WPS6347

Policy Research Working Paper 6347

Integrating Gravity

The Role of Scale Invariance in Gravity Models of Spatial Interactions and Trade

Jean-François Arvis

The World Bank
Poverty Reduction and Economic Management Network

January 2013

Abstract

This paper revisits the ubiquitous bi-proportional gravity model and investigates the reasons why different theoretical frameworks may lead to the same empirical formula. The generic gravity equation possesses scale invariance symmetries that constrain possible theoretical explanations based on optimal allocation principles, such as neoclassical or probabilistic frameworks. These constraints imply that a representative consumer's utilities must be separable, and that an entropy model is the only consistent maximum likelihood allocation of a matrix of flows between origin and destination. The paper explores the feasibility of wider classes of non-scale invariant gravity equations, where gravity is no longer bi-proportional by including nonlinear interactions between trade costs and fundamental country factors such as economic size. It shows that such extensions are feasible but that they do not result in a significant improvement in the explanatory power of the empirical analysis.

This paper is a product of the International Trade Department, Poverty Reduction and Economic Management Network. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The author may be contacted at jarvis1@worldbank.org.

Integrating Gravity: The role of scale invariance in gravity models of spatial interactions and trade

Jean-François ARVIS[1]
The World Bank

Keywords: Gravity models, international trade, trade costs

JEL Codes: C61, F12

EPOL, TRAN

[1] Jarvis1@worldbank.org, International Trade Department

1 Introduction: Positioning the problem

The gravity model is the workhorse of much empirical work in major areas in the social sciences, including economics, especially trade and transportation economics. It provides an intuitive and very effective way to describe bilateral flows between a series of origins (exporters in the case of international trade analysis) and destinations (importers). A flow between i and j takes the form:

Flow i $\rightarrow$ j = "size" of i
x the "size" of j
x function of variables measuring the separation between i and j, such as geographical distance or the cost of transport

The initial models (e.g., Tinbergen) took a rather simple form, with explicit reference to the Newtonian gravitational interaction proportional to each mass and inversely proportional to the square of distance:

$$F_{ij} = \frac{M_i M_j}{Dij^2},$$

In basic economic applications this translates into:

$X_{ij} = \frac{GDP_i * GDP_j}{D_{ij}^{\alpha}}$, where the exponent α is econometrically estimated.

Successor models have used more sophisticated approaches to estimate size and more separation variables in addition to distance (see Section 2).

The tremendous empirical success of gravity modeling in economics has led many authors to look for a theoretical foundation. Pioneers such as Zipf and Savage proposed heuristic explanations, often based on probability theory (random matching between groups of exporters and importers of various size), but could not properly account for the separation effects in terms of cost or distance (Savage & Deutsch 1960); Deardorff 1998).

In the context of modeling transportation and mobility flows, Wilson (1970) proposed an elegant and entirely consistent formulation based not on the paradigm of Newtonian gravity but on the canonical ensemble of statistical physics. He postulated that the most probable flow configuration maximizes entropy subject to budget constraints, and derived an implicit formula for flows that includes separation costs (Wilson 1970) (Roy & Thill 2003).

How trade gravity modeling emerges from trade theory has been the object of intense investigation following the influential work by Anderson and Wincoop (2003), itself following a much earlier proposal by Anderson himself (Anderson 1979)and Bergstrand. As noted by several observers, the problem is that there are almost too many successful explanations (De Benedictis & Taglioni 2010)(James E. Anderson, 2011). However, as recently observed by Novy (2012), at the core of most neoclassical explanations of trade gravity is the notion that a representative consumer will source his or her imports according to destination, following the Armington hypothesis, with his or her preferences derived from a CES utility.

This paper essentially reverses the problem. It does not try to explain the structure of the gravity equation from an economic model but rather looks at how the formal structure of gravity constrains the set of models that can be used to generate the gravity equation. The approach followed amounts to an "integration" of the gravity model, where the utility of a problem is deduced or integrated from the specification of the demand functions.

Sections 2 and 3 discuss how neo-classical models or probabilistic derivations explain gravity as an optimal allocation, the utility of a representative consumer, or the likelihood/entropy of flows that are maximized under budget or market constraints. Section 4 describes how scale invariance of the generic bi-proportional model limits the functional form of the objective function to be optimized to essentially the ones that are used in the literature (i.e. CES). Sections 5 and 6 look at how to construct alternative models where the scale invariance is broken, and gravity is no longer strictly bi-proportional because the impact of trade costs on flows is no longer independent from size.

2 The generic gravity equation and its scale invariances

The generic gravity model takes the general bi-proportional form describing flows between origin i and destination j

$$X_{ij} = A_i B_i K_{ij} \tag{0}$$

where

A_i is a factor that represents the push potential of origin i

B_j is a factor that represents the pull potential of destination j

K_{ij} is the impedance representing the intensity of the interaction between i and j

The potentials A and B incorporate information about the "size" of origin and destination. Potentials are either fixed effects or exogenously determined from country variable explaining the "size" of origin and destination.

The impedance K incorporates bilateral factors that explain the friction between i and j. Following Anderson, this friction is referred to as trade costs.

The bilateral impedance is a parametric function of the otherwise exogenous bilateral trade costs c_{ij}

$$K_{ij} = f(-\beta c_{ij}) \tag{0}$$

where the impedance decreases with higher cost: f is a monotonous increasing function. The parameter β acts as a scale parameter for the trade costs in the impedance. β is either exogenously given or endogenously determined.

Typically the impedance is normalized to one for zero friction or bilateral cost and goes to zero for very large costs. For instance, in Wilson's model the dependence is exponential

$$K_{ij} = \exp(-\beta c_{ij}) \tag{0}$$

In the CES based models, à la Anderson

$K_{ij} = (1 + c_{ij})^{-\sigma}$, where σ is the Armington constant elasticity of substitution. The two models essentially correspond through a logarithmic transformation $c_{ij} \leftarrow \log(1 + c_{ij})$, and $\beta \leftarrow \sigma$

In what follows, except when specified otherwise, we shall refer to the Wilson's exponential dependence.

In early gravity proposals the impedance is simply the negative power of distance (Zipf 1946). More advanced specifications of bilateral trade costs can refer to time or transportation costs and introduce dummy variables. A typical gravity model takes the expanded form:

.

$$\begin{aligned} Log(X_{ij}) = {} & \underbrace{a_1 Log(GDP_i) + a_2 Log(GDP_j)}_{I} \\ & + \underbrace{\begin{array}{l} b_1 Log(GDPpercap_i) + b_2 Log(GDPpercap_j) + c_1 \text{domestic efficiency variables for i} \\ + i \leftrightarrow j \end{array}}_{II} \\ & \underbrace{+ d_1 Log(distance_{ij}) + d_2 \text{other separation variables} + d_3 \text{bilateral dummies}}_{III} \end{aligned} \tag{0}$$

Scale invariance

The gravity specification of bilateral flows is trivially invariant by the scaling transformations.

$$A_i \to \lambda_i A_i$$
$$B_j \to \kappa_j B_j$$
$$K_{ij} \to \frac{1}{\lambda_i}\frac{1}{\kappa_j} K_{ij}$$

or in the case of exponential dependence on costs

$$c_{ij} \to c_{ij} + \frac{1}{\beta}(\log(\lambda_i) + \log(\kappa_j))$$

where the λ and κ are arbitrary real values.

The possibility of such a transformation is obviously consubstantial to the bi-proportional standard gravity structure, and tied to the fact that trade costs and country size factors act independently: trade costs have the same effects on bilateral flows independently irrespective of the size of the partners. The scale invariance means that if a partner i is further away but bigger, the corresponding trade flows are unchanged: twice as big potential yields the same bilateral flows when the corresponding bilateral costs are increased by $\frac{\log 2}{\beta}$.

This phenomenon where rescaling of potential is compensated by a shift in trade costs is not just an algebraic observation, but carries some economic significance. Obviously the inverse problem (inverse gravity) of determining trade costs from trade flows is ambiguous and requires additional assumptions to resolve the ambiguity—e.g., that the diagonal trade costs with oneself is zero or unit impedance $K_{ii} = 1$ or $c_{ii} = 0$ (Novy 2009), (Arvis & Shepherd, 2011). It means also that trade costs are to some extent not directly observable, although individual components of trade costs dependent on directly measurable factors such as distance or transportation cost are.

In the context of trade economics, bilateral trade costs include two categories:

- Bilateral costs that depend on both origin and destination (group III in the above log-linear gravity equation)
- Endogenous costs that represent the thickness of the borders at origin and destination and do not depend on the country pair, typically represented in log-linear trade gravity regressions by country specific dummies or non-dimensional variables where country size does not intervene explicitly.

Endogenous costs are in the group II of variables in equation (4). However this group may also explain the magnitude of the trade potential A, B, along with the size variable in group I. The frontier between variables that explain endogenous trade costs or trade potential is a matter of convention as some factors (labor costs, logistics performance/efficiency, cost of market access) affect both production for export and domestic production, albeit not with the same intensity.

3 Derivations of the bi-proportional gravity equation

Neo-classical models or probabilistic explanations of the gravity equation derive it as an optimal allocation, under one or more budget or market clearance constraints, and depending upon bilateral trade costs. The models found in the literature belong to one or the other category:

- One-dimensional allocation: an asymmetrical model where the allocation is determined separately either for the origin or destination. In trade language, the model derives the optimal allocation for representative importers and exporters for each origin destination (partial equilibrium).
- Two-dimensional allocation: a symmetrical model where the full flow/trade matrix is determined by the optimization problem, under marginal supply and demand constraints.

One-dimensional allocation: Representative trader

The reference asymmetrical model has been proposed by Anderson & Van Wincoop (2003). Following Armington (1969), for each importing country the representative consumer has differentiated preferences according to the country of origin of the goods and according to a CES utility function. The representative consumer in j sources goods from i at price p_{ij}, which include the fob price plus trade costs c_{ij}

$$p_{ij} = 1 + c_{ij}$$

In anticipation of the developments of the next section, we use the language of the indirect utility function. The latter is homogenous function U of degree zero of price and budget. Hence U can be written as

$$U = (\frac{p_{1j}}{m}, \ldots, \frac{p_{ij}}{m}),$$

where m is the budget of the buyer: $m = \sum_i x_{ij} p_{ij}$

By Roy's theorem the quantity of goods bought by j from i x_{ij} is given by

$$x_{ij} = -\frac{\partial U}{\partial p_{ij}} / \frac{\partial U}{\partial m}$$

Following Anderson, most authors posit a CES form

$U = \sum_i a_i (\frac{p_{ij}}{m})^{-\sigma+1}$ where σ is the constant elasticity of substitution, and the a_{i} are share parameters independent of the destination ij.

Applying the previous formula for x_{ij} yields:

$x_{ij} \propto a_i p_{ij}^{-\sigma}$, independently of i

which corresponds to a standard, generic, bi-proportional gravity equation. We refer the reader to De Benetis (2012) for some alternative representative traders explanation of gravity.

Two-dimensional allocation

Two-dimensional allocations are a one-step derivation of the gravity equation with an entirely symmetrical treatment of origin and destination. Wilson (1970) proposed half a century ago a most elegant solution. Wilson determines the flow according to the most likely bilateral values of discrete consignment or movement (of people or vehicles) between origin i and destination j, N_{ij} subject to constraints of marginal total in row and column: $\sum_i N_{ij} = N_{.j}$ $\sum_j N_{ij} = N_{i.}$. N_{ij} is an integer but expected to be large.

The probability or likelihood of a given configuration is given by

$$\frac{\prod_{ij} N_{ij}!}{\prod_i N_i!}$$

Making use of the Stirling's approximation, the log-likelihood is

$$L = \sum_{ij} \log N_{ij}! = \sum_{ij} N_{ij} \log N_{ij} - N_{ij}$$

The most likely configuration maximizes L under line and column constraints as well as a budget constraint (Wilson 1970):

$$\sum_{ij} N_{ij} c_{ij} = Constant$$

Using Lagrange's method, the problem yields an optimal solution which has the standard log gravity form

$$\log N_{ij} = a_i + b_j - \beta c_{ij}$$

where a and b are fixed effects and beta are Lagrange's multipliers which are determined by the constraints of marginal total in row and column.

In the original Wilson model, the gravity exponent β is endogenous to the model and hence not necessarily constant. However the gravity formula can be also deduced by looking for the maximum of the following Lagrangian derived from the previous log-likelihood

$$L1 = \sum_{ij} N_{ij} \log N_{ij} - N_{ij} + \beta c_{ij} N_{ij}$$

under supply and demand line/column constraints, and with the gravity exponent β exogenously given.

Wilson's approach borrows the concept of entropy and draws an analogy from physics of the canonical derivation of the partition function in statistical physics (Kubo 1967), where the most likely probability allocation of states, or partition function, correspond to the most probable configuration (or maximum entropy) for a given average energy (in place of budget constraint). The Lagrange multiplier β, known as the Boltzman's factor, is inversely proportional to the absolute temperature.

Thus, taking also the results of the previous section, it seems that in the context of economic behavior, the equivalent of the temperature would be the inverse of the elasticity of substitution. In the same way that the temperature is a scale for the distribution of energy level, the inverse of β (or of σ) is a scale of the distribution of trade costs. However the analogy may not be appropriate in other respects. For instance, in a transposition of the model to canonical or grand canonical ensemble is not obvious and may not make sense, what would be the equivalent concept in economics of a thermal reservoir?

The global budget constraint for all origins and destinations mirrors the conservation of energy. However, the significance is not entirely obvious. The notion of a global arbitrage between variety of trade linkages and trade costs supported by the trading community is expected (as in the Lagrangian L1), but why would total trade costs be conserved? It amounts to imagining one actor, like a multinational corporation, optimizing a variety of trade between locations under its own budget constraints.

4 Compatible models: One-dimensional allocation case

The scale invariance of the generic bi-proportional gravity models constrains the possible optimization problems that can produce it. The following proposition holds.

Proposition: *The class of models compatible with bi-proportional gravity models are restricted to the following:*

1. *For one-dimensional allocation (representative traders), the corresponding objective function (e.g. utility) must be separable*
2. *For two-dimensional allocation, the corresponding objective function is the entropy formula (Wilson's model)*

For the two-dimensional model, the fact that the entropy model is the only one consistent with the row and column constraints is consistent with the findings by Arvis & Shepherd (2013) that the Poisson Quasi Maximum Likelihood is the only one preserving marginal totals between original and predicted value. Indeed the Poisson QML is essentially an entropy formula.

Compatible models: One-dimensional allocation model

The scaling invariance in the generic gravity model is related to some form of independence of irrelevant opportunities of substitution, where the ratio of substitution between two possible allocations are independent of changes affecting other cases of allocation. Take the case of representative buyer k, then the ratio of allocation of supplies from i and j.

$$\frac{\text{Probability of buying from i}}{\text{Probability of buying from j}} = \frac{X_{ik}}{X_{jk}} = \frac{A_i}{A_j}\frac{K_{ik}}{K_{jk}}$$

will be independent of changes in supplies from other origins or changes in the substitution ratios

$$\frac{X_{lk}}{X_{mk}} \text{ for } l, m \notin i, j$$

The generic indirect utility for trader k takes the form $U(p_{ik}, p_{2k}, \ldots, m)$ of a function homogeneous of degree zero. Hence it can be written as

$$U(\frac{p_{1k}}{m}, \frac{p_{2k}}{m}, \ldots)$$

where p_{ik} is the price of variety y and m the expenditure of the buyer, hereafter treated as exogenous.

Then by Roy's identity the ratios of allocation are

$$\frac{X_{ik}}{X_{jk}} = \frac{U_i}{U_j}, \text{ where } U_i = \frac{\partial U}{\partial p_{ik}} \text{ is the derivative of U in the i-th variable}$$

Furthermore, for the one dimension allocation problem of a representative trader to be invariant to irrelevant opportunities of substitution, it is necessary and sufficient that the utility function is separable and additive (see Annex 1): $U = F(\sum_i f^i(\frac{p_i}{m}))$.

If additional requirements are made that allocation should depend only on price ratios, then the f_i are restricted to power functions with the same exponent $f_i(q) = a_i q^{\alpha}$, yielding the CES utility. However, this is not the only solution. For instance the logit discrete choice equation is derived from the choice $f_i(q) = a_i \exp(-q)$ (Anderson et al. 1992). The later choice corresponds to a case where allocation depends only on price differences.

Compatible models: Two-dimensional symmetric optimization model

In this case we are looking into the maximization of a Lagrangian L of the bilateral flows

$$\mathrm{Max} L(\ldots, X_{ai}, \ldots)$$

under the constraints of conservation of total in line and column,

$$\sum_i X_{ai} = X_a \text{ , and } \sum_a X_{ai} = X_i$$

and a budget constraint. $\sum_{ai} c_{ai} X_{ai}$.

The multiplier method applied to the Lagrangian L yields for each pair or origin destination {a,i} the partial derivative of L is the sum of line and column constants plus costs:

$$\frac{\partial L}{\partial X_{ai}} = a_a + b_i + \beta c_{ai}$$

If the solution of the problem is a generic, scale invariant bi-proportional gravity, then up to a multiplicative constant the cost can be replaced by the logarithm of the flows $c_{ai} \propto \log X_{ai} + \text{fixed effects}$. Hence:

$$\frac{\partial L}{\partial X_{ai}} = a'_a + b'_i + \beta' \log X_{ai}$$

And by identifying the dependence in X, and integrating

$L = \beta' X_{ai} \log X_{ai} + \textbf{linear combination of total of X in rows and columns}$, which is the expected entropy like formula for the Lagrangian.

5 Modified gravity and breaking of scale invariance

Departing from the classical explanations of gravity implies a break in the scale invariance. The easiest and most natural way to do this is to supplement the utility function and Lagrangian with non-linear terms that explicitly break the symmetry of the model. As apparent in the following, the models include at least one additional parameter beyond the scale giving β and become substantially more complex.

One-dimensional allocation: Modified translog gravity

In the representative consumer representation, breaking the symmetry means making the indirect utility function explicitly non-separable by introducing multiplicative interaction between the trade costs. For instance, Novy (Novy 2009) recently tried to describe trade flows starting with a translog utility instead of a CES. Indeed, a translog utility is one of the simplest explicit ways to introduce interactions and break the scale invariance.

$$U(\text{consumer j}) = \exp(\sum_i a_i \log p_{ij} + \beta \sum_{k,l \neq j} \log p_{kj} \mathit{log} p_{lj})$$

Bilateral flows are in the form

$X_{ij} = A_i B_j K^1{}_{ij}$ where

$$K^1{}_{ij} = \frac{1}{p_{ij}}(1 + 2\frac{\beta}{a_i}\sum_{k \neq j} \log p_{kj})$$

As apparent in this structure of the "impedance" K1, this extension of gravity is no longer bi-proportional and breaks a number of symmetries of the traditional equation:

- The bilateral impedance depends not only on the bilateral trade costs but also of other trade costs of the importer.
- The size of the exporting economy matters: K1 is sensitive to the share coefficient. a_i of the exporter, with the direction effect depending on the sign of β: if the latter is positive, the "impedance" is higher for smaller exporter.
- The translog model breaks the formal symmetry between origin and destination.
- Transslog is essentially a nonlinear extension of the Cobb-Douglass utility, which corresponds to a CES of one, while In the context of trade the typical observed CES is much larger (seven).

We propose below a simpler alternative symmetry-breaking extension of gravity, which departs more smoothly from standard gravity.

Two-dimensional allocation: Modified gravity equation

In the case of two-dimensional symmetric gravity, the simplest approach to symmetry-breaking "gravity" consists in adding interaction terms between costs and flows in the Lagrangian yielding the gravity model, so that the scale invariance does not hold. Indeed, the original Wilson's Lagrangian

$$L1 = \sum_{ij} X_{ij} \log X_{ij} - X_{ij} + \beta c_{ij} X_{ij}$$

also reads as a weighted average of a linear combination of $\log X_{ij}$ and the trade costs (in the following angle brackets stands for weighted averages $\langle f_{ij} \rangle = 1/X_{..} \sum X_{ij} f_{ij}$)

$$L1 \propto \langle \log X_{ij} \rangle + \beta \langle c_{ij} \rangle,$$

which is obviously invariant by the scaling transformation

$\log X_{ij} \to \log X_{ij} + a_i + b_j$, and

$$c_{ij} \to c_{ij} - \frac{1}{\beta}(a_i + b_j)$$

Therefore, a minimalist way to break the scaling symmetry is to add to the Lagrangian a non-linear multiplicative second-order interaction between trade costs and logarithm of flows, the most natural choice being their covariance (which is invariant by translation of the log flows and costs):

$$L2 \propto \langle \log X_{ij} \rangle + \beta \langle c_{ij} \rangle + \gamma (\langle \log X_{ij} c_{ij} \rangle - \langle \log X_{ij} \rangle * \langle c_{ij} \rangle)$$

$$L2 = \sum_{ij} X_{ij} \log X_{ij} - X_{ij} + \beta c_{ij} X_{ij} + \gamma \sum_{ij} X_{ij} \log X_{ij} c_{ij} - \gamma \frac{(\sum_{ij} X_{ij} \log X_{ij}) * (\sum_{ij} c_{ij} X_{ij})}{\sum_{ij} X_{ij}},$$

The optimization of this Lagrangian under row and column constraints yields

$$\log X_{ij} + \beta c_{ij} + \gamma (c_{ij} - \langle c \rangle)(\log X_{ij} - \langle \log X \rangle) = a_i + b_j,$$

The previous equations yields an explicit solution for log-linearized flows as

$$\log X_{ij} = a_i + b_j - \beta \bar{c}_{ij} - \gamma (a_i + b_j) \bar{c}_{ij} + \langle \log X \rangle,$$

where

$$\bar{c}_{ij} = \frac{c_{ij} - \langle c \rangle}{1 + \gamma (c_{ij} - \langle c \rangle)}$$

This modified gravity equation

- breaks scale invariance, a positive γ increases the effect of trade costs for large flows (larger than the $\langle \log X \rangle$), while a negative γ further suppresses small flows for the same trade costs.
- keeps symmetry between origin and destination.

6 Implementation

The modified gravity equation is applied to the World Trade matrix (the dataset is the same as the one used in Arvis & Shepherd (2013), zero omitted, using the logarithm of distance $dist_{ij}$ as a proxy for the trade cost. The regression takes the form:

$$\log X_{ij} = a_i + b_j - \beta \log dist_{ij} - \gamma t, \qquad (5)$$

where t is the covariance variable

$$t = (\log dist_{ij} - \langle dist \rangle) * (\log X_{ij} - \langle X \rangle)$$

The following table includes the results of estimation of (5) using:

1. Poisson regression for the model without interaction ($\gamma = 0$)
2. Poisson regression for the model with interaction ($\gamma \neq 0$)
3. OLS for the model without interaction ($\gamma = 0$)
4. OLS for the model with interaction ($\gamma \neq 0$)

The graphs in Annex 2 plot for Poisson and OLS the predicted values for both standards and modified gravity.

Table 1 results

20710 observations	1	2	3	4
R2/ pseudo R2	nd	nd	0.749	0.820
Log distance coeff ($-\beta$)	-0.837	-0.802	-1.739	0.837
Wald Chi2/ F	68035	55339	175	266
,, z statistics	-34.0	-35.9	-77.4	24.2
,, standardized "beta" coefficient	-1.008	-0.965	-2.093	1.008
Interaction coeff ($-\gamma$)		0.114		0.298
,, z statistics		16.1		89.4
,, standardized "beta" coefficient		0.013		0.034

The regression results suggests the following conclusions

1. As expected the Poisson regression has the most consistent result with trade spatial interaction data (Silva & Tenreyro 2006)(Arvis & Shepherd 2013)
2. The improvement in fit provided by the model (R2 or pseudo-R2) is relatively small in both models (also visually apparent in the shape of distributions in Annex 2).
3. In Poisson, the coefficient γ:
   - has a small standardized ("beta") value meaning that the effect of the non-linear covariance term is small as compared with the classical log-linear impact of distance
   - is negative and significant, which means a relatively small effect of stronger suppression of small flows.
   - the predicted values for the trade flows are relatively closed in each models
   - the distance coefficients are close in both Poisson model
4. The OLS regressions has observed by various authors are less robust than the Poisson ones

5. The OLS results (coefficient predicted value) are less consistent between standard and modified gravity model:
   - Loose fit between predicted and actual value of trade
   - The coefficient of log-distance in the modified gravity turns positive with a relatively higher (than in Poisson) coefficient of interaction, which suggests an issue of multi-colinearity better addressed in the Poisson estimate.

7 Conclusions

The use of symmetry principles is not very popular in economics, although the quest for symmetry is central to scientific understanding in other discipline (physics, chemistry). Hence rather than a restriction they are a desirable property that make the description of a phenomenon consistent, less dependent upon extra assumption or parameters, and likely more analytically tractable (the parsimony principle). Scale invariance in gravity constrains the possible explanation to essentially the known one. This should be considered more as a desirable property than a strong restriction.

Possible scale symmetry breaking alternatives are indeed feasible, where trade costs and size of trade flows interact, instead of the standard bi-proportional structure. However even with the simplest modification, modified gravity equations add parameters and complexity including nonlinear equations to estimate econometrically. They do not seem to provide a major improvement over standard gravity. Thus, given the high quality of fit of modern gravity models, especially when using Poisson Quasi Maximum Likelihood estimation, there does not appear to be a very strong case for using an alternative to the bi-proportional standard gravity equation. This result provides additional support for ongoing and recent efforts to better understand and measure the origin and nature of bilateral impedances or trade costs within the standard model.

References

Anderson, J.E., 1979. A theoretical foundation for the gravity equation. *The American Economic Review*, 69(1), pp.106–116..

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

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

Anderson, S.P., De Palma, A. & Thisse, J.-F., 1992. *Discrete choice theory of product differentiation*, MIT Press.

Armington, P.S., 1969. Theory of Demand. *Staff Papers-International Monetary Fund*, pp.159–178.

Arvis, J. & Shepherd, Ben, 2013. The Poisson quasi-maximum likelihood estimator : a solution to the " adding up " problem in gravity models . Applied Economics Letters August 2012, pp.37–41.

Arvis, J.J.F. & Shepherd, B., 2011. The air connectivity index: measuring integration in the global air transport network. *Bank Policy Research Working Paper No.*, (June). A

De Benedictis, L. & Taglioni, D., 2010. The Gravity Model in International Trade. In L. De Benedictis & L. Salvatici, eds. *The Trade Impact of European Union Preferential Policies*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 55–89.

Deardorff, A.V., 1998. Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World? in J. Frankel (ed), *The regionalization of the world economy*, University of Chicago Press.

Kubo, R., 1967. *Statistical mechanics: an advanced course with problems and solutions*, North-Holland.

Novy, D., 2009. Gravity redux: measuring international trade costs with panel data. *Social Research*, (861), pp.0–25..

Novy, D., 2012. International Trade without CES: Estimating Translog Gravity. *Journal of International Economics*, (August), pp.1–37.

Roy, J.R. & Thill, J.C., 2003. Spatial interaction modelling. *Papers in Regional Science*, 83(1), pp.339–361.

Savage, I.R. & Deutsch, K.W., 1960. A statistical model of the gross analysis of transaction flows. *Econometrica: Journal of the Econometric Society*, 28(3), pp.551–572.

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

Tinbergen, Jan. 1962. *Shaping the world economy: Suggestions for an international economic policy. New York: Twentieth Century Fund*

Wilson, A.G., 1970. *Entropy in Urban and Regional Modelling*, Pion.

Zipf, G., 1946. The P1P2/D Hypothesis: On the Intercity Movement of Persons. *American Sociological Review*, 11(1), pp.23–38.

Annex 1

Lemma: The property of independence of ratios of substitution holds if and only if the indirect utility function is separable.

Let be $\{x_i\}$ the vector of the variables to be bought from various sources at prices $\{p_i\}$ and $U = U(\ldots p_i \ldots)$ the indirect Utility (without loss of generality the budget is set to one). Then the ratio of substitution between i and j is by Roy's identity

$\frac{x_i}{x_j} = \frac{U_i}{U_j}$ (0), where $U_i = \frac{\partial U}{\partial p_i}$ is the derivative of U in the i-th variable

The independence property means that that this ratio does not depend of p_k for $k \neq i, j$

A separable utility would be of the form

$U(\{p_i\}) = U(\sum_i f^i(p_i))$ (0)

That the condition is sufficient is immediate as

$\frac{U_i}{U_j} = \frac{U' * f^{i\prime}(p_i)}{U*' f^{j\prime}(p_j)} = \frac{f^{i\prime}(p_i)}{f^{j\prime}(p_j)}$(0)

, which is independent of p_k for $k \neq i, j$.

To prove the necessary condition, let look at the ratio

$\frac{x_i}{x_j} = \frac{\frac{x_i}{x_k}}{\frac{x_j}{x_k}}$, for all $k \neq i, j$ or in log

$\log \frac{x_i}{x_j} = \log \frac{x_i}{x_k} - \log \frac{x_j}{x_k}$(0)

taking the partial derivative in p_k implies that

$0 = \frac{\partial \log \frac{x_i}{x_k}}{\partial p_k} - \frac{\partial \log \frac{x_j}{x_k}}{\partial p_k}$(0)

for all $i, j \neq k$ and hence

$$\frac{\partial \log \frac{x_i}{x_k}}{\partial p_k}$$

must be a function of p_k only for all i, which means by integration that the substitution ratios are necessarily of the form

$\log \frac{x_i}{x_k} = g^i(p_i) - g^k(p_k)$(0), where the g^k are function of one variable.

Let f^k be functions of one variable such that the log of their derivative is g^k. Then

$$\frac{U_i}{U_j} = \frac{x_i}{x_j} = \frac{f^{i\prime}(p_i)}{f^{j\prime}(p_j)} \quad (0)$$

For (11) to hold for every i, j it is necessary that

$U_i = G(\ldots p_k \ldots) * f^{i\prime}(p_i)$, where G is the same for all i. Let G be expressed as a function of the $f^i(p_i)$ t then the equality

$\frac{\partial U_j}{\partial p_i} = \frac{\partial U_i}{\partial p_j}$ means that

$\frac{\partial G}{\partial f^i} f^{i\prime}(p_i) * f^{j\prime}(p_j) = \frac{\partial G}{\partial f^j} f^{j\prime}(p_j) * f^{i\prime}(p_i)$, or

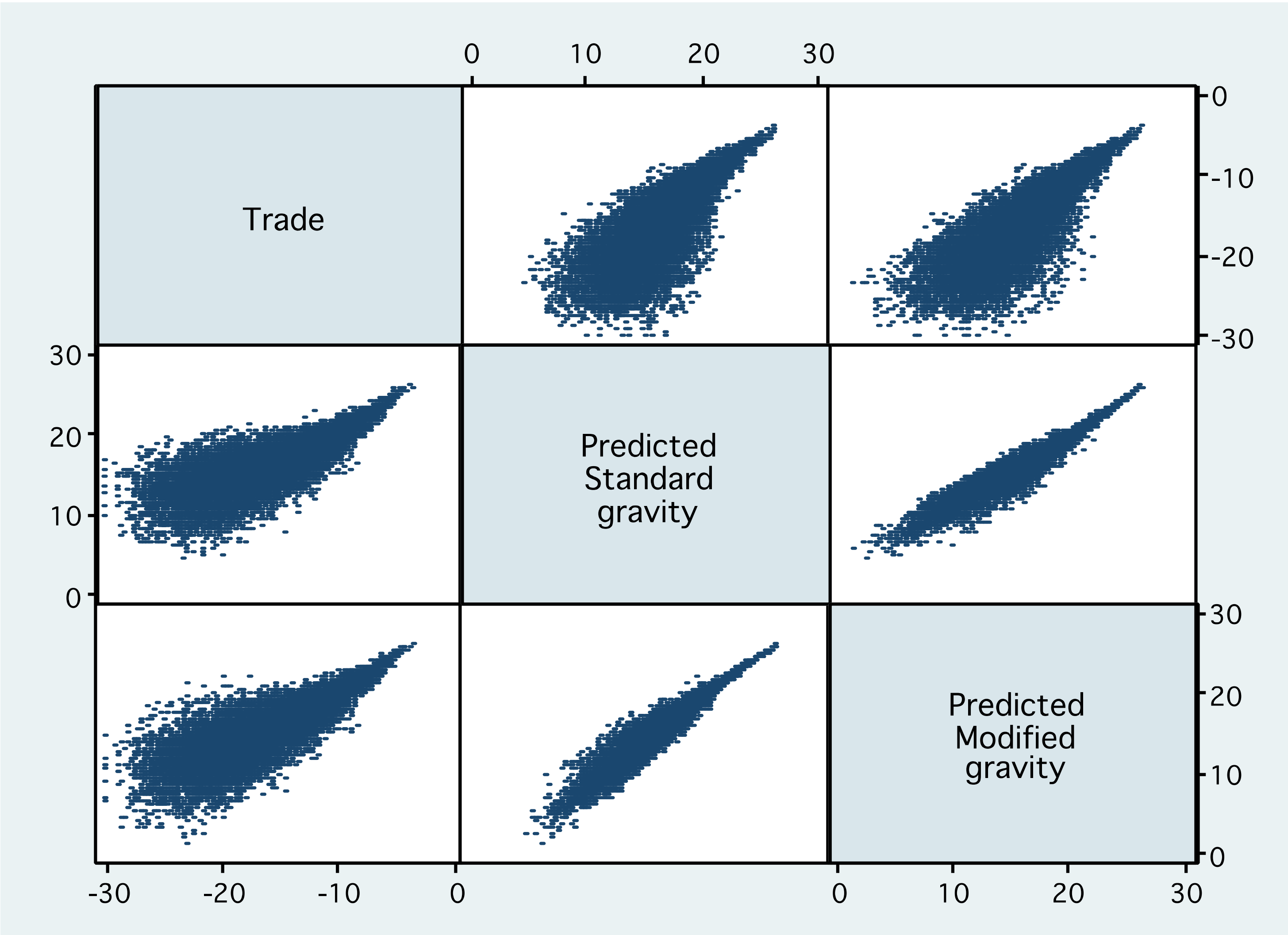
$\frac{\partial G}{\partial f^i} = \frac{\partial G}{\partial f^j}$, for all i, j. This is possible only and only if G is a function of only the sum of the f^i

$G = F(\sum_i f^i)$. Finally, taking U as a primitive of F yields the expected result

$U(\{p_i\}) = U(\sum_i f^i(p_i))$Q.E.D.

Annex 2

Log of Trade vs. predicted value for standard and modified gravity (Poisson Regression).



Log of Trade vs. predicted value for standard and modified gravity (OLS)