \{1, ..., \mathcal{I}\}$ and $l = \{1, ..., i-1, i+1, ..., \mathcal{I}+1\}$ as country 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 $\tau_{il} = d_{il}^{\rho}$ be the iceberg trade costs between them, where $\rho$ 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 triangle inequality is satisfied for any trilateral relationships.[9] 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 (4.7), I allow border costs to impact $\tau_{il}$ so the trade costs are composed of both distance and border effects.

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

## 4.2 Model Formulation

There is a transition of CDE per capita household demand into gravitational import demand. The per capita income is linked by national population and the aggregate income, e.g., $E_l = Y_l/L_l$. The introduction of population terms provides additional variation that helps with identification of structural demand parameters. The CDE-gravity demand variable (e.g., $q_{il} \equiv h_{il}$) has a partial elasticity with respect to population of $-\alpha_i$[10]

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

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

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

## 4.3 General Equilibrium

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 $Y_i = \sum_l X_{il}$. The 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 melts, 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 $i$

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

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 it and prices yields the definition of aggregate income in each region $i$

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

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

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

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

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 $L_i$. 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 onto 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 $\phi_i$. It is consistent with Anderson and van Wincoop (2003) who introduce a free intercept in their empirical model. The purpose of this productivity parameter is to allow the factor income (average wages) to be varied with the GDPs observed from the data, so one unit of produced goods in each origin $i$ requires $\phi_i$ unit of labor, satisfying

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

Under fixed employments, total population must equal the total amount of labor hired for production that meet total demand. 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-country differences* in the gap between average wages and prices in this non-homothetic demand system, which is captured by $\phi_i$.

## 4.4 Identification of Demand Parameters and the Elasticity of Trade Costs with respect to Distance

The expression of 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

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

The bilateral trade flows is 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. The expression inside the square bracket corresponds to the MLR as the average trade barrier in Anderson and van Wincoop (2003).

The structural procedure can identify the elasticity of trade costs with respect to distance $\rho$ from the elasticity of trade with respect to trade costs $\tau_{il}$'s. The variation in the population data is key to identification. Another equally important identification strategy is that the estimation procedure allows solving for cardinal values of $u$.

The key point is that, if we can control for the MLR, the variables $Y$ and $L$ can help pin down $\alpha_i$'s. With the identification of $\alpha_i$'s, the distance, e.g., $d_{il}$'s, can pin down parameter $\rho$. The estimation of $e_i$'s, however, is impossible without evaluation of $u_l$, which cannot be accomplished in conventional methods (e.g., reduced-form regressions). For this reason, I employ MPEC which allows jointly calculating MLR and cardinal values of $u_l$'s (given equilibrium constraints and implicit utility defining equation), even as vectors of $\alpha_i$, $e_i$ and $\tau_{il}$ 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)

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

## 4.5 Parameter Restrictions

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 $\alpha_i$ must be greater than 1 to reflect distance impeding trade flows, whereas $0 < \alpha_i < 1$ is irrational since it would imply positive coefficient on logarithmic distances. However, for estimation purposes, I do not impose any additional restrictions on $\alpha_i$'s except they have to be greater than zero in order to satisfy the properties of a demand function.

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).[11] All other parametric restrictions in the CDE demand system are standard following Hanoch (1975).[12]

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

(i) When $lim_{\alpha_i \to 1}$ or $lim_{\beta_i \to 0}$, zero trade flows may exist between $i$ and $l \ \forall \ l$.

(ii) When $lim_{e_i \to \infty} \ \forall \ \alpha_i \in [1, \infty]$ or $lim_{\alpha_i \to \infty}$, zero trade flows may exist between $i$ and $l \ \forall \ l$.

(iii) When $lim_{\rho \to \infty} \ \forall \ \alpha_i \in [1, \infty]$, zero trade flows may exist between $i$ and $l \ \forall \ l$.

(iv) When $lim_{d_{i,l} \to \infty} \ \forall \ \alpha_i \in (1, \infty] \cap \beta_i \neq 0$, while either condition under (ii) is not satisfied and holding $\rho$ 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 model and estimation will fall apart, invariant to exogenous variables. Case (ii) holds when $\alpha_i$ is infinitesimally large even if $e_i = 0$, as long as there is a trade friction between $i$ and $l$, whereas if $\tau_{i,l} = 1$ (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 $\rho$ approaches to 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, infinitesimally 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.

[11] There are counter cases where evidences show that standard gravity assumption on distance does not hold. Jacques Melitz (European Economic Review, 2007) shows that under North-South bilateral pairs there are proportional increases in trade with distance. Buch, Kleinert and Toubal (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.

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

### 4.7 Asymmetric Border Effects

I now allow bilateral border charges to be included in $\tau_{il}$. Balistreri and Hillberry (2007) propose an estimation framework that measures bilateral border puzzles 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 ($\delta_{il} = \delta_l \; \forall \; i$, $\bar{T}_{il} = \bar{T}_l$ $\forall \; i$); $\text{dummy}_{il}$ 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 $\delta_{il}$, which is linearised from (32) in the transformed estimation equation specialized to (30)

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

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

## 5 Estimation Methods

There are various concerns over standard OLS-based estimations of the linearised gravity. The structural estimation procedure minimizes the least squares using observed trade flow values or import shares in the objective function to predict fitted ones. Some well-known concerns include heteroskedasticity, logarithmic transformation issues, etc. Another important issue is the zero-trade flows which is often a fact in the observed data. Using a conventional log-linear estimator may involve in 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., 2017). Following Gourieroux, Monfort and Trognon (1984) and Silva and Tenreyro (2006), the PPML estimator based on the log-linear expression of bilateral trade flows in Equation (16) is given by

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

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 $\rho$ and $\delta_{il}$), respectively; and $\Upsilon_{il}$ is the new error term.

The constrained optimization problem is to maximize (35)

$$L(b_i) = constant - \sum_i \sum_l exp(x_{il} b_i) + \sum_i \sum_l y_{il} x_{il} b_i \qquad \text{(Objective Function)}, \tag{35}$$

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) $\delta_{il} \geq 0$.

## 5.1 Computation Challenges and Algorithm

The estimation of CDE demand system alone is a problem that is difficult enough to handle. Due to its implicit properties, it is ideal for computation 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 by many trade economist. For the past 30 years, it uses the CDE functional form as preferences of private household 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: $\alpha_i, \beta_i, e_i, u_i, \phi_i$, $\delta_{il}$ and $\rho$.

One 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 eliminates the MLR term, while the second log-difference eliminates utilities. The technique using the first log-difference has been employed by economists who apply CDE as a production function, where $u$ is realized as observable production outputs pulled from the database (e.g., Surry, 1993). There are two major issues for applying this method: (i) not all structural demand parameters can be identified from the transformed equations (only $\alpha_i$ 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 $\rho$ and the border coefficients $\delta_{il}$.

Therefore, 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 used for this model.

#### 5.1.1 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, 1998). The MPEC has been successfully applied to numerous trade gravity literature (Balistreri and Hillberry, 2007; Balistreri, Hillberry and Rutherford, 2011; Tan, 2013). 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 notably burdensome as the trade cost elasticity $\rho$ enters the implicit additivity defining equation. The final prices at home, which is the value of goods adjusted by trade costs, is endogenous because of the unknown elasticity with respect to distance $\rho$. Moreover, as we allow border effects to impact $\tau_{il}$, the program then also needs to determine the asymmetric border coefficients that are represented by $\delta_{il}$. On the positive side, there is also data on population to help evaluate $\alpha_i$, as well as GE system of constraints—Equations (27) and (36)-(39)—to jointly evaluate $u_l$. It is worth also 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.

## 5.2 Inequality Constraints as an MCP

**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

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

and

$$L_i \geq \phi_i K_i^0 \quad \perp \quad PL_i = \phi_i \frac{Y_i}{L_i} \geq 0. \tag{37}$$

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

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

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, 1998). 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 another way of saying, 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).

## 5.3 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

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

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(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 (1). 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 (3.4.1). The final step is to release all parameter values by eliminating both CDE2NHCES and CDE2HCDE as constraints, while fully estimating the CDE-gravity model.

## 5.4 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 spaces. First, it is intuitive to see that any positive constant scaling of the distribution parameters $\beta_i$'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 have

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

which is equivalent to

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

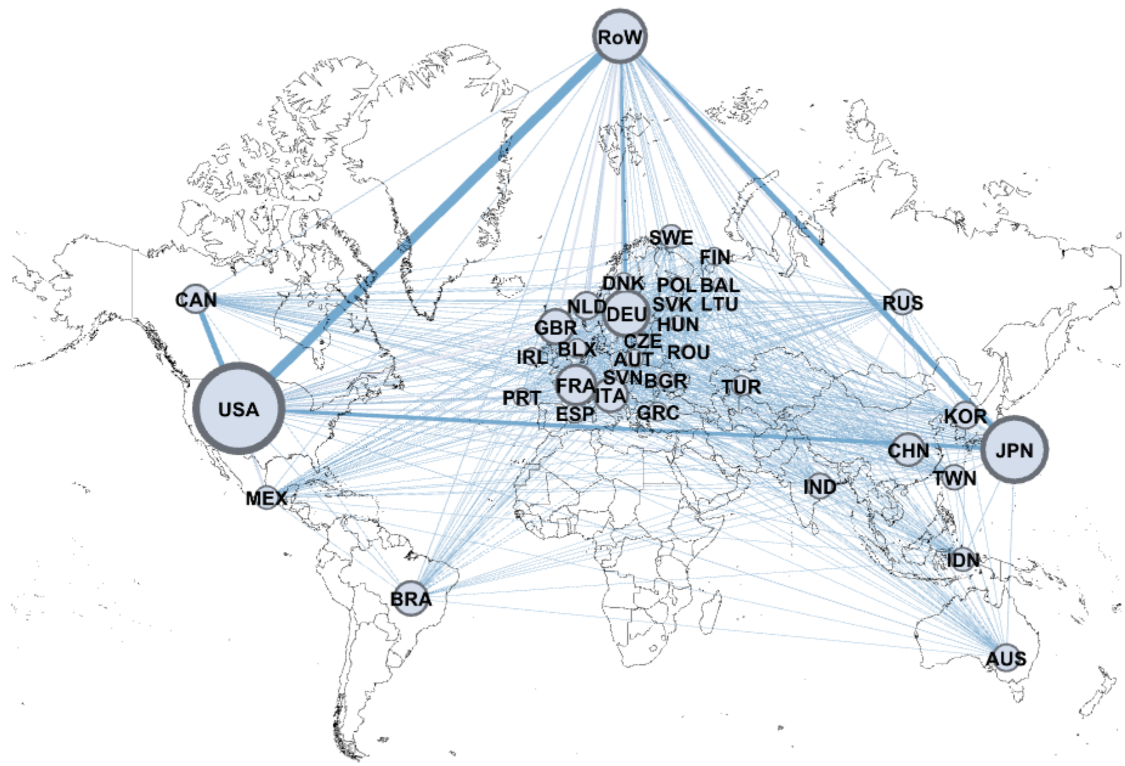
where $\beta_{i1} = \beta_i \gamma^{\mu f_i (1-\alpha)} = \beta_i \gamma^{e_i(1-\alpha)}$, $V_2 = V_1^{\mu} = \left(\frac{u}{\gamma}\right)^{\mu}$. It is easy to check that both $\beta_i$ and $u$ have been absorbed twice by the arbitrary choices of the two scalars, transforming into an identical demand 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 $\beta_i$ parameters and utilities values. Without removing the spare 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 testings.

Similar to Comin, Lashkari and Mestieri (2015), I choose base factors ('goods'), $b$ and $b'$, such that $\beta_b = e_{b'} = 1$. In addition, since the scaling of $u$ is not identifiable from the data, I normalize the cardinal utility for one base country, e.g., $u_{b''} =$ constant. These procedures essentially normalize $\mu$ as an exponent on $u$ for the base factor, while controlling that the *spillover* of $\mu$ does not generate multiple linear possibilities of $e_i(1-\alpha_i)$ on the base of $\gamma$. Furthermore, the choices and values of the numeraire are chosen such that they are innocuous to identification and served to improve the scaling of other structural parameters.

# 6 Data and Preliminary Estimates

The model is estimated using the 1995 observations for 37 countries/regions in the world. The population and distance data are taken from the CEPII gravity database. The data on trade flows is taken from the World Input-Output Database (WIOD). The national GDP data for each country 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 (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.

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




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 (5.3). The estimation program is the General Algebraic Modeling System (GAMS) version 25.1.3, which is solved using the MPEC-NLPEC (non-linear programming with equilibrium constraints) solver with the help of the preprocessor using GAMS-F tool.[13] Table (2) reports preliminary estimates for 37 countries with both distance and asymmetric border effects on bilateral trade costs.

# 7 Out-of-Sample Predictions and Missing Trade

The analytical import demand framework provides us an 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

[13]See `http://www.mpsge.org/inclib/gams-f.htm` (Michael Ferris, Thomas Rutherford, and Collin Starkweather). The tool greatly reduces the computation complexity in calculating the trade costs.

Table 2: Preliminary Estimates of the Gravity Models

| | CDE | | | | NHCES | | | | CES | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PPML (w/o border charges) | 1131.28 | | | | 1097.86 | | | | 1074.71 | | | |
| PPML (with border charges) | 1164.38 | | | | 1163.43 | | | | 1162.88 | | | |
| $\rho$ | 0.269 | | | | 0.129 | | | | 0.113 | | | |
| | $\alpha$ | $\beta$ | $e$ | $\delta$ | $\alpha$ | $\beta$ | $e$ | $\delta$ | $\alpha$ | $\beta$ | $e$ | $\delta$ |
| AUS | 3.24 | 0.08 | 1.06 | 2.05 | 7.85 | 0.41 | 1.01 | 0.57 | 8.78 | 0.45 | 1.00 | 0.50 |
| AUT | 5.07 | 0.20 | 1.03 | 0.92 | | 0.12 | 1.02 | 0.59 | | 0.12 | | 0.51 |
| BAL | 7.87 | 0.27 | 1.04 | 0.52 | | 0.01 | 1.03 | 0.89 | | 0.01 | | 0.83 |
| BGR | 6.62 | 0.13 | 1.06 | 0.62 | | 0.01 | 1.06 | 0.78 | | 0.01 | | 0.76 |
| BLX | 4.46 | 0.15 | 0.95 | 0.71 | | 0.15 | 0.97 | 0.36 | | 0.23 | | 0.34 |
| BRA | 4.35 | 0.24 | 0.99 | 1.56 | | 0.26 | 1.00 | 0.77 | | 0.30 | | 0.67 |
| CAN | 4.53 | 0.37 | 0.93 | 1.09 | | 0.31 | 0.95 | 0.56 | | 0.56 | | 0.49 |
| CHN | 2.67 | 0.06 | 0.94 | 3.18 | | 0.40 | 0.99 | 0.71 | | 0.51 | | 0.57 |
| CZE | 7.51 | 0.68 | 1.05 | 0.45 | | 0.04 | 1.02 | 0.72 | | 0.04 | | 0.64 |
| DEU | 4.53 | 1.00* | 1.03 | 0.79 | | 1.00* | 1.02 | 0.42 | | 1.00* | | 0.36 |
| DNK | 4.57 | 0.11 | 1.05 | 1.07 | | 0.10 | 1.03 | 0.58 | | 0.10 | | 0.49 |
| ESP | 6.32 | 1.77 | 1.01 | 0.65 | | 0.24 | 1.00 | 0.65 | | 0.27 | | 0.58 |
| FIN | 5.86 | 0.38 | 1.03 | 0.63 | | 0.09 | 1.02 | 0.61 | | 0.10 | | 0.53 |
| FRA | 4.58 | 0.63 | 1.03 | 1.02 | | 0.58 | 1.02 | 0.55 | | 0.59 | | 0.48 |
| GBR | 3.95 | 0.30 | 1.01 | 1.08 | | 0.53 | 1.02 | 0.44 | | 0.55 | | 0.38 |
| GRC | 5.43 | 0.12 | 1.01 | 0.99 | | 0.04 | 1.01 | 0.78 | | 0.05 | | 0.69 |
| HUN | 5.44 | 0.08 | 1.03 | 0.96 | | 0.03 | 1.02 | 0.73 | | 0.03 | | 0.65 |
| IDN | 2.92 | 0.03 | 0.93 | 2.89 | | 0.17 | 0.96 | 0.71 | | 0.26 | | 0.52 |
| IND | 4.70 | 0.17 | 0.92 | 1.73 | | 0.12 | 0.96 | 0.98 | | 0.18 | | 0.71 |
| IRL | 4.59 | 0.06 | 0.96 | 0.81 | | 0.05 | 0.97 | 0.46 | | 0.08 | | 0.41 |
| ITA | 4.52 | 0.50 | 1.02 | 1.01 | | 0.50 | 1.01 | 0.53 | | 0.52 | | 0.47 |
| JPN | 3.67 | 0.62 | 1.00 | 1.15 | | 1.59 | 1.01 | 0.43 | | 1.73 | | 0.37 |
| KOR | 2.64 | 0.06 | 1.06 | 2.31 | | 0.39 | 1.03 | 0.41 | | 0.38 | | 0.36 |
| LTU | 7.05 | 0.09 | 1.05 | 0.59 | | 0.01 | 1.04 | 0.85 | | 0.01 | | 0.80 |
| MEX | 5.75 | 0.82 | 0.97 | 0.74 | | 0.16 | 0.97 | 0.69 | | 0.25 | | 0.55 |
| NLD | 4.38 | 0.21 | 1.01 | 0.70 | | 0.25 | 1.01 | 0.34 | | 0.27 | | 0.30 |
| POL | 6.55 | 0.34 | 1.01 | 0.82 | | 0.05 | 1.00 | 0.84 | | 0.06 | | 0.72 |
| PRT | 5.70 | 0.19 | 0.95 | 0.75 | | 0.05 | 0.96 | 0.69 | | 0.08 | | 0.59 |
| ROU | 5.76 | 0.09 | 1.04 | 1.07 | | 0.02 | 1.03 | 0.88 | | 0.02 | | 0.80 |
| RUS | 7.81 | 10.95 | 1.00 | 0.48 | | 0.20 | 1.01 | 0.75 | | 0.22 | | 0.67 |
| ROW | 2.63 | 0.51 | 1.00 | 1.29 | | 4.92 | 1.01 | 0.12 | | 5.38 | | 0.11 |
| SVK | 8.34 | 0.52 | 1.03 | 0.37 | | 0.02 | 1.04 | 0.78 | | 0.02 | | 0.71 |
| SVN | 5.25 | 0.03 | 1.04 | 1.00 | | 0.02 | 1.02 | 0.72 | | 0.02 | | 0.64 |
| SWE | 4.68 | 0.20 | 1.02 | 1.02 | | 0.17 | 1.01 | 0.57 | | 0.18 | | 0.50 |
| TUR | 5.95 | 0.36 | 0.96 | 0.98 | | 0.07 | 0.98 | 0.86 | | 0.09 | | 0.72 |
| TWN | 2.27 | 0.04 | 1.02 | 2.91 | | 0.37 | 1.00 | 0.24 | | 0.41 | | 0.22 |
| USA | 4.77 | 5.39 | 1.00* | 0.69 | | 3.71 | 1.00* | 0.40 | | 4.20 | | 0.36 |

*Normalized factors/countries: $\beta_{DEU} = e_{USA} = 1$, and $u_{JPN} = 10$.

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.

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

$$\begin{aligned}\widehat{q}_{il} = &\Big[e_i(1-\alpha_i) - \sum_j e_j(1-\alpha_j)\omega_{jl}\Big]\widehat{u}_l \\ &+ \Big[\alpha_i + \sum_j (1-\alpha_j)\omega_{jl}\Big]\widehat{E}_l \\ &- \alpha_i \hat{p}_{il} - \sum_j (1-\alpha_j)\omega_{jl}\widehat{p}_{jl},\end{aligned} \tag{42}$$

*Derivation.* See Appendix D.

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

$$\widehat{X}_{il} = \widehat{q}_{il} + \widehat{L}_l + \widehat{FOB}_i + \widehat{\tau}_{il} \tag{43}$$

Combining the two hat equations above, it is intuitive to see that if we can measure the change in wealth (utilities) in identity what follows, 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)

$$\widehat{u}_l\Big[\widehat{w}_l, \widehat{\mathbf{FOB}}_i; \bar{\mathbf{b}}_l, \bar{\delta}_{il}\Big]_{(i,l)\in\mathcal{I}}^{\bar{\mathbf{b}}_l=\{\bar{\alpha}_l,\bar{\beta}_l,\bar{e}_l,\bar{\rho}\}} = \frac{u_l^1 - u_l^0}{u_l^0}, \tag{44}$$

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

$$\widehat{u}_l = \frac{\widehat{w}_l - \sum_i \omega_{il}^0 \widehat{p}_{il}}{\sum_i e_i \omega_{il}^0} = \frac{\widehat{w}_l - \sum_i \omega_{il}^0 \widehat{p}_{il}}{\vartheta_l|_{\omega_{il}^0}}, \tag{45}$$

*Derivation.* See Appendix E.

where $\widehat{u}_l$ is the counterfactual change in welfare response in percentage terms; $\vartheta_l|_{\omega^0_{il}}$ is the aggregate elasticity with respect to utility evaluated at benchmark import shares. Using (42)-(45), 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 (43) 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

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

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.

# 8 Likelihood Ratio Tests

The estimation procedure simultaneously produces the value of objectives, 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, 2011, Sections 14.9.3.c and 18.4.1)

$$\Lambda = 2[\log L_{CDE}(\mathbf{b}_i) - \log L_{CES}(\mathbf{b}_i)], \tag{46}$$

where $\Lambda$ is the log-likelihood ratio.

The hypothesis test then starts with the following assumptions:

**Hypothesis $\mathbf{H_0}$** (test hypothesis): *the less restricted CDE is more consistent with the data.*[14]

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

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

In statistical expressions

$$\begin{aligned} &H_0 : \Lambda > c, && \text{for all } \alpha \in (0, \infty] && \rightarrow \text{do not reject } H_0 \\ &H_a : \Lambda < c, && \text{for some } \alpha \neq 1 && \rightarrow \text{reject } H_0 \end{aligned}$$

for all $\beta, e \geq 0$ and for all $u, \rho > 0$, where $c$ is a specified cutoff value that governs significance level $\varphi$, satisfying: $P\{\Lambda < c|H_0\} + qP\{\Lambda = c|H_0\} = \varphi$, with chosen probability $q$ to reject $H_0$ if $\Lambda = \{c|H_0\}$. The condition under which for some $\alpha \neq 1$ is equivalent as requiring $0 < \alpha_i < 1$ and $\alpha_i > 1$ for some $i \in \mathcal{I}$, which is consistent with Hanoch (1975). In the hypothesis testing where $\Lambda < \{c|H_a; \alpha_i = 1\}\ \forall\ i \in \mathcal{I}$ implies the objective value is obtained from a Cobb-Douglas function, rather than CES that is a more general case.

# References

[1] Adao, Rodrigo, Arnaud Costinot, and Dave Donaldson. Nonparametric counterfactual predictions in neoclassical models of international trade., *American Economic Review*, 107.3: 633-89, 2017.

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

[3] Anderson, James E. The gravity model., *Annu. Rev. Econ.*, 3.1: 133-160, 2011.

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

[5] Balistreri, Edward J. and Russell H. Hillberry. Trade frictions and welfare in the gravity model: how much of the iceberg melts?, *Canadian Journal of International Economics*, 39.1: 247-265, 2006.

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

[7] Balistreri, Edward J., Russell H. Hillberry, and Thomas F. Rutherford. Structural estimation and solution of international trade models with heterogeneous firms., *Journal of International Economics*, 83.2: 95-108, 2011.

[8] Barnett, William A. Definitions of 'second order approximation'and of 'flexible functional form'., *Economics Letters*, 12.1: 31-35, 1983.

[9] Bertoletti, Paolo, Federico Etro, and Ina Simonovska. International trade with indirect additivity., *American Economic Journal: Microeconomics*, 10.2: 1-57, 2018.

[10] Buch, C. M., Kleinert, J., Toubal, F. The distance puzzle: on the interpretation of the distance coefficient in gravity equations., *Economics Letters*, 83(3), 293-298, 2004.

[11] Caron, Justin, Thibault Fally, and James R. Markusen. International trade puzzles: A solution linking production and preferences., *The Quarterly Journal of Economics*, 129.3: 1501-1552, 2014.

[12] Chen, Y-H. Henry. The Calibration and Performance of a Non-homothetic CDE Demand System for CGE Models., *Journal of Global Economic Analysis*, 2.1: 166-214, 2017.

[13] Christensen, L.R. Transcendental logarithmic utility functions., *American Economic Review*, 65: 267-383, 1975.

[14] Christensen, L.R., D.W. Jorgenson and L.J. Lau, Transcendental logarithmic production frontiers., *Review of Economics and Statistics*, 55, 28-45, 1973.

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

[16] Costinot, Arnaud, and Andrés Rodríguez-Clare. The uS Gains From Trade: Valuation using the Demand for Foreign Factor Services., *Journal of Economic Perspectives*, 32.2: 3-24, 2018.

[17] Deaton, Angus, and John Muellbauer. An almost ideal demand system., *American Economic Review* , 70.3: 312-326, 1980.

[18] Diewert, W.E. An application of the Shephard duality theorem: A generalized Leontief production function., *Journal of Political Economy*, 79: 461-507, 1971.

[19] Dirkse, Steven P., and Michael C. Ferris. Modeling and solution environments for MPEC: GAMS & MATLAB., *Reformulation: nonsmooth, piecewise smooth, semismooth and smoothing methods*, Springer, Boston, MA, 127-147, 1998.

[20] Dixit, Avinash K., and Joseph E. Stiglitz. Monopolistic competition and optimum product diversity., *American Economic Review*, 67.3: 297-308, 1977.

[21] Eaton, Jonathan, and Samuel Kortum. Technology, geography, and trade. , *Econometrica*, 70.5: 1741-1779, 2002.

[22] Fajgelbaum, Pablo D., and Amit K. Khandelwal. Measuring the unequal gains from trade., *The Quarterly Journal of Economics*, 31.3: 1113-1180, 2016.

[23] Fieler, Ana Cecilia. Nonhomotheticity and bilateral trade: Evidence and a quantitative explanation., *Econometrica*, 79.4: 1069-1101, 2011.

[24] Gourieroux, Christian, Alain Monfort, and Alain Trognon. Pseudo maximum likelihood methods: Theory., *Econometrica: Journal of the Econometric Society*, 681-700, 1984.

[25] Greene, W. H. *Econometric Analysis,* Harlow: Pearson, 7th edition.

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

[27] Houthakker, Hendrik S. Additive preferences., *Journal of the Econometric Society*, 244-257, 1960.

[28] Lau, L. J. Duality and the Structure of utility Functions., *Journal of Economic Theory*, 1: 374-396, 1969.

[29] Lau, L. J. in *Frontiers of Quantitative Economics, Volume II.*, M. D. Intriligator and D. A. Kendrick (eds.), Amsterdam: North-Holland, 176-199, 1974.

[30] Larch, Mario, Joschka Wanner, Yoto Yotov, and Thomas Zylkin. The currency union effect: A PPML re-assessment with high-dimensional fixed effects., 2017.

[31] Melitz, Marc J. The impact of trade on intra-industry reallocations and aggregate industry productivity., *Econometrica*, 71.6: 1695-1725, 2003.

[32] Melitz, Jacques. North, South and distance in the gravity model., *European Economic Review*, 51.4: 971-991, 2007.

[33] Mukerji, V. V.: "Generalized SMAC Function with Constant Ratios of Elasticities of Substitution., *Review of Economic Studies*, 30: 233-236, 1963.

[34] Rimmer, Maureen T., and Alan A. Powell. An implicitly directly additive demand system: estimates for Australia., No. op-73. Victoria University, Centre of Policy Studies/IMPACT Centre, 1992.

[35] Silva, JMC Santos, and Silvana Tenreyro. The log of gravity., *The Review of Economics and Statistics*, 88.4: 641-658, 2006.

[36] Stern, David I. Elasticities of substitution and complementarity., *Journal of Productivity Analysis*, 36.1: 79-89, 2011.

[37] Surry, Yves. The constant difference of elasticities function with applications to the EC animal feed sector., *Journal of Agricultural Economics*, 44.1: 110-125, 1993.

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

[39] Theil, Henri, Ching-Fan Chung, and James L. Seale. International evidence on consumption patterns., Vol. 1. Jai Pr, 1989.

[40] Yilmazkuday, Hakan. Welfare Gains from Trade in Multi-Sector Models: The Role of Aggregation and Income Elasticities., 2018.

# Appendices

## A Proof of Proposition 1

**Proof.** Let $e_i = e$ and $\alpha_i = \alpha \ \forall i$, then Equation (1) generalizes to

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

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

$$\eta_i = \frac{e_i(1-\alpha_i) + \sum_k e_k \omega_k \alpha_k}{\sum_k e_k \omega_k} + \alpha_i - \sum_k \omega_k \alpha_k, \tag{A.2}$$

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

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

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

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

(3) *Identical to Explicitly Direct CES* Rearranging (A.1) by factoring common terms, denoting $V$ as indirect utility, then it leads to the following explicitly indirect expression

$$V = \left[ \sum_i \beta_i \left( \frac{p_i}{E} \right)^{1-\alpha} \right]^{\frac{1}{e(\alpha-1)}} \quad \text{(Explicitly Indirect Homothetic CES)}, \tag{A.4}$$

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

$$U = \left[ \sum_{i=1}^{N} \beta_i^{\frac{1}{\alpha}} q_i^{\frac{\alpha-1}{\alpha}} \right]^{\frac{\alpha}{e(\alpha-1)}} \quad \text{(Explicitly Direct Homothetic CES)}. \tag{A.5}$$

(A.5 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

$$V = \left[ \sum_i \beta_i \left( \frac{p_i}{E} \right)^{1-\alpha} \right]^{\frac{1}{\alpha-1}} \quad \text{(special case of A.4)}, \tag{A.6}$$

and

$$U = \Big[\sum_{i=1}^{N} \beta_i^{\frac{1}{\alpha}} q_i^{\frac{\alpha-1}{\alpha}}\Big]^{\frac{\alpha}{\alpha-1}} \quad \text{(special case of A.5)}. \tag{A.7}$$

The ordinary demand from utility maximization to (A.7) yields the same result as derived from (A.5), whether or not $e = 1$ is additionally imposed. The demand as a function of $PU \equiv P \equiv E/u$ (as a result from A.5) in this special case can then be written as follows

$$q_i = \frac{\beta_i p_i^{-\alpha}}{PU^{1-\alpha}} E \equiv h_i. \tag{A.8}$$

Hence, a standard CDE can be parameterized to achieve (A.4) 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 (A.5). It further (i) yields standard CES price index in GE; and (ii) satisfies the CES real wealth assumption $\iff e = 1,\ \alpha_i = \alpha\ \forall\ i$. ■

# B Implicitly Indirect NHCES to Implicitly Direct NHCES

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

**Proof.** To be complete, starting from the least restricted CDE demand function

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

Note that if $\alpha_i = \alpha\ \forall\ i$, Equation (B.1) 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 (B.1) (with $\xi_i = p_i/E$)

$$\ln q_i = \ln[\beta_i(1-\alpha_i)] + e_i(1-\alpha_i)\ln u - \alpha_i \ln \xi_i - \ln\big[\sum_j \beta_j u^{e_j(1-\alpha_j)}(1-\alpha_j)\xi_i^{1-\alpha_j}\big]. \tag{B.2}$$

Eliminating the last term in equation (B.2) by using logarithmic ratio

$$\ln\frac{q_i}{q_1} = \ln\frac{\beta_i(1-\alpha_i)}{\beta_1(1-\alpha_1)} + [e_i(1-\alpha_i) - e_1(1-\alpha_1)]\ln u - \alpha_i \ln\xi_i + \alpha_1\ln\xi_1$$
$$= A_i + Z_i \ln u - \alpha_i \ln\xi_i + \alpha_1 \ln\xi_1 \quad \forall i \in [2,\infty) \quad \text{(CDE)}$$
$$= \widetilde{A}_i + \widetilde{Z}_i \ln u - \alpha \ln(\frac{p_i}{p_1}) \quad \forall i \in [2,\infty) \quad \iff \quad \alpha_i = \alpha \quad \forall i, \quad \text{(Implicitly Indirect NHCES)}, \tag{B.3}$$

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)

$$F(\mathbf{q}, u) = \sum_i k_i u^{-e_i(1-g)} q_i^{1-g} \equiv 1 \quad \text{(Implicitly Direct NHCES)}, \tag{B.4}$$

which is parameterized from Mukerji (1963)'s Constant Ratios of Elasticity of Substitution (CRES) model (with $g_i = g\ \forall i$ in B.5)

$$F(\mathbf{q}, u) = \sum_i k_i u^{-e_i(1-g_i)} q_i^{1-g_i} \equiv 1 \quad \text{(Implicitly Direct CRES)}, \tag{B.5}$$

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 (B.5) to be globally valid (monotonic and quasi-concave).

For completeness, from the expenditure minimization problem to (B.5) (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

$$p_i = \lambda k_i (1-g_i) u^{-e_i(1-g_i)} q_i^{-g_i} \qquad \forall i, \tag{B.6}$$

where

$$p_1 = \lambda k_1 (1-g_1) u^{-e_1(1-g_1)} q_1^{-g_1} \tag{B.7}$$

Dividing (B.6) by (B.7) 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

$$q_1^{-g_1}\frac{p_i}{p_1} = \frac{k_i(1-g_i)}{k_1(1-g_1)} u^{e_1(1-g_1) - e_i(1-g_i)} q_i^{-g_i}. \tag{B.8}$$

Solving for $q_i$

$$q_i = (\frac{p_i}{p_1})^{-\frac{1}{g_i}} \left[\frac{k_i(1-g_i)}{k_1(1-g_1)}\right]^{\frac{1}{g_i}} u^{\frac{e_1(1-g_1)-e_i(1-g_i)}{g_i}} q_1^{\frac{g_1}{g_i}}. \tag{B.9}$$

Taking the natural logarithm of both sides of (B.9)

$$\begin{aligned}\ln q_i &= \frac{1}{g_i}\ln\left[\frac{k_i(1-g_i)}{k_1(1-g_1)}\right] - \frac{1}{g_i}\ln(\frac{p_i}{p_1}) + \frac{e_1(1-g_1)-e_i(1-g_i)}{g_i}\ln u + \frac{g_1}{g_i}\ln q_1 \\ &= M_i - s_i\ln(\frac{p_i}{p_1}) + R_i\ln u + \frac{s_i}{s_1}\ln q_1 \quad \forall i \in [2,\infty),\end{aligned} \tag{B.10}$$

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 (B.5) (CRES) to (B.4) (Implicitly Direct NHCES) has arisen by restricting $g_i = g \ \forall i \implies s_i = s \ \forall i$, then (B.10) converges to

$$\ln\frac{q_i}{q_1} = \widetilde{M_i} + \widetilde{R}_i \ln u - s\ln(\frac{p_i}{p_1}) \quad \forall i \in [2,\infty) \quad \text{(Implicitly Direct NHCES)}, \tag{B.11}$$

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 (B.3) $\iff \widetilde{M_i} = \widetilde{A}_i$, $\widetilde{R}_i = \widetilde{Z}_i$, and $s = \alpha \implies \beta_i = k_i^{\alpha} \ \forall i$ and $\alpha = 1/g$.
■

# C Direct Proof of (B)

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

**Proof.** Multiplying both sides of (B.9) by $p_i$

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

Then the total expenditure can be expressed as

$$\sum_i p_i q_i = E = \sum_i p_i^{\frac{g_i-1}{g_i}} p_1^{\frac{1}{g_i}} \left[\frac{k_i(1-g_i)}{k_1(1-g_1)}\right]^{\frac{1}{g_i}} u^{\frac{e_1(1-g_1)-e_i(1-g_i)}{g_i}} q_1^{\frac{g_1}{g_i}}. \tag{C.2}$$

Solving for $q_1$:

$$q_1 = \frac{E^{\frac{g_i}{g_1}}}{\left[\sum_i p_i^{\frac{g_i-1}{g_i}} p_1^{\frac{1}{g_i}} \left(\frac{k_i(1-g_i)}{k_1(1-g_1)}\right)^{\frac{1}{g_i}} u^{\frac{e_1(1-g_1)-e_i(1-g_i)}{g_i}}\right]^{\frac{g_i}{g_1}}} \tag{C.3}$$

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

$$\begin{aligned} q_i &= (\frac{p_i}{p_1})^{-\frac{1}{g_i}} \left[\frac{k_i(1-g_i)}{k_1(1-g_1)}\right]^{\frac{1}{g_i}} u^{\frac{e_1(1-g_1)-e_i(1-g_i)}{g_i}} \frac{E}{\sum_i p_i^{\frac{g_i-1}{g_i}} p_1^{\frac{1}{g_i}} \left(\frac{k_i(1-g_i)}{k_1(1-g_1)}\right)^{\frac{1}{g_i}} u^{\frac{e_1(1-g_1)-e_i(1-g_i)}{g_i}}} \\ &= \frac{\left(\frac{p_i}{p_1}\right)^{-\frac{1}{g_i}} \left[\frac{k_i(1-g_i)}{k_1(1-g_1)}\right]^{\frac{1}{g_i}} u^{\frac{e_1(1-g_1)-e_i(1-g_i)}{g_i}} E}{\sum_i p_i^{\frac{g_i-1}{g_i}} p_1^{\frac{1}{g_i}} \left(\frac{k_i(1-g_i)}{k_1(1-g_1)}\right)^{\frac{1}{g_i}} u^{\frac{e_1(1-g_1)-e_i(1-g_i)}{g_i}}}. \end{aligned} \tag{C.4}$$

If $g_i = g\ \forall i$, then the implicitly Direct CRES generalizes to implicitly direct NHCES. Then we may directly apply this parameterization to (C.4) and solves for $q_i$

$$\begin{aligned} q_i &= \frac{\left(\frac{p_i}{p_1}\right)^{-\frac{1}{g_i}} \left[\frac{k_i(1-g_i)}{k_1(1-g_1)}\right]^{\frac{1}{g_i}} u^{\frac{e_1(1-g_1)-e_i(1-g_i)}{g_i}} E}{\sum_i p_i^{\frac{g_i-1}{g_i}} p_1^{\frac{1}{g_i}} \left(\frac{k_i(1-g_i)}{k_1(1-g_1)}\right)^{\frac{1}{g_i}} u^{\frac{e_1(1-g_1)-e_i(1-g_i)}{g_i}}} \\ &= \frac{p_i^{-\frac{1}{g}} \left(\frac{k_i}{k_1}\right)^{\frac{1}{g}} u^{\frac{(e_1-e_i)(1-g)}{g}} E}{\sum_i p_i^{\frac{g-1}{g}} \left(\frac{k_i}{k_1}\right)^{\frac{1}{g}} u^{\frac{(e_1-e_i)(1-g)}{g}}} \\ &= \frac{p_i^{-\frac{1}{g}} \left(\frac{k_i}{k_1}\right)^{\frac{1}{g}} u^{\frac{(e_1-e_i)(1-g)}{g}} E}{\sum_i p_i^{\frac{g-1}{g}} \left(\frac{k_i}{k_1}\right)^{\frac{1}{g}} u^{\frac{(e_1-e_i)(1-g)}{g}}} \\ &= \frac{k_i^{\frac{1}{g}} u^{\frac{e_i(g-1)}{g}} p_i^{-\frac{1}{g}} E}{\sum_i k_i^{\frac{1}{g}} u^{\frac{e_i(g-1)}{g}} p_i^{\frac{g-1}{g}}} \end{aligned} \tag{C.5}$$

Let $g = 1/\alpha$, then we may express (C.5) as follows

$$q_i = \frac{k_i^{\alpha} u^{e_i(1-\alpha)} p_i^{-\alpha} E}{\sum_i k_i^{\alpha} u^{e_i(1-\alpha)} p_i^{1-\alpha}}. \tag{C.6}$$

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

$$q_i = \frac{\beta_i u^{e_i(1-\alpha)} p_i^{-\alpha} E}{\sum_i \beta_i u^{e_i(1-\alpha)} p_i^{1-\alpha}}, \tag{C.7}$$

which is derived under the same parameterization using the linear approach in Appendix A. ■

# D Derivation and Extension of Equation (42)

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

$$\begin{aligned}\hat{q}_i &= e_i(1-\alpha_i)\hat{u} - \alpha_i \hat{p}_i + \alpha_i \hat{E} - \sum_j e_j(1-\alpha_j)\omega_j \hat{u} - \sum_j (1-\alpha_j)\omega_j \hat{p}_j + \sum_j (1-\alpha_j)\omega_j \hat{E} \\ &= \left[e_i(1-\alpha_i) - \sum_j e_j(1-\alpha_j)\omega_j\right]\hat{u} \\ &+ \left[\alpha_i + \sum_j (1-\alpha_j)\omega_j\right]\hat{E} \\ &- \alpha_i \hat{p}_i - \sum_j (1-\alpha_j)\omega_j \hat{p}_j.\end{aligned} \tag{D.1}$$

Equation (D.1) can be considered as a special case where there are 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 $\alpha_i$, 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 (D.1), the change of quantity consumption can also be expressed as a function of income and cross-price elasticities

$$\hat{q}_i = \eta_i \hat{E} + \sum_j \sigma_{i,j} \hat{p}_j, \tag{D.2}$$

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

$$\hat{q}_i = \hat{E} - \alpha\hat{p}_i - (1-\alpha)\sum_j \hat{p}_j, \tag{D.3}$$

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

# E Derivation of Equation (45)

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

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

By rearranging terms above, we have

$$\begin{aligned}
&\sum_i(1-\alpha_i)\beta_i u^{e_i(1-\alpha_i)}(\frac{p_i}{E})^{1-\alpha_i} \overbrace{\frac{dw}{E}}^{\text{Change of Wealth}} \\
\equiv &\sum_i \beta_i e_i(1-\alpha_i)u^{e_i(1-\alpha_i)}(\frac{p_i}{E})^{1-\alpha_i} \underbrace{\overbrace{\frac{du}{u}}}_{\text{Change of Utility}} \\
+&\sum_i \beta_i(1-\alpha_i)u^{e_i(1-\alpha_i)}(\frac{p_i}{E})^{1-\alpha_i} \underbrace{\overbrace{\frac{dp_i}{p_i}}}_{\text{Change of Price}}
\end{aligned} \tag{E.2}$$

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

$$\omega_i = \frac{\beta_i u^{e_i(1-\alpha_i)}(1-\alpha_i)(\frac{p_i}{E})^{1-\alpha_i}}{\sum_j \beta_j u^{e_j(1-\alpha_j)}(1-\alpha_j)(\frac{p_j}{E})^{1-\alpha_j}} = \frac{\beta_i u^{e_i(1-\alpha_i)}(1-\alpha_i)(\frac{p_i}{E})^{1-\alpha_i}}{T}. \tag{E.3}$$

Note that if we divide both sides by $T$ of Equation (E.2), and using *hat* to denote rate of changes on corresponding terms, then the expression can be further simplified to

$$\sum_i \omega_i\hat{E} \equiv \sum_i e_i\omega_i\hat{u} + \sum_i \omega_i\hat{p}_i, \tag{E.4}$$

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

$$\hat{E} \equiv \sum_i e_i \omega_i \hat{u} + \sum_i \omega_i \hat{p}_i, \tag{E.5}$$

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

$$\hat{u} = \frac{\hat{E} - \sum_i \omega_i \hat{p}_i}{\sum_i e_i \omega_i}. \tag{E.6}$$