European Commission

ISSN 2443-8022 (online)

Solving for Structural Gravity in Panels: Yes We Can

Aurélien Poissonnier

DISCUSSION PAPER 040 | DECEMBER 2016

EUROPEAN ECONOMY

Economic and Financial Affairs

European Economy Discussion Papers are written by the staff of the European Commission's Directorate-General for Economic and Financial Affairs, or by experts working in association with them, to inform discussion on economic policy and to stimulate debate.

The views expressed in this document are solely those of the author(s) and do not necessarily represent the official views of the European Commission.

Authorised for publication by Manfred Bergmann, Director for Economies of the Member States I.

This paper exists in English only and can be downloaded from
http://ec.europa.eu/economy_finance/publications/.

Europe Direct is a service to help you find answers to your questions about the European Union.

Freephone number (*):
00 800 6 7 8 9 10 11

(*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

More information on the European Union is available on http://europa.eu.

Luxembourg: Publications Office of the European Union, 2016

KC-BD-16-040-EN-N (online)
ISBN 978-92-79-54452-1 (online)
doi:10.2765/154608 (online)

KC-BD-16-040-EN-C (print)
ISBN 978-92-79-54453-8 (print)
doi:10.2765/58887 (print)

European Commission
Directorate-General for Economic and Financial Affairs

Solving for Structural Gravity in Panels
Yes We Can

Aurélien Poissonnier

Abstract

Structural gravity models for trade stem from agnostic models of bilateral trade flows. Although more theoretically sound, they are much more complex to estimate. This difficulty is due to the multilateral resistance terms which account for the general equilibrium constraints of global trade and must be inferred from the rest of the model.

In the present paper, I show that solving for these terms explicitly is a valid econometric approach for gravity models, including in panel data. I propose iterative solutions in Stata based on three different techniques.

An example of these solutions on real data is presented. The results from this test confirm the necessity to account for the multilateral resistance terms in the estimation and raise some questions on the alternative solution using dummies.

JEL Classification: C13, F14.

Keywords: structural gravity model, panel data, multilateral resistance, biproportional projection method.

Acknowledgements: I would like to thank an anonymous referee for Economic Systems Research who made me realise the strong similarity between structural gravity's fixed point problem and biproportional projection techniques. The uniqueness result presented here is a direct transposition of results from this field of input-output analysis. I am also grateful to the colleagues from the Joint Research Centre Paolo Paruolo and Gianluca Mazzarella for their comments.

Contact: Aurélien Poissonnier (aurelien.poissonnier@ec.europa.eu), European Commission, Directorate-General for Economic and Financial Affairs, and Crest-LMA.

CONTENTS

LIST OF TABLES

LIST OF BOXES

1. INTRODUCTION

Similarly to the attraction of masses, gravity models of trade explain bilateral exchanges as an increasing function of the trading economies' size and a decreasing function of their distance.

The more theoretically sound structural gravity models have been developed, motivated by the empirical success of simpler gravity models for bilateral trade (Head & Mayer, 2014).

In structural gravity models, one important departure from the analogy with Newtonian gravity is the *multilateral resistance terms* (MRT) capturing general equilibrium forces. As underlined by Anderson & van Wincoop (2003), the more a country is resistant to trade with one country, the more it shall trade with the others (including itself) as a general equilibrium effect. Omitting these MRT in the estimation strategy is what (Baldwin & Taglioni, 2007) call the *gold medal error* in this literature as it puts into the residual, terms which are by construction correlated to the explanatory variables.

The technical issue is that the missing variables depend on the estimation outcome. (Baldwin & Taglioni, 2007) overcome this difficulty with dummies to control for these missing variables. The limitation of the dummy variable solution is that it masks the effect of all variables that are only country, time or even country-time or country pair specific. One can only estimate the effect of factors which are varying with time and exporter and importer. For this reason, one may be interested in an explicit solution to the MRT.

In this paper, I provide solutions to this problem and show that solving for the MRT explicitly is a valid estimation strategy, including for panel data (contrary to what is suggested by Baldwin & Taglioni, 2007). In addition, I provide some new insights about the model drawing a parallel between the MRT and the biproportional projection of matrices in the Input-Output literature.

I propose five Stata routines to embed the explicit solution to the MRT in the regression. My preference goes to the RAS technique borrowed from the Input-Output literature. This technique has the unique advantage of not requiring a measure for trade with self and not using a (poor) proxy for total production and consumption by country. Other solutions extend the SILS solution (Head & Mayer, 2014) to panels, or use a standard Newton's algorithm. They require some approximations and arguably strong assumptions, but a comparison exercise on real data indicates that they confirm a much contrasted picture to the one given by naïve estimation (where the MRT are simply forgotten) and to the one controlling for MRT based on dummies. In short, omitting the MRT in the estimation strongly biases the effects of all kinds of bilateral and multilateral trade agreements. The differences with an estimation using dummies to control for the MRT raises new questions on the structure of the model estimated with dummies or important missing variables that only the dummy approach would account for.

2. THE STRUCTURAL GRAVITY MODEL

Similarly to the attraction of masses, gravity models of trade explain bilateral exchanges as an increasing function of the trading economies' size and a decreasing function of their distance.

Structural gravity can be put into the general form of Equations 2.1 and 2.2 ([1]).

$$X_{nit} = \frac{Y_{it}}{\Omega_{it}} \frac{X_{nt}}{\Phi_{nt}} \varphi_{nit} \quad 2.1$$

$$with\ \Phi_{nt} = \sum_{l} \frac{\varphi_{nlt} Y_{lt}}{\Omega_{lt}}\ and\ \Omega_{it} = \sum_{l} \frac{\varphi_{lit} X_{lt}}{\Phi_{lt}} \quad 2.2$$

with country *n* the importer and country *i* the exporter, *t* the date. Xnit is the export from *i* to *n* at time *t*. Y_{it} is the exporter's production at time t. X_{nt} is the importer's total expenditure at time t. φ_{nit} is the bilateral accessibility term. Ω_{it} and Φ_{nt} are multilateral trade resistance terms (MRT).

The notion of distance (captured by φ_{nit}) goes beyond its physical definition: factors such as the existence of a common border, cultural factors (common language), historical factors (colonial links, conflicts) or participation to a common free trade agreement, custom union, monetary union refine this notion of distance between the two economies. For this reason φ_{nit} and consequently Ω_{it} and Φ_{nt} are time-varying.

The shift from gravity to structural gravity models, i.e. the inclusion of the MRT Ω_{it} and Φ_{nt} in the model, has motivated a change in the estimation strategies to account for missing variables in the original specifications (Anderson & van Wincoop, 2003; Baldwin & Taglioni, 2007).

These missing variables are correlated to the others (φ_{nit}) through a non-trivial fixed point problem described by Equations 2.2. Imbedding a solution to the fixed point problem in the estimation procedure has been proposed by (Anderson & van Wincoop, 2003; Head & Mayer, 2014).

The fixed point formulation of the problem

For a given set of φ_{nit} and dropping the time index for simplification, but without loss of generality as years can be treated independently, the problem is to find

$$\Omega_i = \sum_{l} \frac{\varphi_{li} X_l}{\Phi_l}\ and\ \Phi_n = \sum_{l} \frac{\varphi_{nl} Y_l}{\Omega_l} \quad 2.3$$

or,

$$\Omega_i = \sum_{l} \frac{\varphi_{li} X_l}{\sum_j \frac{\varphi_{lj} Y_j}{\Omega_j}}\ and\ \Phi_n = \sum_{l} \frac{\varphi_{nl} Y_l}{\sum_j \frac{\varphi_{jl} X_j}{\Phi_j}} \quad 2.4$$

which are two symmetrical fixed point problems.

This problem can also be written with vector and matrix notations

$$\Omega = \varphi^t (diag\Phi)^{-1} X\ and\ \Phi = \varphi (diag\Omega)^{-1} Y \quad 2.5$$

([1]) Notations are taken from the Handbook chapter by Head & Mayer (2014).

with X, Y, Ω and Φ the vectors of Y_n, X_i, Ω_i and Φ_n respectively, and φ the matrix of φ_{ni} with *n* the row index and *i* the column index. *diag* is the operator which transforms a vector into a diagonal matrix.

An important constraint

Equations 2.5 can be written $\mathbb{1} = (diag\Omega)^{-1}\varphi^t(diag\Phi)^{-1}X$ and $\mathbb{1} = (diag\Phi)^{-1}\varphi(diag\Omega)^{-1}Y$ with $\mathbb{1}$ a vector of ones of appropriate size. After some substitutions it yields $\mathbb{1}^t\mathrm{Y} = \mathrm{X}^t\mathbb{1}$, i.e. the sum of all productions across the world is the sum of all expenditures, reminding that the supply and use constraint is at the root of the bilateral trade resistance terms. This implies that our fixed point problem will have no solution if X and Y do not satisfy this constraint.

A simplifying approach

The fixed point problem is an imbricated one: $a = f(b)$ and $b = g(a)$. If X = Y and φ is symmetric then $f = g$ and there is a simpler way to find a solution. Indeed any solution to $a = f(a)$ *is a fortiori* a solution to $a = f(f(a))$. This is the approach of (Anderson & van Wincoop, 2003). There are *a priori* more solutions to the imbricated problem than to the simplified one, especially when f is decreasing as here ([2]). Property 3 (below) implies that it is not the case here. Yet I consider the imbricated problem since symmetry is not verified *a priori* ([3]).

The biproportional reading of the problem

One way to look at this issue is related to the abundant literature on Input-Output analysis and the projection of matrices based on biproportional methods (first of which RAS). This reading of the problem offers both new techniques and theoretical results but also another interpretation of the model's structure ([4]).

Formally the RAS projection problem is, starting from an initial matrix A, to find two vectors R and S in order to rescale the rows and columns of A respectively. *Ex post* the columns of the new matrix should sum to a given Y and its rows to some X:

$$\hat{A} = diagR\ A\ diagS$$

$$\text{verifying } X = \hat{A}\,\mathbb{1} \text{ and } \mathrm{Y}^t = \mathbb{1}^t\hat{A} \qquad 2.6$$

We can exactly transpose this problem to our framework by replacing A by φ, R by the vector of $\frac{X_n}{\Phi_n}$, S by $\frac{Y_i}{\Omega_i}$. Then $\hat{A}$ is the matrix of $\left[\frac{Y_i}{\Omega_i}\frac{X_n}{\Phi_n}\varphi_{ni}\right]_{n,i}$ and the constraints in Equation 2.6 are actually a simple reformulation of Equation 2.5. The parallel with this literature provides a comforting uniqueness result (Property 3) but also background knowledge on the convergence of algorithms to solve the problem.

Putting the problem is those terms, the choice of GDP for (X_n, Y_i) in the estimation has a completely different relevance. The parallel with the biproportional projection technique shows that the MRT simply ensure that the matrix φ is rescaled to sum to the vectors X in rows and Y in columns. The terms (X_n, Y_i) are not meant to capture in the estimation of Equation 2.1 the production and expenditure of each economy but to provide a benchmark for the total trade flows from or to an economy in Equations 2.2. In the first case, GDP may be a fine proxy. In the second, using GDP is a rather heroic assumption.

([2]) e.g. f(x)=-x (resp. 1/x) has only 0 (resp. 1 and -1) as a fixed point but f(f(x))=x has an infinity of solutions.
([3]) In particular, being a member of a multilateral trade agreement may have a differentiated effect on exports and imports, which violates symmetry.
([4]) See Miller & Blair (2009), Chapter 7, for a textbook entry to this abundant literature. Bacharach (1970) already provides a uniqueness and convergence result.

Indeed gravity models are estimated with missing reporters, without information of trade with self and usually data on trade in goods only, while GDP covers both goods and services, include trade with self and non-reporting countries and exclude intermediate consumption.

As regards missing trade flows, the biproportional reading of the problem clearly mitigates the importance of this issue. One possibility is to stick to trade with self and missing reporters being out of the model and simply rescale φ on the total exports and imports of each country as reported in the dataset. This method is easier to implement, does not require assumptions for missing trade flows, rescales matrix φ on a correct measure of the proper concept, and makes consistent assumptions between the estimation of 2.1 and the solution of Equations 2.2.

To understand the implications of this direct approach, we shall consider two possible cases. In the first case, missing data correspond to outliers. Had the data been available, one may have controlled for them in the estimation procedure (using dummies for instance because trade with self has different determinants than those included in the regression). This case is an issue neither for the estimation of Equation 2.1 nor for the solution to Equations 2.2. The hypothetical effect on the overall estimation procedure would be that the dummies controlling for those cells in φ would adjust exactly to the right trade flow (after rescaling). In this case the hypothetical full matrix would be benchmarked on the full totals when the matrix with missing flows is benchmarked on the available totals. Thus, solving for the fixed point problem on a matrix with missing flows is a way to compute the solution to a more general problem when missing flows are off the model.

In the second case, censorship is a constraint. Had the data been known, they would have been described by the same model. This case is an issue for the estimation of Equation 2.1 but not for the solution to Equations 2.2. In the absence of some trade flows, one would wish to rescale a completed φ to some adjusted totals. These totals being unknown one would approximate them in proportion of the available information on trade flows relative to the information on φ. A refined way of computing the missing trade flows in proportion of the φ_{ni} *give*n by the model for those cells, is by looking at the total trade on either side of the flow, i.e. total imports in rows and total exports in column. This is precisely the transformation that the RAS technique estimated without the missing trade flows, would imply on the projection of the missing φ_{ni}. In other words, if some relevant trade flows are missing, with the fixed point solution they are implicitly treated as if they were following the estimated model.

The bottom line is, regardless of the way missing data are treated in the estimation of the model, the fixed point problem can be solved on the available information.

The analogy with the biproportional method also sheds a new light on the treatment of the residuals. If one thinks that residuals should be rescaled together with the estimated φ, then no rescaling is needed since they are altogether equal to each trade flow, so *a fortiori* sum in rows and columns to the total imports and exports. In this case, the MRT can be ignored: there is no *gold medal mistake*. This is also the case if the model *perfectly* replicates the data and by continuity, the better is the model, the less serious is the *gold medal mistake*. On the other hand, rescaling φ excluding residuals implies that residuals in rows and columns sum to zero *ex post*. This can also be directly achieved by introducing importer-year and exporter-year dummies. However, as noted by (Fally, 2015), this is true when the model is estimated under its multiplicative form (for instance with the Poisson PML estimator proposed by Santos Silva and Tenreyro, 2006). However, large sets of dummies applied to the log linearized model do not ensure full consistency with the structural gravity models ([5]).

([5]) This approach ensures that the residuals in rows and columns are on average equal to 0% and not 0€ or 0$.

Properties of the solution

Property 1: if (Ω, Φ) is solution to the problem, so is $\left(\theta\Omega, \frac{\Phi}{\theta}\right), \forall\theta \in \mathbb{R}$.

This property implies that $\Omega = 0$ and $\Phi = \infty$ or vice versa, although degenerate, are solutions to our problem. Note that this line of solutions is not problematic for the econometric strategy has the product $\Omega_i\Phi_n$ is independent from θ. Therefore (Baldwin & Taglioni, 2007)'s point is correct: there is a normalisation issue with the fixed point solution; but their argument against (Anderson & van Wincoop, 2003) is not: all normalisations are equivalent ([6]). Solving for the fixed point problem explicitly is thus a valid approach, including with panel data.

Property 2: If (Ω, Φ) is solution to the problem based on matrix φ and vectors (X, Y), it is also solution to the problem based on $\theta\varphi \; or \; (\theta X, \theta Y)$.

The property is clear from Equations 2.4. It implies that whether X, Y and Xin are measured in $, million $, billion $ or any other currency unit yields the same solution.

Property 3: If there is no group of countries operating in autarky (i.e. matrix φ is not block diagonal or cannot be made block diagonal by permutations of its rows and columns) then the pair of MRT which excludes negative trade flows is unique (to a normalisation constant).

This property ensures that the overall estimation procedure is not ambiguous. The demonstration of this property is adapted from the biproportional matrix projections literature and is exposed in appendix B.

When the matrix is block diagonal (in a panel case, each year is a separate block) there is a unique pair of MRT (to a normalizing constant) for each year. As the normalizing constants can be arbitrarily set to different values for different years, there is no unique MRT for the full panel but the products term by term $\Omega_i\Phi_n$ are unique within each year due to Property 3. This is a strong enough result to ensure the soundness of the procedure.

([6]) The real issue is whether this line of solutions is unique on $\mathbb{R}^{+n}$ as if it is not, the method cannot be used at all (panel or cross section). Property 3 clears this uncertainty.

3. PROPOSED ALGORITHMS

Contraction mapping

(Head & Mayer, 2014) propose an iterative solution to the imbricated problem: Structurally Iterated Least Squares (SILS). They start from $\Omega^0 = \mathbb{1}$ and $\Phi^0 = \mathbb{1}$ and a first estimation of 2.1 which provides φ^0. *A simple* contraction mapping on Equations 2.2 (plugging the result of the first equation into the second and its outcome into the first…) allows to compute the updates Ω^1 and Φ^1, to then use them in a new estimation of 2.1 and find an update φ^1, etc. until convergence is obtained.

A priori, this strategy may not converge because the function for which to find a fixed point is not necessarily well behaved. In practice, simulations show that convergence is obtained. This observation is directly related to the proximity with the RAS technique. Experience from the RAS practitioners shows that convergence is generally obtained in a few dozen iterations. Convergence is theoretically ensured unless matrix φ is particularly problematic, for instance by having too many zeros (Miller & Blair, 2009). One of the advantages of this contraction mapping is that starting from $\mathbb{R}^{+n}$ ensures that if a solution is obtained, it is on $\mathbb{R}^{+n}$ as well. Also, if in two sequences Ω^n and $\Omega^{n\prime}$ only differ by a normalisation factor, Ω^{n+1} and $\Omega^{n\prime+1}$ will differ by this same factor. As a consequence, forcing the contraction mapping to stay on the unit hypersphere (i.e. forcing the solution's normalisation from Property 1) does not alter nor accelerates the contraction mapping algorithm.

Newton algorithm

This strategy is robust to the possibility of the function $f_\omega(\Omega) = \varphi^t\left(diag(\varphi(diag\Omega)^{-1}Y)\right)^{-1}X$ not being contracting and has a faster convergence in theory. Newton's algorithm replaces the contraction mapping in the SILS algorithm above. Its update follows $\Omega^{k+1} = \Omega^k - (I - f'_\omega(\Omega^k))^{-1}(\Omega^k - f_\omega(\Omega^k))$ with $f'_\omega(\Omega) = \varphi^t(diag\Phi)^{-2}diagX\varphi(diag\Omega)^{-2}diagY$ the derivative of $f_\omega(\Omega)$. This algorithm unfortunately converges to zero, which is a fixed point for our problem. A solution to this unfortunate convergence is to solve the following $\Omega = f_\omega(\Omega)/\|f_\omega(\Omega)\|$ which provides solutions to the same fixed point problem restricted to the unit hypersphere. Defining $\widetilde{f_\omega}(\Omega) = f_\omega(\Omega)/\|f_\omega(\Omega)\|$ Newton's algorithm's update is $\Omega^{k+1} = \Omega^k - \left(I - \widetilde{f'_\omega}(\Omega^k)\right)^{-1}\left(\Omega^k - \widetilde{f_\omega}(\Omega^k)\right)$ with $\widetilde{f'_\omega}(\Omega^k) = \frac{f'_\omega(\Omega)}{\|f_\omega(\Omega)\|} - \frac{f_\omega(\Omega)\Omega^t}{\|f_\omega(\Omega)\|^3}$.

RAS

This method is directly transposed from the Input-Output literature. Since the multilateral resistance terms are such that matrix $\left[\frac{Y_i}{\Omega_i}\frac{X_n}{\Phi_n}\varphi_{ni}\right]_{n,i}$ sums in rows and columns to the same totals as matrix $[X_{ni}]_{n,i}$, I do not compute Ω and Φ as verifying Equations 2.2 but more simply the rescaling vectors $R = (X_n/\Phi_n)_n$, $S = (Y_i/\Omega_i)_i$ (7) . The RAS iterative procedure replaces the contraction mapping. It first rescales the rows of φ to the total imports in the dataset, then rescales the outcome's columns to the total exports, and so on until convergence. R is then the cumulated rescaling of the rows while S is the cumulated rescaling of the columns.

This method is valid on rectangular matrices (when some countries are not reporting exports or imports) and the outcome in terms of S and R is unique. Similarly to the contraction mapping, starting from $\mathbb{R}^{+n}$ ensures that the solution is obtained on $\mathbb{R}^{+n}$ as well.

(7) The sole difference between the contraction mapping and the RAS method is to search for a solution in terms of (R,S) instead of (Ω, Φ).

Uniqueness of the solution to the fixed point problem ([8])

Uniqueness of the solution to the fixed point problem is a crucial property for the validity of the empirical strategy. If multiple solutions arose, there would be no way of choosing the *true* one to properly control for missing variables in the regression. Multiplicity of the solutions would allow for sun-spot fluctuations, a theoretically interesting property but an empirical curse.

For the 2 country case we can find the formal solution to the non-imbricated problem (see Appendix). There are four solutions, one on each quadrant of the $\mathbb{R}^2$ plan, i.e. only one in $\mathbb{R}^{+2}$ which is the solution of interest. Without symmetry this approach is no longer suitable and only a numerical solution can be found.

As the solutions to the imbricated problem come in lines crossing the origin of axes (Property 1), one may look for solutions for the two country case on the unit circle. Numerical investigations show that there is only one such solution on $\left[0, {}^{\pi}/_{2}\right]$. Uniqueness of the solution line on $\mathbb{R}^{+3}$ is verified for three countries as well (I use a double loop and polar coordinates to map the positive eighth of the unit sphere). For higher dimensions, instead of mapping a higher dimension set, I run both algorithms starting from draws in a uniform distribution. As predicted by Property 3, all solutions when convergence is obtained in $\mathbb{R}^{+n}$ are identical.

Comparison of fixed point solution algorithms in simulations ([9])

For *nice* matrices (no empty cell in particular), convergence is very fast, even for a number of countries in the hundred. In addition, convergence is of comparable speed for the contraction mapping and Newton algorithms. If the contraction mapping naturally stays on positive solutions, Newton's algorithm may lean toward negative ones. If the solution found is simply the opposite of the one wanted, this normalisation factor is not an issue. However, if the solution has both positive and negative coordinates, it must be rejected (as it would predict negative flows).

Constraining the contraction mapping to stay on the unit circle does not accelerate convergence as anticipated. For Newton's algorithm, not controlling for the norm of the solution yields a solution equal to 0. Modifying the algorithm to solve the fixed point problem on the unit hypersphere avoids this issue. However forcing the algorithm solution to stay at each step on this sphere prevents convergence by altering the recurrence. For this reason, renormalizing must be avoided with Newton's algorithm but can be used with the contraction mapping.

The RAS algorithm being just a reformulation of the contraction mapping is just as fast. Not having data for trade with self or not the same list of countries as importers and exporters is, as expected, not an issue.

Comparison of algorithms on real data ([10])

I test the algorithms on the small dataset of (Head, Mayer, & Ries, 2010). The solutions are reported in Table 3.1 together with the relative time for convergence of the iterative solution I propose and the number of iterations needed to obtain convergence.

The algorithms consist of two loops, an outer loop for the convergence of the estimation (common to all solutions) and an inner loop for the resolution of the fixed point problem (with different solutions). For each iteration of the outer loop, the inner loop is initialized with the solution from the previous outer loop. Hence, as the estimation converges, the solution to the fixed point problem is obtained faster.

([8]) Matlab code available upon request
([9]) Matlab code available upon request
([10]) Stata-Mata code available upon request

Table 3.1: **Gravity model estimation on real data**

	no MRT	Contract. mapping Stata	Contract. mapping Mata	Newton	RAS Stata	RAS Mata	FE	PPML
distance	**-1.304*** **	**-1.123*** **	**-1.123*** **	**-1.124*** **	**-1.417*** **	**-1.417*** **	**-1.764*** **	**-0.879*** **
	(0.0205)	(0.0238)	(0.0238)	(0.0238)	(0.0204)	(0.0204)	(0.0219)	(0.0143)
regional trade agreement	**0.822*** **	**2.112*** **	**2.112*** **	**2.111*** **	**1.599*** **	**1.599*** **	**0.433*** **	**0.621*** **
	(0.0505)	(0.0567)	(0.0568)	(0.0567)	(0.0483)	(0.0483)	(0.0515)	(0.0302)
gatt_b	**0.335*** **	**1.965*** **	**1.965*** **	**1.961*** **	**3.159*** **	**3.159*** **	**0.557*** **	**0.529*** **
	(0.0320)	(0.0375)	(0.0375)	(0.0376)	(0.0339)	(0.0339)	(0.0955)	(0.101)
Currency union	**0.409** **	**-0.0980**	**-0.0980**	**-0.0987**	**0.0825**	**0.0825**	**0.266* **	**0.0130**
	(0.131)	(0.151)	(0.151)	(0.151)	(0.117)	(0.117)	(0.135)	(0.0285)
Constant	**-8.864*** **	**-52.30*** **	**-52.29*** **	**-52.28*** **	**-17.83*** **	**-18.93*** **		
	(0.178)	(0.208)	(0.208)	(0.208)	(0.180)	(0.180)		
Relative speed		3.45	1.42	68.63	1.04	1.00		
Iterations of outer loop		112	112	121	29	30		

Standard errors in parentheses, * p<0.05, ** p<0.01, *** p<0.001
Source: (Head, Mayer, & Ries, 2010), 2000 to 2006, 137552 observations

Within the outer loop, the estimation of the gravity model is a simple regression with country pair clustering. One may easily modify this step to a Poisson model for instance or possibly account for censorship.

The convergence criterion of the inner loop is expressed in relative terms and normalised by the number of trade flows (so that convergence is robust to changes in the sample size). The convergence criterion can either be expressed in terms of norm one (absolute values) or norm two (quadratic, case reported here). A modification with respect to (Head & Mayer, 2014) is that convergence is not that of Ω and Φ as separate vectors but of matrix of $\Omega\,\Phi^t$ in logarithm, that is the coefficients of interest for the gravity model and the outer loop.

The first solutions elaborate on the companion code from (Head & Mayer, 2014). Modifications are made to treat panel data. I also introduce assumptions to fill the missing cells in matrix φ, in particular the diagonal elements of trade with self. The diagonal elements are filled to be slightly higher than the elements on the corresponding line and column (to ensure that accessibility to self is higher than accessibility to any other trade partner). For non-reported trade flows between otherwise reporting countries, the underlying assumption is that the flow is null so that the corresponding cell is as well. Another possibility is commented within the code. Assuming the panel is censored, the alternative would be to compute the model forecasted value for the cell to solve the fixed point problem including for non-reported trade pairs. This possibility requires that the dataset includes explanatory variables for missing trade flows. In Mata I make an extensive use of bloc diagonal matrices to treat all years at the same time ([11]). The procedure is robust to not all countries being present each year. The three methods treat years orthogonally but within the same loop (no need for a loop over each year). There are also possibilities of rescaling in the Mata algorithms (Property 2). This rescaling can easily be adapted and may be used when φ tends to be very small to avoid numerical precision issues.

([11]) Running the corresponding lines of codes in a *sandbox* with a small number of years and countries is probably more informative than long textual explanations.

Building on the Input-Output projection techniques, I also implement the RAS procedure both in Stata and Mata. Contrary to the generalisation of the SILS method above, I do not make assumptions on trade with self, and use reported total exports and imports instead of GDP as vectors (X_n, Y_i). As I exposed above, this approach limits the number of assumptions and remains theoretically consistent.

Finally, I use three estimations as references, first a naïve estimation where the *gold medal error* is made, an estimation correcting for these missing variables by use of dummy variables for exporter-year and importer-year (Guimarães & Portugal, 2010; Correia, 2014) and a Poisson Pseudo-Maximum Likelihood estimator (Guimarães, 2014) advocated by (Santos Silva and Tenreyro, 2006).

On the real data example, convergence is quite fast for the contraction mapping and RAS algorithms, including without resorting to Mata. As with the Newton algorithm, the cost of computing the inverse of a matrix largely outweighs the gains of a theoretically powerful algorithm.

A first contraction mapping (the most directly linked to SILS by (Head & Mayer, 2014), second column of Table 3.1) relies solely on Stata, however using Mata (third column) makes a better use of faster matrix computing (more than doubles the speed). Using Mata is also necessary for Newton's algorithm which requires the computation of a first derivative matrix. This algorithm (fourth column) converges only in almost 2 hours. This algorithm despite its theoretical soundness has two major drawbacks, first the computation of the inverse of a matrix in the inner loop is time consuming, but also it can converge towards non positive solutions which can be penalized but slows down convergence even more.

The solutions based on the RAS routine (fifth and sixth columns) are faster than the SILS algorithm (which needs extra steps to implement some assumptions on trade with self in particular). Based on Stata, convergence is obtained only 4% more slowly than with Mata. More interestingly, the number of iterations needed to obtain convergence is much less with the RAS routine than with the SILS routine (30 against 112 in Mata).

For these reasons the contraction mapping and RAS techniques using Mata may be preferred (keeping in mind that the first makes some assumptions which the second avoids).

Estimation results

Comparing the estimation results with a regression without accounting for the multilateral resistance terms (Table 3.1, first column) shows the importance of these terms ([12]) for the estimation and confirms the recommendation of (Baldwin & Taglioni, 2007) to correct for the *gold medal mistake* in gravity model estimation. In addition, all three algorithms adapted from SILS on the one hand and the two algorithms implementing the RAS method converge toward the same estimation.

With the iterative solutions proposed here, the effect of regional trade agreement or participation to the GATT-WTO are considerably larger. Also, the effect of being in a currency union is found null once corrected for the *gold medal mistake*. This result in particular contrasts with the 2WFE estimation. This estimation with dummies is however consistent with (Glick & Rose, 2016) latest findings with the same method and comforts the past estimates of a large effect of being in the EMU. By construction, this technique is both quite agnostic about the structure of the model and also econometrically powerful. Correcting for the *gold medal mistake* in the non-agnostic way proposed here, which is more linked to the specificities of structural gravity models, conveys a radically different message: the effect of being in a currency union on trade is more limited than usually argued. Fally (2015) argues that the PPML estimate with dummies is consistent with the structural gravity model while the 2WFE is not. In line with the method I propose, this estimation also finds a null effect on trade of being in a currency union. These

([12]) From (Head, Mayer, & Ries, 2010) I take only the distance, regional trade agreement, GATT and currency union variables. Sample starts in 2000 and regression includes country pair clustering.

quite contrasted estimation outcomes suggest that either the 2WFE technique estimates a model other than structural gravity or it controls for important variables missing in the other estimations.

Comparing the RAS estimation with the other estimation methods controlling for the MRT shows the effect of using GDP as a proxy for total consumption and production and making assumptions for bilateral resistance to trade with self. These assumptions have a secondary but sizeable effect on the estimation in particular of the effect of regional trade agreements and participation to the WTO.

4. CONCLUSION

The main purpose of the present paper is to introduce new tools to correct for the *gold medal mistake* in gravity models (missing multilateral resistance terms, MRT).

Solving for MRT explicitly is proved to be a valid approach for the estimation of structural gravity models, including with panel data. In addition, the parallel between the MRT and the biproportional projection of matrices offers some new insights about the model. In particular, missing trade flows may be a problem for the estimation but are not for computing the multilateral resistance terms.

Several algorithms are proposed based on Stata and Mata. Out of the five proposed algorithms, the contraction mapping (adapted from Head & Mayer, 2014) and RAS (taken from the Input-Output literature) using Mata perform the best.

An estimation example on real data confirms that solving for the *gold medal mistake* is quite influential on the estimated trade effect of some of the usual suspects. In addition, the method proposed seems to challenge the alternative solution to the *gold medal mistake.* The approach based on high dimension sets of dummies is efficient but agnostic. These comparatively quite different estimation outcomes suggest that this technique estimates a model other than structural gravity (a claim which is theoretically sound) or controls for important variables which are otherwise missing.

Given the limited size of the sample used for the example (2000-2006) and the non-inclusion of other relevant variables and controls, the numerical results must be seen as illustrative and not fully tested for robustness.

REFERENCES

Anderson, J. E., & van Wincoop, E. (2003). Gravity with Gravitas: A Solution to the border puzzle. The American Economic Review, 170–192.

Bacharach, M. (1970). Biproportional Matrices and Input-Output Change. Cambridge: Cambridge University Press.

Baldwin, R., & Taglioni, D. (2007, December). Trade Effects of the Euro: a Comparison of Estimators. Journal of Economic Integration, 22(4), 780-818.

Correia, S. (2014). REGHDFE: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects. Statistical Software Components S457874, Boston College Department of Economics, revised 25 Jul 2015.

Dietzenbacher, E., & Miller, R. E. (2009, August). Ras-ing the transactions or the coefficients: It makes no difference. Journal of Regional Science, 555-566. doi: 10.1111/j.1467-9787.2008.00598.x

Fally, T. (2015). Structural gravity and fixed effects. Journal of International Economics, 97, 76-85, doi: http://dx.doi.org/10.1016/j.jinteco.2015.05.005

Glick, R., & Rose, A. K. (2016). Currency Unions and trade: a post-EMU reassessment. European Economic Review. doi: http://dx.doi.org.10.1016/j.euroecorev.2016.03.010

Guimarães, P. (2014). POI2HDFE: Stata module to estimate a Poisson regression with two high-dimensional fixed effects, Statistical Software Components S457777, Boston College Department of Economics, revised 16 Sep 2016.

Guimarães, P., & Portugal, P. (2010). A simple feasible procedure to fit models with high-dimensional fixed effects. The Stata Journal, 10(4), 628–649.

Head, K., & Mayer, T. (2014). Gravity equations: Workhorse, toolkit, and cookbook. In E. Helpman, K. Rogoff, & G. Gopinath (Eds.), Handbook of International Economics (Vol. 4, pp. 131-195). Oxford: North Holland.

Head, K., Mayer, T., & Ries, J. (2010). The erosion of colonial trade linkages after independence. Journal of International Economics, 1-14. doi:10.1016/j.jinteco.2010.01.002

Miller, R. E., & Blair, P. D. (2009). Input–Output Analysis Foundations and Extensions. Cambridge : Cambridge University Press.

Santos Silva, J.M.C. & Tenreyro, S. (2006), The Log of Gravity. The Review of Economics and Statistics, 88(4), pp. 641-658.

ANNEX 1

Some intuitions from a two country simple solution

I consider two countries (a and b) and the problem in a symmetric from as (Anderson & van Wincoop, 2003), that is X=Y and $\varphi_{ni} = \varphi_{in}$ or equivalently matrix φ is symmetric ([13]). The symmetry implies that a sufficient condition for a solution to exist, is that it solves

$$\Phi_i = \sum_l \frac{\varphi_{li} X_l}{\Phi_l} \tag{1}$$

a simpler, non-imbricated fixed point problem. It is then true that $\Omega_i = \Phi_i$.

In the two country case, there are four real solutions to the problem in the four quadrants of the $\mathbb{R}^2$ plan. The system of two equations is

$$\begin{aligned} \Phi_a &= \frac{\varphi_{aa} X_a}{\Phi_a} + \frac{\varphi_{ba} X_b}{\Phi_b} \\ \Phi_b &= \frac{\varphi_{ab} X_a}{\Phi_a} + \frac{\varphi_{bb} X_b}{\Phi_b} \end{aligned} \tag{2}$$

and its solutions in $\mathbb{R}^{*2}$ are

$$\begin{aligned} &{\Phi_a}^+ \\ &= \pm \sqrt{\frac{\varphi_{ba} X_b - \varphi_{ab} X_a + 2\frac{\varphi_{bb}\varphi_{aa}}{\varphi_{ba}} X_a + \sqrt{(\varphi_{ba} X_b - \varphi_{ab} X_a)^2 + 4\varphi_{bb}\varphi_{aa} X_a X_b}}{2\frac{\varphi_{bb}}{\varphi_{ba}}}}; \\ &{\Phi_b}^+ \\ &= \pm \sqrt{\frac{\varphi_{ab} X_a - \varphi_{ba} X_b + 2\frac{\varphi_{aa}\varphi_{bb}}{\varphi_{ab}} X_b + \sqrt{(\varphi_{ba} X_b - \varphi_{ab} X_a)^2 + 4\varphi_{bb}\varphi_{aa} X_a X_b}}{2\frac{\varphi_{aa}}{\varphi_{ab}}}} \end{aligned} \tag{3}$$

$$\begin{aligned} &{\Phi_a}^- \\ &= \pm \sqrt{\frac{\varphi_{ba} X_b - \varphi_{ab} X_a + 2\frac{\varphi_{bb}\varphi_{aa}}{\varphi_{ba}} X_a - \sqrt{(\varphi_{ba} X_b - \varphi_{ab} X_a)^2 + 4\varphi_{bb}\varphi_{aa} X_a X_b}}{2\frac{\varphi_{bb}}{\varphi_{ba}}}}; \\ &{\Phi_b}^- \\ &= \mp \sqrt{\frac{\varphi_{ab} X_a - \varphi_{ba} X_b + 2\frac{\varphi_{aa}\varphi_{bb}}{\varphi_{ab}} X_b - \sqrt{(\varphi_{ba} X_b - \varphi_{ab} X_a)^2 + 4\varphi_{bb}\varphi_{aa} X_a X_b}}{2\frac{\varphi_{aa}}{\varphi_{ab}}}} \end{aligned} \tag{4}$$

([13]) This is not necessarily the case as the effect of being in a multilateral agreement may differ from an importer or an exporter perspective.

Box A1.1: **calculation of the solution**

$$\frac{X_b}{\Phi_b} = \frac{1}{\varphi_{ba}}\left(\Phi_a - \frac{\varphi_{aa} X_a}{\Phi_a}\right)$$

$$\Phi_b = \frac{\varphi_{ab} X_a}{\Phi_a} + \frac{\varphi_{bb}}{\varphi_{ba}}\left(\Phi_a - \frac{\varphi_{aa} X_a}{\Phi_a}\right)$$

$$\varphi_{ba} X_b = \left(\Phi_a - \frac{\varphi_{aa} X_a}{\Phi_a}\right)\left(\frac{\varphi_{ab} X_a}{\Phi_a} + \frac{\varphi_{bb}}{\varphi_{ba}}\left(\Phi_a - \frac{\varphi_{aa} X_a}{\Phi_a}\right)\right)$$

$$\varphi_{ba} X_b = \varphi_{ab} X_a + \frac{\varphi_{bb}}{\varphi_{ba}}{\Phi_a}^2 - 2\frac{\varphi_{bb}\varphi_{aa}}{\varphi_{ba}} X_a - \varphi_{aa}\varphi_{ab}\frac{{X_a}^2}{{\Phi_a}^2} + \frac{\varphi_{bb}\ {\varphi_{aa}}^2}{\varphi_{ba}}\frac{{X_a}^2}{{\Phi_a}^2}$$

$$\alpha{\Phi_a}^4 - \beta{\Phi_a}^2 + \gamma = 0$$

$$\Phi_b = \frac{\varphi_{ab} X_a}{\Phi_a} + \frac{\varphi_{bb}}{\varphi_{ba}}\left(\Phi_a - \frac{\varphi_{aa} X_a}{\Phi_a}\right)$$

$$\varphi_{ba} X_b = \left(\Phi_a - \frac{\varphi_{aa} X_a}{\Phi_a}\right)\left(\frac{\varphi_{ab} X_a}{\Phi_a} + \frac{\varphi_{bb}}{\varphi_{ba}}\left(\Phi_a - \frac{\varphi_{aa} X_a}{\Phi_a}\right)\right)$$

$$\varphi_{ba} X_b = \varphi_{ab} X_a + \frac{\varphi_{bb}}{\varphi_{ba}}{\Phi_a}^2 - 2\frac{\varphi_{bb}\varphi_{aa}}{\varphi_{ba}} X_a - \varphi_{aa}\varphi_{ab}\frac{{X_a}^2}{{\Phi_a}^2} + \frac{\varphi_{bb}\ {\varphi_{aa}}^2}{\varphi_{ba}}\frac{{X_a}^2}{{\Phi_a}^2}$$

$$\alpha{\Phi_a}^4 - \beta{\Phi_a}^2 + \gamma = 0$$

with

$$\alpha = \frac{\varphi_{bb}}{\varphi_{ba}}\ ;\ \beta = \ \varphi_{ba} X_b - \ \varphi_{ab} X_a + 2\frac{\varphi_{bb}\varphi_{aa}}{\varphi_{ba}} X_a\ ;\ \gamma = \left(\frac{\varphi_{bb}\ {\varphi_{aa}}^2}{\varphi_{ba}} - \varphi_{aa}\varphi_{ab}\right){X_a}^2$$

$${\Phi_a}^2 = \frac{\beta \pm \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha}$$

$${\Phi_a}^2 = \frac{\varphi_{ba} X_b - \ \varphi_{ab} X_a + 2\frac{\varphi_{bb}\varphi_{aa}}{\varphi_{ba}} X_a \pm \sqrt{(\ \varphi_{ba} X_b - \ \varphi_{ab} X_a)^2 + 4\varphi_{bb}\varphi_{aa} X_a X_b}}{2\frac{\varphi_{bb}}{\varphi_{ba}}}$$

Note that as long as $\varphi_{ab} = \varphi_{ba} < \varphi_{aa}$ or φ_{bb}, that is trade accessibility to oneself is higher than accessibility to others $\beta > 0$ *and* $\gamma > 0$. In addition, $\beta^2 - 4\alpha\gamma = (\ \varphi_{ba} X_b - \ \varphi_{ab} X_a)^2 + 4\varphi_{bb}\varphi_{aa} X_a X_b$ is always strictly positive, even in the *super symmetry* case where $X_b = X_a$. Hence there are two real and positive solutions to ${\Phi_a}^2$.

The condition for one of the two solutions to be negative (and hence ineligible) is $\varphi_{ab}\ \varphi_{ba} > \varphi_{aa}\varphi_{bb}$.

Let ${\Phi_a}^{2+}$ *and* ${\Phi_a}^{2-}$ denote these solutions. Let's give an idea of their neighbourhood. If $(\ \varphi_{ba} X_b - \ \varphi_{ab} X_a) < 0$, i.e. a is the largest country, ${\Phi_a}^{2+} > \varphi_{aa} X_a$ and ${\Phi_a}^{2-} < \varphi_{aa} X_a + (\ \varphi_{ba} X_b - \ \varphi_{ab} X_a)\frac{\varphi_{ba}}{\varphi_{bb}}$. If $(\ \varphi_{ba} X_b - \ \varphi_{ab} X_a) > 0$, i.e. a is the largest country, ${\Phi_a}^{2+} > \varphi_{aa} X_a + (\ \varphi_{ba} X_b - \ \varphi_{ab} X_a)\frac{\varphi_{ba}}{\varphi_{bb}}$. and ${\Phi_a}^{2-} < \varphi_{aa} X_a$.

For a convergence algorithm, $\Omega_i = \Phi_i = \sqrt{\varphi_{ii} X_i}$ could be a better starting point than 1 which ignores the unit of Ω_i *and* Φ_i ($\$^{½}$).

(Continued on the next page)

Box (continued)

$$\Phi_a{}^2 = \varphi_{aa} X_a + \varphi_{ba} X_b \frac{\Phi_a}{\Phi_b}$$

$$\Phi_b{}^2 = \varphi_{ab} X_a \frac{\Phi_b}{\Phi_a} + \varphi_{bb} X_b$$

$$\Phi_b{}^2 = \frac{\varphi_{ab} X_a \varphi_{ba} X_b}{\Phi_a{}^2 - \varphi_{aa} X_a} + \varphi_{bb} X_b$$

Which is equal by symmetry between the two countries to

$$\Phi_b{}^2 = \frac{\varphi_{ab} X_a - \varphi_{ba} X_b + 2\frac{\varphi_{aa}\varphi_{bb}}{\varphi_{ab}} X_b \pm \sqrt{(\varphi_{ba} X_b - \varphi_{ab} X_a)^2 + 4\varphi_{bb}\varphi_{aa} X_a X_b}}{2\frac{\varphi_{aa}}{\varphi_{ab}}}$$

The pairs of solutions being $\left(\pm\sqrt{\Phi_a{}^{2+}}, \pm\sqrt{\Phi_b{}^{2+}}\right)$ *and* $\left(\pm\sqrt{\Phi_a{}^{2-}}, \mp\sqrt{\Phi_b{}^{2-}}\right)$.

ANNEX 2

Uniqueness of the multilateral resistance terms

As shown in the present paper, the multilateral resistance terms are solutions to a problem which is identical to the biproportional rescaling of matrix φ on the total exports and imports of each country.

The validity of the iterative solution proposed here relies on the uniqueness of the solution to this problem which we can prove thanks to this transposition of the problem.

In the Input-Output literature, (Bacharach, 1970) proves that the rescaled matrix is unique. This does not mean however that the rescaling vectors $R = (X_n/\Phi_n)_n$ and $S = (Y_i/\Omega_i)_i$ (and by extension the MRT) are unique or that their cross product are unique. This is however the Property 3 needed to ensure that there is a unique solution for the MRT included in the estimation process. Without this uniqueness result, the validity of the estimation outcome is questionable.

(Dietzenbacher & Miller, 2009) prove that the rescaling vectors are unique (to a normalizing constant, c.f. Property 1) under the following condition:

(i) φ is a square matrix (i.e. each country is recorded both as an importer and an exporter), (ii) its diagonal elements are strictly positive (i.e. every country trades with itself) (iii) it is not block diagonal (or cannot be made block diagonal by permutations of its rows and columns, i.e. there is no group of countries operating in autarky).

Actually, their result is more general as condition (iii) is enough to ensure the uniqueness of the rescaling vectors. This implies that the estimation procedure can be applied when some countries are not reporting as exporters or importers (rectangular matrix) and when trade with self is unknown (zeros on the diagonal).

We know from (Bacharach, 1970) that the biproportional projection is unique. Let's now assume that there are two different sets of vectors (R1,S1) and (R2,S2) which yield this unique solution, with the (armless) restriction for instance that the first element of R1 and R2 is equal to one. We can prove that R1=R2 and S1=S2.

Recall that the biproportional rescaling is simply the multiplication term by term of φ by M1= R1*S1' and M2=R2*S2' respectively. As a consequence if the projection is unique, then M1=M2 for all non-zero cells of φ (actually (Bacharach, 1970) demonstrates unicity of the rescaled matrix by proving a closely related condition). Since the first element of both R1 and R2 is 1, it follows that for all the non-zero cells in the first row of φ we have S1=S2. Changing rows in these columns, since S1=S2, it must be also the case that R1=R2 for all non-zero cells in these columns. Changing columns in these rows, we find that S1=S2 for all the non-zero cells in these rows…

In the end, this wandering procedure from non-zero cells to non-zero cells along the lines and columns of φ covers all rows and columns if and only if φ is not block diagonal (or cannot be transformed into a block diagonal matrix by permutation of rows and columns).

EUROPEAN ECONOMY DISCUSSION PAPERS

European Economy Discussion Papers can be accessed and downloaded free of charge from the following address:
http://ec.europa.eu/economy_finance/publications/eedp/index_en.htm

Titles published before July 2015 under the Economic Papers series can be accessed and downloaded free of charge from:
http://ec.europa.eu/economy_finance/publications/economic_paper/index_en.htm

Alternatively, hard copies may be ordered via the “Print-on-demand” service offered by the EU Bookshop: http://bookshop.europa.eu.

HOW TO OBTAIN EU PUBLICATIONS

Free publications:

- one copy:
  via EU Bookshop (http://bookshop.europa.eu);

- more than one copy or posters/maps:
  - from the European Union's representations (http://ec.europa.eu/represent_en.htm);
  - from the delegations in non-EU countries (http://eeas.europa.eu/delegations/index_en.htm);
  - by contacting the Europe Direct service (http://europa.eu/europedirect/index_en.htm) or calling 00 800 6 7 8 9 10 11 (freephone number from anywhere in the EU) (*).

  (*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

Priced publications:

- via EU Bookshop (http://bookshop.europa.eu).

KC-BD-16-040-EN-N

Publications Office

ISBN 978-92-79-54452-1