A New Theoretical Approach to the Gravity Trade Model*

Richard Barrett[a], Somnath Sen[b] and Anca Voicu[c]

June 16, 2011

Abstract

Microfoundations for the well-known gravity trade model are investigated. A new framework is adopted, inclusive of transport costs and tariffs, together with assumptions that while strong are rather less restrictive than those typically employed in this context, thus enabling rather more complex and interesting results to be obtained than are available in the literature.

[a,b]Economics Dept., Birmingham University, Edgbaston, Birmingham B15 2TT, UK and
[c]Economics Dept., Rollins College, 1000 Holt Ave., Winter Park, FL32789-4409, USA

1 Introduction

The gravity trade model is essentially an empirical model that explains bilateral trade between two regions in terms of their populations and incomes, and barriers between the two, such as distance apart or tariffs. In aggregate form, a typical such model (Frankel, Stein and Wei, 1995) is:

$$\log T = a + b\log(Y_A Y_B) + c\log\left(\frac{Y_A}{N_A}\frac{Y_B}{N_B}\right) + d\log S, \tag{1}$$

where T is trade between A and B, Y_A is A's GDP, Y_B is B's GDP, N_A is A's population, N_B is B's population and S is distance between A and B. Henceforth, we call A and B 'countries'.

In the literature, diverse ways of theoretically deriving such an aggregate trade model are presented. This diversity, showing the model's robustness, is the literature's strength. A weakness is failure to go convincingly beyond a result even simpler than equation (1): The trade between two countries is proportional to the product of their incomes divided by world income. Understandably, given the complexity otherwise involved, these papers typically start from a (uniform) homothetic utility function; less understandably, they often neglect barriers to trade such as transport costs! Thus the theoretical foundations of the gravity trade model remain relatively unexplored. The model that is typically derived is not the one econometricians estimate, motivating us to attempt here to close some of this gap between theory and applied practice.

To summarise the literature briefly, after early work by Tinbergen (1962) the seminal paper is that of Anderson (1979). Some of the papers which follow are: Helpman and Krugman (1985), Bergstrand (1989), Deardorff (1998), Everett and Keller (2002), and Anderson and van Wincoop

*We would like to express our indebtedness to the referees for their very helpful comments.

(2003, 2004). Anderson (1979) assumes a uniform Cobb-Douglas utility function. While he extends the model described in the previous paragraph, this is only in a very ad hoc manner. Anderson and van Wincoop (2003, 2004) impose a (uniform) CES utility function. The importance of the other papers is in the focus on structure. Everett and Keller (2002), for example, find interesting equivalences among several alternative structures (the diversity mentioned above), while neglecting trade costs.

Two further papers that deserve mention are: Mitra and Trindade (2005) and Dalgin, Trindade and Mitra (2008). In building trade models, these depart from homothetic preferences in allowing the shares in expenditure at given prices of food and manufactures to decrease and increase, respectively, with income. The interesting effect is that inequality is introduced as a factor into the gravity equation.

In our paper we adopt a fresh approach, which like the latter two papers primarily focuses on the demand side. A particular innovation is that, instead of starting from a (uniform) utility function, we assume (uniform) Cobb-Douglas *demand* functions for tradables and apply a condition required for their integrability[1] in obtaining comparative statics results. Cobb-Douglas demand functions enable us to relax homothetic preferences. We also introduce a non-traded good and, while following others in assuming specialisation in trade, make 'locations' rather than countries the fundamental unit, so that a single country, consisting of many locations, exports many tradables. The approach enables us to provide a general treatment of transport costs as well as tariffs in a unified multi-regional framework.

Let us straightaway state a result: The symmetry between exporter and importer expressed when trade is proportional to the product of their incomes divided by world income, or as in equation (1), is not justified.

2 The Model

Assume there are N 'locations' and $N+2$ goods. Goods 1 to N are traded consumption goods, good $N+1$ is a non-traded consumption good, and good $N+2$ is 'shipping'[2]. Each location has a population of size 1. Location i produces, competitively, goods i, $N+1$ and $N+2$. No other locations produce good i, i.e. traded goods have some local characteristics. Separate local markets exist for the non-traded consumption good, but there is a single market for each traded good, and for shipping. All markets clear.

Let location i produce positive amounts of good k for $k = i$, $N+1$ and $N+2$. All marginal rates of transformation are unity. Thus all goods have the same price at point of origin, which we set equal to unity.

Traded goods differ in price between locations, both through the cost of shipping and because of tariffs. Let θ_{kj} be the cost of shipping a unit of traded good k from location k to location j, let τ_{kj} be the tariff applied to good k at location j, and let p_{kj} be good k's price at location j. Then, for $k, j = 1, ..., N$:

$$p_{kj} = (1+\theta_{kj})(1+\tau_{kj}). \tag{2}$$

According to (2), the price of traded good k at location j is the price (i.e. unity) at location k plus the cost of shipping a unit from k to j, scaled up by the tariff applied at j.

[1]Integrability means there exists a utility function which implies these demand functions.

[2]As shipping is of course a service, the term 'good' is used here in the generic sense inclusive of services that is found in the theory of general equilibrium.

Assume location j's demand for consumption good k is:

$$x_{kj} = F_{kj}(p_j, y_j), \tag{3}$$

where $p_j = (p_{1j}, ..., p_{Nj}, 1)$, and y_j is location j's income.

We suppose the N locations are partitioned into 'countries' and focus on any two of these, A and $A*$, in order to investigate bilateral trade. Let A include N_A locations and $A*$ include N_{A*} locations.

Suppose countries exclusively provide shipping for their own exports.[3] The value of total exports of goods from A to $A*$ (or imports of $A*$ from A), including shipping costs, is:

$$X = \sum_{i \in A} \sum_{j \in A*} (1 + \theta_{ij}) F_{ij}(p_j, y_j). \tag{4}$$

2.1 *Cobb-Douglas Demand*

Referring to (3), we suppose that the demand functions for traded goods take the specific form:

$$x_{kj} = c p_{1j}^{\beta} ... p_{k-1,j}^{\beta} p_{kj}^{\beta-1} p_{k+1,j}^{\beta} ... p_{Nj}^{\beta} p_{N+1,j}^{\alpha} y_j^{\delta}, \tag{5}$$

where c is a scale parameter,$\beta < 1$ and $\delta > 0$. I.e. the own-price effect is negative, cross-price effects among traded goods are equal but indeterminate in sign, and traded goods are normal.

In (5), note that for traded goods:

(a) The symmetry condition requires the own-price index to be one less than corresponding cross-price indices.

(b) The homogeneity condition requires:

$$\alpha = 1 - \delta - N\beta. \tag{6}$$

It can be shown that the integrability condition is satisfied when β is small or negative and there is a sufficiently dominant non-traded consumption good; i.e. there then exists a utility function from which these demand functions can be derived. A necessary condition is:

$$N\beta < 1 \tag{7}$$

(see Appendix). An important feature of Cobb-Douglas demand is it allows traded goods to be luxury (or necessary) goods.

Substituting (2) into (5), noting $p_{N+1,j} = 1$, we obtain:

$$\begin{aligned} x_{kj} &= c(1+\theta_{1j})^{\beta}(1+\tau_{1j})^{\beta} ... (1+\theta_{k-1,j})^{\beta}(1+\tau_{k-1,j})^{\beta} \\ &\quad (1+\theta_{kj})^{\beta-1}(1+\tau_{kj})^{\beta-1}(1+\theta_{k+1,j})^{\beta}(1+\tau_{k+1,j})^{\beta} \\ &\quad ...(1+\theta_{Nj})^{\beta}(1+\tau_{Nj})^{\beta} y_j^{\delta}. \end{aligned} \tag{8}$$

Substituting (8) into (4), we obtain:

$$\begin{aligned} X &= \sum_{i \in A} \sum_{j \in A*} (1+\theta_{ij}) c (1+\theta_{1j})^{\beta}(1+\tau_{1j})^{\beta} ... \\ &\quad (1+\theta_{ij})^{\beta-1}(1+\tau_{ij})^{\beta-1} ... (1+\theta_{Nj})^{\beta}(1+\tau_{Nj})^{\beta} y_j^{\delta}. \end{aligned} \tag{9}$$

[3]Note that in the concluding section we discuss relaxing this assumption.

In order to express exports in terms of aggregates, we assume:

(1) Locations in $A*$ have uniform income, i.e. for all $j \in A*$, $y_j = y_{A*}$.

(2) $A*$ applies a uniform tariff to imports from a subset S of locations, where $S \frown A* = O$, either $A \subset S$ or $A \frown S = O$ and S includes N_S locations. I.e. for all $i \in S$ and $j \in A*$, $\tau_{ij} = \tau_{A*}$, while $\tau_{ij} = 0$ otherwise.

Let $\omega = 1$ if $A \subset S$ and $\omega = 0$ if $A \frown S = O$. (9) is replaced by:

$$X = (1 + \tau_{A*})^{N_S\beta - \omega} \sum_{i \in A} c y_{A*}^{\delta} \sum_{j \in A*} (1 + \theta_{1j})^{\beta} ... (1 + \theta_{Nj})^{\beta}. \tag{10}$$

Define the average trade cost, θ_j, for shipping goods to location j by:

$$(1 + \theta_j)^N = (1 + \theta_{1j})...(1 + \theta_{Nj}). \tag{11}$$

Then, substituting (11) into (10):

$$X = (1 + \tau_{A*})^{N_S\beta - \omega} \sum_{i \in A} c y_{A*}^{\delta} \sum_{j \in A*} (1 + \theta_j)^{N\beta}. \tag{12}$$

Define the average trade cost, θ, for shipping goods to $A*$ by:

$$N_{A*}(1 + \theta)^{N\beta} = \sum_{j \in A*} (1 + \theta_j)^{N\beta}. \tag{13}$$

Substituting (13) into (12):

$$X = (1 + \tau_{A*})^{N_S\beta - \omega} (1 + \theta)^{N\beta} c N_A N_{A*} y_{A*}^{\delta}. \tag{14}$$

Let

$$\gamma = (1 + \tau_{A*})^{N_S\beta - \omega} (1 + \theta)^{N\beta} c \tag{15}$$

$$Y_{A*} = \sum_{j \in A*} y_j = N_{A*} y_{A*}. \tag{16}$$

Substituting (15) and (16) into (14) gives:

$$X = \gamma N_A N_{A*}^{1-\delta} Y_{A*}^{\delta}. \tag{17}$$

We note that, since by (7) $N\beta < 1$, it follows that $N_S\beta < 1$, and so by (15) an increase in the tariff, ω_{A*}, lowers exports if $\omega = 1$ (A outside the tariff wall); and if $w = 0$ (A inside the tariff wall) raises exports if $\beta < 0$ (tradables are substitutes), and lowers exports if $\beta > 0$ (tradables are complements).

A way to extend the model is to relax the assumption that all locations produce shipping. This allows the prices of tradables at point of production to differ across locations. The price of shipping still provides a lower bound to prices, but no longer anchors the prices of all goods (apart from shipping) to unity. Some locations now find it uneconomic to produce shipping, since the prices of the tradables they produce are greater than unity. Without going into details, one can then derive a gravity equation which differs from (17) in two ways: Firstly, N_A is replaced by Y_A; secondly, γ now depends in addition on the prices of tradables at point of production. Econometrically, these can be measured by real exchange rates (including those of third countries).

3 Conclusion

As stated in the introduction, a primary object of our paper has been to close some of the gap between theory and application. We also noted our conclusion that symmetry between exporter and importer is not justified. For both countries, importer and exporter, measures of 'size' matter, but for different reasons.

Focus on the demand side has meant that closing the gap between theory and application is achieved for the importer. (17) shows that if, for example, tradables are luxury goods then, in log-linear form, the coefficient on $A*$'s income (the importer) is greater than unity, and the coefficient on $A*$'s population is negative. (17), however, gives the coefficient on A's income (the exporter) as zero and on A's population as unity (reflecting the way the model relates variety of goods to population). By focusing more on the supply side, first relaxing the assumption that all countries produce shipping (see Section 2), similar complexity for the role of the exporter to that of the importer can, we hope, be obtained, but this remains unfinished business.

Our framework has enabled us to include a general treatment of transport costs and tariffs. There are three interesting results here: (1) All shipping costs to locations in $A*$ feature symmetrically in the gravity equation. This elegant result is perhaps no more than a curiosity, driven by the simplifying assumption that all accompanying shipping costs are included in A's exports. If none were included (implying a more complex equation), then an increase in these accompanying shipping costs would lower rather than raise exports. This assumption does not, however, affect our second and third results: (2) An important result, which runs counter to that of Anderson and van Wincoop (2003, 2004), is that transport costs from the exporter to third countries are not relevant to bilateral trade. The reason here is, flexible rather than fixed production is assumed, where flexible production does, however, seem the appropriate assumption to make for use in empirical cross-section studies. In the model, flexibility arises from the substitutability of production of tradables for that of shipping at constant marginal cost. (3) The third result of interest is that tariffs have the effect of raising A's exports if A is within $A*$'s tariff wall and tradables are substitutes ($\beta > 0$) but not if they are complements ($\beta < 0$), and lowering A's exports if A is outside $A*$'s tariff wall.

Other barriers to trade, such as cultural differences or the absence of a common frontier, can be dealt with within the above framework, but deserve detailed treatment.

4 Appendix

Lemma

Define A of order n by

$$A = \begin{bmatrix} a & b & b & .. & b \\ b & a & b & .. & b \\ b & b & a & .. & b \\ : & : & : & : & : \\ b & b & b & .. & a \end{bmatrix}.$$

A is negative semi-definite if

$$b \geq a$$

and

$$b \leq -\frac{1}{n-1}a.$$

Subtract the second row from the first, the third from the second, etc., to obtain

$$B = \begin{bmatrix} a-b & b-a & 0 & .. & 0 \\ 0 & a-b & b-a & .. & 0 \\ 0 & 0 & a-b & .. & 0 \\ \vdots & \vdots & \vdots & .. & \vdots \\ b & b & b & .. & a \end{bmatrix}.$$

Then add the first column of B to the second, the second to the third, etc., to obtain

$$C = \begin{bmatrix} a-b & 0 & 0 & .. & 0 \\ 0 & a-b & 0 & .. & 0 \\ 0 & 0 & a-b & .. & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ b & 2b & 3b & .. & a+(n-1)b \end{bmatrix}$$

$$\begin{aligned} & \det(A) \\ = & \det(C) \\ = & (a-b)^{n-1}[a+(n-1)b] \end{aligned}$$

In the model, let

$$e = dp_1^{\beta} p_2^{\beta} ... p_N^{\beta} p_{N+1}^{\alpha} y^{\delta}.$$

Note that e is expenditure on a traded good. The substitution matrix for traded goods is:

$$\begin{bmatrix} \frac{\partial x_1}{\partial p_1} + x_1 \frac{\partial x_1}{\partial y} & \frac{\partial x_1}{\partial p_2} + x_2 \frac{\partial x_1}{\partial y} & .. & \frac{\partial x_1}{\partial p_N} + x_N \frac{\partial x_1}{\partial y} \\ \frac{\partial x_2}{\partial p_1} + x_1 \frac{\partial x_2}{\partial y} & \frac{\partial x_2}{\partial p_2} + x_2 \frac{\partial x_2}{\partial y} & .. & \frac{\partial x_2}{\partial p_N} + x_N \frac{\partial x_2}{\partial y} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial x_N}{\partial p_1} + x_1 \frac{\partial x_N}{\partial y} & \frac{\partial x_N}{\partial p_2} + x_2 \frac{\partial x_N}{\partial y} & .. & \frac{\partial x_N}{\partial p_N} + x_N \frac{\partial x_N}{\partial y} \end{bmatrix}$$

$$= \begin{bmatrix} (\beta-1)\frac{e}{p_1^2} + \delta \frac{e^2}{p_1^2 y} & \beta \frac{e}{p_1 p_2} + \delta \frac{e^2}{p_1 p_2 y} & .. & \beta \frac{e}{p_1 p_N} + \delta \frac{e^2}{p_1 p_N y} \\ \beta \frac{e}{p_2 p_1} + \delta \frac{e^2}{p_2 p_1 y} & (\beta-1)\frac{e}{p_2^2} + \delta \frac{e^2}{p_2^2 y} & .. & \beta \frac{e}{p_2 p_N} + \delta \frac{e^2}{p_2 p_N y} \\ \vdots & \vdots & \vdots & \vdots \\ \beta \frac{e}{p_N p_1} + \delta \frac{e^2}{p_N p_1 y} & \beta \frac{e}{p_N p_2} + \delta \frac{e^2}{p_N p_2 y} & .. & (\beta-1)\frac{e}{p_N^2} + \delta \frac{e^2}{p_N^2 y} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{p_1} & 0 & .. & 0 \\ 0 & \frac{1}{p_2} & .. & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & .. & \frac{1}{p_N} \end{bmatrix} \begin{bmatrix} (\beta-1)e + \delta \frac{e^2}{y} & \beta e + \delta \frac{e^2}{y} & .. & \beta e + \delta \frac{e^2}{y} \\ \beta e + \delta \frac{e^2}{y} & (\beta-1)e + \delta \frac{e^2}{y} & .. & \beta e + \delta \frac{e^2}{y} \\ \beta e + \delta \frac{e^2}{y} & \beta e + \delta \frac{e^2}{y} & .. & \beta e + \delta \frac{e^2}{y} \\ \beta e + \delta \frac{e^2}{y} & \beta e + \delta \frac{e^2}{y} & .. & (\beta-1)e + \delta \frac{e^2}{y} \end{bmatrix}$$

$$\begin{bmatrix} \frac{1}{p_1} & 0 & .. & 0 \\ 0 & \frac{1}{p_2} & .. & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & .. & \frac{1}{p_N} \end{bmatrix}.$$

By the lemma, the substitution matrix is negative semi-definite if:

$$\beta + \delta\frac{e}{y} \leq -\frac{\beta - 1 + \delta\frac{e}{y}}{N-1}$$
$$\beta + \delta\frac{e}{y} \leq \frac{1}{N}.$$

Thus, for β small or negative, and a sufficiently dominant non-traded good, for example $N = 10$, $\beta = 0.05$, $\delta = 1$ and $e \leq 0.05y$, the system of demand functions for traded goods is integrable. A necessary condition is:

$$N\beta < 1.$$

References

[1] Anderson, J.E., 1979, A theoretical foundation for the gravity equation, American Economic Review 69, 106-116.

[2] Anderson, J.E., and E. van Wincoop, 2003, Gravity with gravitas: A solution to the border puzzle, American Economic Review 93, 170-192.

[3] Anderson, J.E., and E. van Wincoop, 2004, Trade costs, Journal of Economic Literature 42, 691-751.

[4] Bergstrand, J., 1989, The generalized gravity equation: Monopolistic competition and the factor proportions theory in international trade, Review of Economics and Statistics 71, 143-153.

[5] Deardorff, A.V., 1998, Determinants of bilateral trade: Does gravity work in a neoclassical world?, in J.A. Frenkel (ed.), The Regionalisation of the World Economy (University of Chicago Press, Chicago) 7-22.

[6] Dalgin, M., V. Trindade and D. Mitra, 2008, Inequality, nonhomothetic preferences and trade, Southern Economic Journal 74, 747-774.

[7] Everett, S.J., and W. Keller, 2002, On theories explaining the success of the gravity models, Journal of Political Economy 110, 281-316.

[8] Frankel, J., E. Stein and S. Wei, 1995, Trading blocs and the Americas: The natural and unnatural and the super-natural, Journal of Development Economics 47, 61-95.

[9] Helpman, C.B., and P. Krugman, 1985, Market Strategies and Foreign Trade: Increasing Returns, Imperfect Competition and the International Economy (MIT Press, Cambridge).

[10] Mitra, D., and V. Trindade, 2005, Inequality and trade, Canadian Journal of Economics 38, 1253-1271.

[11] Tinbergen, J., 1962, Shaping the World Economy (Twentieth Century Fund, New York).