Does the Gravity Model Explain India's Direction of Trade? A Panel Data Approach

Ranajoy Bhattacharyya
Department of Business Management
University of Calcutta, India

Tathagata Banerjee
Indian Institute of Management
Ahmedabad, India

W.P. No.2006-09-01
September 2006

INDIAN INSTITUTE OF MANAGEMENT
AHMEDABAD-380 015
INDIA

Does the Gravity Model Explain India's Direction of Trade? A Panel Data Approach

Ranajoy Bhattacharyya
Department of Business Management, University of Calcutta, India

Tathagata Banerjee
Indian Institute of Management, Ahmedabad, India

Abstract

In this paper we apply the gravity model to the panel consisting of India's yearly bilateral trade data with all its trading partners in the second half of the twentieth century. The main conclusions that emerge from our analyses are: (1) The core gravity model can explain around 43 per cent of the fluctuations in India's direction of trade in the second half of the twentieth century (2) India's trade responds less than proportionally to size and more than proportionally to distance (3) Colonial heritage is still an important factor in determining India's direction of trade at least in the second half of the twentieth century (4) India trades more with developed rather than underdeveloped countries, however (5) size has more determining influence on India's trade than the level of development of the trading partner.

1. Introduction

The possible set of factors affecting a country's direction of trade (DOT) can be quite large, touching many aspects of its existence. Some of the important economic factors are comparative advantage relative to the other countries, economies of scale, the aggregate income of home as well as partner residents, government (especially related to trade and exchange rate) policies, membership to currency unions and/or customs unions and finally, participation in bilateral, regional and multilateral agreements. Some of these factors are intricately related to issues beyond simple DOT and need disaggregated data at finer levels than what we find in a DOT handbook. For example, data on commodity composition of trade at the disaggregated level is essential for investigating the role of comparative advantage, either in terms of technology or resource endowments. Similarly, to calculate the components of trade that is theoretically associated to economies of scale, data at the disaggregated level are necessary.

There is however an extremely successful empirical economic theory that is readily amenable to the simple aggregated DOT data[1]: the theory of gravity. The theory, borrowed from the Newtonian law of gravity, states that trade flows between two countries increases with the product of their gross domestic products (GDP) and decreases with the distances between them. In the literature there have been several attempts to derive this relationship between trade, GDPs and distance from theoretical considerations. These are met with moderate success. It has been shown that the positive relationship with the GDPs and the negative relationship with distance can be established from a variety of assumptions regarding the production structure and preferences[2]. However such relationships have not been 'pure' in the sense that the reduced form could not always be solved down to the level where only the GDPs and the proxy for distance remain[3] or where all the three variables are simultaneously present (the usual casualty is the distance variable[4]). Notwithstanding this theoretical deficiency there have literally been hundreds of empirical estimates of the 'pure' gravity equation in the literature[5]. One reason for the astounding volume of this literature is its empirical 'success' with the usual 'R^2 s in the range of .65 to .95' (Harrigan (2004)). But the more important reason is that we can use the model as a benchmark to estimate a variety of hypotheses regarding trade. It is common to ask questions like: do we really know that the WTO increases trade[6], does a currency union affect trade[7], and how important is the Linder hypotheses (trade increases with similarity of per capita incomes) in deciding bilateral trade flows by using the gravity model as a benchmark. It should be noted that using the model as a benchmark

---

[1] There have been several attempts with disaggregated data as well (see, for example, Feenstra et al (2001)).

[2] Anderson (1979), Anderson and van Wincoop (2003), Bregstand (1985, 1989), Helpman and Krugman (1985) use monopolistic competition and the Armington assumption (goods are differentiated by location of production) either exogenously or endogenously, Feenstra et al. (1998, 2001) and Eaton and Kortum (2002) use a homogeneous products framework, Evenett and Keller (2002) shows why the gravity equation is a "fact of life" for the Heckscher-Ohlin model or the increasing returns to scale model or a marriage between them while Feenstra et al (1998) derive the gravity equation in a reciprocal dumping model.

[3] See, for example, Anderson (1979), Bergstand (1985) and Anderson and Wincoop (2003).

[4] See, for example, Helpman and Krugman (1985), Helpman (1987), Eaton and Kortum (2002),

[5] See Sanso et al (1993) for some experimentations with the functional form of the 'pure' gravity model.

[6] See, for example, Rose (2004).

[7] See Glick and Rose (2002)

implicitly implies that it accounts for some of the more important 'natural' reasons for trade. Additional factors come in to explain what remains unexplained after these natural factors have taken their course.

In addition to its role as a benchmark, another important contribution of the gravity model is the importance that it assigns to distance. Mainstream trade theory typically consigns distance to the margins[8]. Is that a good strategy for modeling purposes? During the years of the formulation of the neo classical models of international trade from Smith to Heckscher-Ohlin and their extensions, transport costs were substantial and there could be no doubt that these costs played a role in determining the flow of commodities between countries. But has distance died in the modern era? The answer is by no means unambiguously positive. In fact, there is a lot of evidence to the contrary[9]. Thus the issue of distance as a determinant of trade flows is still an open issue and the gravity model is perhaps the only model that addresses the issue directly.

Given the above perspective in this paper we try to evaluate the influence of the three variables suggested by the gravity theory in determining India's DOT in the second half of the twentieth century[10]. We also attempt some simple extensions of, what we call, the 'core' gravity model by incorporating dummies and additional variables for size in the basic structure.

An important point needs to be noted regarding the empirical estimation of the gravity model in the present context. To estimate such a model we ideally need bilateral trade data on pairs of countries chosen with certain criteria. The model is not naturally geared to explain the DOT of a particular country. If we attempt to use it for that purpose with purely cross section data one of the three variables in the model (GDP of the home country in this case) becomes constant. One alternative is to work with panel data where the variable at least varies over data points recorded over time. However, one should interpret the home GDP variable with caution in this case. If the time series component is not long, the variable may be confounded with the intercept term thus resulting in multi-collinearity.

The rest of the paper is arranged as follows: In the next section we discuss the basic theory of the gravity equation and the equations that are actually estimated. Section 3 explains the data sources and the basic layout of the data. Section 4 estimates the gravity equation for India both in its core and augmented forms. Section 5 concludes the paper.

[8] As Leamer and Levinsohn (1995) point out "(w)hy don't trade economists "admit" the role of distance into their thinking?" (p 1385).
[9] See Brun et al (2005)
[10] Batra (2004) and Nag and Nandi (2006) are two of the many examples of the application of the gravity model to the Indian case for studying India's trade potential and dynamics vis-vis particular regions of the world. However to our knowledge, this is the first attempt to estimate a gravity model for India vis-à-vis all its trading partners for a such a long time series.

2. The Gravity Model

2.1 Theory

According to the Newtonian law of gravity if two objects A and B, D_{ab} distance apart, have masses M_a and M_b respectively then the force of mutual attraction F is given by

$$F = \frac{GM_a M_b}{D_{ab}},$$

where G is a constant. Replacing F by the total volume of trade between two trading partners i and j (TV_{ij}), M_a and M_b by the GDP of i and j (Y_i and Y_j) and D_{ab} by the distance between i and j (D_{ij}) we get the gravity model of international trade:

$$TV_{ij} = G\frac{Y_i Y_j}{D_{ij}}. \qquad (1)$$

Anderson (1979) gives a simple theoretical derivation of the equation. He shows that for cross section studies the Cobb-Douglas expenditure system yields a gravity type equation minus the distance factor (called the 'frictionless gravity model'). His formulation is as follows: suppose we are considering balanced trade between two trading partners i and j in a world that consists of only these two countries. Suppose preferences are Cobb-Douglas so that a constant fraction of income is spent on goods originating from each country. Also suppose that θ_i is the proportion of income spent on goods from country i. Finally suppose E_i is the export of country i (to j) and I_i is the import of country i (from j). Then:

$$E_i = \theta_i Y_j$$

$$I_i = \theta_j Y_i$$

Due to balanced trade[11]:

$$TV_{ij} = E_i + I_i = 2\theta_i Y_j (= 2\theta_j Y_i) \qquad (2)$$

However,

$$Y_i = \theta_i Y_j + \theta_i Y_i = \theta_i (Y_j + Y_i) = \theta_i Y_W \qquad (3)$$

Putting the value of θ_i from (3) into (2)

$$TV_{ij} = 2\frac{Y_i Y_j}{Y_W} \qquad (4)$$

Since for cross section studies the world income (Y_W) is a constant (4) exactly looks like the frictionless version of (1)[12]. Anderson (1979) then generalizes (4) to include transport cost of the iceberg type (that is the cif price is pt for a fob price of p, t>1)[13]. Anderson and Wincoop (2003) has reformulated the model to yield the following reduced form:

---

[11] See Kim et al (2003) for a method to relax this assumption in the actual estimation process.

[12] One problem with (4) or (5) in our case is that they are completely cross sectional in their specification in the sense that world income cannot be considered as a constant over time. The solution is simple: take into account the variability of world income over time. The other problem is more fundamental, especially for India. This concerns the assumption of balanced trade.

[13] Anderson and Wincoop (2003) derives another reduced form that is more intuitive.

$$TV_{ij} = 2t_{ij}^{1-\sigma}\frac{Y_iY_j}{Y_W}\times(P_iP_j)^{1-\sigma} \qquad (5)$$

Where σ (>1) is a parameter that is constant across countries[14] and P_i and P_j are the CES price indices of counties i and j. Assuming transport cost t_{ij} to be an increasing function of distance[15], (5) yields a version of the gravity equation (1) with the price terms as add-ons[16]. To estimate this equation P_i can be interpreted as:

$$P_i^{1-\sigma} = \sum_j (Y_i / Y_W)(t_{ij} / P_j)^{1-\sigma}$$

The price terms in (5) play an important role. It says that controlling for 'relative distance' is crucial for correctly specifying a gravity equation. Relative distance captures the fact that if there are two countries A and B at an equal distance from another country C then C will buy more from A than B if A is more isolated with the rest of the world than B. This is because the isolated country will have lower price for its export goods due to lower demand of their export goods from other countries (due to the distance).

2.2 Estimation

Estimating the gravity equation has two distinct traditions. The first and the older tradition were to estimate modified versions of (1) without any reference to economic theory. The reason was obvious; there was no economic theory then to fall back upon. The recent tradition, which began after economists found theoretical underpinnings of the equation, was to estimate the theoretically derived reduced form. However, remnants of the older tradition remained dominant even after theoretically derived versions made their way into the literature. The reasons, possibly, are the simplicity and the empirical success of the model in explaining the bilateral trade flows.

Let us consider the older tradition first. Given the nature of the data that we are working with in this paper, in what follows we consider a setup where (a) we model bilateral trade of a particular country (the 'home country') with the rest of the world and (b) we assume that the data have both time series and cross section components. Indexing the home country by I (for India, which is the home country in the empirical exercise) and taking logarithm of (1) we get:

$$\ln TV_{Ijt} = \alpha + \ln Y_{It} + \ln Y_{jt} - \ln D_{Ij} + \varepsilon_{Ijt}$$

Where TV_{Ijt} is the total value of trade (export plus import) between India and its bilateral trading partner j at time period t, Y_{It} is India's Gross Domestic Product (GDP) at time t, Y_{jt} is the GDP of the trading partner at time t, D_{Ij} is the distance between India and country j, $\alpha = \ln G$ and ε_{ijt} is the error term. Since unit elasticity for Y_{It}, Y_{jt} and D_{Ij} need to be confirmed from the data a more general form for the basic equation is considered:

$$\ln TV_{Ijt} = \alpha + \beta_1 \ln Y_{It} + \beta_2 \ln Y_{jt} + \mu \ln D_{Ij} + \varepsilon_{Ijt}\,.$$

Though the above specification is sufficient for empirical purposes the following additions are common: (a) the distance variable is expanded to include dummies such as common

---

[14] σ comes from the utility function. The generalized utility function is a variant of the CES functional form: $U^i = \left[\sum_j \beta_j x_{ij}^{(\sigma-1)/\sigma}\right]^{\frac{\sigma}{1-\sigma}}$ where β_j is a parameter

[15] See Hummels (1999) for the evidence of the actual functional form

[16] See Harrigan (2003) for the interpretation of the price terms

borders and land locked countries (b) since GDP is a measure of size other variables determining the size of countries such as population are also considered. Including them expands the above equation as follows:

$$\ln TV_{Ijt} = \alpha + \beta_1 \ln Y_{It} + \beta_2 \ln Y_{jt} + \gamma_1 \ln N_{It} + \gamma_2 \ln N_{jt} + \mu \ln D_{ij} + \rho^k DUM_{Ij}^k + \varepsilon_{ijt} \qquad (6)$$

Where N_i (N_j) denotes population of country i (j) and DUM denotes dummies as defined above[17]. Note that the expected sign for the γ's is negative. This is because small economies (like Singapore) tend to be more open than large economies (like USA). The following two equivalent modified versions of (6) are particularly popular (see, for example, Frankel (1997)):

$$\ln TV_{Ijt} = \alpha + \beta \ln Y_{It} Y_{jt} + \gamma \ln N_{It} N_{jt} + \mu \ln D_{Ij} + \rho^k DUM_{Ij}^k + \varepsilon_{ijt} \qquad (7)$$

$$\ln TV_{Ijt} = \alpha + \beta \ln Y_{It} Y_{jt} + \gamma \ln (Y_{It} / N_{It})(Y_{jt} / N_{jt}) + \mu \ln D_{Ij} + \rho^k DUM_{ij}^k + \varepsilon_{ijt} . \qquad (7a)$$

There are two additional interpretations of these versions. First, if β <1 then we can conclude that as the combined size of the two countries increases trade increase less than proportionately. This is another way of confirming the fact that smaller economic entities (consisting in this case of India and its trading partner j) are more open to trade and as the size of the economic entity increases trade does not increase that fast. On the other hand a positive and significant γ in (7a) implies that trade increases with the economic development of the combined entity defined by the two countries.

How (6) to (7a) are further modified for actual estimation depends on the purpose of the study. For example if the intention is to address particular issues such as 'does membership to a customs union increase India's trade' then the dummy variables are expanded to address this issue. The dummy variables are also routinely expanded in some works to capture such effects as culture, language, colonial links etc.

As far as actual estimation of (6) to (7a) is concerned, if the data is pure cross section then OLS is used with or without fixed effects for the bilateral trading partners. The Hausman criterion is used to choose between fixed and random effects if the data is a panel with a relatively small time span. If the time span is long then unit root tests are conducted to determine whether OLS is spurious or cointegration is present. The following factors have also been sometimes considered to improve the quality of estimates by different authors in the estimation procedure: unobservable heterogeneity, omitted variables bias, bilateral specific effects, whether some explanatory variables are correlated with bilateral random effects, whether some of the variables are endogenous, taking account of missing values, sub period regressions.

Coming to the recent tradition of taking theoretical underpinnings literally, the most common reduced form, that is considered, is (5). Even though (5) is not exactly the gravity specification

[17] Sanso et al (1993) uses Box-Cox transformations on this basic form to arrive at a data determined functional form.

(Anderson and Wincoop (2003) call it the 'generalized' gravity equation), as we have seen there is a clear interpretation of the price terms in terms of 'relative distance'. Three methods are followed in estimating (5). First, if data on prices are available then they are used directly (Baier and Bergstrand (2001)). Second the prices can be estimated from the structural model by using non-linear least squares (Anderson and Wincoop (2003)). Third, if data on prices are not available, as pointed out by Anderson and Wincoop (2003), country specific effects can be used to proxy for the price terms.

3. The Data

3.1 Data Sources and Notes

India's export and import data with all its trading partners has been taken from the Direction of Trade Statistics, IMF. The trade data is in US $. To bring it to real terms we have divided it with the consumer price index (CPI) of the US. The US CPI (for all urban consumers) data is taken from the US Department of Labour (Bureau of labour statistics). Data for GDP is taken from the PENN world table. We have multiplied CGDP (real per capita GDP in current US $ i.e. current GDP divided by current price) from PENN by population (taken from the same source) to arrive at the aggregate GDP of each country. The data for distance is taken from Centre D'etudes Prospectives Et D'informations Internationales[18].

There are a total of 177 countries with which India had trade relations *at least once* between 1950 and 2000. Therefore the total number of potential observations for the present work is 9027. The total number of observations used in regression is 3990. Therefore there are 5037 observations that are potentially missing. However it should be noted that the 177 countries include those like Yugoslavia, USSR and West Germany, which ceased to exist after a point of time as well as the names of all the countries into which they (dis)integrated. Thus the actual potential was slightly lower than 9027. A large portion (about 70 per cent) of the missing observations is due to the fact that India had trading relations with many countries in a limited number of years between 1950 and 2000. The rest are due to non-availability of data on the independent variables.

As far as the trade data is concerned, all countries whose names appear in the DOT yearbook but have no figures reported corresponding to it have been taken as 'not available'. But for countries whose names do not appear in the yearbook at all (corresponding to a particular set of years), we have taken total trade as zero. A total of 1012 observations out of the 3990 used in regressions have zero value for the dependent variable (trade). We have replaced these by .00001 for models where log values were used.

3.2 The Trade Data

India was one of the least open economies of the world in the second half of the twentieth century[19]. Export pessimism, a policy of import substitution and a declared goal of self-

---

[18] http://www.cepii.fr/anglaisgraph/bdd/distances.htm, under file name dist_cepii.xls and dist_cepii.dta.
[19] India was the 7th least open economy of the world in 1996 with the ratio of total trade volume (export plus import) to GDP being .24 as against a world average of .44 (International Financial Statistics data quoted in Rodriguez (2000))

reliance throughout most of the period resulted in an environment that can easily be called anti-trade. This, however, does not mean that all trade was stifled out. In fact, notwithstanding the government's anti trade policies both exports and imports in India increased rapidly throughout the period. Since only those trade transactions that were absolutely necessary were encouraged and since imports of unavailable necessary goods were crucial to its sustenance, the growth rate of import was far greater than the growth rate of export. Thus almost throughout the period, India had a large and rapidly increasing balance of trade deficit. The composition of trade was typically that of an extremely underdeveloped country slowly finding its feet. Apart from petroleum, the main import items were capital goods like machinery for its upcoming industries. Earlier food was also an important import item but its importance declined as the green revolution took shape. As far as exports are concerned, the importance of labour intensive industrial items like jute, tea and cotton textiles sharply declined over the period and the importance of more capital intensive light engineering goods rose sharply. Capability of value addition in the jewelry industry resulted in an important re-export item over the years with the raw material for this industry being imported as 'Pearls and semi precious stones' and the end product being exported as 'Gems and Jewelry'. Given the nature of import items (as necessities) and export items (as the ones that faced competition from other exporters) the price elasticities of demand and supply of export and import followed the textbook pattern: exports were more price sensitive than imports, the export elasticity being quite high[20].

As far as the direction of trade is concerned, the number of export destinations more than doubled between 1950 and 2000. The importance of the top country or even the top two countries had waned considerably. Thus looking merely at the numbers it would seem that India diversified its export destinations to a large extent in the second half of the twentieth century. This, however, is a misleading picture as far as the value of exports is concerned. In terms of value, the change had been mainly cosmetic: (1) the top five countries continued to account for around fifty percent of exports throughout the period[21], (2) by the end of the century about 86% of countries were, what we may call, 'smalltime' export partners absorbing less than 1% of total exports. Moreover the proportion of these 'smalltime' trading partners rose significantly (by more than 50%) during the period 1950-2000. United States had been a very important trading partner of India in this period both in exports and imports while the importance of United Kingdom, which was very important in the first two decades since 1950, waned over the years. USSR was important during the 1970s and the 1980s.

3.3 Correlations

Is there an association between the variables suggested by the gravity theory and India's bilateral trade flow? In this section we take an initial look at the data in terms of correlations to try and understand the basic character of the data in the Indian case. First let us look at the simple correlations of the variables. Denoting the natural logarithm of a variable by 'L' table 1 reports these correlations.

Two important facts related to our context are evident from the table. First, India's bilateral trade (LTRADE) is positively correlated with both the GDP of the trading

[20] For example Joshi and Little (1994) calculated the short run and long run relative price (of Indian and competitor country exports) elasticity of export demand as 1.1 and 3.0 respectively from 1962 to 1987.
[21] There was some decline though, from 62% in 1950 to 44% in 2000

partner (LAGGGDP) and that of itself (LINDGDP) (all the variables are in natural logarithmic scale). Further both are significant at the 1% level (.276) though the magnitude of the former is more than double that of the latter. Next, LTRADE is negatively correlated with distance (in natural logarithmic scale) and the correlation is significant at 1% level. In a way, this substantiates the basic premise of gravity model.

Since, following the suggestion of the gravity model, we'd study the dependence of trade on the combined mass or, as in this case, the combined GDP of the two trading partners on trade, we combine the GDPs in estimating the core gravity model. There are two straight routes of combining the GDPs: adding them or multiplying them. The gravity model clearly supports the latter over the former, as the two GDPs appear in multiplicative form in this model.

One possible problem with using GDPs in multiplicative form (that, to our knowledge, does not find its place in the literature) is that the multiplicative term does not distinguish between two equally large countries and one small and another large country. This is likely to be an important distinction to make when we are dealing with country pairs, but this is not clearly relevant in our case as we are dealing with a particular country vis-à-vis its trading partners. Following the notations of the previous section Table 1 thus gives us a firm indication regarding the choice of the functional form of the gravity model that is relevant for India's bilateral data:

$$\ln TV_{Ijt} = \alpha + \beta \ln Y_{It} Y_{jt} + \mu \ln D_{Ij} + \rho^{k} DUM_{Ij}^{k} + \varepsilon_{ijt} \quad (8)$$

4. Estimation Results

Before we go into the results it should be noted that the panel we are considering is unbalanced. Thus each country's effect on the estimated coefficients is not equal. However the data is complete (that is, from 1950 to 2000) for the set consisting of the union of the sets of the first 10 trading partners in any year[22]. Thus the unbalanced nature of the panel is mainly due to the absence of data regarding particular years of countries that can be considered as 'smalltime' trading partners of India. In fact trade with them, in the years they occur, is less than one per cent of India's total trade volume in that particular year.

4.1 The Core Gravity Model

Following the discussion in sections 2 and 3 let us begin by estimating the following model:

$$\ln TV_{Ijt} = \alpha + \beta \ln Y_{It} Y_{jt} + \mu \ln D_{Ij} + \varepsilon_{ijt} \quad (9)$$

Table 2 reports the estimates of the parameters of the equation using a variety of estimation methods. Let us first look at the estimates by using OLS. The first point to note is that the core gravity model (without any augmentation through the use of dummies) can explain about 43 percent of India's direction of trade. Secondly, since β is significant but less than one in magnitude, India's trade increases with size but it increases less than

[22] The only exception is the USSR prior to its breakup after 1991, which is entirely excluded due to the non-availability of data for per capita GDP in the PENN.

proportionally. On the other hand (since μ is negative and greater than one) it decreases with the distance of the trading partner and it does so more than proportionally. As we had noted in the Introduction, most gravity models estimated over pairs of countries often for major portions of the world have an R^2 of .65 or above. Most of these models use a large number of dummy variables to capture effects far beyond these basic factors. Though it is hard to guess what the R^2s would be without them, the fit of the model to Indian data does not seem to be far from the benchmarks set by the literature. Overall therefore the gravity model performs well for India and both size and distance play important roles in shaping India's direction of trade. It should be noted that the period under consideration is a period when policies played a very important role in shaping India's trade. Indeed throughout most of the period India followed an extremely conservative trade regime and, as we have already noted India was one of the least open economies of the world. To the extent that the gravity model is policy neutral, the relative success of the model is not entirely on the expected lines. The result highlights the fact that in spite of the stringent regimentation of policies, India's trade continued to be guided by natural factors to a large extent throughout the second half of the twentieth century.

Turning now to the model with country dummies, the value of the Hausman (1978) statistic for testing fixed versus random country effects is found to be 62.43, thus suggesting that the fixed effects model is more suitable than the random effects model. The analysis with fixed effects model then shows that *all* the country dummies are statistically significant and they together explain as much as 70 percent of the variation in India's trade. Thus individual country factors are very important in modeling India's direction of trade and they need to be given a serious look to improve the performance of the core gravity model. Secondly, accounting for these country specific factors, the estimates have changed significantly. The effect of size has diminished by about 44 percent and the effect of distance has increased more than three times. Thus distance rather than size is important for India's trade, once the country characteristics are brought to the fore.

Finally, since, as we have already reported, about 25 percent of the values of the dependant variable are zero, we have also provided estimates by using the Tobit model (last column in table 2). Though it is common in the literature to follow this method whenever the dependent variable has a disproportionately large number of zeros irrespective of whether a clear interpretation of the latent variable is available or not (see Maddala, 1992), we can fortunately do better here. This is because there is always a potential for trade between any two countries of the world. However lack of information regarding products, trade policies, and difference in culture, climate, and topology etc. does not always allow the realization of the potential. In the cases where the potential is realized we have mutual exchange of goods between the countries however if such mutual exchange fails to materialize then there is no way of measuring the potential. Thus "the potential of trade" is a convenient latent variable here. The assumption is that if the potential surpasses a threshold value then trade is noted in reality, otherwise, for all potentials below the threshold value observed trade is zero. Since the Tobit estimates do not use country dummies therefore the values of the coefficient for size variable is higher and that for LDIST is lower in absolute value than in the fixed effects case. However the coefficient of the size variable is still lower than the OLS case and the value of the distance variable is higher implying that most of the unfulfilled potential is with countries that are physically distant from India. Distance rather than size is more important if we consider the effect of the unrealized potential explicitly.

Overall therefore we can conclude from an analysis of the core gravity model that both size and distance play an important role for Indian bilateral trade. The effect of distance becomes sharper if we bring in individual country factors explicitly into the model as well as when the countries and years for which the trade potential is not realized is explicitly modeled.

4.2 Some Augmented Gravity Models for India

The most popular extension of the gravity model is by incorporating dummies in the basic equation (that is, estimation of (8) instead of (9)). We will begin with the dummies that are used almost as a default mechanism in almost all augmented gravity models and hence may be considered as a part of *any* gravity model[23]: 1. Whether the two countries are contiguous (Contig) 2. Share a common official language, (Comlang) 3. Have had a common colonizer (Comcol) 4. Was the same country at any previous point in time (Smctry). However Comlang in our case is very similar to Comcol (major exceptions include Egypt and Malaysia which were British colonies but do not currently have English as an official language). So is Contig to Smctry (exceptions: Nepal, Bhutan and Mainland China which are included in the first but not in the second). So we proceed with the two following dummies: Contig and Comcol. The first dummy adds a new dimension to the geography and trade relationship imperfectly captured by the distance variable in equation (9). The second dummy, though explicitly political in its definition spills over into several other inter-disciplinary areas including culture and history. Let us turn to the result first (to avoid cluttering the analysis we report simple OLS results only, there are major changes in the other estimates reported in table 2):

Ltrade =	-15.64	+.95 $\ln Y_{It}Y_{jt}$	-2.56 $\ln D_{Ij}$	-3.88 Contig	+1.27 Comcol
	(-11.29)	(49.02)	(-21.44)	(-10.84)	(10.13)

$\bar{R}^2 = .46$

First note that the adjusted R^2 has improved a little compared to (9). Thus some of the conventional dummies included in the gravity model do not significantly help in explaining India's trade further. The results are nonetheless extremely interesting. Being contiguous actually *hurts* rather than helps India's trade! The reason is obvious. India's trade with its neighbors is small (except China, which clearly happened so recently that it could not influence the result sufficiently) compared to the majority of its other trading partners[24]. It is thus way below potential given their proximity. Interestingly, this result seems to rather strongly suggest that distance is not such an important factor in India's trade as being near means nothing in terms of trade volume. But the magnitude and the significance of the distance variable in this equation is actually higher than in equation (9)

[23] There are other such 'default dummies' like 'the number of landlocked countries', 'number of island nations' etc. which are not obviously relevant for studying the bilateral trade of a particular country.
[24] It should be noted that all trade between the neighbors and India are not reported, as there is a huge amount of illegal trade across the porous borders. Thus there is bound to be a lot of reporting bias in the data. The negative sign can thus also be put down to errors in observations. India does not have very cordial relations with many of its neighbors, especially Pakistan. Thus political factors also have their due share in the sign.

implying that the smallness of the volume of trade with distant countries in Latin America and Africa have overwhelmed this result in the aggregate. However India had very strong trade links with the former colonies of Britain in the second half of the twentieth century and this is clearly reflected by the sign and magnitude of the coefficient of Comcol and also its statistical significance.

Since the relationship with Contig is inconsistent with expectations we drop this dummy for the rest of the analysis. After dropping Contig (8) reduces to:

Ltrade =	-22.34	+.96 $\ln PY_{It}PY_{jt}$	-1.83 $\ln D_{Ij}$	+1.25 Comcol
	(-17.87)	(48.47)	(-18.28)	(9.90)

$\bar{R}^2 = .44$

A second set of possible augmentations of the basic gravity model is the inclusion of more explanatory variables. One popular extension that has an interesting interpretation is (7a). As we have already noted if γ in (7a) is positive and significant then it can be said that (given India's per capita GDP) India trades more with richer countries than with poorer countries. Ideally, as in (7a), the product of the per capita incomes should be considered together with the products of the GDPs. However in our case the correlation between the two is as high as .84. To avoid multicolinearity we reframe (7a) by excluding the product of the per capita incomes:

$$\ln TV_{Ijt} = \alpha + \gamma \ln(Y_{It}/N_{It})(Y_{jt}/N_{jt}) - \mu \ln D_{Ij} + \rho^k DUM_{ij}^k + \varepsilon_{ijt}\,. \qquad (10)$$

Note (10) is not quite a gravity equation, however, it uses the same framework as the gravity equation. The estimates of the coefficients of the equation are:

Ltrade =	13.77	+.97 $\ln PY_{It}PY_{jt}$	-2.72 $\ln D_{Ij}$	+0.56 Comcol
	(13.38)	(29.86)	(-24.21)	(3.96)

$\bar{R}^2 = .28$

The equation clearly confirms the well-known fact regarding India's trade - India trades more with richer countries than with poorer countries (P in front of the variable implies per capita income). The interesting point to note about the estimate is the rather sharp decline in the adjusted R^2 for the equation compared not only to (8) but also to (9). Aggregate size and not per capita income explains India's trade better.

On the other hand if the augmentation is done with an additional determinant of size, viz. population, as in (7) then the estimated coefficients change to:

Ltrade =	-37.91	+.63 $\ln Y_{It}Y_{jt}$	+.79 $\ln P_iP_j$	-1.40 $\ln D_{Ij}$	+1.56 Comcol
	(-23.59)	(21.70)	(14.83)	(-13.75)	(12.47)

$\bar{R}^2 = .47$

Thus India has a higher tendency to trade with more populous countries than with less populous countries. The significant rise in the adjusted R^2 compared to (8) reemphasizes the strength of the size variable in explaining India's trade. Comparing the adjusted R^2s in the estimates of equations (7) and (10) the following conclusion clearly emerges: size is more important than the level of development in explaining India's direction of trade.

5. Conclusion

The main conclusions that emerge from our analysis are as follows: (1) The gravity model can explain about 43 to 50 per cent of the fluctuations in India's direction of trade in the second half of the twentieth century (2) India's trade responds less than proportionally to size and more than proportionally to distance (3) Colonial heritage was still an important factor in determining India's direction of trade in the second half of the twentieth century (4) India trades more with developed rather than underdeveloped countries, however (5) size has more determining influence on India's trade than the level of development of the trading partner.

This paper has just made a beginning as far as the analysis of India's direction of trade is concerned. The analysis is obviously far from complete. The most important task would be to model the un-modeled part of individual country effects (reflected by the country dummies). The effect of trade policy, membership to trading blocks and the WTO, economic liberalization of the 1990s and especially a more through analysis of the distance variable are some of the issues that need urgent attention. Also the panel data analysis needs to be supplemented by a pure cross section analysis and a cointegration analysis to sharpen the conclusions. All these will have to be taken up in the future.

Acknowledgements:

We are deeply indebted to Sarmila Banerji of Department of Economics, University of Calcutta for her detailed comments on an earlier version of the paper that improved it substantially. We are also indebted to Atri Mukherjee, Susmita Roy, Satadru Chakraborty and Subhabrata Sarkar for help in collection of data and computations. The second author also gratefully acknowledges the funding support in the form a seed money project provided by Indian Institute of Management Ahmedabad.

Table 1: Correlation Between the Gravity Variables and Trade for India (1950 to 2000)

	LTRADE	LINDGDP	LAGGGDP
LINDGDP	.276*		
LAGGGDP	.688*	.515*	
LDIST	-.335*	-.109	-.176

Notes: '*' implies Correlation is significant at the 0.01 level (2-tailed).

Table 2: Estimates of a Gravity Model for India (1950 to 2000)[1]

Variables	OLS[2]	FE[3]	Tobit[6]
$\ln Y_{It} Y_{jt}$	.93 (47.02) [54.26]	.52 (30.65) [26.04][4]	.79 (31.55) [24.22][7]
$\ln D_{Ij}$	-1.96 (- 9.43) [-16.39]	-6.39 (-7.95) [-58.74][4]	-3.19 (-30.19) [-21.11][7]
$\bar{R}^2$	.43	.76[5]	–

Notes: 1. See text for definition of variables. Number of observations 3990. 2.Figures in first brackets are t-values. Figures is third brackets are t-values after using White Heteroscedasticity corrected covariance matrix. The Breusch and Pagan statistic is 375.55 (2 d.f) 3. Constants are omitted wherever applicable. 3. FE is OLS with country fixed effects. Fixed effect dummies, *all* of which are significant are not reported. 5. $\bar{R}^2$ for fixed effect dummies only is .70. 4. t-values using White Heteroscedasticity corrected covariance matrix. 5. Random effects. Hausman Statistic for Fixed vs Random effects is 62.43. 6. The primary index coefficients of the model are outside the brackets in this column. LM test for Tobit = 3387.32. 7. Coefficient/standard error assuming disturbance to be heteroscedastic.

References

Anderson, J.E.(1979) A Theoretical Foundation for the Gravity Equation. *American Economic Review*, 69, 106-16.

____________. and E. van Wincoop (2003) Gravity with Gravitas: A Solution to the Border Puzzle. American Economic Review,93, 1, pp.170-192.

Batra, A. (2004) India's Global Trade Potential: The Gravity Model approach, Working Paper No. 151, Indian Council for Research on International Economic Relations, New Delhi, India.

Baier, Scott. L. and Bergstrand, J. (2001) The Growth of World Trade:Tariffs, Transport Costs & Income Inequality, Journal of International Economics, 53, 1-27.

Bergstand, J. (1985) The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidance, *The Review of Economics and Statistics*, 67, 474-81.

___________ (1989) The Generalized Gravity Equation, Monopolistic Competition and the Factor-Proportions Theory in International Trade. *The Review of Economic and Statistics*, 71, 143-53.

Brun, J.F., C. Carrere, P. Guillaumont and J. de Melo (2005) Has Distance Died? Evidence From A Panel Gravity Model, *The World Bank Economic Review*, 1-21.

Deardorff, A.V. (1995), "Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World?", NBER Working Paper # 5377, Cambridge MA

Eaton J. and S. Kortum (2002) Technology, Geography and Trade, *Econometrica*, 70, 5.

Evenett S. J. and W. Keller (2002) On Theories Explaining The Success of The Gravity Equation, *Journal of Political Economy,*110, 2, 281-316.

Feenstra, R.C., J.A. Markusen, and A. K. Rose (1998) Understanding the Home Market Effect and the Gravity Equation: The Role of Differentiated Goods. NBER Working Paper No. 6804.

_______________________________________ (2001) Using The Gravity Equation to Differentiate Among Alternative Theories of Trade. *Canadian Journal of Economics*, 34, 2, 430-47.

Frankel, J.A. (1997) *Regional Trading Blocks in the World Economy*, Institute of International Economics, New york.

Glick, R. and Rose, A.K. "Does a Currency Union Affect Trade? The Time Series Evidence", *European Economic Review*, 46, 6, 1125-51. (2002)

Harrigan J. (2003) Specialization and The Volume of Trade, in E. Kwan. Choi aand J. Harrigan (-eds) *Handbook of International Trade*, Blackwell Publishers.

Hummels, D. (1999) Toward a Geography of Trade Costs, Purdue University, Manuscript.

Helliwell, J.F. (1997) National Borders, Trade and Migration, NBER Working Parer No. 6027.

Helpman, E. and Krugman. P. R. (1985) Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition, and the International Economy. MIT Press. Cambridge.

Helpman (1987) Imperfect Competition and International Trade: Evidence From Fourteen Industrialized Countries. *Journal of Japanese and International Economics*.Vol.1.

Joshi, V. and Little, I.M.D. (1994) *India, Macroeconomics and Political Economy, 1964-1991*.Oxford University Press. Delhi.

Kim, M., G. D. Cho and W. W. Koo (2003) Determining Bilateral Trade Patterns Using A Dynamic Gravity Equation, Agribusiness & Applied Economics Report No. 525, Centre for Agricultural Policy and Trade Studies, North Dakota State University.

Leamer, E.E. and Levinson, J. (1995) International Trade Theory: The Evidence in G.M. Grossman and K. Rogoff (Ed.) *Handbook of International Economics*. Vol.3. 1339- 1394.

Maddala (1992) *Introduction To Econometrics*. Maxwell Macmillan Publishing Company. New York.

Nag, B. and Nandi, A. (2006) Analysing India's Trade Dynamics vis-à-vis SAARC Members Using The gravity Model, South Asia Economic Journal, 7, 83-98.

Rodriguez A. C., (2000)) *On the Degree of Openness of an Open Economy* mimeo, Universidad del CEMA Buenos Aires, Argentina

Rose A.K. (2004) Do We Really Know That the WTO Increases Trade? *American Economic Review*, 94, 1,98-114.

Sanso M., R. Cuairan, F. Sanz (1993) Bilateral Trade Flows, The Gravity Equation, and Functional Form *The Review of Ecinimics and Statistics*, 75, 2, 266-275.

Savage, I. R. And K. W. Deutsch (1960), "A Statistical Model of the Gross Analysis of Transactions Flows", Econometrica 28, (July), pp. 551-572.