Business School / School of Economics

UNSW Business School Working Paper

UNSW Business School Research Paper No. 2017 ECON 05

A Gravity-Based Revealed Comparative Advantage Estimator

Scott French

A Gravity-Based Revealed Comparative Advantage Estimator

Scott French*

February, 2017

Abstract

I propose a method of moments estimator of revealed comparative advantage based on a flexible specification of trade flows that is consistent with a large class of gravity models of international trade. I show that this estimator has many desirable properties. It is theoretically consistent with the classical notion of Ricardian comparative advantage and is easily computed, even for very large samples. Statistical inference is straightforward, and it is closely related to a commonly-used estimator in the gravity literature that is known to be robust to various forms of heteroskedasticity and measurement error common to trade data.

JEL Classification: F10, F11, F14, C13, C21, C55

Keywords: Method of moments; pseudo-maximum likelihood; Poisson; Ricardian; RCA; index

*School of Economics, University of New South Wales. scott.french@unsw.edu.au.

1 Introduction

The Ricardian model of international trade is perhaps the most influential and celebrated theory in the discipline of economics. However, it has been employed in empirical analyses with mixed success.[1] Meanwhile, gravity models have been an important empirical tool since Tinbergen (1962), but only more recently were they underpinned by rigourous theoretical foundations.[2] In this paper, I propose a simple method of moments estimator of "revealed comparative advantage" (RCA) that is consistent with a wide range of structural multi-sector and product-level gravity models and with the classical Ricardian defintion of comparative advantage. This estimator has a number of desirable properties which make it a useful complement to the regression-based estimator of Costinot et al. (2012) and the simple but atheoretic RCA indexes following Balassa (1965).

2 Specification

A large class of trade models imply that sectoral bilateral trade flows follow a gravity equation.[3] In keeping with this literature, I make the following assumption:

Assumption 1. *Bilateral trade flows of product k to destination n from source i take the following form:*

$$E[X_{nik}] = \psi_{ik} \cdot \tilde{\phi}_{nk} \cdot d_{nik}^{-\eta}, \tag{1}$$

where ψ_{ik} and $\tilde{\phi}_{nk}$ represent exporter-product- and importer-product-specific variables, $d_{nik} > 1$ is an "iceberg" trade cost, and η is the elasticity of trade flows with respect to trade costs.

This specification allows X_{nik} to be observed with error, either due to measurement errors or unobserved shocks to bilateral trade flows.[4]

French (forthcoming) shows that, in a class of models consistent with (1), ψ_i^k constitutes a measure of RCA in that

$$\frac{\bar{P}_{ik}}{\bar{P}_{ik'}} < \frac{\bar{P}_{i'k}}{\bar{P}_{i'k'}} \iff \frac{\psi_{ik}}{\psi_{ik'}} > \frac{\psi_{i'k}}{\psi_{i'k'}},$$

where $\bar{P}_{ik}$ is the autarky price of good k in country i. In words, whether country i has a comparative advantage in producing product k, compared to country i' and product k', according to the classical Ricardian definition, is determined by its relative value of ψ_{ik}. Thus, a valid estimator for these parameters is a theoretically-consistent indicator of RCA.

As specified, the parameters of (1) cannot be uniquely identified without imposing some additional structure on the form of trade costs. In the gravity literature, trade costs are often taken vary only bilaterally ($d_{nik} = d_{ni}$) and are parameterized as a function of observable variables, such as

[1] See Costinot et al. (2012) for a discussion of the empirical applications of Ricardian theory.

[2] See Head and Mayer (2014) a review of history of gravity models.

[3] See Anderson and van Wincoop (2004) for a rigourous development of the economic theory of gravity and Anderson (2011) and Head and Mayer (2014) for summaries of recent developments in the literature.

[4] Note that these shocks do not include unobserved trade costs, which Egger and Nigai (2015) point out may cause ψ_{ik} and $\tilde{\phi}_{nk}$ to be endogenous.

bilateral distance and indicators for a shared border, common language, etc. I take a less restrictive approach and make the following assumption:

Assumption 2. *Trade costs can be separated into a bilateral component and an importer-product-specific component:*

$$d_{ni}^k = d_{ni} d_{nk}, \quad \forall n \neq i. \tag{2}$$

The first component captures all country-pair-specific trade costs, including the effects of most variables included in a typical gravity estimation. The second component captures product-specific trade barriers in each import market, such as customs requirements and import tariffs. While (2) is unlikely to hold precisely in the data, it is reasonable to suggest that it captures the vast majority of variation in trade costs across countries and products. It is much less restrictive than is typically assumed in gravity estimations, and it is consistent with the Most Favored Nation principle of the World Trade Organization regarding import tariffs.[5]

Together, assumptions 1 and 2 imply that trade flows of product k are given by

$$E[X_{nik}] = \psi_{ik} \cdot \phi_{nk} \cdot \delta_{ni}, \quad \forall n \neq i, \tag{3}$$

where $\phi_{nk} = \tilde{\phi}_{nk} d_{nk}^{-\theta}$ and $\delta_{ni} = d_{ni}^{-\theta}$.

3 A Method of Moments Estimator

The form of (3) suggests an intuitive method of moments estimator based on the the moment condition $E[X_{nik} - \psi_{ik}\phi_{nk}\delta_{ni}] = 0$, which implies the following sample moment conditions:

$$\begin{aligned}
\bar{m}_{ik} &\equiv \frac{1}{N-1} \sum_{n \neq i} \left(X_{nik} - \hat{\psi}_{ik} \hat{\phi}_{nk} \hat{\delta}_{ni} \right) = 0, \quad \forall i, k \\
\bar{m}_{nk} &\equiv \frac{1}{N-1} \sum_{i \neq n}^{N} \left(X_{nik} - \hat{\psi}_{ik} \hat{\phi}_{nk} \hat{\delta}_{ni} \right) = 0, \quad \forall n, k \\
\bar{m}_{ni} &\equiv \frac{1}{K} \sum_{k=1}^{K} \left(X_{nik} - \hat{\psi}_{ik} \hat{\phi}_{nk} \hat{\delta}_{ni} \right) = 0, \quad \forall n \neq i.
\end{aligned} \tag{4}$$

Because this estimator is exactly identified – comprising a system of $J = 2NK + N(N-1)$ equations and unknowns, up to $2N + K - 1$ normalizations – there is a set of parameters, $\hat{\boldsymbol{\theta}} = \{\hat{\boldsymbol{\psi}}, \hat{\boldsymbol{\phi}}, \hat{\boldsymbol{\delta}}\}$, that gives and exact solution to the set of J moment equations, $\bar{\boldsymbol{m}} = 0$.[6]

[5] This specification does allow bilateral trade costs to be asymmetric – i.e., $d_{ni} \neq d_{in}$.

[6] It is easy to verify that, if $\{\hat{\psi}_i^k, \hat{\phi}_n^k, \hat{\delta}_{ni}\}$ is a solution to (4), then $\{\tilde{\psi}_i^k, \tilde{\phi}_n^k, \tilde{\delta}_{ni}\}$ is also a solution to (4) if $\tilde{\psi}_i^k = (\alpha_i/\alpha^k)\hat{\psi}_i^k$, $\tilde{\phi}_n^k = (\alpha^k/\alpha_n)\hat{\phi}_n^k$, and $\tilde{\delta}_{ni} = (\alpha_n/\alpha_i)\hat{\delta}_{ni}$, for any set of positive constants $\{\alpha_n\}$, $\{\alpha_i\}$, and $\{\alpha^k\}$.

3.1 Relation to Gravity

I refer to the estimator defined by (4) as the "gravity-based" estimator of $\boldsymbol{\theta}$ because it imposes the same adding-up constraints as structural gravity models. To see this, note that by combining (3) and (4), we can express the fitted values of trade flows as $\hat{X}_{nik} = \hat{\psi}_{ik}\hat{\phi}_{nk}\hat{\delta}_{ni}$, where estimated parameter values solve the following system of equations:

$$\hat{\psi}_{ik} = \frac{E_{ik}}{\sum_{n\neq i}\hat{\phi}_{nk}\hat{\delta}_{ni}}, \qquad \hat{\phi}_{nk} = \frac{M_{nk}}{\sum_{i\neq n}\hat{\psi}_{ik}\hat{\delta}_{ni}}, \qquad \hat{\delta}_{ni} = \frac{X_{ni}}{\sum_k \hat{\psi}_{ik}\hat{\phi}_{nk}}, \tag{5}$$

where $E_{ik} = \sum_{n\neq i} X_{nik}$, $M_{nk} = \sum_{i\neq n} X_{nik}$, and $X_{ni} = \sum_k X_{nik}$.[7]

To make the equivalence with structural gravity clear, take the value of δ_{ni} as given and define $\Phi_{nk} \equiv M_{nk}/\hat{\phi}_{nk}$ and $\Psi_{ik} \equiv E_{ik}/\hat{\psi}_{ik}$. Then, we have that

$$\hat{X}_{nik} = \frac{E_{ik}}{\Psi_{ik}}\frac{M_{nk}}{\Phi_{nk}}\delta_{ni},$$

where

$$\Psi_{ik} = \sum_{n\neq i}\frac{\delta_{ni}M_{nk}}{\Phi_{nk}} \quad \text{and} \quad \Phi_{nk} = \sum_{i\neq n}\frac{\delta_{ni}M_{nk}}{\Psi_{ik}}$$

are the "multilateral resistance" terms defined by Anderson and van Wincoop (2003). Note that, while structural gravity equations are more typically expressed in terms of output and expenditure, rather than exports and imports, French (2017) shows that this specification is fully consistent with structural gravity which allows for missing data on domestic trade flows.

3.2 Relation to Poisson PML

It is straightforward to show that the gravity-based estimator is identical to a Poisson pseudo-maximum likelihood (PML) estimator with (only) importer-product, exporter-product, and importer-exporter fixed effects. To see this, note that first-order conditions of the Poisson PML estimator are

$$\sum_{n=1}^{n}\sum_{i=1}^{N}\sum_{k=1}^{k}\left(X_{nik} - e^{\mathbf{x}'_{nik}\boldsymbol{\beta}}\right)\mathbf{x}_{nik} = 0. \tag{6}$$

If $\mathbf{x}_{nik}$ is a vector of importer-product-, exporter-product-, and importer-exporter-specific dummy variables, then the solution to these first-order conditions is identical to $\hat{\boldsymbol{\theta}}$. Gourieroux et al. (1984) show that, if a solution to (6) exists, it is unique. Thus, $\hat{\boldsymbol{\theta}}$ is the unique solution to (4). Further, this implies that, when the dataset is not too large, this estimator can be computed by estimating a Poisson regression with the appropriate set of dummy variables, which is straightforward using a number of standard statistical programs.[8]

[7] This result is closely linked to the findings of Fally (2015) and Arvis and Shepherd (2013) that the Poisson PML estimator with importer and exporter fixed effects implicitly imposes the structural gravity constraints.

[8] Santos Silva and Tenreyro (2011) suggest using Stata's user-written **ppml** command or the **glm** command (and not the **poisson** command).

3.3 Relation to OLS

If, instead of (1), we assume that

$$E[\ln X_{nik}] = \ln \psi_{ik} + \ln \tilde{\phi}_{nk} + \ln d_{nik}^{-\eta}, \tag{7}$$

the empirical moment conditions analogous to (4) are

$$\bar{m}_{ik} \equiv \frac{1}{N-1} \sum_{n \neq i} \left(\ln X_{nik} - \ln \hat{\psi}_{ik} - \ln \hat{\phi}_{nk} - \ln \hat{\delta}_{ni} \right) = 0, \quad \forall i, k$$

$$\bar{m}_{nk} \equiv \frac{1}{N-1} \sum_{i \neq n}^{N} \left(\ln X_{nik} - \ln \hat{\psi}_{ik} - \ln \hat{\phi}_{nk} - \ln \hat{\delta}_{ni} \right) = 0, \quad \forall n, k$$

$$\bar{m}_{ni} \equiv \frac{1}{K} \sum_{k=1}^{K} \left(\ln X_{nik} - \ln \hat{\psi}_{ik} - \ln \hat{\phi}_{nk} - \ln \hat{\delta}_{ni} \right) = 0, \quad \forall n \neq i.$$

These moment conditions are equivalent to the normal equations associated with a linear regression of the form

$$\ln X_{nik} = \ln \psi_{ik} + \ln \phi_{nk} + \ln \delta_{ni} + \varepsilon_{nik}. \tag{8}$$

As a result, I refer to this as the "regression-based" estimator of $\hat{\boldsymbol{\theta}}$. This regression is identical to the one proposed by Costinot et al. (2012) to estimate relative productivity across countries and industries. If (7) is correct, then (8) will provide a consistent estimate of $\boldsymbol{\theta}$.

This estimator also has the advantage of being easy to implement. Any standard statistical software can estimate (8) as a linear regression of $\ln X_{nik}$ on the appropriate set of dummy variables. However, there is reason to suggest that the gravity-based estimator is weakly superior to this regression-based estimator. Most importantly, Santos Silva and Tenreyro (2006) point out that (7) assumes a very particular form of heteroskedasticity and that estimates of $\boldsymbol{\theta}$ will be biased if the data are inconsistent with this assumption. In addition, the log-linear specification of (8) requires that observations with zero-valued trade flows be dropped from the estimation, leading to sample-selection bias. In Monte Carlo experiments meant to replicate the features of international trade data, Santos Silva and Tenreyro (2006) find that these biases are substantial. By contrast, they find that the Poisson PML estimator performs very well in the presence of many forms of heteroskedasticity and a large number of zeros due to rounding error.

3.4 Relation to Other Estimators

While (4) is the obvious method of moments estimator based on (3), there are many other potentially valid moment conditions. For example, consider the moment condition $E[(\psi_{ik}\phi_{nk}\delta_{ni})^{1-\rho} \times (X_{nik} - \psi_{ik}\phi_{nk}\delta_{ni})] = 0$, for a predetermined value of ρ. This is clearly valid, given that (3) holds. Moreover, it is the optimal moment condition if $V[X_{nik} - \psi_{ik}\phi_{nk}\delta_{ni}] \propto (\psi_{ik}\phi_{nk}\delta_{ni})^{\rho}$. The method of moments estimators based on this moment condition nest several PML estimators based on linear-exponential family distributions. Clearly, $\rho = 1$ corresponds to the gravity-based estimator and Poisson PML. Other notable cases are $\rho = 0$, which corresponds to Guassian PML (nonlinear least squares), and $\rho = 2$, which corresponds to gamma PML.

However, estimators based on this generalized moment condition have several drawbacks compared to the gravity-based estimator. In addition to the gravity-based estimator's consistency with structural gravity, the fact that each of the equations of (4) is linear in all variables means that solving this system is quite computationally efficient. This is not the case when $\rho \neq 1$. In addition, Poisson PML is special in that, like OLS, it does not suffer from an incidental parameters problem with a large number of fixed effects, meaning that consistent estimation of one parameter does not require consistent estimation of all parameters.[9] This not the case with other PML estimators. By analogy, the gravity-based estimator is likely to have better finite-sample properties than estimators based on this generalized moment condition.

In addition, Santos Silva and Tenreyro (2006) argue that $\rho = 1$ is likely to be a good assumption regarding trade data. They argue that, because $X_{nik} \geq 0$, its variance must approach zero as $E[X_{nik}]$ approaches zero, meaning that the optimal value of ρ is likely positive. On the other hand, measurement error is likely more prevalent for small values of X_{nik}, so the estimator should not place too much weight on small values, meaning that ρ should not be too large, and $\rho = 1$ is a reasonable compromise. Santos Silva and Tenreyro (2006) also perform a series of Monte Carlo experiments to evaluate the relative performance of these PML estimators and find Poisson PML to be the most robust under various forms of heteroskedasticity and measurement error.

It is worth noting that, because $\boldsymbol{\theta}$ is exactly identified by these estimators, it is not possible to estimate ρ without adding additional moment conditions. However, for those interested in testing whether a particular value of ρ is appropriate, the guidance provided by Head and Mayer (2014) regarding PML estimators is relevant to the generalized gravity-based estimator. To summarize, they find that the Park-type tests proposed by Manning and Mullahy (2001) and applied by Santos Silva and Tenreyro (2006) to be useful indicators (though not robust estimators) of the optimal value of ρ. They also suggest that gamma PML ($\rho = 2$), along with the analogue of the regression-based estimator, be used as a robustness test for the Poisson PML estimator ($\rho = 1$).

4 Statistical Inference

Because the gravity-based estimator is a standard method of moments estimator, statistical inference based on the estimates is also standard. The standard estimator for the asymptotic variance of $\hat{\boldsymbol{\theta}}$, assuming that observations are independent over (n, i, k), is

$$\hat{V}[\hat{\boldsymbol{\theta}}] = \frac{1}{N^2K}\hat{\boldsymbol{G}}^{-1}\hat{\boldsymbol{S}}\hat{\boldsymbol{G}}',$$

[9] Fernández-Val and Weidner (2016) show this for Poisson PML with two-way fixed effects.

where

$$\hat{\boldsymbol{G}} = \frac{\partial \bar{\boldsymbol{m}}}{\partial \boldsymbol{\theta}'},$$

where $\bar{\boldsymbol{m}}$ is the set of moment equations, and

$$\hat{\boldsymbol{S}} = \frac{1}{N^2 K} \sum_{n,i,k} \bar{\boldsymbol{m}}_{nik} \bar{\boldsymbol{m}}'_{nik},$$

where $\bar{\boldsymbol{m}}_{nik}$ is the $J \times 1$ vector of the contribution of observation (n, i, k) to $\bar{\boldsymbol{m}}$.

However, with international trade data, the independence assumption is a strong one. In particular, it is likely that errors are clustered for trade flows to or from a common country. As a result, an appropriate cluster-robust covariance matrix should be estimated. This can done by replacing $\hat{\boldsymbol{S}}$ by

$$\hat{\boldsymbol{S}}_{clus} = \frac{1}{N^2 K} \sum_{n,i,k} \sum_{n',i',k'} \bar{\boldsymbol{m}}_{nik} \bar{\boldsymbol{m}}'_{n'i'k'} \mathbf{1} \left[(n, i, k) \text{ and } (n', i', k') \text{ share any cluster} \right],$$

where $\mathbf{1}[\cdot]$ is the indicator function.[10]

5 Practical Considerations

Several practical considerations are likely to be of interest in applications using the gravity-based estimator defined by (4). In particular, I discuss a straightforward method for computing $\hat{\boldsymbol{\theta}}$ and a method which ameliorates the technical difficulty of statistical inference for large sample sizes. I also propose a simple statistic that tests for comparative advantage, and I discuss a method for identifying d_{nk} when data on domestic trade flows are available.

5.1 A General Solution Algorithm

For large numbers of countries and products, it is not computationally feasible to implement the gravity-based estimator via Poisson regression with dummy variables. It turns out that the gravity-like structure of the estimator provides simple and effective alternative. We have already seen that $\hat{\boldsymbol{\theta}}$ is the solution to (5). Because we know that this solution is unique, we can solve for $\hat{\boldsymbol{\theta}}$ by iterating on these equations until convergence.[11]

Algorithms such as this have a long history in various literatures. For example, this algorithm is analogous to the "zig-zag" algorithm proposed by Heckman and Macurdy (1980) in the context of a fixed-effects Probit. It is also a generalization of the solution algorithm for the gravity model of Anderson and van Wincoop (2003), where δ_{ni} is treated as an unknown, rather than being taken as given. It is extremely simple to implement, requiring only matrix multiplication and element-wise

[10] See Cameron and Miller (2015) for a recent survey of the literature on cluster-robust inference.

[11] While it is beyond the scope of this paper to prove that convergence is guaranteed, in practice the system appears always to converge monotonically to $\hat{\boldsymbol{\theta}}$.

division, and while convergence may take many iterations, each iteration is very computationally efficient so that convergence can be achieved quickly and with minimal memory requirements.

Extending this algorithm to the regression-based estimator yields exactly the iterative solution for OLS employed by Guimarães and Portugal (2010). Despite the similarity between the two solution algorithms, computation of the regression-based estimator is much slower in practice. This is because the regression-based estimator drops zero-valued trade flows, leading to an unbalanced sample. This means that standard matrix operations cannot be employed, increasing the computation time considerably.

5.2 Bootstrapped Cluster-Robust Standard Errors

While it is straightforward to compute $\hat{\boldsymbol{\theta}}$ for large samples, it may be infeasible to compute the $J \times J$ matrices $\hat{\boldsymbol{G}}$, $\hat{\boldsymbol{S}}$, and $\hat{V}(\hat{\boldsymbol{\theta}})$. One solution is to employ the pairs cluster bootstrap discussed by Cameron and Miller (2015). This approach allows the elements of interest of $\hat{\boldsymbol{V}}$ to be computed without needing to construct $\hat{\boldsymbol{G}}$. It does require that (5) be solved many times. However, because this can be done relatively quickly, especially once $\hat{\boldsymbol{\theta}}$ has been found and can be used as a starting value, the additional computational burden is not prohibitive.

Because trade data typically take the form of a balanced sample, it is particularly straightforward to use a two-way cluster by both importer and exporter using the following procedure:

1. Draw N importers and N exporters, with replacement.
2. Construct the sample of trade flows, $X_{n^* i^* k}$, where n^* and i^* are the n^{th} importer and i^{th} exporter drawn.
3. Compute $\hat{\boldsymbol{\theta}}^b$ by solving (5) for the constructed sample.
4. Repeats steps 1-3 for $b = 1, ..., B$.

Then, compute the elements of interest of $\hat{\boldsymbol{V}}$ as

$$\hat{\boldsymbol{V}}_{jj'} = \frac{1}{N_{jj'}} \sum_{b=1}^{B} (\hat{\theta}_j^b - \bar{\hat{\theta}}_j^b)(\hat{\theta}_{j'}^b - \bar{\hat{\theta}}_{j'}^b) \mathbf{1}[\hat{\theta}_j^b \text{ and } \hat{\theta}_{j'}^b \text{ both identified in sample } b],$$

where $N_{jj'} = \sum_{b=1}^{B} \mathbf{1}[\hat{\theta}_j^b \text{ and } \hat{\theta}_{j'}^b \text{ both identified in sample } b]$.

Cameron and Miller (2015) recommend using $B = 400$. However, because all countries do not appear in each bootstrap sample, a larger B may be warranted. A parameter involving country i will be estimated in a given sample with probability $1 - (1 - N^{-1})^N$, which converges to $1 - e^{-1} \approx 0.63$ as N becomes large. Thus, when computing the diagonal elements of $\hat{\boldsymbol{V}}$, $B = 650 > 400/0.63$ should be sufficient for any N. Some of the off-diagonal elements of $\hat{\boldsymbol{V}}$ are estimated less precisely because a bootstrap sample is only informative if *both* countries are sampled. Therefore, an even larger B may be warranted if one or more of these elements is needed. The probability of sampling two different countries in a given sample converges to $(1 - e^{-1})^2 \approx 0.4$, so $B = 1,000 = 400/0.4$ should be sufficient for any N.

One advantage of the bootstrap approach is that, once the bootstrap procedure has been run, each element of $\hat{\boldsymbol{V}}$ can be calculated without re-running the bootstrap. Thus, the bootstrap can be performed at the time of estimation and the results saved so that the relevant elements of $\hat{\boldsymbol{V}}$ can be calculated as necessary for any statistical test conducted in the future.

5.3 Testing for Comparative Advantage

Testing whether a country has a comparative advantage for a particular product, vis-à-vis a reference product and country can be done by testing the following null hypothesis:

$$H_0 : \frac{\psi_{ik}}{\psi_{i'k}} - \frac{\psi_{ik'}}{\psi_{i'k'}} = 0.$$

It is tempting to perform a standard Wald test of this hypothesis, but this is potentially problematic because the Wald test is not invariant to algebraically equivalent specifications of the null hypothesis, meaning it could lead to contradictory conclusions. Other tests, such as likelihood ratio and Lagrange multiplier tests, do not suffer from this issue but are not straightforward to compute for the gravity-based estimator because they require estimating restricted versions of the empirical model.

A reasonable compromise is to estimate $\hat{\boldsymbol{V}}$ based on the logged values of the model parameters. Then, it is straightforward to perform a Wald test of the linear restriction

$$H_0 : (\ln \psi_{ik} - \ln \psi_{i'k}) - (\ln \psi_{ik'} - \ln \psi_{i'k'}) = 0.$$

This calculation is simple when done as part of the pairs cluster bootstrap.

5.4 Identifying Border Costs

One appealing feature of the gravity-based estimator is that it does not require data on domestic trade flows, X_{nnk}, to identify $\boldsymbol{\theta}$. This is important because such data are typically available for many fewer countries and at much higher levels of aggregation than are international trade data. However, if such data are available, then it is also possible to identify the effect of border costs, $d_{nk}^{-\theta}$, which may be of interest in certain applications.

If we assume that $d_{nnk} = d_{nn}$, for all k, then given estimates of $\hat{\psi}_{ik}$ and $\hat{\phi}_{nk}$, based on (4), the gravity-based estimator for border costs is

$$\hat{d}_{nk}^{-\theta} = \frac{\hat{\psi}_{nk}\hat{\phi}_{nk}\hat{\delta}_{nn}}{X_{nnk}},$$

where $\hat{\delta}_{nn} = X_{nn} \Big/ \sum_k \hat{\psi}_{nk}\hat{\phi}_{nk}$.

If domestic trade data are available only at a higher level of aggregation than international trade data, it is not possible to identify $d_{nk}^{-\theta}$, but it is possible to identify an index of these border costs. Suppose that data on domestic trade flows are available for sector j, which includes products $k = 1, ..., K^j$. Then, the following index of $d_{nk}^{-\theta}$ is identified:

$$\hat{d}_{nj}^{-\theta} = \frac{\hat{\delta}_{nn}}{X_{nnj}} \sum_{k=1}^{K^j} \hat{\psi}_{nk} \hat{\phi}_{nk},$$

where $X_{nnj} = \sum_{k=1}^{K^j} X_{nnk}$. French (2017) shows that, because the values of $\hat{\psi}_{nk}$ and $\hat{\phi}_{nk}$ do not depend on d_{nk} in (5), $\hat{d}_{nj}^{-\theta}$ can be interpreted as the hypothetical uniform value of $\hat{d}_{nk}^{-\theta}$ for which predicted sector-level domestic trade flows equal their values in data, when product-level imports and exports, M_{nk} and E_{ik}, are held constant.

6 Conclusion

This paper proposes a gravity-based method of moments RCA estimator. This estimator has a number of desirable properties: (1) it is theoretically consistent with the classical notion of Ricardian comparative advantage, (2) it is consistent with structural gravity models, (3) its relation to Poisson PML implies that it has the same robustness properties, and (4) it is easy to compute the parameter estimates and conduct statistical inference. This suggests that the gravity-based estimator meets the necessary requirements to be a workhorse estimator of RCA.

References

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

Anderson, James E. and Eric van Wincoop (2004), “Trade Costs,” *Journal of Economic Literature*, 42 (3), 691–751.

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

Arvis, Jean-François and Ben Shepherd (2013), “The Poisson Quasi-Maximum Likelihood Estimator: A Solution to the ‘Adding Up’ Problem in Gravity Models,” *Applied Economics Letters*, 20 (6), 515–519.

Balassa, Bela (1965), “Trade Liberalisation and ‘Revealed’ Comparative Advantage,” *The Manchester School*, 33 (2), 99–123.

Cameron, A. Colin and Douglas L. Miller (2015), “A Practitioners Guide to Cluster-Robust Inference,” *Journal of Human Resources*, 50 (2), 317–372.

Costinot, Arnaud, Dave Donaldson, and Ivana Komunjer (2012), “What Goods Do Countries Trade? A Quantitative Exploration of Ricardo’s Ideas,” *Review of Economic Studies*, 79 (2), 581–608.

Egger, Peter H. and Sergey Nigai (2015), “Structural Gravity with Dummies Only: Constrained ANOVA-Type Estimation of Gravity Models,” *Journal of International Economics*, 97 (1), 86–99.

Fally, Thibault (2015), “Structural Gravity and Fixed Effects,” *Journal of International Economics*, 97 (1), 76–85.

Fernández-Val, Iván and Martin Weidner (2016), “Individual and Time Effects in Nonlinear Panel Models with Large N, T,” *Journal of Econometrics*, 192 (1), 291–312.

French, Scott (2017), “Comparative Advantage and Biased Gravity,” UNSW Business School Research Paper No. 2017-03.

——— (forthcoming), “Revealed Comparative Advantage: What Is It Good For?” *Journal of International Economics.*

Gourieroux, C., A. Monfort, and A. Trognon (1984), “Pseudo Maximum Likelihood Methods: Applications to Poisson Models,” *Econometrica*, 52 (3), 701–720.

Guimarães, Paulo and Pedro Portugal (2010), “A Simple Geasible Procedure to Fit Models with High-Dimensional Fixed Effects,” *Stata Journal*, 10 (4), 628–649(22).

Head, Keith and Thierry Mayer (2014), "Gravity Equations: Workhorse, Toolkit, and Cookbook," in Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff (Eds.), *Handbook of International Economics*, Vol. 4, Elsevier, 131–195.

Heckman, James J. and Thomas E. Macurdy (1980), "A Life Cycle Model of Female Labour Supply," *The Review of Economic Studies*, 47 (1), 47–74.

Manning, Willard G and John Mullahy (2001), "Estimating log models: to transform or not to transform?" *Journal of Health Economics*, 20 (4), 461–494.

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

——— (2011), "poisson: Some convergence issues," *Stata Journal*, 11 (2), 207–212(6).

Tinbergen, Jan (1962), *Shaping the World Economy: Suggestions for an International Economic Policy*, New York, Twentieth Century Fund.