Currency unions and gravity models revisited

Christie Smith

November 2002

Discussion Paper Series

DP2002/07

Currency unions and gravity models revisited

Abstract[1]

Gravity models have been shown to be fairly effective in modelling bilateral trading patterns, explaining more than 50 per cent of the variation in trade. This paper examines bilateral trade patterns using a data set provided by Rose and van Wincoop (2001). Rose (2000) has suggested that forming a currency union has a dramatic effect on the volume of intra-union trade. A number of econometric issues are identified with respect to this claim. There is some evidence that Rose's (2000) empirical results are not entirely robust to the sample of countries used, and to the estimation method. In particular, some of the regressors may be endogenous, which casts doubt on the magnitude of the parameter estimates.

1   Introduction

Newton's theory of gravity asserts that the force exerted by two objects is a function of their respective masses and the square of the distance between them. Analogues of this theory have been applied in a number of different contexts, primarily to explain interactions with spatial dimensions, such as trade or migration. Gravity models represent trade between two economies as a function of their respective economic masses, the distance between the two economies, and a variety of other factors. Gravity models of bilateral trade were originally developed by Pöyhönen (1963) and Tinbergen (1962), but with little in the way of theoretical justification.

Gravity models have been used to investigate a number of empirical regularities. One avenue of this empirical research has been to investigate whether border effects inhibit international trade. For example, Helliwell (1996) and Helliwell (1998) find that trade between Canadian provinces far outweighs trade between the provinces and US states, taking into account distance and economic mass. McCallum's (1995) initial investigation suggested that intra-Canada trade was approximately 22 times larger than trade between Canadian provinces and US states; Helliwell's (1998) estimate was similar in magnitude.[2]

In a similar piece of work, Fitzsimons et al. (1999) use a gravity model to investigate the border effect between Ulster and the Republic. They find that Ulster and the Republic trade together more than would be predicted by their gravity model, taking into account the Republic and Ulster's common language and common land border. This is an interesting result, given that Ulster and the Republic use different currencies – the British pound and Irish punt respectively (though up

until March 1979 they were exchanged costlessly at par; more on Ireland below).[3]

Following on from the border-effects literature, Andrew Rose and a number of co-authors (e.g. Rose, (2000), Frankel and Rose (2002), Rose and van Wincoop (2001), Glick and Rose (2002)) have attempted to show the impact that currency arrangements have on bilateral trade. To be more specific, Rose (2000) and his co-authors argue that two countries sharing the same currency trade roughly three times as much as they would if they used different currencies. In order to capture the partial impact of currency arrangements Rose (2000) also takes into account structural and institutional features – such as common language, common colonial history, the presence of trade agreements, etc. – features that might also be correlated with a 'common currency' dummy variable. The impact of currency union is also found to be distinct from currency volatility.[4]

Rose (2000) conducts numerous robustness checks of his overall results. He finds that his conclusions about the statistical significance of parameters are robust to the choice of estimation techniques, the sample period, and choice of sub-samples of countries. Using essentially the same data, Rose and van Wincoop (2001) reach the same conclusion.

Gravity models were originally introduced as atheoretical, albeit plausible, empirical models. Subsequent work established micro-foundations for the basic building blocks that appear in gravity models. Deardorff (1998) argues that the basic regressors of gravity models – distance and income – are actually implied by a wide variety of theoretical models. In contrast, Feenstra, Markusen and Rose (2001) argue that empirical gravity models can be used to discriminate between alternative trade theories.

---

[3] Over half of the data sample used by Fitzsimons et al. (1999) is post-1979, during which Ireland was part of the European Exchange Rate Mechanism (ERM) and the United Kingdom was not.

[4] Rose does not explain what the precise benefits of currency union are, beyond a fairly general reference to transactions costs.

The currency union literature that exploits gravity models is not explicitly grounded in theory; rather it has an empirical focus. Consequently, it behooves investigators to take econometric issues very seriously, since it is the empirical content of currency union gravity models that makes them interesting.

Andrew Rose has accumulated a large panel data set on bilateral trade flows, and has made this data available on the Internet.[5] In this paper we use the data from Rose and van Wincoop (2001). The main goal of the paper is to reconsider how gravity models have been used by Rose and his co-authors to investigate currency union. Are the data appropriate? Are gravity models plausible representations of the relationships between various pieces of data? Are the inferences that are drawn appropriate? Are the econometric problems solvable? To foreshadow the conclusions – personal prejudices will largely determine whether you believe that Rose and his co-authors have adequately resolved the econometric issues. One has to acknowledge, however, that they have tried very hard to do so.

2 Analytical framework

The basic gravity equation for trade is:

$$Trade_{ijt} = A d_{ij}^{\beta_1} \left(y_{it} \cdot y_{jt} \right)^{\beta_2} \qquad [1]$$

In equation [1] trade between country i and j at time t is a function of the distance between the two countries, d_{ij},[6] and the product of the incomes of countries i and j at time t (y_{it} and y_{jt} respectively). At various times below this log of the product of GDPs will be referred to as the 'GDP' regressor or as the 'income' or 'output' regressor. A, β_1 and β_2 are coefficients; β_1 is assumed to be negative and β_2 is assumed to be positive.

---

[5] See http://www.haas.berkeley.edu/~arose/.

[6] The distance between two countries does not vary over time, and hence does not require a time subscript.

It is generally more convenient to deal with this equation in logs,[7]

$$\ln(Trade_{ijt}) = \beta_0 + \beta_1 \ln(d_{ij}) + \beta_2 \ln(y_{it} \cdot y_{jt})$$ [2]

To investigate issues such as currency union additional variables are included in the specification,

$$\ln(Trade_{ijt}) = \beta_0 + \beta_1 \ln(d_{ij}) + \beta_2 \ln(y_{it} \cdot y_{jt}) + \beta_3' \mathbf{X}_{ijt} + u_{ijt}$$ [3]

In this equation X_{ijt} is a column vector of the additional variables (all vectors are column vectors unless otherwise specified); β_3' represents a vector of coefficients; and u_{ijt} is a stochastic error term. The parameters for the log variables above can be interpreted as elasticities (provided that the changes in regressors are small).

The 'additional variables' that have been considered include: time dummies to account for possible institutional changes that may have affected trade (time dummies can also be interpreted as a time trend); common currency dummies; common border dummies; regional trade agreements; various geographical characteristics such as contiguous borders; and a number of others associated with common heritage (e.g., common colonisers, common language, and the like). See the data appendix or Rose (2000) for more details.

The functional form implies a great deal of symmetry. To be more specific, the functional form is essentially blind to the direction of trade, since trade is measured as the (natural log of the) sum of exports and imports. Among the regressors, incomes enter as a product, so that they have a common parameter.[8] In principle, this is a testable restriction, one that Tinbergen's original analysis on exports rejected, though it does appear to be fairly reasonable in much of Helliwell's

---

[7] Coe and Hoffmaister (1999) argue that using logs is inappropriate in considering African data, since it results in censoring a large number of zero-trade observations (since the log of zero is undefined).

[8] This restriction has also been justified by noting that if one of the countries has zero income we expect it to undertake zero trade (e.g. Deardorff, 1998).

(1998) analysis of the United States and Canada. Ideally, one would prefer to model the individual trade flows directly, without imposing this type of restriction, in order to capture possible asymmetries. (For example, a country pair might impose different average tariffs on each other's exports, or a colony might trade more with its coloniser than vice-versa.) However, this avenue has not been explored here.[9]

The gravity model of Newtonian physics also embodies no third-party effects. In other words, the presence of a third particle C does not affect the gravitational force between A and B. This is not necessarily appropriate when one thinks of trading a given endowment of resources, because of the opportunity cost involved. For example, a country may exchange butter for guns or, alternatively, it may exchange butter for the latest consumer electronics. Thus, a model of trade between country A and B must also take into account the conditions and opportunities that prevail between A and C and B and C (given that trade is defined as the sum of imports and exports). This is perhaps the most important insight provided by Anderson and van Wincoop (2001a,b).

3 Econometric issues

3.1 Currency union and the endogeneity of income

Frankel and Rose (2002) are not simply trying to establish that currency unions improve trade. Rather, their research agenda is to establish that currency unions would be a welfare-improving monetary framework – see also Rose and van Wincoop (2001). The line of argument is approximately that currency unions increase trade, and trade increases output, and thus welfare.

---

[9] Peter Kenen has noted that it is possible to back out most of the individual income data by identifying US GDP in the corresponding years. The remaining income data can usually be found by taking one intermediate step – finding the income of a country that trades with the United States and thence the income for the country that does not trade directly with the United States. A similar approach could be used to divide trade into its export and import components.

Paradoxically, this line of reasoning highlights that at least one of the regressors in equation [2] is endogenous. In assessing the impact of trade on output, Frankel and Rose (2002) use the exogenous regressors of the gravity model to form a predicted trade variable, which is then used as an instrument for a regression where the dependent variable is income. In essence, the income regressor has shifted across to the LHS of the equation, and an instrumented version of trade has shifted to the RHS. Frankel and Rose (2002) interpret this exercise as suggesting that trade does facilitate growth. In contrast, Rodrik has suggested that the causality might not necessarily be so straightforward: he suggests that geography might affect both trade and output simultaneously.[10] Theil and Galvez (1995), for instance, document that per capita GDP is indeed correlated with latitude.

Frankel and Rose's (2002) approach suggests that output – a regressor in the gravity model – is not actually exogenous. This begins to cast doubt on the estimation of the gravity model and its conclusions about the relationship between currency union and trade. (The empirical details of this problem will be elaborated further in the following sections.)

The theoretical foundations of Anderson and van Wincoop (2001a) do not provide a great deal of support for the Frankel-Rose endeavour either, since an underlying assumption of the structural model is that the supply of each good is fixed. The Anderson and van Wincoop (2001a) model is about allocating a given level of output between domestic and foreign regions, taking into account preferences, prices, and barriers to trade. Naturally, an increase in trade might still be welfare-enhancing as a result of comparative advantage, but it cannot increase the underlying level of output in a given country.

---

[10]   See http://ksghome.harvard.edu/~.drodrik.academic.ksg/shortpieces.html.

3.2 Is currency union exogenous with respect to trade?

The previous section has argued that income is an endogenous variable in the gravity models that are used to assess the impact of currency union. This problem may be compounded by the fact that forming a currency union – or staying within a currency union – is also an endogenous choice. Rose and van Wincoop (2001) argue that:

> "Reverse causality also does not explain away the findings [that currency union affects trade]; there is little evidence in the political science literature that countries join currency unions to increase trade, and instrumental variables only increase the impact of currency unions on trade".

To make this endogeneity issue concrete consider the case of Ecuador. There are certain elementary considerations that Ecuador had to address in 2000 when choosing to adopt a foreign currency. Obviously Ecuador was not going to adopt a currency that had an unstable internal value, since they already had their own unstable, high-inflation currency. There are, of course, quite a few currencies that have stable internal values (as opposed to stable external prices/exchange rates, which may be another story). Countries such as Canada, the United Kingdom, the Eurozone countries, Sweden, Switzerland, New Zealand, Australia and the United States, among others, all have currencies that could have been adopted as Ecuadorian legal tender.

How then to discriminate between the currencies of these different countries? Given that the United States is Ecuador's biggest export market, taking 38 per cent of Ecuadorian exports in 1997, it is perhaps not particularly surprising that Ecuador decided to US-dollarise, since dollarising would also minimise trade transactions costs with the United States. Europe, by way of comparison, received 19 per cent of Ecuador's trade, while 'other' unspecified countries took 28 per cent, making other currencies less desirable than the US dollar. The US market is also one of the largest in the world, and the Ecuadorians

probably felt confident that there was scope for their trade with the United States to continue to grow. It is hard to believe that Ecuador's decision to US-dollarise was independent of trade considerations.[11]

Rose (2000) discusses another currency union transition – Ireland's abandonment of the peg against the pound sterling. He argues that reverse causality – trade causing currency union – does not explain Ireland's entrance into the Exchange Rate Mechanism (ERM) in 1979. Indeed, Thom and Walsh (2001) note that political and inflation concerns were the major motivations for the change in the exchange rate regime. Nevertheless, it is worth noting that Ireland's trade with the European mainland had increased substantially during the years prior to joining the ERM. And, at least initially, there was some expectation that the United Kingdom would also enter the ERM (the United Kingdom did eventually join the ERM in 1990).

Thom and Walsh (2001) investigate Irish-Anglo trade in some detail. If currency union is trade-enhancing then breaking a currency union should be trade retarding. In their case study, Thom and Walsh find that Ireland's entry into the ERM did not retard Irish-Anglo trade. They suggest that Rose's (2000) results for Ireland reflect the measurement of Irish trade in US dollars deflated by US price indices.[12]

The Bahamas provides another explicit example where trade patterns have affected the exchange rate regime. Up until 1968 the (Bahamian) Commissioners of Currency acted as a currency board, exchanging local currency for pounds sterling at parity. However, as noted by DaCosta et al. (1999):

> "In light of the shift in the country's trading pattern in favor of the United States, and the increasing reliance of the Bahamian economy on the United States for tourism revenue and

---

investment, in 1967 the government decided to switch the currency peg to the U.S. dollar at the rate of B$1 to US$1.''

Similarly, Hargreaves and McDermott's (1999) discussion of currency union issues for New Zealand explicitly considered New Zealand's trading pattern and the lack of concentration in New Zealand's trade. In other words, notwithstanding the earlier quote from Rose and van Wincoop (2001), trade concentration does appear to influence both the discussion and decision regarding the benefits of currency union (or its cousin, currency-board arrangements).[13] Together, these examples cast some doubt as to whether currency unions can be treated as exogenous to trade.

The econometrics of this problem is comparatively straightforward. Rose argues that the currency union dummy is not endogenous. This suggests that, in the context of a single equation model, the currency union dummy does not appear on the left hand side. However, a standard simultaneous equations perspective is perhaps rather more appropriate. The parameter estimates that are commonly presented rest on the assumption that trade, currency-union and output are exogenous variables. Suppose that the relationships are linear, and let X represent the exogenous variables (such as distance, perhaps),[14] with Z = [output, CU, trade]', where output, CU, and trade are $T \times 1$ vectors (CU, for example, is the vector of dummy variables indicating whether a country pair is in a currency union or not). Such a model can be represented as:

$$Z \cdot \Gamma = X \cdot \beta + U$$

where U is the $T \times 3$ matrix of stochastic error terms. Assuming that the diagonal elements of Γ are ones, the third equation can be re-written as:

---

[11] Rose (personal communication) identifies financial considerations as being important too.

[12] Rose's response to the Thom and Walsh paper is Glick and Rose (2002), see http://haas.berkeley.edu/~arose/RecRes.htm. This paper is subject to problems that are discussed below.

[13] In the Rose data set the Bahamian currency board arrangement is treated as being equivalent to currency union.

[14] Distance can be interpreted as an exogenous proxy for transportation costs, which might also be endogenous.

$$trade = -[output \; cu] \cdot \begin{bmatrix} \Gamma_{1,3} \\ \Gamma_{2,3} \end{bmatrix} + X \cdot \beta_3 + U_3 \qquad \qquad [4]$$

In this equation $\Gamma_{i,j}$ is the i,j-element of Γ, β_3 is the third column of the matrix β, and U_3 is the third column of U.[15] However, one cannot get unbiased estimates for $[\Gamma_{1,3} \; \Gamma_{2,3}]'$ from such a regression because the endogenous variables, output and currency union, are correlated with the error term U_3 (assuming that the error terms from the three equations are cross-correlated). In other words, if the elements of U_3 are correlated with elements in U_1 and U_2, then one has econometric difficulties, because these error terms affect the first and second equation, which in turn affect output and trade – regressors in equation [4] above.

The usual limited-information approach to dealing with this simultaneous equations problem is to use instrumental variables. But what is an appropriate instrument for currency union? If U exhibits only contemporaneous correlation then lagged output and lagged currency union could be used as instruments. Output is well-known to exhibit substantial inertia, thus satisfying the requirement that the intended regressor and instrument be strongly correlated. (It seems likely that the currency union dummy exhibits a similar property – though whether this holds for observations separated by a period of five years is perhaps more debatable.) Lagged variables will also not be correlated with the contemporaneous error, assuming that the error is in some sense well-behaved. To put this more concretely, $output_{t-1}$ and u_{3t} will not be correlated, where $output_{t-1}$ and u_{3t} are the $(t-1)^{th}$ and t^{th} observations of output and U_3 respectively. The use of lagged variables, however, does not work if U exhibits serial correlation. This is likely to be an issue if there are country-specific omitted variables.

The problem of finding good instruments for currency union is further complicated by the fact that the CU variable is binary. Binary variables can only be perfectly correlated with other binary variables. Rose (2000) uses various functions of inflation rates as instruments for the

---

15   The notation is similar, though not exactly comparable, to that of equation [3].

variance of exchange rates, and for the currency union parameter. However, one of the key requirements for an instrument is that it should be correlated with the endogenous regressor. What one finds is that the inflation measures that Rose (2000) uses are poorly correlated with the currency union dummy variable (of the three inflation instruments the most highly correlated, for the full sample of data, has a correlation coefficient of −0.012).

Unfortunately, it seems unlikely that a good instrument actually exists for currency union. Thus, the endogeneity problem for currency union may not have a solution.

3.3   Omitted variables

Many currency unions, though not all, reflect a historical relationship, such as colonisation. Suppose for the moment that the currency union variable, CU, and other regressors in the gravity model are exogenous. Rose (2000) and his co-authors seek to ensure that the coefficient estimated for the CU variable is appropriate by including controls that might be correlated with both the CU dummy variable and trade. To be confident that the coefficient estimated for the CU variable is appropriate one would have to assume that the selection of controls adequately encompasses all of the other factors that positively affect trade. One would also have to assume that the relationship between trade and the control variables has been appropriately specified.

Consider the Ecuadorian case once more. Ecuador is a beneficiary of the generalised system of preferences (GSP) with respect to the United States, as well as the European Union and Japan. Potentially, then, the GSP may increase Ecuadorian exports to the United States, at least relative to the 'expected' exports derived from a gravity model estimated on the 'world' (or some similarly large group of countries). In other words, there is a variable that helps to explain high Ecuadorian-US trade and, in turn, this higher-than-predicted trade is likely to have affected the choice to US-dollarise.

The 'world', of course, includes countries that are not beneficiaries of the GSP, countries such as Australia and New Zealand. To the extent that it is countries like Ecuador – countries whose weak economic institutions encourage them to dollarise or adopt the currency of some other area – the currency union dummy may also be capturing the impact of the Generalised System of Preferences. Grether and Olarreaga (1998) note that preferential trade represents approximately 40 per cent of world trade. Failing to account for these agreements may thus have a material impact on the estimate of the currency union coefficient. Furthermore, even if currency union does improve trade, extrapolating the behaviour of the GSP-advantaged countries to countries such as New Zealand and Australia may be inappropriate.

To deal with the omitted variable problem, Glick and Rose (2002) use time series evidence in the post World War II era to examine the effect of transitions into and out of currency unions. They argue that there is a substantial loss of trade associated with exiting a currency union. However, as Patrick Honohan (2001) notes in his discussion of Persson (2001) and Rose (2001), and as reiterated by Thom and Walsh (2001), many of these transitions represent the breakup of colonial relationships, the adoption of autarkic policies in the wake of colonialism, and oftentimes civil war. In these circumstances it is not surprising that one would subsequently witness a decline in trade. Thus, even the time series evidence based on exchange rate regime transitions can be contested.

3.4   Are the currency union results general or unique?

It is instructive to consider the data from the currency union and non-currency union countries respectively. Table 1 records the means of the data from currency union and non currency union sub-samples of data, and also provides standard errors for the means of the currency union sample. The standard errors for the data means have been constructed by treating the full-sample of data as the population of interest.[16] One

---

can then assess whether the sub-sample means are statistically similar to the expected values of the whole population (the full sample means).

This is an important comparison to make if one wants to generalise the impact of currency union to those countries that have independent currencies. What one finds in table 1 is that currency union country pairs tend to have lower incomes, they tend to be closer (are much more likely to have a common border), are more likely to have a regional trade agreement, and are much more likely to have historical ties, such as language, the same colonizer, or indeed are much more likely to be the same country (such as France and its overseas dependencies). In short, the currency-union pairs of countries may not be representative of the wider sample. Rose (2001) performs a similar exercise on the IMF data set that Glick and Rose (2002) later used. He claims that the means of variables for currency union and non-currency union countries are the same in that data set, with the exception of the common language dummy. Still, this does not really address the issue presented by the data set used here and by Rose and van Wincoop (2001), and by the data used in Rose (2000), which is largely the same.

Table 2 reports parameter estimates from regressions on data where the country-pairs are in currency unions (a sample with 252 observations) and on data where the countries are not in currency union (26356 observations). The specification of the regression is as close as is possible to that in Rose and van Wincoop (2001). There are sizable differences in the point estimates of the parameter values from the two sub-samples, though for the most part the parameter estimates obtained from the currency union data (CU = 1) are not significantly different from the parameters estimated on the whole population of countries. The exception appears to be the distance parameter, which, rather curiously, is smaller for the currency union countries. (We will see below that the lower distance coefficient is true of trade with African and Caribbean countries, which provide most of the currency union observations.)

---

[16]   A sample of 252 is actually quite large, so the central limit theorem implies the means of the currency union sub-sample are, to a reasonable approximation,

normally distributed. The variances of the means were also calculated by repeatedly re-sampling from the population (the full sample) with replacement. This exercise confirmed that the variance implied by the normal distribution was appropriate.

Persson (2001) attempts to cope with the problems raised by table 1 (and to a lesser extent table 2) by using a treatment effect approach adopted from health and labour economics. The essence of this approach is to find a control group that has similar characteristics to the 'treated' group (i.e., the group of countries that has been 'treated' with a currency union). One then compares the mean trade of the treated group with the mean trade of the (appropriately similar) control group. Persson (2001) finds that the effect of currency union is much reduced in this comparison, and is not statistically significant. Of course, questions remain about the degree of comparability between the control group and the treated group.

4   Panel data regressions

4.1   Cleaning the data

The variety of 'countries' that are included in the Rose data set may cause sceptics to be concerned about the validity of the empirical results. To some degree, scepticism about the empirical results can be allayed by using diagnostics to search the data set for influential observations, i.e., observations that have a material effect on the parameter estimates. Krasker, Kuh and Welsch (1983) provide econometric tools that can be used to identify observations that have a material effect on parameters.

Krasker, Kuh and Welsch (1983) show that one cannot identify influential observations simply by looking at the magnitude of the residuals; one also has to take account of the 'leverage' that such residuals have, which is a function of the regressors. They show that the difference between $\hat{\beta}$ and $\hat{\beta}(i)$, where $\hat{\beta}(i)$ is the parameter estimate one obtains from deleting the ith row in the data, can be represented as:

$$\hat{\beta} - \hat{\beta}(i) = \frac{(X'X)^{-1} x_i e_i}{1 - h_i} \qquad [5]$$

where xi is the ith row of the regressor matrix, and $h_i = x_i(X'X)^{-1} x_i'$. If X is perfectly balanced then $h_i = \sum_{i=1}^{T} h_i / T$.

Krasker, Kuh and Welsch (1983) argue that observations that have leverage greater than $3\sum_i h_i / T$ are potentially influential. We take a conservative approach and look at leverage points that are greater than $5\sum_i h_i / T$. (Another approach would be to focus on data rows that change the currency union parameter by ±0.01.) Using this metric on the full Rose data set (for the 5-yearly data from 1970 to 1990) yields 523 influential rows, from a data set with 26608 rows of data. In the terminology of Krasker, Kuh and Welsch (1983), these rows can be examined for 'gross errors', numbers that may have been mis-identified or mis-entered. This enables one to establish observations that have been mis-coded (e.g., Suriname was formerly part of the Netherlands and never part of Portugal), as well as identifying other rows of data that might also be rather questionable.

The techniques used to identify influential observations also show how difficult it is to represent geographical characteristics using dummy variables. For example, Indonesia-Malaysia and Indonesia-Papua New Guinea are both identified as being influential. If one looks at these data one finds that Indonesia and Malaysia are regarded as having a common border, but not a common language (though Malay and Bahasa Indonesia are generally considered to be mutually intelligible). The common border dummy for Indonesia-Malaysia is perhaps a trifle ironic, since the common land border is in Borneo, between Kalimantan (Indonesia) and Sabah and Sarawak (Malaysia). Borneo is of course one of the more remote regions of the world – having a common land border in Borneo is unlikely to greatly facilitate trade. Similarly, the common Indonesia-Papua New Guinea border in New Guinea is also unlikely to stimulate trade greatly, given that it connects two remote areas with low population density and given that the infrastructure connecting the two countries is actually very poor.

The South Pacific Regional Trade Agreement (Sparteca) is also revealed as another data inconsistency. Sparteca opens Australian and New Zealand markets to small South Pacific countries, but does not imply free trade among the Pacific Island countries themselves. Thus, the regional trade agreement dummy variable for Tonga-Kiribati, Kiribati-Samoa, Niue-Samoa, and Niue-Fiji, etc., appears to be mis-coded.[17]

As one can see from equation [5] above, the effect that a residual has on the parameters depends on the regressors. When the CU and the sameco (same country) dummy variables are included in the regression all currency-union and same-country pairs are considered to be influential. This is not surprising given how few observations there are of such variables.

Nevertheless, bar removing all such observations, one finds that influential observations have an approximately neutral effect on the estimate of the currency union parameter – there are as many positive adjustments as there are negative. Thus, there is no single observation, and no small sub-sample of currency union observations, that drives the currency union estimate. Taking greater care with the language variable also does not appear to make a great deal of difference.[18] We shall see in the next section that this does not necessarily imply that the estimate of the currency union effect is robust. The impact of the currency union effect depends crucially on the parameters that one estimates, particularly the estimate of the income coefficient.

---

[17] Article II(a) states that the objective is "to achieve progressively in favour of Forum Island countries duty free and unrestricted access to the markets of Australia and New Zealand over as wide a range of products as possible".

[18] There are quite a few data errors associated with French as a lingua franca. Nevertheless, replacing the CIA language data with a more authoritative source, such as www.ethnologue.net, does not make a material difference to the analysis.

4.2 Regional subsets

One of the sensitivity checks that Rose (2000) has performed is an examination of various sub-groups of countries. For example, he has examined the results on purely intra-Less Developed Country (LDC) trade. He also considered parameter estimates when Australia, France, New Zealand and the United States are excluded from the sample. In a third experiment Rose (2000) excludes African countries; a fourth experiment excludes Europe and countries in the Pacific. A fifth experiment excludes countries from the Caribbean, and the Americas.

Here we take a slightly different tack. Rather than excluding subsets of countries, we focus on regional subsets (see table 3). In general, one finds that many of the influential observations arise from trade between African countries, as well as Belize, Panama, Kiribati, and a number of Caribbean countries. Since the CFA Franc area and the member countries of the Eastern Caribbean Central Bank provide many of the currency union observations (approximately 100 observations in each) it seems sensible to consider Africa, the Caribbean and 'the rest-of-the-world' separately, to examine whether the parameters of the equation are in fact stable over different geographical areas. These samples overlap, since a data point is included if one of the two countries is from a particular region. Thus, for example, Bahamas-Liberia trade affects both the Caribbean and African regressions.

Table 3 presents the regression parameters associated with regressions run on these geographical sub-samples. The results indicate that the effect of currency union varies sharply across the different regions. Most notably, the point estimate of the currency union parameter for the Caribbean is actually negative, though it is not statistically significantly different from zero.

Many of the other regressors are also quite different across the various geographical regions. Interestingly, the parameter estimates on the income variable are markedly lower for both Africa and the Caribbean in comparison to the world as a whole. Rather curiously, distance appears to be somewhat less important for both Africa and the

Caribbean than for all countries on average. An earlier version of this paper documented that the impact of distance varies considerably depending on the region being considered. It was suggested that this result reflects the heterogeneous distribution of income and hence transportation networks; other indicators of trade costs support the view that geographical differences are important. South Pacific countries, for example, appear to be more severely affected by distance than European countries.

Having borders appears to be much more important for African countries than it does for the rest of the world. (The Caribbean border effect is imprecisely estimated because there is only one land border in the Caribbean, between Haiti and the Dominican Republic.) Interestingly, language does not appear to be a very important determinant of trade in Africa, when compared with its impact on trade elsewhere in the world. The constants in these regressions also differ fairly substantially and this suggests that either a fixed or random effects model might be appropriate.

The main point of this exercise is to indicate that the estimate of the currency union parameter is not entirely robust to the sample of countries that is used in the analysis. Furthermore, there are substantial differences in the parameter estimates, which is suggestive of important heterogeneity. And it does appear that the estimate of the currency union effect depends in a material way on the parameter estimates that one obtains. The types of regional experiments that Rose performed – excluding countries or groups of countries – ensure that the great weight of data provide some 'average' perspective on the parameter estimates, even when the various countries were excluded. The key issue was touched on earlier: are the cross-sectional results general? Can different countries be treated as though they were similar, so as to be informative about each other?

4.3 Endogeneity revisited

As noted above, distance and income provide most of the explanatory power in all of the regressions. The various dummy variables, and other variables such as per capita GDP and the product of areas, provide little additional explanatory power. If these so-called 'nuisance factors' are uncorrelated with the currency union variable then omitting them from the regression would not bias the estimate of the currency union coefficient. It is interesting to consider what happens when a stripped-down regression is estimated, with trade as the dependent variable and distance, GDP, and the currency union variables as regressors. What one finds using ordinary least squares (see table 5) is that the distance and GDP coefficients are much as expected, and the currency union coefficient is large and positive (much as in the more elaborate regressions).

Suppose that the currency union variable can be treated as exogenous (as is required by Rose's (2000) research programme). It is reasonable to assume that the distance variable is also exogenous.[19] However, even in a stripped down regression this still leaves questions about the income variable – which should not be regarded as exogenous, since Frankel and Rose (2002) argue that trade increases income. What then can be used to instrument for income? Fortunately, Rose (2000) provides us with a suitable instrument: the log product of land areas. The correlation between this variable and income is approximately 0.65. Unfortunately, this variable fails to capture the growth of income over time, which means that one has to abandon the advantages of having a panel data set. However, one can still examine each of the cross-sections to gauge the impact of the currency variable. The regression results obtained from the different cross-sections prove to be fairly similar (see table 5).

---

[19] The location of firms – and hence their distance from other countries is not necessarily invariant (consider the development of Northern Mexico post-NAFTA) – but typical measures of distance will not capture such developments.

The impact of using instrumental variables is dramatic. It is reassuring that the impact of distance remains roughly the same (give or take a 15 per cent impact on the elasticity). However, the estimated coefficient for income drops away dramatically, nearly halving in size, to a level closer to that obtained from the Caribbean sub-sample (see table 3). Similarly, the coefficients estimated for the currency union dummy variable also drop substantially.

It turns out that the estimate of the income coefficient is highly material in determining the magnitude of the currency union effect, irrespective of whether or not other variables besides income and distance are included (excluding of course land area, which is being used as an instrument, and GDP per capita, which must be endogenous if income is endogenous). One can establish this quite easily by conducting a grid-search over 'plausible' income elasticities, and looking at the currency union parameter estimates that arise given the restriction on the income elasticity. There is quite clearly a positive relationship: a higher income elasticity implies a larger currency union effect. The estimates of the income elasticity obtained from instrumental variables are much lower, and thus imply a much lower currency union effect. If the income elasticity is below about 0.45 then the impact of the currency union effect appears to be negative.

The reason for this result is obvious once one plots the income-trade data, highlighting the observations involving the currency union countries. Most currency union-pairs have comparatively low incomes. Thus, by reducing the income coefficient one raises the predicted values for low-income country pairs in an ex-currency union regression (since the regression hyper-plane still passes through the vector of data corresponding to the sample means). Consequently, the regressions errors for the currency union countries are correspondingly smaller – since the currency union effect is simply a dummy variable, the coefficient value is determined by the average regression error for the currency union countries. Thus, the currency union effect is materially related to the size of the income coefficient.

One might question whether these lower parameter values are plausible. Naturally, there are theoretical reasons that might predispose us to believe that the coefficient on GDP is close to 1. However, there are reasons why the coefficient might be less than 1 for some countries. Suppose that British income goes up by 10 per cent for example. Does this necessarily imply that demand for Caribbean bananas will also go up by 10 per cent? It seems arguable that this might not, in fact, be the case.

Of course, there is an inherent econometric tension in what has been done here. The areaprod variable was already included in the regression. By excluding areaprod (which is necessary to be able to use it as an instrument) and by excluding log per capita GDP (LNGDPP), for which we do not have a second instrumental variable, the regression may suffer from an omitted variable bias.

Additionally, pre-testing of the exclusion restrictions (for example using an F-test based on the R^2's from the two regressions), indicates that the zero-ing out restrictions are not valid. However, this result is clearly being driven by the fact that the sample size, n = 26,608, is very large, blowing up the denominator of the statistic. But again, questions remain about whether the data come from the same generating process for all countries.

One can illustrate this point by considering regional regressions once more. Is it, for example, legitimate to assume that the coefficient on the currency union dummy is the same across the Caribbean and across the rest of the world? This can be tested by interacting the currency union dummy variable with a dummy variable indicating whether a country pair has a member in the Caribbean, and testing whether this interacted variable has a zero coefficient. The answer is no – such a restriction does not hold. The F-statistic is F[1,n-k] = 16.0, which is far in excess of the critical value (which is approximately 3.84). The data also reject the restriction that the GDP regressor is the same for the Caribbean countries too, even more strongly (the statistic is circa 390).

This section has tried to provide a flavour of some of the econometric problems that affect the estimation of the currency union parameter. Once again, there is some suggestion that the estimated parameter is not entirely robust. But the econometric problems are such that one's prejudices will largely determine whether one believes currency union has a material effect on trade.

5   Conclusion

The main goal of this paper was to re-examine the use of gravity models to assess the impact of currency union. The paper emphasised that the examination of currency union using gravity models is primarily an econometric exercise, one without explicit theoretical foundation. Consequently, this paper focussed on some of the empirical and econometric issues that arise with respect to currency union effects.

Using econometric techniques we have established that the empirical estimates of the currency union effect, as typically calculated, are not driven by a single trade relationship or a small number of trade relationships. These techniques also help identify influential data observations, some of which represent data coding errors. While these techniques support the importance of the currency union effect some of the other issues cannot be resolved so easily.

The first of these issues is that the currency union dummy variable and income regressors may not actually be exogenous. Endogeneity raises serious questions about the possibility of inference. Instrumental variables (IV) estimation was used to take account of income's endogenous character. The IV estimates of the income elasticity and the currency union coefficient were substantially smaller than the corresponding OLS estimates. It turns out that the currency union effect is closely related to the magnitude of the income coefficient, and the IV results raise question about the income estimates that are usually obtained. (Though these IV estimates are also not above reproach.)

The second problem is that the currency union country pairs may not be representative of the rest of the sample. Extrapolating the impact of

currency union from those countries to the rest of the sample may lead to inappropriate conclusions. The currency union and other parameters were found to vary substantially across regional sub-samples. If Rose's results are to be convincing then one must feel comfortable extrapolating from the currency union countries to the rest.

Rose and others have concentrated on the currency union coefficient, and have paid less attention to the parameter estimates for language, free-trade areas, coloniser-colony, common coloniser, common borders, and the like. Such structural features all prove to be important and are interesting in their own right. In aggregate, these factors dwarf the importance of currency union. Clearly, many of these factors are predetermined with respect to trade (one cannot swap one's colonial history for another). However, these results do suggest that structural efforts to promote trade, such as the closer economic relations (CER) agreement between New Zealand and Australia, do have a material impact on trade. This has implications both for trade policy, and potentially for immigration policy as well. For example, it is likely to be valuable to further develop trade agreements, and it may also be useful to encourage immigrants to maintain their linguistic ties with their countries of origin.

Still, it is worth re-emphasising that the bulk of the explanatory power for trade in gravity models comes from income and distance (perhaps best regarded as a proxy for transportation costs). In terms of explanatory power, these two factors dwarf all of the others mentioned here.

References

Anderson, J. E. & van Wincoop, E. (2001a), 'Borders, trade and welfare', *NBER Working Paper* 8515.

Anderson, J. E. & van Wincoop, E. (2001b), 'Gravity with gravitas: A solution to the border puzzle', *NBER Working Paper* 8079.

Coe, D.T. & Hoffmaister, A.W. (1999), 'North-South trade: Is Africa unusual?', *The Journal of Financial Research* 8(2), 228-56.

DaCosta, M., Gómez-Oliver, A., Itam, S., Piñerúa, C. & Shah, S. (1999), 'The Bahamas: Selected issues and statistical appendix', *IMF Staff Country Report* 99(106).

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, Chicago.

Feenstra, R. C., Markusen, J. R., & Rose, A. K. (2001), 'Using the Gravity Equation to Differentiate among Alternative Theories of Trade', *Canadian Journal of Economics*, 34(2), 430–47.

Fitzsimons, E., Hogan, V. & Neary, P. J. (1999), 'Explaining the volume of North-South trade in Ireland: A gravity model approach', *Economic and Social Review* 30(4), 381–401.

Frankel, J. A. & Rose, A. K. (2002), 'An estimate of the effect of common currencies on trade and income', Mimeo. http://haas.berkeley.edu/~arose. *The Quarterly Journal of Economics*, 117(2). Forthcoming.

Glick, R. & Rose, A. K. (2002), 'Does a currency union affect trade? The time series evidence', *European Economic Review*. Forthcoming. See also http://haas.berkeley.edu/~arose.

Grether, J.-M. & Olarreaga, M. (1998), 'Preferential and non-preferential trade flows in world trade', *WTO Staff Working Paper* 98-10.

Hargreaves, D. & McDermott, C. J. (1999), 'issues relating to optimal currency areas: theory and implications for New Zealand', *Reserve Bank of New Zealand Bulletin* 62(3), 16–29.

Helliwell, J. F. (1996), 'Do national borders matter for Quebec's trade?', *Canadian Journal of Economics* 29(3), 507–22.

Helliwell, J. F. (1998), How Much Do National Borders Matter?, The Brookings Institution, Washington D.C.

Honohan, P. (2001), 'Discussion', Economic Policy*: A European Forum*, 0(33), 457–59.

Krasker, W. S., Kuh, E. & Welsch, R. E. (1983), Estimation for dirty data and flawed models, in Z. Griliches & M. D. Intriligator, eds., 'Handbook of Econometrics', North-Holland Publishing Co., Amsterdam, pp. 651–698.

McCallum, J. (1995), 'National borders matter: Canada-U.S. regional trade patterns', *American Economic Review* 85(3), 615–623.

Persson, T. (2001), 'Currency Unions and Trade: How Large Is the Treatment Effect?', Economic Policy: A European Forum 0(33), 433–48.

Pöyhönen, P. (1963), 'A tentative model for the volume of trade between countries', *Weltwirtschaftliches Archiv* 90(1), 93–99.

Rose, A. K. (2000), 'One money, one market: Estimating the effect of common currencies on trade', *Economic Policy: A European Forum* 0(30), 7–33.

Rose, A. K. (2001), 'Currency Unions and Trade: The Effect Is Large', *Economic Policy: A European Forum* 0(33), 449–57.

Rose, A. K. & van Wincoop, E. (2001), National money as a barrier to international trade: The real case for currency union, The *American Economic Review: Papers and Proceedings* 91(2), 386-390.

Theil, H. and Galvez, J. (1995), 'On latitude and affluence: The equatorial canyon', *Empirical Economics* 20, 163–66.

Thom, R. & Walsh, B. (2001), 'The effect of a currency union on trade: Lessons from the Irish experience', *European Economic Review*. Forthcoming.

Tinbergen, J. (1962), Shaping the World Economy: Suggestions for an International Economic Policy, New York.

Wei, S.-J. (1996), 'Intra-national versus international trade: How stubborn are nations in global integration?' *NBER Working Paper*, No. 5531.

Tables

See the data appendix for explanation of variable names.

Table 1: Means of trade and regressors for sub-samples

	Data Means				
	Full sample	Sub-samples			
Column	(2)	(3)	(4)	(5)	(6)
				Standard	Significant
	CU = {1,0}	CU = 0	CU = 1	Errors*	Difference?**
CONSTANT	1	1	1	0.000	No
CU	0.01	0	1	0.006	Yes
LNDIST	8.18	8.2	6.47	0.052	Yes
LNGDP	34.38	34.43	28.87	0.173	Yes
LANGUAGE	0.15	0.19	0.98	0.022	Yes
BORDER	0.03	0.03	0.11	0.010	Yes
REGTA	0.02	0.02	0.3	0.009	Yes
COMCOL	0.09	0.08	0.71	0.018	Yes
COLONY	0.01	0.01	0.03	0.007	No
SAMECO	0	0	0.14	0.003	Yes
LNGDPP	16.23	16.24	15.25	0.086	Yes
LOCKED	0.18	0.18	0.16	0.025	No
AREAPROD	24.08	24.12	20.01	0.211	Yes
No. of Obs.	26608	26356	252		

* The standard errors are for the CU=1 data means; the standard errors were calculated by re-sampling (from the whole sample) 30,000 sub-samples with 252 observations.

** Col. 6 indicates whether the CU = 1 data means are greater than 3 standard deviations from the col. 2 means

*** The specification is comparable to that of Rose and van Wincoop (2001)

Table 2:
Currency union/not-union sub-sample regressions

	Regression Results					
Column	(1)	(2)	(3)	(4)	(5)	(6)
	ALL Countries		CU = 0 Countries		CU = 1 Countries	
	Parameters	Std Errors	Parameters	Std Errors	Parameters	Std Errors
CONSTANT	-19.07	0.26	-18.99	0.26	-13.55	2.27
CUNION	1.37	0.14	#N/A	#N/A	#N/A	#N/A
LNDIST	-1.04	0.02	-1.05	0.02	**-0.63**	0.17
LNGDP	0.92	0.01	0.92	0.01	0.82	0.12
LANGUAGE	0.54	0.04	0.54	0.04	**1.01**	1.20
BORDER	0.61	0.08	0.56	0.08	0.98	0.46
REGTA	1.02	0.07	1.11	0.07	1.48	0.56
COMCOL	0.37	0.06	0.38	0.06	-0.54	0.49
COLONY	2.10	0.07	2.07	0.07	2.11	0.47
SAMECO	0.95	0.32	1.43	0.57	0.54	0.44
LNGDPP	0.48	0.01	0.48	0.01	0.13	0.16
LOCKED	-0.25	0.03	-0.26	0.03	-0.66	0.41
AREAPROD	-0.14	0.01	-0.14	0.01	-0.07	0.11
D1970	1.66	0.04	1.66	0.04	1.42	0.51
D1975	1.40	0.04	1.40	0.04	1.37	0.44
D1980	1.02	0.04	1.02	0.04	0.84	0.29
D1985	0.25	0.04	0.24	0.04	0.40	0.28
No. Obs.	26608		26356		252	

1. Columns (1) and (2) are from a regression on data from all countries in the data set.
2. Columns (3) and (4) are for a regression on countries *not* in a currency union, i.e. the data are from country-pairs with CU = 0.
3. Columns (5) and (6) are for a regression using data from country-pairs *in* a currency union, i.e., CU = 1.

Table 3: Geographical sub-region regressions

Countries	ALL		Not-Caribbean		Not-Africa		Caribbean		Africa	
	Parameters	Std Errors	Parameters	Std Errors	Parameters	Std Errors	Parameters	Std Errors	Parameters	Std Errors
CONSTANT	-19.07	0.26	-19.14	0.28	-21.91	0.19	-14.52	0.34	-14.86	0.59
CUNION	1.37	0.14	1.79	0.16	0.90	0.21	-0.21	0.26	-0.76	0.24
LNDIST	-1.04	0.02	-0.99	0.02	-0.99	0.02	-0.67	0.05	-0.76	0.04
LNGDP	0.92	0.01	0.94	0.01	0.91	0.01	0.13	0.01	0.35	0.02
LANGUAGE	0.54	0.04	0.56	0.04	0.74	0.04	0.53	0.11	0.11	0.07
BORDER	0.61	0.08	0.70	0.08	0.32	0.08	1.55	0.46	1.49	0.19
REGTA	1.02	0.07	0.63	0.07	1.16	0.07	2.09	0.22	#N/A	#N/A
COMCOL	0.37	0.06	0.33	0.06	0.44	0.06	-0.21	0.17	-0.10	0.09
COLONY	2.10	0.07	2.07	0.08	1.63	0.08	3.18	0.25	3.15	0.11
SAMECO	0.95	0.32	1.22	0.40	1.19	0.35	0.22	0.42	2.62	0.23
LNGDPP	0.48	0.01	0.47	0.01	0.63	0.02	1.04	0.04	0.90	0.03
LOCKED	-0.25	0.03	-0.28	0.03	-0.37	0.04	-0.98	0.16	-0.30	0.05
AREAPROD	-0.14	0.01	-0.18	0.01	-0.14	0.01	0.23	0.02	0.11	0.01
D1970	1.66	0.04	1.66	0.04	1.71	0.05	1.85	0.05	1.75	0.08
D1975	1.40	0.04	1.40	0.04	1.40	0.05	1.52	0.13	1.68	0.08
D1980	1.02	0.04	1.00	0.04	1.01	0.04	1.27	0.12	1.35	0.07
D1985	0.25	0.04	0.23	0.04	0.54	0.04	0.54	0.11	0.78	0.07

1. Not-Caribbean and Not-Africa: regression on data where *neither* member of a country pair is in the Caribbean or Africa (as appropriate).
2. Caribbean and Africa: regression on data where *one* (or more) of the *two* countries is from the Caribbean or Africa (as appropriate).

Table 4: Correlation of trade and regressors

	LNTRADE	CUNION	LNDIST	LNGDP	LANGUAGE	BORDER	REGTA	COMCOL	COLONY	SAMECO	LNGDPP	LOCKED	AREAPROD
LNTRADE	1	-0.02	-0.19	0.65	0.01	0.15	0.10	-0.14	0.12	0.00	0.45	-0.07	0.24
CUNION	-0.02	1	-0.21	-0.2	0.19	0.05	0.19	0.22	0.01	0.28	0.07	0	-0.12
LNDIST	-0.19	-0.21	1	0.18	-0.22	-0.38	-0.15	0	-0.15	-0.05	0.04	-0.07	0.16
LNGDP	0.65	-0.2	0.18	1	-0.21	0.03	-0.1	-0.32	0.04	-0.1	0.36	-0.03	0.63
LANGUAGE	0.01	0.19	-0.22	-0.21	1	0.17	0.15	0.37	0.17	0.1	-0.04	-0.04	-0.11
BORDER	0.15	0.05	-0.38	0.03	0.17	1	0.11	0.05	0.16	-0.01	-0.05	0.05	0.08
REGTA	0.10	0.19	-0.15	-0.1	0.15	0.11	1	0.16	0.05	0.1	-0.06	0.09	-0.16
COMCOL	-0.14	0.22	0	-0.32	0.37	0.05	0.16	1	-0.04	0.17	-0.03	-0.22	-0.23
COLONY	0.12	0.01	-0.15	0.04	0.17	0.16	0.05	-0.04	1	-0.04	-0.06	0.05	-0.01
SAMECO	0.00	0.28	-0.05	-0.1	0.1	-0.01	0.1	0.17	-0.04	1	-0.02	0.05	-0.16
LNGDPP	0.45	0.07	0.04	0.36	-0.04	-0.05	-0.06	-0.03	-0.06	-0.02	1	-0.08	0.03
LOCKED	-0.07	0	-0.07	-0.03	-0.04	0.05	0.09	-0.22	0.05	0.05	-0.08	1	0.03
AREAPROD	0.24	-0.12	0.16	0.63	-0.11	0.08	-0.16	-0.23	-0.01	-0.16	0.03	0.03	1

Shaded areas have correlations ≥ 0.3.

Table 5: Regressions on individual years

Year	1970		1975		1980		1985		1990		ALL YEARS	
	Parameters	Std Error	Parameters	Std Error	Parameters	Std Error	Parameters	Std Error	Parameters	Std Error	Parameters	Std Error
OLS												
CONSTANT	-10.04	0.60	-9.92	0.54	-10.67	0.46	-10.77	0.47	-11.67	0.47	-9.91	0.23
CUNION	1.63	0.43	2.23	0.40	1.56	0.46	2.34	0.29	2.34	0.28	1.81	0.14
LNDIST	-1.17	0.04	-1.25	0.04	-1.21	0.03	-1.22	0.03	-1.31	0.03	-1.24	0.02
LNGDP	0.86	0.01	0.87	0.01	0.88	0.01	0.86	0.01	0.91	0.01	0.86	0.01
IV												
CONSTANT	1.74	0.53	-1.15	0.46	-0.37	0.43	-0.48	0.41	-0.48	0.45	-0.65	0.20
CUNION	0.51	0.32	0.92	0.27	-0.09	0.20	0.69	0.22	0.40	0.21	0.38	0.10
LNDIST	-1.04	0.03	-1.11	0.03	-1.02	0.03	-1.07	0.03	-1.16	0.03	-1.09	0.01
LNGDP	0.48	0.02	0.58	0.01	0.53	0.01	0.55	0.01	0.55	0.01	0.55	0.01

Data Definitions

See Rose (2000) and Rose and van Wincoop (2001) for more extensive descriptions of the data. All logs below are natural logs. The sample consists of data from 1970, 1975, 1980, 1985, and 1990. (The 1995 data were not used due to the rebasing of the 1995 income data.)

- LNTRADE: The log trade term is the natural log of the sum of exports and imports between two countries.
- CU: The currency union dummy variable assumes a value of 1 when two countries use the same currency, and is 0 otherwise.
- LNDIST: Distance is measured in logged kilometres.
- LNGDP: The GDP term is the log of the product of the GDPs of the two countries.
- LNGDPP: The per capita GDP term is the same as the GDP term except it accounts for the respective population sizes.
- SAMECO: The same-country dummy is 1 if the two regions are part of the same country.
- COMCOL: The common colonizer dummy is 1 if the two countries were colonies after 1945 and had a common colonizer.
- LANGUAGE: The language dummy is 1 if countries have a language in common. This variable has been modified from the original Rose data set. A language is held in common if it is commonly understood or is an official language (for example Nauru and Kiribati).
- COLONY: The colony dummy is 1 if a country pair represents a colony and a colonizer.
- LOCKED: The landlocked dummy is 1 if one of the two countries is landlocked, 2 if both are, and 0 otherwise.
- BORDER: The border dummy is 1 if the two countries share a common land border.
- AREAPROD: This variable is the log product of the areas of the two countries.