The euro's trade effects

Richard Baldwin[1]

Graduate Institution of International Studies, Geneva

First draft 8 May 2005; Second draft 29 May 2005; Prepared for ECB Workshop "What effects is EMU having on the euro area and its member countries?" Frankfurt 16 June 2005.

1. INTRODUCTION

The euro must be the world's largest economic policy experiment. Six years ago, European nations accounting for 20% of world output, 30% of world trade and 300 million people found themselves using the same currency. Given the importance that monetary regimes have on the course of human events, this should have had effects all across the board – affecting everything from union's wage bargaining to educational exchanges and corporate investment strategies. Every problem looks like a nail when you have a hammer in your hand and I am a trade economist, so I am naturally drawn to the euro's trade effects.

In this paper, I review the existing empirical literature on the trade effects of currency unions, for non-European and European cases. This is done in sections 2 and 3. My bottom line summary of this literature is that the euro probably did boost intra-Eurozone trade by something like five to ten percent on average, although the estimated size of this effect is likely change as new years of data emerge. Then I collect together the clues in section 4 and use them in section 5 speculate on the sorts of economic mechanisms that might be driving the euro's affect trade. I come up with a set of competing hypotheses in section 6 and, in Section 7, propose a battery of diagnostic tests that could help reject some or all of the theoretical explanations. The final section contains my concluding remarks.

Right up front I should apologise to the reader for the length of the paper. The ECB asked me to write a short paper on the subject but I didn't have time for that so I wrote a long paper instead.

[1] I would like to thank Nadia Rocha for assistance with data-wrestling. Andy Rose, Volker Nitsch, Howard Wall, Alejandro Micco and Hakan Nordstrom provided excellent comments and answered my many questions about their data and regressions. They also saw early drafts of this paper and eliminated of the mistakes that must still be in this version.

2. THE ROSE VINE: A LITERATURE REVIEW

Andy Rose of the University of California-Berkeley opened a new chapter in international economics research with an Economic Policy article that was published in 2000. Rose (2000) asked a simple question and got a simple answer:

"What is the effect is a common currency on international trade? Answer: Large."

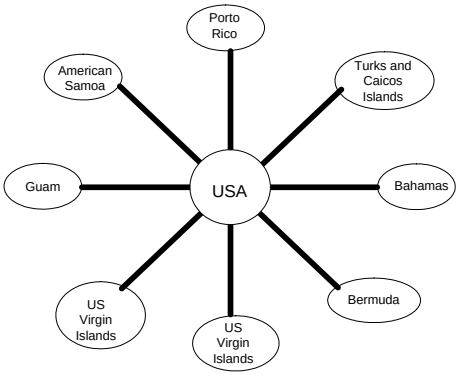
Rose's paper generated a huge response. It has been cited over 550 times according to Google Scholar as of May 2005. (Barro's 'Growth in a Cross-Section of Countries' paper, for comparison, has been cited two thousand times.) Much of the response has been critical with many authors trying to reduce the size of the impact. This section reviews the evidence on the trade effects of currency union – what I'll call the Rose effect for brevity's sake – on the pre-euro data. The next section reviews the evidence on the euro itself.

As an aside, I want to thank Andy for introducing a tradition of jocular writing into this literature. The final version that Andy turned in to Charles Wyplosz (the Managing Editor who did the final editing,; David Begg handled the manuscript in its early stages) was chock-block full of exuberant usages of the English language. Charles tempered some of the most avant-garde constructions, but the published version of Rose (2000) is still a lot of fun to read. Volker Nitsche seconded this with his papers entitled: "Honey I shrunk the currency union effect" and "Have a Break"; I shall struggle to uphold the tradition in this paper.

Scientific rectitude

Andy Rose is also responsible for another remarkable feature of this literature – transparency and scientific rectitude. All of Rose's data sets and regressions are posted his web site. This has permitted scholars from around the world to check his data and results, tinker with specifications and challenge his findings. Subsequent contributors to this literature have generally followed this stellar example.

Figure 1: Schematic depiction of hub and spoke common currency arrangement



2.2. Roots: the world through Rose coloured glasses

Rose (2000) started the debate with his finding that countries in a currency union traded 3 times more with each other than one would expect. He arrived at that astonishing result using a gravity-equation approach on data for bilateral trade among 186 nations. His cross-section regression was:

$$\ln(RV_{od})=a_0+\beta_1\ln(RY_oRY_d)+\beta_2\ln(Distance_{od})+\beta_3(CU_{od})+controls$$

where RV is the real value of bilateral trade, the RY's are real GDPs of the origin nation ('o' is a mnemonic for origin) and destination nation ('d' is a mnemonic for origin), and CU is a dummy that switches on when nations o and d share a common currency. In his favourite regression, $\beta_3=1.21$ which implies trade between common-currency pairs was $e^{1.21}=3.35$ times larger than the baseline model would suggest, i.e. a common currency boost trade by over 200%.

The size of this common-currency effect was just far too large to be believed and the profession's assault on this claim began even before he presented it at the October 1999 Economic Policy Panel hosted by the Bank of Finland. There were three main themes in these critiques:

- Omitted variables: Biases stemming from omitted variables that are pro-trade and correlated with the CU dummy;
- Reverse causality: big bilateral trade flows causes a common currency rather than vice versa; and
- Model misspecification.

Almost all turned on the fact that most of the common currency pairs involved nations that were very small and very poor. One very nice, early critique can be found in Nitsch (2002).

Table 1: The Rose Garden, currency unions considered in Rose (2000)

Hub and Spoke arrangements		**Multilateral currency unions**	**Misc.**
√ **Australia**	√ **USA**	**CFA**	√ **India**
Christmas Island	American Samoa	√ Benin	√ Bhutan
Cocos (Keeling) Islands	Guam	√ Burkina Faso	√ **Denmark**
Norfolk Island	√ US Virgin Islands	√ Cameroon	Faroe Islands
√ Kiribati	Puerto Rico	√ Central African Republic	√ Greenland
√ Nauru	Northern Mariana Islands	√ Chad	**Turkey**
√ Tuvalu	√ British Virgin Islands	Comoros	N. Cyprus
Tonga (pre '75)	√ Turks & Caicos	√ Congo	**Singapore**
√ **France**	√ Bahamas	√ Cote d'Ivoire	Brunei
√ French Guyana (OD)	Bermuda	Equatorial Guinea (post '84)	**Norway**
√ French Polynesia	√ Liberia	√ Gabon	Svalbard
√ Guadeloupe (OD)	Marshall Islands	Guinea-Bissau	**South Africa**
Martinique (OD)	Micronesia	√ Mali (post '84)	Lesotho
Mayotte	Palau	√ Niger	Namibia
√ New Caledonia (OT)	√ Panama	√ Senegal	Swaziland
√ Reunion (OD)	√ Barbados (? 2:1)	√ Togo	**Switzerland**
Andorra	√ Belize (? 2:1)	**ECCA**	Liechtenstein
√ St.Pierre & Miquelon	√ **Britain**	√ Anguilla	**Spain**
Wallis & Futuna Islands	√ Falkland Islands	√ Antigua and Barbuda	Andorra
Monaco	√ Gibraltar	√ Dominica	**Singapore**
√ **New Zealand**	Guernsey	√ Grenada	Brunei
√ Cook Islands	Jersey	√ Montserrat	**Italy**
√ Niue	Isle of Man	√ St. Kitts and Nevis	San Marino
Pitcairn Islands	√ Saint Helena	√ St. Lucia	Vatican
Tokelau	Scotland	√ St.Vincent	**Morocco**
	√ Ireland (pre '79)		Western Sahara

Source: Rose (2000) appendix table and footnotes.

In his revisions, Rose produced a battery of robustness checks that he claimed had repulsed each of these critiques, leaving his central result essentially unaltered. As the Editors' Introduction to the issue in which Rose (2000) appears says: "The Panel admired the paper and the author's thoroughness but retained an uneasy feeling that something had eluded them."

Much of the subsequent literature on the Rose effect can be thought of as a search for that elusive something. Before reviewing the 'rose vine' that has grown from Rose's roots, it is critical to have an idea of the currency unions that this literature investigated. For the most, they involve bizarre partnership (to put it mildly).

2.2.1. **Pre-euro currency unions**

Rose (2000) lists all the currency unions (CU) and CU-like monetary arrangements from 1970 onwards. This is reproduced in Table 1. There are three types of CUs in his table. The first two columns show the hub-and-spoke CU arrangements. As Figure 1 shows with a schematic diagram for the US; hub and spoke CU's involve small nations (the spokes) adopting the currency of their dominate trade partner (the hub). The hubs are the USA, France, Britain, Australia and New Zealand.

There are two types of bilateral trade flows in hub and spoke arrangements, flows between the hub and a spoke and flows between the spokes. For the most part the hub-spoke flows involve the exchange of extremely different goods (so-called Heckscher-Ohlin trade). For example, the US sells machinery to Barbados while Barbados sells rum to the US. The spoke-spoke flows are typically very small, as is true of trade among most poor nations.

The third column of the table lists the second type, multilateral currency unions. The two major multilateral currency unions that existed before the euro are the West African CFA arrangement and the Caribbean arrangement, the ECCA. These CUs are among nations that are tiny economically by world standards. The fourth column lists a series of highly idiosyncratic CU pairs often involving a very local hegemony, like Switzerland and Liechtenstein, or Italy and San Marco.

Rose (2000) does not have data for all these. I have checked in the table the ones he includes in his study.

Another way to look at the oddness of the non-European currency union pairs is to plot their openness ratios. The openness ratio is just the sum of trade divided by real GDP, were the trade is the bilateral trade data from Rose (2000) summed across all of each nation's trade partners. The results are displayed in **Error! Not a valid bookmark self-reference.**. The top panel shows all 141 nations with data. The bottom panel includes only nations that have openness ratios of less that 2.0 (i.e. trade which is less than 200% of GDP).

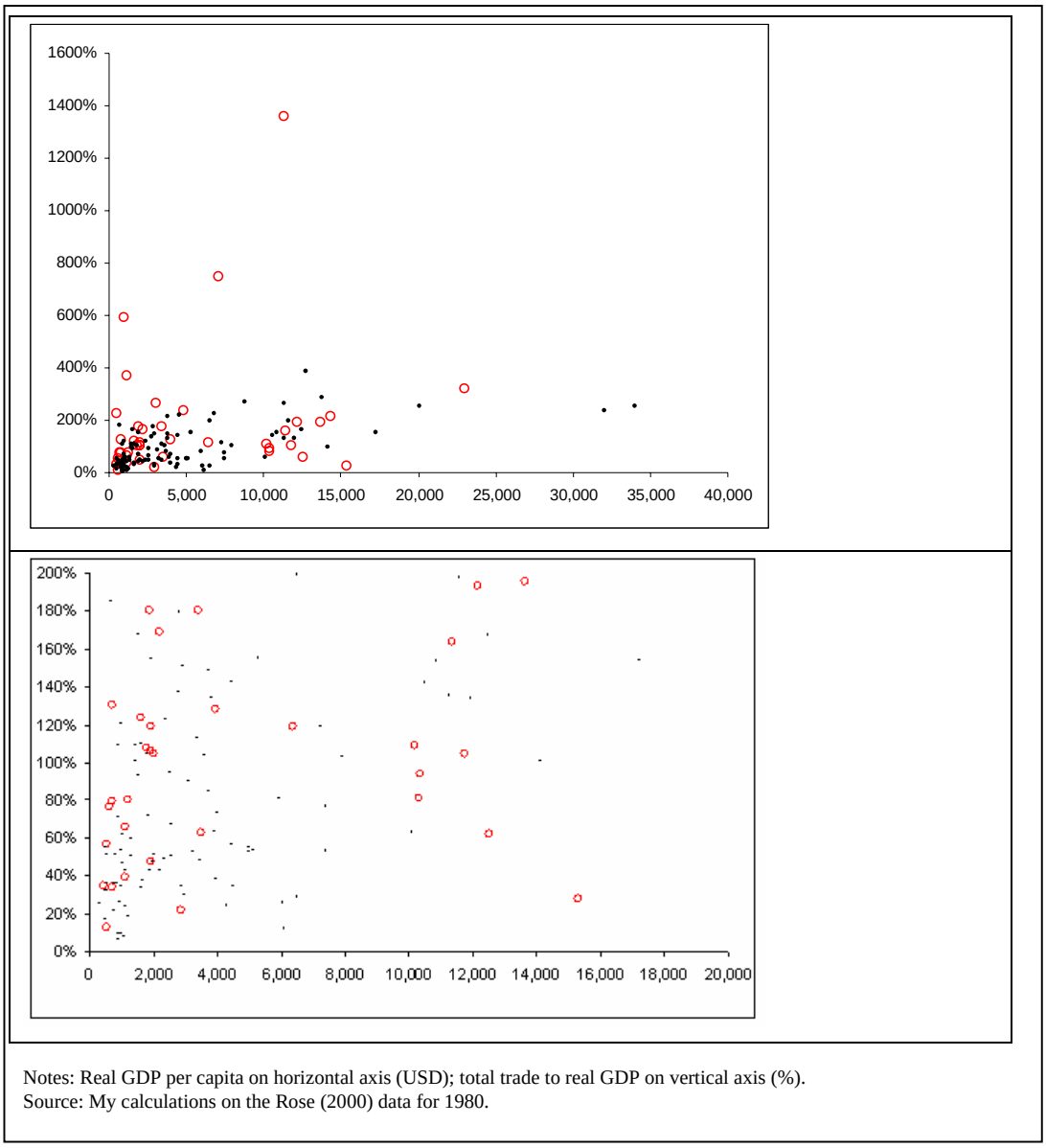
The top panel shows that there are some extremely open nations that also share a currency with some other nation.[2] These nations' openness is so unusual that it is hard to see what is going on with the rest. There are 6 nations with openness above 200%, Bahamas (1400%), Singapore (750%), Liberia (600%), Bahrain (400%), Kiribati (370%) and Belgium-Luxembourg (320%). All but one are involved in a currency union. Eyeballing this list we see that many of these nations are know as centre for transit trade.

The bottom panel excludes these nations so as to better see the others. Here we clearly see the spokes (the circles for poor nations to the left) and the hubs (rich nations to the right). The nine rich nations participating in CU's are (by decline order of GDP per capita), US, Bermuda, Australia, Norway, France, Denmark, New Zealand, Italy and UK. Note that Rose (2000) does not use data for all of these. For example, Bermuda, Denmark, Italy and Norway have no trade data with their CU partners so they are not included.

Given simple combinatorics, there are many, many more spoke-spoke pairs than hub-spoke pairs (e.g. Rose 2000 lists 16 nations using the US dollar, which implies $16^2/2=128$ spoke-spoke bilateral flows and 16 hub-spoke flows). Thus most the CU pairs in the data will be between the nations with circles in the left part of the bottom panel.

[2] I believe the Rose trade data has a systematic bias in it, what I call the silver-medal mistake below.

Figure 2: CU nations tend to be very small, poor and open.



Notes: Real GDP per capita on horizontal axis (USD); total trade to real GDP on vertical axis (%).
Source: My calculations on the Rose (2000) data for 1980.

The main point of these graphics is that nations involved in currency unions are a long way from average nations.

2.3. Garden pests: biases in gravity model estimations

"Without theory, practice is but routine born of habit." ~ Louis Pasteur

Rose (2000) employed a naïve version of the gravity model for his preferred specification, a version that was had been widely used by policy analysts in the 1980s and 1990s (including by me in my 1994 book on Eastern EU enlargement). The naïve version omits two variables that severely bias results.

To see this point, however, we need to work through a bit of theory. Although the theory does involve a small amount of 'heavy lifting.' the work is handsomely rewarded. It helps us understand all the mistakes in Rose (2000) and the subsequent literature, and why only a handful of the hundreds of

estimates of the Rose effect are worth paying attention to for policy purposes. In any case, who ever said studying the Rose effect would be a bed of rose?

2.3.1. **Gravity for dummies and dummies for gravity (equation)** [3]

Which falls faster, a cannonball or a feather? Every high school student has been taught the Galilean lesson that they fall at the same rate in a vacuum. To practical, empirically oriented people however this has got to be the silliest bit of knowledge tucked into humankind's collective wisdom.

There is no perfect vacuum on earth, and where there is a perfect vacuum (deep space) there is no gravity. Nevertheless, the thought experiment is useful since it allows students to get their minds around the 'headline news' when it comes to falling bodies – gravity – without confusing it with secondary factors, like air resistance. Eventually, the students find out what happens with one lets some air into the vacuum chamber, but only after they are really convinced that gravity affects all objects equally, ceteris paribus.[4] This section works through the gravity model 'without air resistance' so as to illustrate the common mistakes involved in gravity model estimation.

The value version of the gravity equation: a demand equation with social pretensions

The gravity model has more theoretical foundations than any other trade model I can think of, but these foundations are usually ignored (Jim Anderson published the first theoretical foundation in 1979 and despite it appearing the AER, economists continue invent new ones, claiming that the model has none).[5] This failure to consider its theoretical foundation has produced a string of errors – errors that have been repeated so often that they have become accepted practice.

Most researchers estimate what I call the 'value version' of the model – i.e. the dependent variable is the *value* of bilateral trade deflated by a common price index, so that's where we'll start.

Step 1: The expenditure share identity

The first step is the expenditure share identity:

$$p_{od}x_{od} \equiv share_{od}E_d; \tag{1}$$

where x_{od} is the quantity of bilateral exports of a single variety from nation 'o' to nation 'd' (the 'o' is a mnemonic for 'origin' and 'd' for 'destination'), p_{od} is the price of the good inside the importing nation also called the 'landed price.'[6] Of course, this makes $x_{od}p_{od}$ the value of the trade flow measured in terms of the numeraire. E_d is the destination nation's nominal expenditure (again measured in terms of the numeraire). By definition, $share_{od}$ is the share of one good nation-o good share of expenditure in nation d.

Step 2: The expenditure function: shares depend on relative prices

Basic microeconomics tells us that expenditure shares depend upon relative prices and income levels, but we postpone consideration of the income elasticity in this new-math presentation, so the expenditure share is assume to depend only on relative prices. Adopting the CES demand function and assuming that all goods are traded, the imported good's expenditure share is linked to its relative price by:

$$share_{od} \equiv \left(\frac{p_{od}}{P_d}\right)^{1-\sigma}, \quad P_d^{1-\sigma} = \Delta_d, \quad \Delta_d \equiv \sum_{k=1}^{m} n_k \left(p_{kd}\right)^{1-\sigma}, \sigma > 1 \tag{2}$$

---

[3] For more on this, see Baldwin and Taglioni (2005).

[4] Little remembered fact: Commander David Scott tried Galileo's experiment on the moon a hammer and a feather, 2 August 1971. It worked.

[5] Jim Anderson published the first theoretical foundation in 1979 and despite it appearing the AER, economists continue invent new ones, claiming that the model has none. For example, Deardoff refers to the gravity model in his 1984 Handbook of International Economics chapter as having "somewhat dubious theoretical heritage (p.503)," (despite Anderson having published his foundations 5 years before) and then ten years late saying, "It is certainly no longer true that the gravity equation is without theoretical foundations, since several of the same authors who noted its absence went on to provide one," in his 1997 paper "Determinants of bilateral trade."

[6] Roughly speaking this would be the cif price not the fob price (cif stands for cost, insurance and freight, while fob stands for free on board, i.e. on the boat in the exporting nation's port).

where p_{od}/P_d is the relative price, P_d is nation-d's CES price index for all goods that compete with the imported good, 'm' is the number of nations from which nation-d buys things (this includes itself), and σ is the elasticity of substitution among all varieties (all varieties from each nation are assumed to be symmetric for simplicity); n_k is the number of varieties exported from nation k. The symbol Δ is a mnemonic for 'denominator' of the CES demand function. Combining (2) and (1) yields product specific import expenditure equation that could be estimated directly if we had the data. Lacking data on the landed prices of individual goods, we compensate by putting more structure on the problem.

Step 3: Adding the simplest pass-through equation

Assuming full pass-through (this is consistent with both Dixit-Stiglitz monopolistic competition and perfect competition) all manner of trade costs – tariffs, transport costs, foreign exchange costs, etc. – are passed on to the consumer, so the landed price in nation-d is linked to the producer price in nation-o and the bilateral trade costs via:

(3) $$p_{od} = p_o \tau_{od}$$

where p_o is the producer price of nation-o exports, τ_{od} reflects all trade costs, and we have used the symmetry of varieties to drop the variety index.

Step 4: Aggregating across individual goods

So far we've focused on per-variety exports. To get aggregate bilateral exports from 'o' to 'd', we multiply the expenditure share function by the number of varieties nation 'o' has to offer, denoted as n_o. Using upper case V to indicate to total value of trade, $V_{od} \equiv n_o s_{od} E_d$, i.e.:

(4) $$V_{od} = n_o p_o^{1-\sigma} \frac{\tau_{od}^{1-\sigma}}{\Delta_d} E_d$$

We could estimate this, and maybe one day when governments spend more on gathering trade data, we shall. For now, however, we continue substituting assumptions for information.

Step 5: Using general equilibrium in the exporting nation to eliminate the nominal price

Lacking data on the number of varieties n_o and producer prices p_o, we compensate by turning to nation-o's general equilibrium condition. This is the only part that is even slightly tricky, so I'll illustrate with a diagram. The producer price in the exporting nation, nation-o, must be such that it can sell all its output, either at home or abroad. Taking its nation-o's output as given, we can use the CES expenditure share function for all of nation-o's destination markets to work out what nation-o's producer price must be.

Nation-o's expenditure must equal the total value of its output (ignoring current account imbalances), so E_o is the amount that nation-o must sell. To make this happen, nation-o's producer prices must adjust to ensure that:

(5) $$E_o = n_o \sum\nolimits_k (s_{ok} E_k)$$

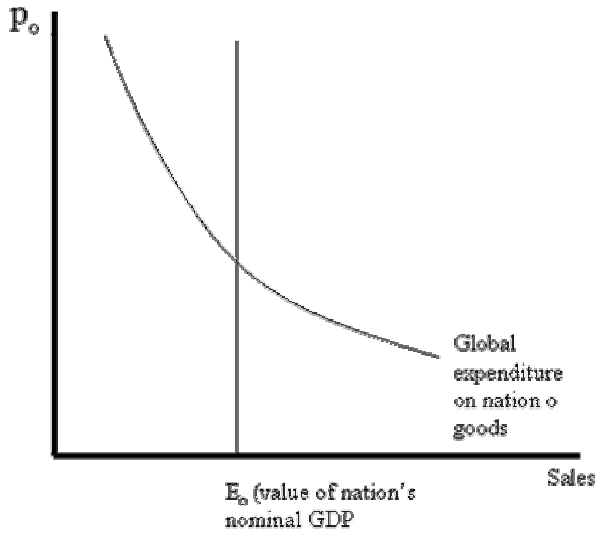
where the summation is over all markets (including o's) and the producer prices are hidden in the 's_{ok}' terms.

Figure 3 show this general equilibrium condition for the exporting nation diagrammatically; the right-hand side is the downward sloped line; it is downward sloped since expenditure on nation-o's exports in every market fall with a ceteris paribus rise in nation-o producer prices. Using the CES expenditure share function in (2), and solving for $n_o p_o^{1-\sigma}$:

(6) $$n_o p_o^{1-\sigma} = \frac{E_o}{\Omega_o}; \qquad \Omega_o = \sum\nolimits_k \tau_{ok}^{1-\sigma} \frac{E_k}{\Delta_k}$$

Here Ω_o can be thought of as the de facto openness of the world to nation-o's exports, or, in other words as nation-o's market access.

Figure 3: Determination of the exporter's producer price.



Step 6: Twelve times twenty-eight equals 300 – a 'new maths' gravity equation[7]

Substituting (6) into (4), we get our first-pass, or new-maths version of the gravity equation:

$$V_{od,t} = \left(\frac{\tau_{od,t}^{1-\sigma}}{\Omega_{o,t}\Delta_{d,t}}\right)E_{o,t}E_{d,t}; \quad (7)$$

This is a microfounded gravity equation. I have added time subscripts to stress the point that all of the variables are time varying.

Multilateral resistance, remoteness, bilateral openness, relative-prices-matter

The novel term here (at least they were novel in 1979 when Jim Anderson first derived a similar, but more general, expression) is the $1/\Delta\Omega$ term. Of course (7) cannot be directly estimated since we do not have data on Δ or Ω, nor do we have data on the total trade cost between nations, but this is what we should estimate.[8]

As mentioned, Ω can be thought of as a measure of the exporting nation's market access. The Δ term is the denominator of the demand function, but in the international context it can be thought of as a measure of the importing nation's openness – greater openness lowers the landed prices of imports in nation d and thus raises Δ_d since $\sigma>1$.

[7] In the 'old maths' one worked out, e.g. 12 times 28 by direct calculation, while in the 'new math' one says it's about 10 times 30=300, and then sets about refining the estimate if need be.

[8] Anderson and van Wincoop (2001), a classic article, make a big deal of the fact that the Δ and Ω are not price indices. This is true, but just barely. The CES price index that is relative to the import demand equation must be $\Delta^{1-\sigma}$. For example, if home-bias (what Anderson-Wincoop call 'non-pecuniary' effects) is important, then a properly constructed CES price index – for example, one that uses expenditure shares as weights – would pick this up. What is true is that one needs to assume the range of goods with which imports compete in order to know which price index to use. My guess is that an index of import prices would do a very good job of standing in for the ideal Δ. It would be important, however, to enter it separately rather than, e.g., divide the importer's nominal GDP by it and then putting it in. The reason is that $P=\Delta^{1-\sigma}$ should get a different coefficient in a log-log regression than does nominal GDP, E. The issue is more complex for Ω, but this is proportional to the E-weighted sum of nation-o's market shares in each market (including its own). Nitsche (2000) and De Souza and Lamotte (2004?) implement this.

Some authors call the new term the 'remoteness' term; Anderson and van Wincoop (2001) call it the 'multilateral trade resistance' term.[9] But from a purely description perspective, it should be called the relative-prices-matter term – that phrase doesn't exactly roll off the tongue but it says what it means. After all, the Δ is there to reflect the relative price of nation-o's exports to nation-d and E_o/Ω_o is there since it affects nation-o's the producer prices. Yet another way to view the term in parentheses is to think of it as bilateral relative openness. The numerator reflects direct openness of the bilateral relationship (trade costs are raised to a negative power so the term gets larger as bilateral trade gets cheaper, i.e. more open). The denominator reflects both the openness of the importing nation to all goods (recall Δ rises when the landed price of imports from anywhere falls) and the openness of the world to the exporter's goods (recall Ω rises when the exporter's share of any market rises).

Why the relative-prices-matter term affects statistical inferences, 2 examples

A classic example of the importance of relative-prices-matter terms is the trade between New Zealand and Australia. Although these nations are far apart (it takes a jet airliner 3 hours to get from Sydney to Wellington) both nations are very far from other industrial nations (the Sydney flight to Tokyo, Los Angles and London take 9, 13 and 23 hours respectively). Thus the naïve gravity model's prediction is terrible. It under-predicts the bilateral trade. The reason is that it ignores the fact that the bilateral Australian-New Zealand trade costs are not high relative to the trade costs facing exporters from the rest of the world. As an aside, this is one of the reasons that the ANZ preferential trade agreement looks like it has such a big effect in naïve gravity models; the ANZ dummy picks up the big error term.

In Europe, where distances between major industrial nations are extremely small by world standards, the point works in reverse. Frankfurt is only an hour flying time from Amsterdam, but the Dutch-German bilateral trade costs are not particularly low compared to the trade costs faced by other exporters to the Netherlands. This is probably why new naïve gravity model estimates tend to show that the EU has no impact on trade. In the old days, when it was hard to manipulate large data sets, the naïve equation was estimated on bilateral flows for European trade only so the distance coefficient was typically estimated as -0.7 and this left some room for the EU to matter. In broader data sets, the distance coefficient is estimated to be something like-1.0, so the predicted intra-EU flows are expected to be very high and this leaves nothing for the EU dummy.

Why it fits so well

Having slogged it through till here, I hope you'll agree that the gravity model can be thought of as demand curve with 'social ambitions.' This is, of course, why it fits so well. Demand functions generally do well in the data, and nations' output and expenditure generally match. This perspective also makes it clear that it is not a model in the traditional sense of the word. It is a regression of endogenous variables on endogenous variables.

2.3.2. The biases in what Rose estimates

All well trained economists know they must use real variables, so when most hear about the gravity model, they presume that the model must be about real GDPs and real trade. This is what occurred to Rose and most of the international macroeconomists who used the model before him, although there are some exceptions (de Sousa and Lochard 2003). This response, however, is incorrect. The V and E's in (7) are in nominal terms. Just to be clear, there is no money illusion in (7). The E's and V must be, and are, measured in terms of a common numeraire, but US dollars would do just fine. For example, if the prices were suddenly switched from dollars to cents, nothing would change in (7) – it is homogeneous of degree zero in nominal prices. Of course, the econometrician would have to worry about regressing trends on trends but trends can be eliminated in other ways.

Be that as it may, Rose (2000)'s preferred regression estimates (simplifying for clarity's sake):

---

[9] Anderson and van Wincoop 2001 assume each nation produces the same number of varieties (i.e. only one variety), and then one can show that $\Omega_i=\Delta_i$ for each nation. Take a two country case and write down what all the Ω's look like; then write down the Δ's. You'll see that the two are isomorphic so $\Delta=\Omega$ is a solution; see Anderson and Wincoop for why it is the only solution.

(8) $$\frac{V_{od}}{P_{USA}} = \tau_{od}^{1-\sigma}\left(\frac{E_o}{P_o}\right)\frac{E_d}{P_d}; \qquad \tau_{od} = f(dist_{od}, other\ stuff)$$

where $dist_{od}$ is the distance between o and d. In words, he deflates the bilateral trade value with the United States' GDP deflator, and uses real expenditure rather than nominal expenditure deflated by Δ or Ω. Rose follows a long tradition of modelling τ as depending upon natural barriers (bilateral distance, adjacency, land border, etc.), various measures of manmade trade costs (free trade agreements, etc.), and cultural barriers (common language, religion, etc.). His original contribution was to add a common currency dummy to the list – hard to imagine that no one had thought of it before 2000, but that's always the case with 'low hanging fruit' research – you don't have to be tall, you just have to be the first one in the orchard.

Rose (2000) estimates this on various cross-sections of his data as well as the full panel. This procedure is a mistake, a mistake that would get the gold medal in the race for the most frequent mistake in gravity equation estimations.[10]

Now if one estimates (8) when (7) is correct, what one is really doing is estimating:

(9) $$\frac{V_{od}}{P_{USA}} = \tau_{od}^{1-\sigma}\left(\frac{E_o}{P_o}\right)\left(\frac{E_d}{P_d}\right)\left\{\frac{P_d}{\Delta_d}\frac{P_o}{\Omega_o}\frac{1}{P_{USA}}\right\}; \qquad \tau_{od} = f(dist_{od}, other\ stuff)$$

but omitting the terms in brackets.

What is wrong with this? One big problem – the gold medal of classic gravity model mistakes – and one small problem – the bronze medal winner in the mistake race. The big problem is that the omitted terms are correlated the trade-cost term, since τ_{od} enters Ω and Δ directly (see (2) and (6)). This correlation biases the estimate of trade costs and all its determinants including, the currency union dummy.

The small problem – what might be called the bronze-medal mistake – is that the inappropriate deflation of nominal trade values by the US' aggregate price index. Since there are global trends in inflation rates, inclusion of this term probably creates a spurious correlation. Fortunately, Rose (2000) and other most researchers offset this error by including time dummies. Since every bilateral trade flow is divided by the same price index, a time dummy corrects the mistaken deflation procedure. Note that when Glick and Rose (2001) run their regression without the time dummies, their estimated coefficient on the CU dummy is one standard deviation larger than it is with time dummies, so it can be important to correct the small problem.

2.3.3. **The rose has thorns only for those that would gather it: the gold-medal of gravity mistakes**

"He who loves practice without theory is like the sailor who boards ship without a rudder and compass and never knows where he may cast." ~ Leonardo da Vinci

Time for more heavy lifting. Coefficients estimated by regressing (8) on pooled data contain all the biases that the subsequent literature has sought to correct. It is impossible to fully understand the sequence of biases in the subsequent literature, including studies on the euro, without thinking a bit more precisely about the source of the biases.

Gravity model with common kluges[11]

What Rose (2000) estimates is a bit more complex than (9). Before getting to estimation, he uses two fixes that are common in the literature. First, he does not work with bilateral trade, exports from France to Germany for instance, but rather the average of bilateral trade, namely the average of French exports to Germany and German exports to France. Second, he does not estimate separate parameters for the real GDP figures, he constructs the product of the GDPs and uses this in his regression. Using our new-math version of the gravity equation (assuming trade costs are bilaterally symmetric), he is estimating:

---

[10] See Anderson and van Wincoop (2001) for the original and a more detailed version of this critique.

[11] Google definition: A kludge (or kluge) is a 'solution' for accomplishing a task, originally a mechanical one and usually an engineering one, which consists of various otherwise unrelated parts and mechanisms, cobbled together in an untidy or downright messy manner.

(10) $$\sqrt{V_{od}V_{do}} = \tau_{od}^{\,1-\sigma}\,\frac{E_o E_d}{P_o P_d}\left\{\frac{P_o P_d}{\sqrt{\Omega_o \Delta_o \Delta_d \Omega_d}}\right\}; \qquad \tau_{od} = f(dist_{od}, other\ stuff)$$

Here and subsequently, I assume that the broze-medal mistake – deflation by US price index – has been offset by time dummies so we can ignore P_{USA}. Again simplifying for the sake of illustration, suppose the true model of bilateral trade costs is:

$$\tau_{od} = dist_{od}^{\,b_1} CU_{od,t}^{\,-b_2} Z_t^{\,b_3}; \qquad b_1, b_2 > 0$$

where CU is the currency union dummy and Z is the other (omitted) determinants of bilateral trade costs (suppose there is only one for simplicity's sake). Then the true gravity model (in logs) is:

$$y = X_1\beta_1 + X_2\beta_2 + \varepsilon$$

where y is the trade flow, X_1 includes the product of the real GDP's, bilateral distance and the CU dummy, and X_2 includes the relative-prices-matter terms, Δ and Ω, as well as that omitted determinants of trade costs Z. Rose, however, estimates:

$$y = X_1\beta_1 + u_t$$

where

$$u_t = \ln\left\{\frac{P_{o,t}}{\sqrt{\Omega_{o,t}\Delta_{o,t}}}\right\} + \ln\left\{\frac{P_{d,t}}{\sqrt{\Delta_{d,t}\Omega_{d,t}}}\right\} + \ln\{Z_{od,t}\} + \varepsilon_t$$

The biases from OLS on pooled data will be:

(11) $$\hat{\beta}_1 = \begin{pmatrix} 1 \\ -b_1(\sigma-1) \\ b_2(\sigma-1) \end{pmatrix} + \begin{bmatrix} (RY_{o,t}RY_{d,t})X_{21t} & (RY_{o,t}RY_{d,t})X_{21t} \\ dist \cdot X_{21t} & dist \cdot Z_t \\ CU_t \cdot X_{21t} & CU_t \cdot Z_t \end{bmatrix} \begin{pmatrix} 1 \\ b_3 \end{pmatrix}$$

where the RY's are shorthand for the E/P's, and

$$X_{21} \equiv \frac{P_o P_d}{\sqrt{\Omega_o \Delta_o \Delta_d \Omega_d}}$$

Here we assume the true error ε is uncorrelated with anything.

The biases

Many researchers call the first two rows 'nuisance' parameters. That is a strange name, but I guess that means they don't want you to ask too many questions. Adopting that attitude for a moment, I focus solely on the bottom row of the matrix, the one dealing with the 'variable of interest', i.e. the currency union dummy.

The currency union dummy coefficient is biased: Part 1

We can be absolutely sure that CU and X_{21} are correlated since X_{21} contains CU and all other determinants of bilateral trade costs. In short omission of the relative-price-matters term produces biased results. But what is the likely direction of the bias?

Stepping outside the model for a moment, suppose that not all goods are traded, so the GDP price deflators include non-traded goods prices. Since the Δ and Ω include only traded goods prices, X_{21} is proportional to the ratio of non-traded prices to traded prices in the trading nations. If non-traded goods are idiosyncratically high in these nations, they will be idiosyncratically open (consumers substitute away from nontraded goods). Next, suppose that nations that are idiosyncratically open are more likely

to engage in pro-trade policies, like a currency union or FTA. If both of these conjectures are true, there will be a positive correlation between CU and the relative-prices-matter term. In this case, the coefficient on the CU dummy is upward biased.

As a matter of fact, most estimates of the Rose effect that ignore the relative-prices-matter term estimate the pro-trade effects of a currency union to be amazingly high. For example, the pooled and cross-section estimates in Rose (2000) ignore that relative-prices-matter term and these suggest that a currency union boosts bilateral trade by 235%. When he roughly corrects for this in his difference-in-differences regression in the same paper, he finds that the coefficient is about 8 standard deviations lower, namely that the Rose effect boosts trade by 17% instead of 335%.

The currency union dummy coefficient is biased: Part 2

Expression (11) shows a second bias in the CU dummy estimator, the CU·Z term. Bilateral trade costs are determined by many factors, ranging from personal relationships among business leaders that were developed as school children on cultural exchange programmes to convenient flight schedules. David Hummels has a series of papers pushing forward the range of these factors that can be measured, but there will always be omitted variables in the regression. This is not a problem unless the omitted variables are correlated with X_1 variables. However, it is very likely that the omitted pro-bilateral trade variables are positively correlated with the 'variable of interest', i.e. the CU dummy in this case. The point is that the formation of currency unions is not random but rather driven by many factors, including many of the factors omitted from the gravity regression. Indeed, below we discuss a number of papers on the determinants of CU's.

If the omitted variables Z and CU are positively correlated, the estimated trade impact will be upward biased for this second independent reason.

Nuisances with nuisance variables

Many researchers have found that the 'nuisance' variables in the gravity model are indeed a huge nuisance. The new-math version of the theory tells us the GDP variables should have coefficients of unity, and while slightly more sophisticated theory explains why the elastiticites may deviate from unity, most people get suspicious when the point estimates on GDP fall below, say 0.5. It is plain from the reasoning above that we should expect the GDP elastiticites to be biased downward, since X_{21} contains the price index that is used to deflate nominal GDP. The correlation is not -1, however, since X_{21} includes the Δ and Ω terms as well. Since the ratio of traded to nontraded goods will vary across country samples and time periods, the biases on the GDP coefficient need not be systematic.

2.3.4. **More thorns: the silver-medal of gravity mistakes**

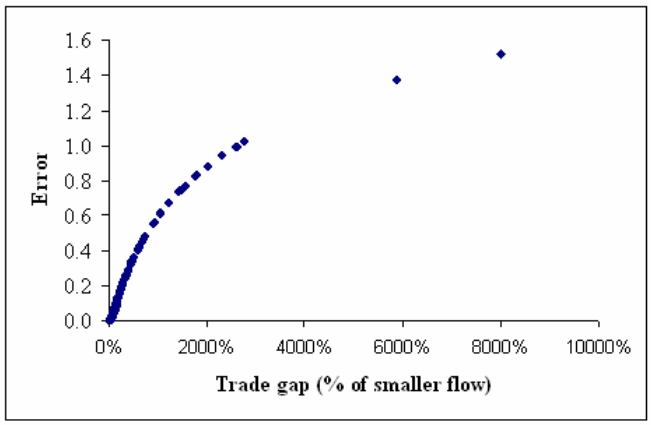
The basic theory tells us that the true relationship is for bilateral exports. Many gravity models – including most in the Rose effect literature – are estimated on 'bilateral trade' namely, the average of the two-way exports. There is nothing intrinsically wrong with this, but since it was done without reference to theory, most researchers commit a simple, but grave error. They mistake the log of the average for the average of the logs.

This can seriously bias the results. The sum of the logs – the right way – is approximately the log of the sums, but the approximation gets worse as the two flows summed become increasingly different.[12] In plain English, the error will not be too bad for nations that have bilaterally balanced trade, but it can be truly horrendous for nations with very unbalanced trade. In fact, unbalanced trade is a huge issue. The biggest exporters, German, Japan and the US, for example, sell something to most nations around the world. However, many small nations sell nothing in return, at least not to all of the big-3. Thus the problem is systematically worse for North-South trade than it is for North-North trade. Note that this mistake is quite systemic, it is even in one of the most common references for the gravity model, Chapter 5 in Feenstra (2004); see

[12] If $x=y\delta$, $\ln[(x+y)/2]=\ln x+\ln(1+\delta)-\ln 2$, while $\ln(xy\delta)/2=\ln(x)+\ln(\delta)/2$. The wrong way minus the right way is $\ln(1+\delta)-\ln\delta/2-\ln 2$.

Box 1: Log of the sums

In his chapter on the gravity model, Feenstra shows an equation with the log of the sums, rather than the sum of the logs. The theory leading up to this, however, is developed in the context of the simplest trade model – i.e. the Krugman trade model without trade costs. In this model, all bilateral flows are identical – because there are no trade costs, inter alia – so the sum of the logs does equal the log of the sums. However, when trade costs are introduced, the theory does not predict balance trade bilaterally, and indeed real-world trade flows are often very unbalanced; see Helpman, Melitz and Rubinstein (2004) for some figures on this.



Note: The error is the incorrect (log of average) minus the correct (average of logs), for Germany's 200 or so trade partners in the IMF DOT data for the year 2000; trade gap is the absolute value of difference in bilateral exports (to and from Germany) as percent of the smaller of the two flows. The two extreme outliers are nations with extremely unbalanced trade, West Bank/Gaza Strip and Niger. The incorrect averaging makes Germany-Niger bilateral trade about 150% higher than it actually is when averaged correctly.

Figure 4: Log averaging mistake for Germany, 2000, IMF DOTS data.

To see the sorts of bias this mistake can induce, look at what the mistake does to Germany's bilateral trade data (IMF DOTS data for the year 2000). For nations with which Germany has perfect bilateral trade balance the log of the sums is exactly equal to the sum of the logs. But when the two flows to be averaged are quite different, then the approximation becomes very wrong as Figure 4 shows. The extreme outlier in the figure is German-West Bank trade. The proper measure is 1.2 in logs, while the mistaken calculation yields 2.7 in logs. In short, the mistaken measure is always bigger and the mistake is extra big for unbalanced bilateral trade relations. I also calculated this for Germany's trade with EU15 and other OECD partners for which bilateral trade is more balanced, but still I find errors on the order of 15% even for these fairly similar nations.

By the way, the error always makes the bilateral trade look to bigger (Jensen's inequality).

The difference between theory and practice is different in theory than it is in practice

Of course this silver medal mistake only matters if the error is especially bad for currency union trade flows. To look at this quickly, I calculated the bilateral imbalance for all the hub and spoke CU pairs around the US dollar. I used IMF DOTS data for 2000, so not all of the islands in the Rose (2000) list are present. Table 2 shows that most of the spoke-spoke trade flows are zero and the non-zero entries all have imbalances on the order of 100%, so the trade flow will be severely upward biased. The hub-spoke flows are less likely to be zero, and the trade imbalances are less severe, but in most cases they are over 50% and so also severely overestimated due to the silver medal mistake.

Table 2: Bilateral Imbalance as % of 1-way flow, US dollar currency pairs

	Am. Samoa	Bahamas	Belize	Bermuda	Guam	Liberia	Palau	Panama	USA
Am. Samoa									
Bahamas									

Belize	-1420%				
Bermuda	-120%				
Guam					
Liberia	100%				
Palau					
Panama	100%	89%			
USA	76%	52%	91%	12%	78%

Source: My calculations on IMF DOTS for year 2000, export data.

Altogether only one of the ten non-zero pairs has less than a 50% imbalance. Plainly, someone needs to re-do the Rose effect estimates on data that is correctly averaged and see whether this really matters.

2.3.5. **Summing up on Rose (2000)**

Just to sum up, this reasoning explains why the pooled estimates in Rose (2000) – the most famous of which is the +200% estimate – should be ignored for policy purposes. They are based on an estimation technique that has subsequently been proved to be wrong by several authors, including Andy Rose himself. More on this below.

2.4. Rose Branch #1: Rose and van Wincoop (2001)

"I pass with relief from the tossing sea of Cause and Theory to the firm ground of Result and Fact." ~ Winston Churchill

Once the gold-medal mistake became clear (Jim Anderson's work was rediscovered once again), Andy Rose immediately teamed up with Eric van Wincoop to try to correct it.[13] Rose and van Wincoop (2001) was the result and it shows that forgetting about the relative-prices-matter term – as Rose (2000) did – leads to a severe upward bias in the Rose effect.

Rose and van Wincoop address the model misspecification issue in two ways. The simplest is to include origin-nation and destination-nation dummies in a cross section regression. They continue Rose's practice of using real GDPs and US price deflators for trade flows. Thus they estimate:

$$\textbf{(12)}\quad \ln\left(\frac{V_{od}}{P_{US}}\right) = \ln f(dist_{od}, CU_{od}, other\ stuff)^{1-\sigma} + \beta_1 \ln\left(\frac{E_o E_d}{P_o P_d}\right) + \beta_2 \left(\frac{P_d P_o}{\Delta_d \Omega_o}\right) + u$$

on panel data using the usual proxies for trade costs – most notably the common currency dummy, CU, that equals one if nation-o and nation-d use the same currency. Of course they do not have data for the terms involving Δ and Ω, but they use country specific dummies instead.

With these country dummies, the estimated Rose effect is radically lowered; it falls by 2.7 standard deviations. However, this diminished Rose effect is still mighty; without the country dummies a common currency is estimated to boost trade by 3.97 times; with them by 2.48 times.

2.4.1. **A mistaken correction of the mistake**

Putting in time-invariant country-specific fixed effects is wrong, as simple theory laid out above shows clearly. The omitted terms reflect factors that vary every year, so the country dummies need to be time varying. If the researcher forgets about this and includes time-invariant country dummies, as Rose and van Wincoop do, part of the bias may be eliminated. But since there will be a time varying residual in the error, the results will still be biased to the extent that the trade costs are also time varying. This problem may be relatively minor in the Rose-VanWincoop data since there is very little time variation in the CU dummy (more on this below).

Just to be perfectly clear, the Rose-van Wincoop country dummy procedure ameliorates what was called "the currency union dummy coefficient is biased: Part 1" above. The offending correlation was

[13] They don't deal with the silver medal mistake, no way has to my knowledge.

$CU \cdot X_{21}$. This has a cross-section element to it, and the time-invariant dummies eliminate this. However, it also has a time-series element to it and this is not eliminated.

This point probably explains why the second, harder way of correcting for the relative-prices-matter effect in Rose-Van Wincoop yields such a different result. Given all the structure imposed on the demand system – I mean the gravity equation – the econometrician can actually generate data for the Δ's (Anderson-Wincoop makes assumption that imply that a nation's openness to imports equals the world's openness to its exports, so the Ω goes away). When they do, the estimated Rose effect is again radically reduced. Doing some rough calculations on the numbers in the paper suggests the coefficient on the common currency dummy fell to 0.65, or about one more standard deviation; the Rose effect is then that a common currency boosts trade by 1.91 times.

2.4.2. **Lessons: Still a rosy scenario**

What are the lessons?

1) The estimates in the preferred regression in Rose (2000) are just plain wrong. They are overestimates. They are overestimated because the naïve gravity model is mis-specified and this mis-specification matters hugely in the dataset of Rose (2000).[14]

History divides neat into two parts: pre Rose-van Wincoop and post Rose-van Wincoop. Pre-RvW, we believed the 200% currency-union effect might have been correct. Post-RvW, we know better. More generally, one should never pay attention to estimates of the Rose effect that come from the naïve gravity model, i.e. one without fixed effects a la Anderson and van Wincoop or an equivalent fix (see below).

2) The Rose effect was still blooming after this correction; the best estimate is that it boosts trade by 1.9 times.

3) We can probably deduce something about the nature of bilateral trade flows that get the currency union dummy.

4) The omitted variable bias discussed above as 'The currency union dummy coefficient is biased: Part 2' is still in the Rose-van Wincoop numbers, so they are still too high.

2.5. Rose branch #2: Omitted variables

If there were such a thing as the 'Gravity model for fun and profit handbook,' page one would give this advice: "To amaze your friends with another important trade effect, develop a new proxy for trade costs and use a really big dataset; success is not guaranteed, but your likely to find significance (standard errors involve the inverse of the square root of number of observation) and you'll have loads of fun in any case." This is too cynical, but the basic point is that the gravity model omits an incredible range of factors that are likely to affect bilateral trade – I've seen people get statistically significant coefficients on time zones, language proximity, membership in the Austro-Hungarian Empire, presence of a China Town in the two capitals just to name a few. More to the point consider the trade among the nations listed Table 1 and ask yourself: "Can we be sure that Rose has not left out some key trade-boosting factor that operates between many CU pairs?"

This matters for the Rose effect estimate since many of those omitted factors may be correlated with the CU dummy in a way that boosts the estimated coefficient. There are ways for addressing this econometrically, and we'll get to them soon. It is useful, however, to get and idea of just what sort of omitted variables we are talking about.

[14] Formally, if one asks the statistics whether the country dummies should be excluded from Rose-Wincoop, the answer is no. They belong. Therefore, the Rose effect estimates performed without them are null and void.

2.5.1. **Exceptions that prove the rose**[15]

When the Economic Policy Panellists at the presentation of the original Rose paper said they had "an uneasy feeling that something had eluded them", one of the many things that bothered them was that they did not know enough about the particulars of the CU pairs that drove Rose's results. Maybe if one were an expert on West African trade, one Panellist suggested to me, one would know exactly what omitted variable explained the overestimate of the Rose effect. The literature has followed up on this in two ways. The first is to looks at particular cases where we do really understand what was going on. The second was to play with the CU dummy. I address these in order

A parable

Imagine an economist asserted that the growth of the money supply was the main cause of long-run inflation and estimated the link using a huge international dataset. Using money supply growth and a handful of other variables that were available for 150 nations, she estimated the money-prices elasticity to be unity and every other explanatory variable had a negligible effect on inflation. Then suppose another economist showed that Ireland's money supply grew at 300% for decades but its inflation rate was zero. This would make one pause. It would make one think that maybe something else was going on. That maybe the original regression had omitted an important variable.

Of course, a sample of one has infinite standard error, but the counter-example investigator can consider a much more subtle model of the phenomenon since much more information is available than is the case for a sample of 150 nations.

Importantly, the sorts of variables that are available for 150 nations are the sorts of variables that matter for average nations. But Rose looked at a phenomenon that – until the Eurozone – was limited to distinctly non-average nations. Thus, maybe using the 150 nation dataset approach guarantees that no one can find the 'silver bullet' pro-trade variable that would make the Rose effect disappear because no one bothered to gather internationally comparable information on a factor that matters only in for a couple dozen very unusual nations. This is why I think the counter examples consider below must be taken very seriously.

Revolutions, economic chaos and asymmetric inflation lead to CU dissolution

Many currency union break ups are done in the context of, or as the result of massive social, economic and/or political turmoil, many in the context of revolutions. As Thom and Walsh (2002) write:

> "[m]any of these unions ended as part of a bloody decolonising process followed by the adoption of Marxist/autarkic policies, bilateral trade deals with the Soviet Union or China, and a descent into economic chaos – France and Algeria, whose independence was granted only after a bitter struggle; India and Pakistan, who ended their currency union after the war of 1965; Pakistan and Bangladesh, who split up after the war of 1971; South Africa and Southern Rhodesia (Zimbabwe) who were ejected from the Commonwealth and had trade sanctions imposed as they broke with sterling; the five Portuguese colonies in Africa, that broke with Portugal after wars of liberation followed by civil wars. In all these cases and many others it is very likely that trade between the former currency union partners would have collapsed regardless of the currency regime in force."

If this is right, the silver bullet is the mysterious third-cause that drove the revolution, currency union break up and decline in trade. The same in true of currency union joiners. The decision to adopt a currency – for example, to dollarise – takes place in the context of big political and economic changes, many of which could be expected to affect trade. Plainly no one has good data on such factors – oh sure there are proxies to be found in hyperspace – but such things really cannot be accurately measured with a one-dimensional variable; if it could, we wouldn't need historians and political scientists. Be that as it may, my point is that one cannot see whether the Rose effect survives the inclusion this sort of mysterious third effect in the Rose, Rose-Wincoop or Glick-Rose dataset. We also need case studies.

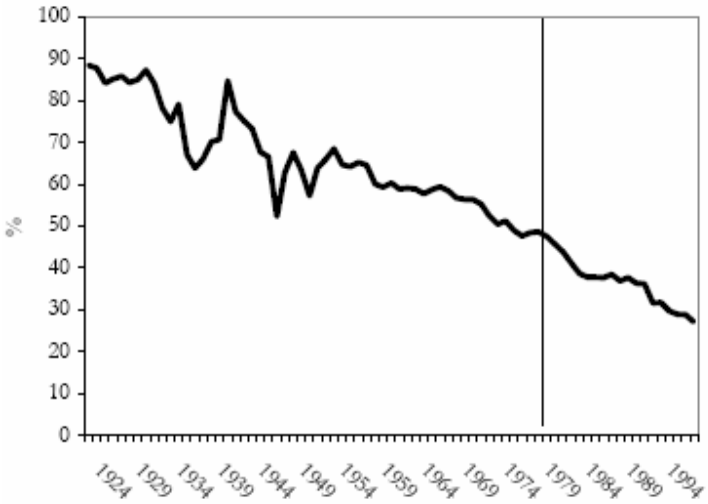
---

[15] Bill Bryson in his book "Mother Tongue" claims that this aphorism is ancient, so that we should read 'proves' using its archaic meaning, 'tests', as in proving grounds.

Thom and Walsh (2002): The bloom on 'my wild Irish rose' is not for the taking

Ireland used the British pound before its independence. After independence and the introduction of the Irish pound in 1927, the pound-punt exchange rate was held at 1-to-1 with no margins. Talks leading up the European Monetary System suggested that this peg would remain in the context of the ERM since everyone initially expected Britain to join. When Thatcher said no in 1979, Ireland was forced to choose between 'Europe' (as they call it in Britain) and the ERM on one hand, and Britain and sterling on the other. Ireland chose Europe and the 1-to-1 peg was abandoned. Market forces lifted the rate rapidly away from the level it had been at for 50 years. What happened to Anglo-Irish trade?

Figure 5: UK's share of Irish trade, 1924-98 (Thom and Walsh 2002).



Since Ireland and the UK were both embedded in the EEC, the termination of the currency union did not and could not raise bilateral trade barriers. Moreover, both nations were run by stable, predictable governments and although there certainly were a number of idiosyncratic factors affecting bilateral trade, one has a very good idea of what they were and very good data that allows one to control for them. In short, we should be able to learn a lot about the Rose effect by studying the Irish case. One recent investigation of this example, Thom and Walsh (2002), finds no evidence from time series or panel regressions that the change of exchange rate regime had a significant effect on Anglo-Irish trade. Should we be shocked?

Let's set out the priors. If the Rose effect discussed in Rose (2000) is roughly right – the currency regime switch should have reduced Anglo-Irish trade to about a third of its initial level. The impact on Ireland should have been massive since the UK absorbed about half Ireland's exports at the time. Even if there were countervailing forces generated by the break up, it is hard to imagine any such forces that would – all else equal – raise Anglo-Irish trade by enough to substantially offset a Rose effect of -200%. By contrast, if the lower ranges of the Rose effect are right – say the effect is 15%, then we might miss the Rose effect in the Irish experience – especially if one thinks the 15% would take a number of years to be realised. My point here is that the Irish experience might help us reject a big Rose effect, but not a modest one.

Inspection of Figure 5 shows that the initial Rose effect just could not have been right. OK, one should run some regressions and talk about the standard errors (Thom and Walsh do), but really, would you ever believe a regression that said the data in figure was generated by a model where trade would have dropped by 200% in 1979 were it not for some offsetting effect?[16]

I think there are other lessons in Figure 5.

[16] Glick and Rose (2003) show a figure for Anglo-Irish hat looks quite different; they use the level of trade which drops due the second oil-shock regression. Thom and Walsh look at the bilateral trade as a share of all Irish exports and thus control somewhat for the global recession.

The gradual decline of Anglo-Irish trade was due to structural changes, in my opinion, mainly in changes in the Irish economy. As Ireland developed from a potato-exporting agrarian economy into the Celtic Tiger it is today, its trade pattern naturally eroded from its historical overdependence on its nearest market. This sort of thing is not in any version of the gravity model. The closest would be to allow for a separate GDP per capita variable for exporter and importer nations (for the exporter it would reflect structural shifts, in the importer an income elasticity – see Anderson 1979), but Rose only includes the product of the two. Now suppose one threw into the gravity equation, the 1965 Anglo-UK free trade agreement, the 1974 adhesion to the EEC and a CU dummy. Moreover, suppose one did this in a panel where it is not really possible to check for serial correlation in the errors. Plainly, the CU dummy would pick up most of the action of the omitted variables that explained Irelands historic over dependency on the British economy. One could throw in proxies for colonial relations in various guises, but none of this would pick up the structural transformation of the Irish economy. Moreover, the history related by Thom and Walsh make it clear that the reduced dependency on the British market – which was driven by factors that are unobservable to the gravity model – is one of the factors that caused the Currency Union to break-up (more on reverse causality below).

Fidrmuc and Fidrmuc (2001): Central and Eastern European break ups

The eyeball evidence for a big Rose effect looks much better for the recent break ups of currency unions in Central and Eastern Europe. Figure 6, taken from Fidrmuc and Fidrmuc (2001) shows that the break ups were followed by dramatic drops in bilateral trade.

The top left panel makes that best case for a big, negative Rose effect due to a currency union break up. In 1993, Czechoslovakia went through a 'velvet divorce' just a few years after its 'velvet revolution.' The two parts of the nation separated into the Czech Republic and the Slovak Republic. They maintained a customs union (no tariffs between them and a common external tariff) until they simultaneously joined the EU's customs union. On the face of it, this is just the sort of natural experiment one should study. The figure plots year-by-year estimates of the above normal level of trade between the partners (e.g. this is the exponent of the pair dummy, e.g. the Czech and Slovak dummy in the Czechoslovak case). But even here one must raise a note of caution. As Fidrmuc-Fidrmuc paper concludes:

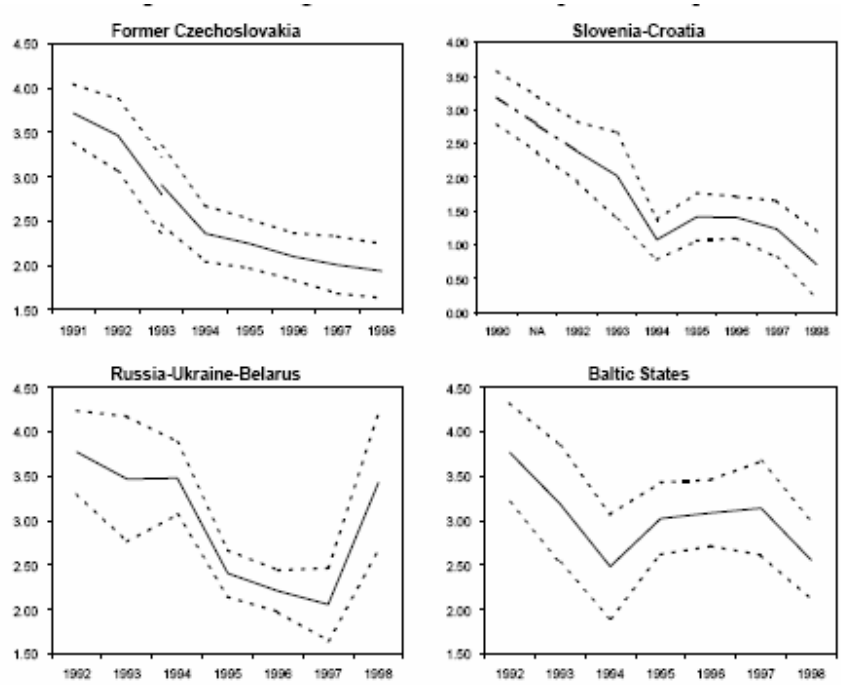
> "Our findings are broadly consistent with earlier findings on currency unions. In particular, Rose (2000) shows that a common currency increases bilateral trade flows approximately three times. Indeed, we found a decline of bilateral trade intensity by about this factor during the first years of independence. However, we cannot separate the effect of the currency separation from that of the political disintegration as both effects occurred (more or less) simultaneously in the countries under scrutiny."

Implications for the size of the Rose effect

The total drop was less than the size of Rose's first estimate of 200%, with the size of the pair dummy falling 100% from about 4.0 to about 2.0. At one extreme, we could claim that the only thing affecting this trade was the loss of a common currency. This is rather naïve, but it give a Rose effect of 2.0, which is similar to many estimates. Yet one suspects that political and economic disintegration also lowered trade. This means that a 100% currency union trade effect is too high; he Rose effect in isolation would be smaller. To explore this conjecture, it would be interesting to revisit the Fidrmuc-Fidrmuc data using some of the more sophisticated methods discussed above to sort out the two effects. De Souza and LaMotte (2004?) have some work on this, but they never include a currency union dummy in their regressions.

Further observations follow from this work. Fidrmuc and Fidrmuc provide a qualitative discussion of the changes that accompanied the currency union dissolutions. Their discussion makes it clear that many time-varying, pair-specific omitted variables affect trade yet were spawned by the same forces that lead to the CU break up. To list just one of a dozen stories, the Czechs and Slovaks maintained free trade after the currency split, but they set up border controls that some businesses claimed acted as a trade barrier. None of these stories could be included in regression like Rose estimates since there would be no way to gather such data for 100+ nations.

Figure 6: Trade collapses in Central and Eastern Europe.



The lessons from these two cases are unclear in terms of specifics, but crystal clear in terms of generalities – lots of other complicated stuff matters. And it is the sort of factors one which we will never have good, internationally comparable data. In short, gravity equations will always have omitted variables. (Thank goodness for that; think how boring international trade economics would be otherwise.)

2.5.2. **Pair dummies: Glicks 'N Roses**

Andy Rose was, of course, well aware of the omitted variable bias critique even before it was echoed many times by Economic Policy referees and panellists in Helsinki. He was also well aware that using pair-specific dummies would wipe out ALL idiosyncratic level effects between ALL pairs of nations. The only sticking point is that this tends to throw the roses out with the vase water. It eliminates all cross-section variation from the residual, so the identification comes solely from time series variation.[17] In plain English, we need lots of data to do this. As he explains it, he didn't do it in Rose (2000) since there was too little time variation in his original dataset. In Rose (2001) he shows what this means. Using pair fixed effects on his original dataset, the Rose effect wilts (the raw estimate on the CU dummy is minus 0.38 and the standard error is 0.67).

2.5.3. **Pakko and Wall (2001)**

Pakko and Wall (2001) independently obtain the same results using a more general approach in terms of fixed effects and data. They use the Rose (2000) data set but instead of averaging the two way bilateral flows (i.e. Germany's exports to Denmark and Denmark's exports to Germany), they preserve the directional flows. This allows them to impose direction-specific pair dummies, i.e. two different dummies per bilateral flow – a technique that is more general than in Rose (2001). Although they get Rose-like estimates of the Rose effect without pair dummies, they find that the Rose effect droops and withers away completely with pair dummies. The point estimate is negative and not significantly different than zero.

---

[17] Note that Rose (2000) does roughly this with his difference-in-difference regression that is reported in the text but not in a table; the Rose effect this yields is only 17% more trade!

Rather than pushing quickly on to the next dataset and empirical technique as does Rose (2001), Pakko and Wall take the time to crush the rose petals one-by-one. Here is how they put it:

> "Independently, Rose (2001) obtains these same results using the general fixed-effects model. However, he rejects the findings on the grounds that the statistical insignificance of the common-currency dummy is due to a small number of switches in common-currency status. While it may well be true that the statistical insignificance of the common currency dummy should not be taken to mean that the effect is not positive, this misses the point. A comparison of the two sets of results suggests that pooled cross-section estimates are not reliable because they are biased by the exclusion or mismeasurement of trading pair–specific variables. This is evident in the dramatically different coefficients on the GDP and per capita GDP variables that are found when using the two methods. In other words, the restrictions necessary to obtain the pooled cross-section specification from the fixed-effects specification are rejected, indicating that the fixed-effects specification is preferred.
>
> The difference between the two methods in their estimates of the trade-creating effect of a common currency is a separate issue. The proper conclusion to draw is that, when the statistically preferred fixed-effects specification is used, there is no statistically significant evidence of large trade effects (positive or negative). Although this means that Rose's results cannot be supported statistically, the small number of switches precludes us from saying much about the effects of common currencies on trade, although the tripling of trade found by Rose is well outside of a 95 percent confidence interval."

This is a critical point that should not be overlooked by researchers. If you can show that the pooling assumptions are false, then you should ignore all pooled estimates for policy purposes.

2.5.4. **Rose revival**

O MY Luve 's like a red, red rose/ …/ And I will luve thee still, my dear/ Till a' the seas gang dry/ …/ And I will come again, my Luve,/ Tho' it were ten thousand mile. - Robert Burns, 'A Red, Red Rose'

Andy Rose is not a man to shy from a challenge. He saw the wilting of the Rose effect as a lack of data, so he set about collecting an enormous panel dataset. He was, so to speak, trying to graft the old flowering stem on to a healthy new dataset, and guess what? The flower continued to blossom. The massive dataset he collected included annual data from 1948 to 1997data on bilateral trade between 217 countries. Theoretically, that's $50(217^2)/2$=2,354,450 data points, but with missing observations and zero flows (lots of little nations sell nothing to each other), the new Rose dataset has 219,558 observations.

Glick and Rose (2001) exploit this data in a number of ways. They throw in pair-specific dummies that soak up any sort of idiosyncratic omitted variables that do not vary between 1948 and 1997. This, of course, mimics the impact of country-dummies as in Rose-vanWincoop, but it goes further. The result was, as we should have expected in the post-RvW world, that the size coefficient drops dramatically – about 5 standard deviations from an estimated coefficient of 1.3 to 0.65. This brings down the Rose effect from 3.7 to 1.9 times more trade among CU pairs (both estimates are statistically significant at any conceivable level of confidence).

What is going on here? The estimates are still biased

Pretend, for a moment, that Glick and Rose did their regression in two stages. First they regressed the left-hand side variable on the time-invariant pair dummies. Second, they regressed the residuals from that regression on the main right hand side variables, distance, CU and the joint real GDP variable. This procedure is terribly inefficient in the econometric and practical sense, but it is very efficient from an intuitive stand point.

The first stage, strips out all time-invariant features of each bilateral trade flow. This completely removes the bias stemming from the cross-section correlation between CU and Z. It does the same for the cross-section correlation between what we called X_{21} (the relative-prices-matter term) and the CU dummy. However, there is almost surely a bias left. We know that the relative-prices-matter term varies

over time, so there will still be a correlation with CU, after all the theory tells us that the relative-prices-matter term contains CU and CU itself is time varying.

Second, there may still be a bias stemming from time-series correlation with omitted variables. For example, if currency union encourages nations to deepen other forms of integration that are unobservable to the econometrician, then CU and Z could be correlated over time as well as in a cross-section sense. This is especially true since each pair gets only one dummy for the full 1948-1997 data period.

Stop and smell the roses

What do we learn from this?

First, the impact of allowing for country-specific idiosyncrasies (either via the Rose-Wincoop procedure or by the Glick-Rose technique of pair dummies) reduces the Rose effect massively. This confirms yet again that the original +200% Rose effect was over-estimated. Of course, 90% more trade is still a huge number, but 200% is huge-er.

Second, the Glick-Rose result should have a large 'caveat emptor' stamped across its forehead. The pair dummies mean that ALL the identification is coming from the way in which trade between CU pairs changes over the sample, compared to the way it changes for the non CU pairs, controlling for other factors. The Glick-Rose data has lots of pairs that leave monetary or currency unions, but very few that join (16 joiners and 130 leavers, with almost all of having happened before the post-war independence wave ended around 1970). This means that the results are being driven by how much trade DROPPED after a nation leaves a monetary union, not by how much trade is created by a currency union. Nitsch (2005) estimates that effects of currency union joiners separately and find the point estimate is small (about +8%) and statistically indistinguishable from zero.

To put it differently, if you want to know how much a small nation's trade might drop if something happens such that it has to, or wants to leaves its currency union, the then Glick-Rose numbers are what you need. If you want to know how much trade a small nation would gain from abandoning its own currency and adopting someone else's – the Glick-Rose numbers are not what you need. I think we have to admit that there just haven't been enough new currency unions to answer the question. Or at least not until the Eurozone came along, but I'm saving that part of the story for later.

Why are the country and pair dummy results so similar?

Third, the point estimates from the Rose-Wincoop and Glick-Rose approaches are amazingly similar. With the naïve gravity model the Rose effect is 3.97 in the Rose-Wincoop dataset and it is 3.66 in the Glick-Rose data set. Allowing country-specific idiosyncrasies drops the estimate to 1.9 times more trade among CU pairs in both datasets.

This, I believe is reassuring on one hand, but worrying on the other. The reassurance is obvious; different datasets, same results. The worry is that it seems to make no difference whether one controls only for country-specific idiosyncrasies or one controls for country-specific idiosyncrasies and all other pair-specific idiosyncrasies. Why worry?

My priors are that there are omitted variables correlated with the CU dummy – e.g. the quality of FTAs, informal ethnic networks, foreign direct investment flows (trade and FDI are complements empirically), and many more – for which the data is non-existent or too poor to use in a regression. If my priors are right, the pair dummies are not doing their job properly. The why-part is easy.

Many of the pair-specific omitted variables probably varied over the five decades in Glick-Rose data set. Thus putting in a time-invariant pair dummy leaves a times-series trace in the residual and this trace is probably still correlated with the CU dummy. In particular, there are probably pair-specific factors that caused nations to leave currency unions and these are probably time-varying. This is certainly the lesson to draw from the case studies. More vigorously, Nitsch (2004) uses a large panel of currency unions pairs to identify factors involved in the break up. Inter alia, he finds that finds that departures

from currency unions tend to occur when there are large inflation differences among member countries, and when there is a change in the political status of a member.[18]

Glick and Rose try out an admirable range of robustness checks, but they obviate most of the merit of the exercise by trying them one by one. For example, they use data from tiny nations in 1950 in the same regression as data from the United States in 1995. It would take a brave soul to assert that the income elasticity of imports was the same number in these two cases. Tenreyro (2004) is particular strong on this idea that one must address all the problems together. Sure, that's a lot of regressions to try, but your have to water the thorn to harvest the rose.

2.6. Rose branch #3: Complicated mis-specification

There are two ways of correcting for omitted variable biases. The Glick-Rose approach works by throwing lots of dummies into the regression. The alternative works by throwing lots of observations out of the regression. That sounds strange, but it has many merits. Torsten Persson's 2001 paper in Economic Policy introduced this technique into the Rose effects literature. The technique is subtle and complex; Persson explains it in technical terms (and you should have seen how technical it was before Giuseppe Bertola rewrote it in his role as Managing Editor of the Persson's paper).[19] Allow me to relate a parable that may make the nature and intractability of the problem clearer.

A parable

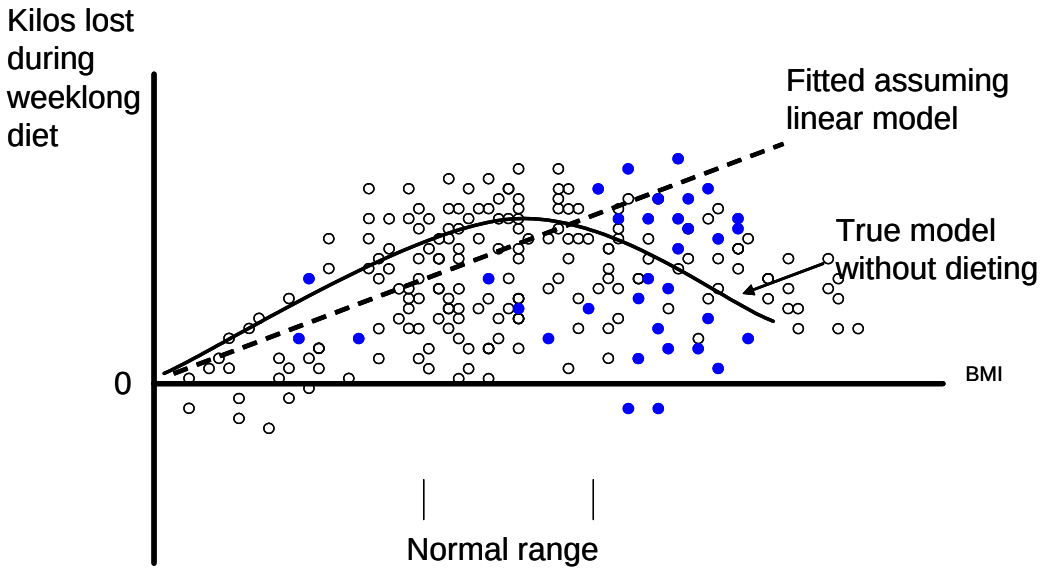
A few years ago, middle-age surprised a 'friend of mine' and it chose to focus on his middle; he developed a little belly. He decided to do something about it and, being an egghead, he started reading studies on the effectiveness of dieting. One study found that a week's worth of dieting was astoundingly effective. I have plotted the data in Figure 7 (at least the data as my friend remembers it). Crucial background: People tend to gain weight as they get older (their metabolism slows), so there is an empirical link between weight and gaining weight. A proper account of the effectiveness of dieting must take account of this. Assuming the link is linear, the study fitted the curve shown with the dashed line. However medical science (as my friend remembers it) tells us that the true weight-weight gain link is bell shaped. (Once you reach middle age you pile on an extra 10 kilos and this rapidly pushes you just beyond the normal Body Mass Index, or BMI, range but then the process slows down.)

With this background we can see how the study overestimated the dieting effect. The solid dots are the weight gain of dieters and you can plainly see that they are below the linear dashed line. The study claimed therefore that dieting was very effective, controlling for other factors. Obviously this is a spurious finding since the dieters' weight gain is white noise around the true-model prediction without dieting. Why the incorrect inference? The subtle interaction between nonlinearity and self-selection.

[18] By the way, these suspicions of mine were raised by the similarity of the two different techniques applied on two separate dataset. It would be interesting to make a more direct comparison, to see what the Rose-Wincoop country-dummy technique would yield on the Glick-Rose dataset, and what the Glick-Rose pair dummy technique would yield on the Rose-Wincoop dataset. Such comparisons would help us to judge the importance of the omitted variable critique, and the validity of the Glick-Rose solution of throwing in one pair dummy for the whole period.

[19] Actually you can see it, since Economic Policy posts the Panel drafts on its web site www.economic–policy.org.

Figure 7: Model mis-specification and overestimation of treatment effect



First, if the study had estimated the correct nonlinear model, it would have found that dieting was useless. Second, if the true relationship had been linear, then the deduction would have been valid. Finally, if dieting were randomly distributed across all weight classes, the model mis-specification would not have mattered since there would have been an equal number of dieters above and below the fitted line (that's what OLS does). But, dieting is self-selected. The people who are most likely to start a diet are ones, like my friend, who have just crossed into the 'jolly but not yet jelly' category.

2.6.2. **War of Roses**

The Persson critique was presented at the Paris meeting of the Economic Policy Panel, which was, by the way, was hosted by Jean-Claude Trichet in his role of Governor of the Bank of France.

Persson (2001) employs a matching technique that can control for this sort of nonlinearity-with-self-selection to the Rose (2000) dataset and finds that the point estimate for the Rose effect is much lower – Persson's estimate of the Rose effect range from 1.13 to 1.66 – and they are not significant statistically. Kenen (2001) confirms part of the basic result using a different matching technique, but obtains very different result in the regression analysis.

What's going on here? It is useful to think about matching in terms of the parable. The matching technique would throw out most of the non-dieter observations since they do not match those of the dieters. If one compares the mean weight-loss of dieters and non-dieters in the narrow range just to the right of normal, there is no difference in mean, so in the case of the parable, matching would yield the correct inference. In this way, matching automatically eliminates the impact and any sort of non-linearity by neutralising the interaction between self-selection and non-linearity.

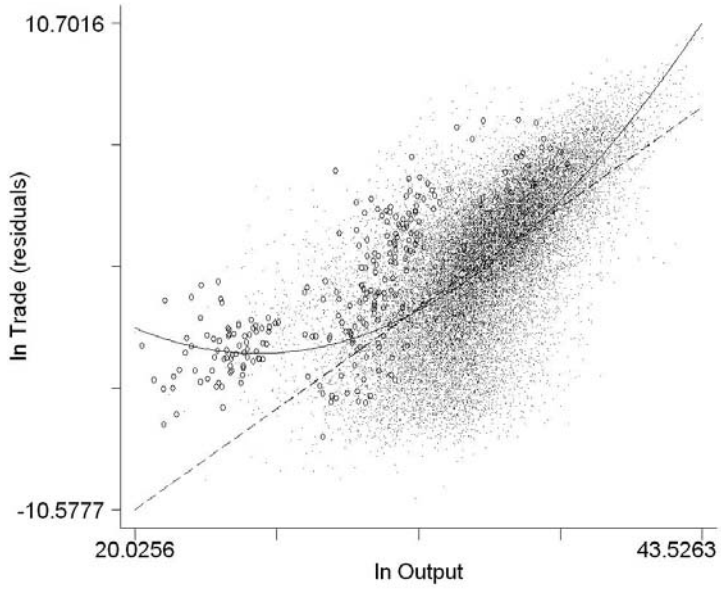
What is the nature of the nonlinearity-with-self-selection in Rose's study? Persson rightly points out that while there is only one way to be linear, there are an uncountable infinity of ways to be nonlinear. One cannot check them all, but the Persson thinks he may have found one important nonlinearity – a nonlinearity that concerns the openness and output link.[20] Figure 8 shows the suggestion.

The figure plots all residuals from Rose's preferred linear regression with the residual plotted against their corresponding log of GDP (pair product as usual). Non-CU observations are shown with black dots, CU observations with circles. The straight line shows the estimated linear relation between bilateral trade and output – i.e. the linear model imposed by Rose. The curved line shows the best fit

[20] By the way, this nonlinearity is consistent with the Krugman's famous Home Market Effect whereby a nation's exports are affected by its size.

allowing for a nonlinear relationship between openness and output. Just as in the parable above, the non random distribution of CU pairs team up with the 'true' model's nonlinearity to produce an overestimation of the effect. The point is that if one compares the positions of the circles to the straight line, it looks like they have far greater trade than they should have had. If one compares them to the curved line, the circles are, on average, above predicted relationship, but much less so than if one takes the straight line as the true model. Thus the linear regression substantially over estimates the impact of a common currency on trade because is underestimates how much trade they would have had without a common currency.

Figure 8: Persson's hypothesis for why the Rose effect is overestimated.



Persson's punchline

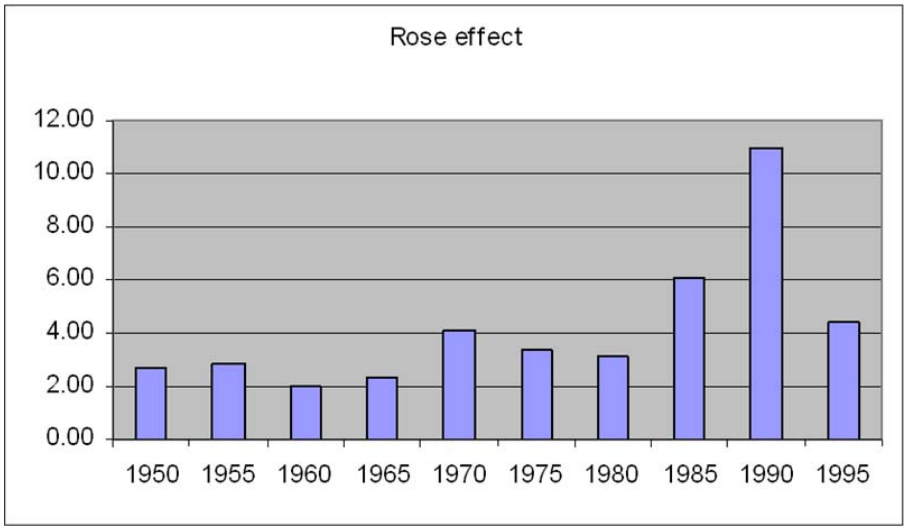
In short, Persson asserts that Rose (2000) overestimated the effect since he was comparing the actual trade to a mis-specified model of what trade should have been absent the common currency.

Further evidence comes from the fact that allowing a quadratic term in Rose's regression (i.e. pooled cross-section without country or pair dummies), the Rose effect estimate drops radically. Rose (2000) included a squared output and per-capita output terms in one of his dozens of regressions. When he does, he finds that the Rose effect drops dramatically, from 3.39 times more trade to 1.95 times more; this is a four standard deviation drop in the coefficient.

Further evidence for this interpretation – albeit very indirect evidence – can be found in Glick and Rose (2001). Glick and Rose (2001) estimate the naïve gravity model on cross-section data for a handful of years reaching back to 1950. The estimated Rose effects from a selection of years are plotted in Figure 9. It is interesting that the size of the effect rises over time. What could this mean? One cannot know for sure, but the Persson-Kenen finding suggests a story. In 1950, many nations participated in currency unions. Most nations were still colonies and many of these used the currency of the coloniser. Or, to put it differently, the group of nations sharing common currencies was much more randomly spread. As the decade of independence arrived, many nations adopted their own currency as a symbol of sovereignty.

In the Glick-Rose data there are 130 CU leavers but only 16 entries.[21] The roll-call of CU dummy pairs thinned out, but the decision to quit the coloniser's currency was surely not random. Really tiny, really open economies like New Caledonia decided they could not afford their own currency, while nations like Algeria went their own way. In this chronicle, the self-selection part of the nonlinearity-with-self-selection bias gets more severe as time passes. Thus if you believe the Persson-Kenen account, the rising Rose effect is completely in line with your priors.

Figure 9: The Rose effect over time.



2.6.3. **The bloom is off the rose, or is it? Rose redux**

Andy Rose is not one to let a new econometric technique lay in bed till noon. When Economic Policy invited Andy to present a live rejoinder to Persson at the April 2001 Panel in Paris, he leapt to the wall with bow strung and arrow nocked. He used a new, bigger data set and applied Persson's matching technique. Guess what? Rose (2001) confirms that matching lowers the point estimate on the CU dummy (the Rose effect drops form 3.39 to 1.21 in the strictest match to 1.43 in the most relaxed; that is 21% and 43% more trade predicted for CU pairs). However, and critically, he finds that the Rose effect is statistically significant in the bigger dataset even with matching. In the end, the rose was still blooming. 20%, after all, is a pretty big number, certainly far bigger than prior in the minds of most international economists in 2001.

2.6.4. **Lessons: take nonlinearity and selection seriously**

Decades of gravity model research tell us that the naïve model does pretty well in most cases. The vast majority of the thousands of gravity equations estimated over the past 40 years assumed linearity without objection. But in the old days, we could not handle large datasets, so most gravity estimators, including me in 1993, used data from nations that were pretty homogeneous, like European nations plus the USA, Canada and Japan. Or, all Latin American nations. In these cases, linearity – even if it is wrong – is not a big deal. After all, any continuous model is approximately linear.

But the linear approximation gets worse the further one is from the point of approximation. Since estimations is, in effect, approximating the functional relationships for the average country in the sample, the problem gets worse as the sample includes more extremely big, small, open or closed

[21] 21 It would be interesting to see the share of trade pairs with common currencies by year, but this is not reported in Glick-Rose, only the full panel average is reported.

nations. The problem is extra severe in the hunt for the Rose effect since nations that are members of monetary or currency unions are extremely far from average; see Figure 2. One way of thinking about what Persson and Kenen did is to say that they were trying to get a more homogeneous sample so that whatever nonlinear does exist is not a big deal.

I believe is extremely important to take seriously the Persson-Kenen lesson in any gravity equation study that uses data from a very heterogeneous group of nations. The econometric theory tells us that if the true model is nonlinear yet a linear model is estimated, then the estimated coefficients are biased if the policy under consideration is not randomly distributed across all observations. Both of these premises hold for the gravity model on the Rose (2000), Rose and Van Wincoop (2001) and Glick and Rose (2001) data sets, so we know the standard gravity-model estimate of the Rose effect is biased. To wit,

- We know that CU pairs are not random. The first stage matching regressions confirm the suspicion raised by Table 1 and this has been confirmed many times over by authors such as Alesina, Barro and Tenreyro (2003), and Nitsch (2004).
- We know that the true gravity is nonlinear (Rose 2000 finds a t-statistic on 24 on the GDP squared term and there may be many other nonlinearities).

Again, history bifurcates. Before the 2001 Persson-Kenen-Rose papers, we didn't think nonlinear was an issue. Now we know it is. We are not exactly sure how best to address the nonlinearity, but we know it is a problem. Two more lessons:

5) For policy purposes, we should ignore all Rose effect estimates on large dataset that do not address the nonlinearity-cum-selection issue. Researchers would be wise to address it in both ways: 1) try out various nonlinearities. In the context of Rose effect regressions, be sure to try a quadratic terms for GDP and GDP per capita; 2) try matching procedures like those suggested by Persson or Kenen (Honohan 2001 suggests another in his discussion of Persson).

6) The Rose effect on multilateral data is about on the order of 20% to 40%, but this figure basically reflects the extent to which bilateral trade *dropped* between nations when a currency union pair involving a small poor nation is dissolved.

2.7. Rose branch #4: Roster-makes-the-sun-rise reasoning

While Andy Rose was declaring 'Eureka' for having shown that a common currency bumps up bilateral trade, other researchers were declaring a Eureka for showing the reverse. Devereaux and Lane (2002), inter alia, showed that nations tend to stabilise their bilateral exchange rates against nations with whom they trade a lot, with a common currency being an extreme form of stabilisation. There are many sophisticated reasons for this reverse causality ('reverse' that is from the Rose effect perspective), but my favourite is a political economy story. If a nation's currency depreciates against its major trade partner, cheers arise from exporting firms but screams are heard from firms that import components and materials. The roles are reversed for appreciations. Since losers lobby harder[22] there is strong political pressure on a Central Bank is to keep the exchange rate steady against the currency of its major trading partner, especially in very small nations where the importers and exporters dominate local politics. In extreme case, this means adopting the partner's currency unilaterally or multilaterally.

The reverse causality problem is a thorn in the side of would-be Rose-effect estimators. Although one could invent elaborate stories for why the simultaneity bias might be negative, common sense tell is that the bias should lead to an upper bias. That is, at least part of the reason the circles in Persson's diagram are above the predicted trade line is because unusually open nations, especially tiny ones located near big markets are more likely to join currency unions. In this case the conditional correlation between a common currency and trade is boosted in part by the impact of currency on trade and part by the impact of trade on currency. This leads me to believe that all of the Rose effects discussed up to this point are too large.

[22] See Baldwin and Robert-Nicoud (2002) for an explanation based on sunk costs.

Of course this possibility occurred to Andy and in his Economic Policy article, he tries to control for this with instrumental variable techniques. His choice of instruments (inflation rates), however, was regrettable and forgettable. When he instruments for the CU dummy, the Rose effect becomes "wildly and implausibly bigger" in Andy's own words; to wit, the Rose effect is 1.1 times 10^{36}. To put that in context, it implies that Fiji's adoption the Australian dollar would raise its bilateral exports to a level that is several times larger than the value of world trade. I think we can agree that his instrumental variables strategy was not successful; he abandoned the effort in subsequent papers.

2.7.1. **Other instrumenting strategies**

Three other instrumenting strategies have been mentioned in the Rose-effect literature and two have been implemented: one is based on money supplies along the lines suggested by Frankel and co-authors, and one by Tenreyro (2003). The final one, by Devereaux and Lane (2002), has not been tried to my knowledge.

Tenreyro (2002) estimates the likelihood of a nation joining a hub-and-spoke currency arrangement with the US dollar, British pound, French Franc or Australian dollar. She explains this decision using a dozen of so variables that are very closely related to right-hand side variables in the gravity equation. She finds that the probability is increasing in common language, common border, former colonial status and the smallness and poorness of the nation. The fit of this first stage regression is not very good; the pseudo-R^2 is only 0.473, so roughly speaking she only explains about half of hub-and-spoke currency pairs in her data. Since CU pairs account for less than 1% of all pairs in her data, this could be a real problem in the sense of amplifying the nonlinear-and-selection biases. Her first stage explanatory variables are a sub-set of the second stage explanatory variables, so the probability itself cannot be used as an instrument. To get around this, she constructs an artificial probability of any two spokes sharing the same currency by multiplying each spoke's probabilities of adopting each anchor currency and summing over the four products.

This seems a clever idea, but there are two problems. First, I would like to see what this procedure yields. For example, how well does her newly-minted fitted CU variable line up with the real CU variable? I would bet that this procedure invents lots of CU ties between very small, very open economies and thus exacerbates the Persson-Kenen problem.

Second, she writes in Alesina, Barro and Tenreyro (2002), "The underlying assumption for the validity of this instrument is that the bilateral trade between countries i and j depends on bilateral gravity variables for i and j but not on gravity variables involving third countries." As the simple gravity model theory laid out above shows, this identifying assumption is false. All you need to do is remember that the gravity model is essentially a demand equation and you know that each bilateral flow is affected by the trade costs of every partner of the importing nation – via Δ -- and every partner of the exporting nation, via Ω. The Tenreyro identifying assumption is essentially saying that the trade between two nations depends only upon the nominal price, not the nominal price divided by an index of prices from all sources. In fact, this third country dependence is exactly what the whole Anderson-Wincoop paper is about. Curiously, she includes Anderson-Wincoop dummies in some of her regressions and thus implicitly admits that her identification strategy is based on a false premise. But maybe I missed something. Other things seem strangely upside down in that paper; adding country dummies increases the OLS estimate of the Rose effect, in contrast to what many others have found.

In any case, she finds that instrumenting raises the Rose effect to levels that would make Andy blush as red as a rose – e.g. she gets a Rose effect of 6.77 in her paper with Barro and 14.87 in her paper with Alesina and Barro (for the record, that means joining a currency union would increase a nation's trade with its CU partners by 577% and 1387% respectively). If this right, then Finland with its 5 million people will export more to Germany than the United States exports to the whole world; since Finland has 10 Eurozone partners and the effect applies to each, Finland's decision to join Euroland will double world trade – according to Tenreyro's upper estimate. Her lowest IV estimate predicts that formation of the Eurozone will more than double world trade.

Of course, you may think that it is inappropriate, maybe even unfair, to extrapolate from the results of small nations. But if you think that, then you don't believe the linear gravity model works well for

nations that are extremely different from the average nation. In other words, you don't believe what one must believe to think the Tenreyro identification strategy makes sense.[23]

Given these problems, I think we can conclude that Tenreyro procedure failed. Probably prudent to consign it to the regrettable and forgettable bin along with Rose's instrumenting strategy.

I hasten to note that the theoretical points in Alesina, Barro and Teneryo (2004) are interesting and useful. The empirical implementation, however, is a failure IMHO.

2.7.2. **Aphides: All instruments will be bad instruments**

It is hard to know exactly how it all went so wrong with the Rose and Tenreyro approaches. It helps, however, to think about the instrumenting technique. Basically, the econometrician has to invent a new data series that is looks something like the CU dummy but isn't the CU dummy. This newly invented variable is thrown into the gravity model and the coefficient on the new invented variable is taken to be the currency union effect. There are several problems with this procedure in my opinion:

- **Schrodinger's cat and amplified nonlinearity-selection problems**. The CU dummy is digital by the nature of the policy under investigation. The instrument, by construction, will be a continuous variable. What this means is that the Rose effect estimate will be influenced by the sample covariance between the invented variable and the bilateral trade of all pairs. If the invented variable doesn't resemble the CU dummy very closely, it is likely to display many features that the CU dummy does not and may therefore generate an estimated coefficient that has nothing to do with the Rose effect. This point is amplified if nonlinearities are importance, as I believe they are (at least at the extremes of country size). As discussed above, part of the overestimate in Rose (2000) stemmed from self-selection and nonlinearities. The first stage of the instrumenting process guarantees non-random selection – in fact it probably makes it much worse. Just take a look at Persson's diagram and imagine what the invented variable would look like. I haven't seen the data, but my bet is that Tenreyro's invented CU variable places lots more circles (in a probabilistic sense) far above the straight line in Figure 8.
- **All instruments will be bad instruments**. A good instrument is correlated with the offending explanatory variable, but uncollected with the regression error. I believe all CU instruments will fail on both criteria. No instrument is going to fit well. Having a common currency is a really unusual outcome that is in reality governed by factors that can never be quantified in variables that exist for 150+ nations. This means a bad fit, but things are worse. No instrument will be uncorrelated with the error due to 'double omitted variables'. The first stage of fitting a currency-union model will, of course, omit many variables given the fuzzy political, social and cultural factors involved in such a choice. The instrument will not be orthogonal to these factors. Thus if some of the variables omitted first-stage variables are also omitted variables in the gravity equation, then the instrument will not serve its purpose. Lots of things on which we have no reliable data could promote both CU membership and bilateral trade among CU members. Here are couple of ones that occurred to me: personal ties developed while studying in the 'hub' nation, and greater than usual mutual understanding of each others' legal systems, but the list is almost endless.

2.7.3. **Lessons: Briar Rose**

The Brothers Grimm tell the tale of a beautiful princess who came to be known as Briar Rose since she slept a hundred years in a castle surrounded by an impenetrable hedge of briars (thorns). Princes from far and wide sought to claim the great prize that lay within, but all found it impossible to get through the thorny hedge, "for the thorns held fast together, as if they had hands, and the youths were caught in them, could not get loose again, and died a miserable death." Then one day a prince came near to the thorn-hedge and "it was nothing but large and beautiful flowers, which parted from each other of their own accord, and let him pass unhurt."

---

[23] Tenreyo has a recent paper on the impact of exchange volatility on trade that applies the same IV strategy. Her conclusion is that volatility has no effect on trade which is strange given her early findings on the CU dummy. Strangely, this paper, Tenreyro (2004), excludes the common currency variable altogether and indeed never mentions it. Maybe it would have been too jarring to have a common currency boosting trade many times over, but lower volatility having no impact. Or, maybe she, like Rose, decided that IV estimation of the Rose effect was a dead end.

Plainly, instrumental variable techniques are what we need to properly control for the reverse causality that must be biasing the Rose effect upward. Maybe there is a beautiful instrument lying inside the thorn bush of practical problems. Up to now, however, attempts to find that an instrument have failed miserably. What might we do better? Well surely the idea would be to get some financial variables a la Devereaux and Lane since these may influence the CU decision without influencing bilateral trade, at least at trade data at 5 year intervals.

Table 3: Rose effect estimates arrange by estimator.

AUTHOR	*OLS (Pooled)*		*Country Fixed Effect Time-Invariant*	*Pair Specific*		*Matching*	
	Linear	*Non linear*	*Linear*	*Linear*	*Non linear*	*Linear*	*Non linear*
Rose (2000)[a)]	1.21 (0.14)		0.77 (0.16)	-0.38 (0.6)			
Rose & van Wincoop[a)]	1.38 (0.19)		0.91 (0.18)				
Glick & Rose[b)]	1.30 (0.13)			0.65 (0.05)	0.61 (0.05)		
Tenreyro	0.09 (0.14)						
Persson[a)]		0.937 ; 0.69 (0.15); (0.15)				0.52 (0.320)	0.37 (0.320)
Rose response[b)]				0.74 (0.052)	0.66 (0.05)	1.47 – 2.19 (0.09) – (0.14)	1.4 – 2.1 (0.09) – (0.14)
Pakko & Wall[a)]	1.17 (0.143)			-0.378 (0.529)			
Kenen	1.7 (0.310)					1.2 ; 1.4 (0.30) ; (0.32)	

Notes: a) Rose's 5 year dataset. 1970-1990. UN data. b) Glick-Rose dataset. 1948-1997. IMF data.
To find Rose effect in terms of % increase in trade, take exponent of coefficient and subject 1.

2.8. Meta-Analysis: A rose is a rose is a rose

Readers who have made it this far may have a muddled impression of the many estimates discussed above. Wouldn't it be great to have one summary statistic, the number as it were? There is such a number but I do not believe it is useful for policy purposes.

2.8.1. **Weighted average of all point estimates**

Rose and Stanley (2004) perform a sophisticated analysis on thirty-four studies of the Rose effect that yield 754 point estimates. They reject the hypothesis that the true number is zero. The range they arrive at is 30% and 90%. Surely, this is taking things too literally. Or more precisely, it throws away too much information by treating all estimates as having been generated by the same process. As the authors note: "While we have strong views about the quality of some of these estimates, each estimate is weighted equally; alternative weighting schemes might be regarded as suspect." Please, suspect! That's what empirical researchers get paid for. All the estimates in Rose (2000), for example, should be ignored except the difference-in-difference estimator that roughly controls for the gold-medal mistake of gravity models. Andy Rose himself showed that all of them were incorrect since the pooling

assumptions necessary for them to make sense have been rejected by his papers with van Wincoop and Glick.

Moreover, the patently incorrect pooled estimates of the Rose effect – all of which are at least twice too big – are generally repeated in the literature as a way of showing that the author's data set is sound in that it can reproduce the mistaken estimates in Rose (2000). In other words, authors repeat them as a form of benchmarking, not for policy relevance. The meta-analysis statistical techniques are fascinating, but I don't believe it adds to our knowledge since deep down they are basically a weighted average of all point estimates. As we have seen, many of the published estimates are patently overestimated for reasons that are quite clear.

2.9. Lessons for the Eurozone from non-European experiences

I believe the cleanest estimates we have on non-European currency unions are those that use matching since these go a long way towards controlling for the omitted variable bias, and any bias from model mis-specification. All the other estimates seem to have serious flaws that would bias them upward. On the original Rose dataset, Person (2001) found the effect to be between 15% and 66%. Rose (2001) found it to be between 21% and 43% on a much larger dataset. An informal meta-analysis on these four estimates would suggest a number like 30%. And we can be pretty sure that this is an overestimate since the matching procedure cannot control for reverse causality, and it seems extremely likely that the je-ne-sais-quoi factors that lead nations to adopt each others currency also tend to promote bilateral trade.

I believe, however, that the non-European evidence has essentially zero informational content for the Eurozone – apart from the fact that it is worth looking for a Rose effect in Europe. The basic problem is that the non-European data are driven by nations that are very small, very poor and very open. Exactly because the currency union pairs in the data are so strange, we cannot use the 30% to predict the currency union affect for any nation that is not strange in the same way.[24] I guess this falls into the category of common sense. If you study the trade effects of a currency union on very small, very poor and very open nations, then what you learn is how much currency unions affect the trade of very small, very poor and very open nations. Did I repeat myself? Well, I guess that is why they call it common sense.

3. EMPIRICAL FINDINGS ON THE EUROZONE

"What matters is not the length of the wand, but the magic in the stick" ~ Hagrid to Harry Potter

Most of the Rose effect literature treats currency unions as magic wands – one touch and intra-currency-union trade flows rise between 5% and 1400%. The only question is: "How big is the magic?" This approach was understandable when the literature was dealing with hundreds of pairs of trade among CU members. Given the amazing range of peculiar situations under study – everything from France's trade with its overseas departments to trade between two tiny Pacific islands Nauru and Tuvalu – one was naturally attracted generalisations.

In Europe, however, the big-magic approach is most definitely not good enough. We know an awful lot about the affected countries, far too much to pretend that the euro will affect all their trade flows in the same way. Moreover, the euro matters far too much for easy generalisations to be appropriate. 290 million people use the euro, and euro monetary policy quite directly touches the lives of another 200 million people living the non-Euroland EU nations and near neighbours.

[24] Actually, since France includes some small, poor, open and remote islands (Outré-Mers), we could test whether the euro boosted trade between these islands and, the nations that were not in the DM bloc-franc fort complex, say, Greece, Portugal, Spain, Finland and Ireland.

Empirical studies of the euro's trade impact[25]

Given the roaring interest sparked by Rose's papers, it was inevitable that someone would try to estimate the Rose effect for the Eurozone. In April 2002, Andy Rose alerted the Managing Editors of Economic Policy of what appears to be the first paper on the subject.[26] When I got this message, I looked up the paper and saw it was very much in the how-big-is-the-magic line – not at all what we thought the world needed, but we asked the authors to write a paper that would go much further and after some iteration, commissioned what was eventually published as Micco, Stein and Ordnez (2003), or MSO for short. After a massive first revision between the first draft and the Panel draft, the authors presented their paper at the October 2000 panel meeting, which was hosted by the Bank of Greece.

By the way, Lucas Papademos attended the Panel, and gave us an excellent dinner speech[27], but I am not sure he sat through this particular paper. The Economic Policy Panellists were quite positive towards the paper and its main conclusion that the euro had already boosted trade, but they had many questions and suggestions. In their second revision, the authors addressed almost all these concerns. After a thorough edit, MSO was published in the April 2003 issue of Economic Policy – and issue in which all articles dealt with the euro's impact.[28]

3.2. MSO (2003)

The estimates in the earliest version of MSO presented estimates of the Rose effect on the order of about 25%. These were so big that the referees and editors asked them to plot the data and see if it appeared even without conditioning for other trade-enhancing factors. The result is shown in Figure 10.

It is somewhat hard to see what is going on when the data is plotted from 1980 to 2002 (the latest data point MSO had), but in the right panel, the same data is plotted from 1993 to 2002. Note that there was a major change in the way the EU collected its trade data from 1993, so it is difficult to compare pre and post 1993 figures. (Much more on this below.) In the right panel, there does indeed seem to be some sort of break in the data. But again, it is not that the touch of the euro's magic wand made trade jump up by 25%. What happened was that the Eurozone's trade with everyone fell – as did trade among other developing nations at the time – but the intra-Eurozone trade fell by less. Gomes et al (2004) show a very similar graph, adding the data for 2003, and find similar results. Flam and Nordstrom (2003) plot more detailed data, showing the Eurozone's exports to other nations (8 other OECD nations) separately from the other-8's exports to the Eurozone. The results in Figure 10 confirm the basic message but rely solely on export data.

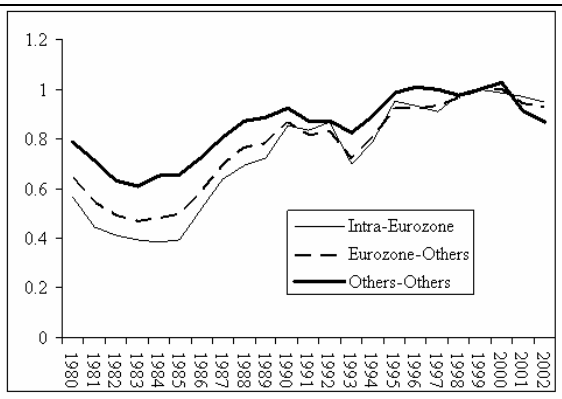
[25] This literature review draws on Gomes et al (2004).

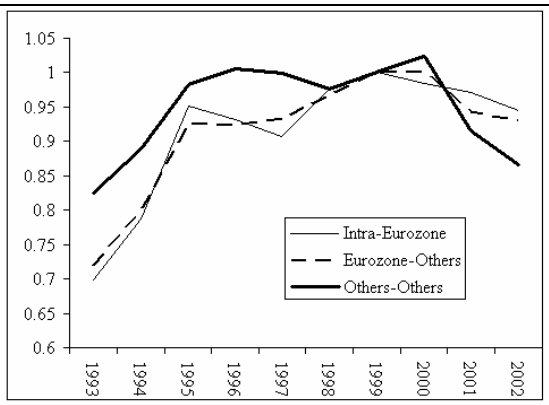
[26] Here is Andy's email: "Respected Editors, Ernesto Stein (and co-authors) at the IADB has just started to circulate a short paper which analyzes the effect of EMU on intra-EMU trade using data from the first couple of years of EMU. He shows the effect is significant (about 15-25% after just two years), using only data from the EU-15 and also a larger sample of developed countries. I'm obviously biased (though I should say that I'm trying to escape this particular sub-literature). But it's of obvious policy relevance for Europe and the ancestors if his work are in Economic Policy, so I think it's of potential interest to you. Anyway, now that I've alerted you to it, I've done my duty to God and the Queen. "

[27] See chapter 1 of Baldwin, Bertola and Seabright (2003).

[28] Full disclosure: I was the Managing Editor who did the rewriting.

Figure 10: Intra-EZ trade, EZ trade with others, and trade among others, 1980-12002.





Source: MSO (2003) with some re-calculations by me (100=1999)
Notes: The series (1997=100) show the trade evolution between classified country pairs. Specifically, for every country in the sample, we calculate a trade index with Eurozone (EZ) countries and one with non-EZ countries. The EZ-EZ series is the un-weighted average of the EZ country's EZ trade indices. The non-EZ-non-EZ series is the un-weighted average of the non-EZ country's non-EZ trade indices. The EZ-non-EZ series is the average of all 'cross group' indices. Nations in the sample: Australia, Austria, Belgium-Luxembourg, Canada, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Japan, New Zealand, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, Britain, and the United States.

The time-series are suggestive, but since bilateral trade is influenced by many things that vary over time, especially incomes, so MSO estimate a gravity model on data from 1992 to 2002 for two samples, one that is quite homogenous (only the 15 EU nations) and one with 22 developed countries that is less so.

3.2.1. **Rose effect estimates for the Eurozone, 6% more trade**

The cleanest results in the MSO paper are estimates done with pair fixed effects on the EU15 sample for the 1992 to 2002 period. Using this technique, MSO find that the Rose effect induced 6% more trade among Eurozone members.[29] This estimation strategy can be thought of as a difference-in-differences estimate. Using terminology that comes from medical studies, one group of trade pairs gets 'treated', while the control group of pairs gets a placebo; here, the treatment is Eurozone membership and the placebo is non-membership. Using the gravity model to control for observable differences between the control group and the treatment group, the estimate tells us how much bilateral trade rose in the 'treatment' group relative to the rise in the 'control' group. This is called difference in differences, since it compares the before-and-after difference for the treatment group to the before-and-after difference for the control group.

In doing this sort of exercise, it is important to get a control group that is as comparable as possible to the treatment group as far as unobservable factors are concerned (any observable factors can be controlled for via regression analysis). Given that EU membership is an extremely complex thing – one that involve literally thousands of laws, regulations and practices that affect trade within the EU and with third nations, most of which are unobservable to the econometrician since they are difficult or impossible to quantify[30] – limiting the control group to EU members is very useful. Moreover, limiting the data to post 1992 data is useful since the EU changed the way it collects trade statistics in 1993.[31]

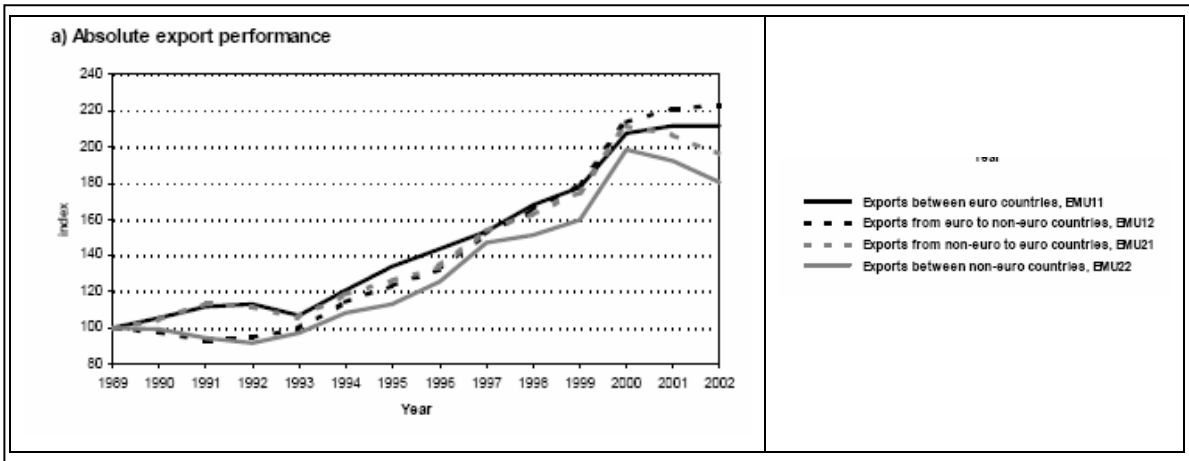
---

[29] MSO, like many authors in this literature, use EMU to stand for European monetary union; unfortunately EMU stands for Economic and Monetary Union – at least since the Maastricht Treaty that implemented EMU in both senses. Since all EU members are part of EMU, writers who are familiar with European integration use the terms Eurozone or Euroland to refer those EU members who have adopted the euro. Also EZ is shorter than EMU.

[30] Just to take one example, the EU signed dozens of preferential trade agreements during the 1992-2002 period. Since each of these erodes the preference margin of EU members, they should alter intra-EU trade flows.

[31] In fact, MSO should probably have used 1993-2002 data since the new data collection systems started with 1993 data.

Figure 11: Intra- and extra-Eurozone trade, 1989 to 2002.



Source: Flam and Nordstrom (2003).

The difference in difference estimator on the EU sample takes seriously the lessons of Persson-Kenen (use a sample where the treatment and control groups are as homogenous as possible), and the lessons of Anderson-Van Wincoop (control for omitted variables and model misspecification with dummies).[32]

Finally, the intra-EU trade data, but especially the intra-Eurozone trade data, may have some serious data problems – intra-EU imports are under-reported and intra-EU exports may be over-reported due to VAT fraud; there is some hope that MSO's averaging the two-way flows helps with this.

What does this technique not control for? Probably the main thing is difference between Eurozone and non-Eurozone members' implementation of EU-wide reform. The EU is continuously 'deepening' its integration, removing various barriers to the free movement of goods, people, capital and services. All EU members must adopt these measures, but many EU members delay – sometimes for years – and so the Single Market is not really a single market at any given moment. If the delays are systematically more important for the 'out's', i.e. non-Eurozone members than they are for the 'in's', then the Eurozone dummy may be biased upwards. In fact, the fastest implementers include all three of the out's (Britain, Denmark and Sweden) while the three laggards (Italy, Portugal and Ireland) are in's, so the MSO 6% may be an underestimate due to this point.[33]

Estimates with other control groups

MSO also perform the difference-in-difference technique using a broader sample that includes 8 extra rich nations (Iceland, Norway, Switzerland, Australia, Canada, Japan, New Zealand and the US). Trying to control for other forms of EU integration with dummies and proxies, they find the Rose effect is 8%. This result is likely to be subject to more biases than the EU sample since many omitted factors affect trade with and among these extra nations. Just to take one example, the US-Mexico free trade agreement (in the guise of NAFTA) was phased in slowly during the MSO period. Classical trade theory tells us that this preferential liberalisation should have reduced all third nation exports to the US and Canada. If NAFTA were a one-time thing, the pair dummies would control for this, but NAFTA was phased in slowly, so the trade-diversion effect is not fully controlled for. This matters since all non-Eurozone flows are used to establish a basis for what intra-Eurozone trade would have done were it not for the monetary union. Similarly, New Zealand and Austria deepened their trade ties during this period and the EU signed many free trade deals with third nations – some of which were shadowed by Iceland,

[32] The pair dummies are time-invariant and thus miss part of the Anderson-Van Wincoop point of using time-varying country dummies, but given the short period, one can hope that the omission is not too important. Moreover, someone should redo MSO's estimates with time-varying country dummies (obviously in this case one cannot also include pair dummies).

[33] Note that MSO try to control for the observable part of this, but the measure they use is extremely crude and so surely fails to fully control for this. See http://europa.eu.int/comm/internal_market/score/index_en.htm.

Norway and Switzerland, but not the others. Given this, it is easy to see why limiting the sample to the EU is a useful way to control for abundance of unobserved factors.

Biased estimates and exchange rates

Additionally, MSO estimate the Rose effect without country or pair dummies, for comparison with the early literature. They get a bigger estimate, 28%, but we know that we should ignore this for policy purposes since it is biased upwards – for the same reason most of the Rose (2000) numbers were biased, namely omitted variables and model misspecification.[34]

The authors also do the regressions including real exchange rate variables between the US dollar and the origin nation, and the US dollar and the destination nation. The inclusion of exchange rate variables is fairly rare in gravity equations. MSO justify it on the basis of a 'valuation bias'.[35] Their specification seems quite wrong to me – I'll explain in depth when discussion the Flam-Nordstrom paper – but in any case, it doesn't make much difference even though these variables turn up as highly significant. My guess is that the statistical significance of these variables arises from a correlation between their real exchange rate variables and the time residual for the relative-prices-matter term.[36]

3.2.2. **Trade diversion?**

The simplest stories behind the Rose effect are that they reduce bilateral trade costs – transaction costs are a standard suspect. If this is the case, then the euro's introduction is like a discriminatory trade liberalisation among Eurozone members and this should lead to supply switching from non-Eurozone to Eurozone suppliers. Thus, if the simple transaction cost story is correct, all other bilateral trade flows should be reduced by anything that boosts intra-Eurozone trade.

MSO look for trade diversion by including a dummy that switches on when either partner is in the Eurozone in addition to the standard currency union dummy. They call this the EMU1 dummy, but it should have been called the EZ1 dummy (see footnote 29). In any case, they find no evidence of trade diversion. Indeed, in the developed country sample, they estimate a positive impact on the Eurozone's trade with the rest of the world. On the EU sample, the point estimate is bigger, but it is not significantly different from zero. Moreover, the Rose effect jumps up somewhat from 4% to 13% in the big sample and from 6% to 9% in the EU sample. Alho (2002) confirms the basic finding of no trade diversion and a positive Rose effect.

Why does the Rose effect rise when they include a dummy for trade between Eurozone nations and third nations? Recalling the difference in difference interpretation, we know that the size of the Rose effect depends on what happened to intra-Eurozone trade compared to what would have happened without the euro. By including their EZ1 dummy, essentially take trade among non-Eurozone nations as their control group, instead of all non-intra-Eurozone trade flows. Since MSO do not properly control for free trade agreements among the third nations (e.g., they lump NAFTA, ANZCER, and EER all together in one dummy called FTA), we it difficult to know what is really happening. It would useful to redo the MSO with more attention paid to time-varying trade arrangements among third nations.

3.2.3. **Timing is everything**

MSO also study the timing of the Rose effect. What they do is interact the CU dummy with year dummies, so they can estimated the Rose effect year by year. Of course this means they are identify the Rose effect off of the cross section variation – and we know this results in point estimates that are too high due to the 'gold-medal' mistake discussed above – but it is nonetheless instructive to look at what

---

[34] More formally, MSO statistically reject the pooling hypothesis that would be necessary for the estimates without dummies to make sense.

[35] They write: "If dollar prices of goods produced in the euro zone fall as a result of depreciation, the value of trade between two EMU countries will fall as well, relative to trade between other countries, and the EMU effect on trade could potentially be underestimated. One way to deal with this issue would be to control for bilateral unit value indices in order to capture the change in import and export prices. Unfortunately, these indices are not available. For this reason, in order to control for these valuation effects we include in most regressions an index of the real exchange rate for each of the countries in the pair (the index is the ratio between the nominal exchange rate of each country vis-à-vis the US dollar and the country's GDP deflator). Reassuringly, the inclusion of these indices does not change the results significantly."

[36] Their pair dummies correct for the relative-prices-matter term on average but Anderson-VanWinccop showed us that this terms should vary over time; moreover, the relative-prices-matter term definitely includes bilateral exchange rates between the US and each importing nation so a spurious correlation is assured.

they find. The Rose effect first appears in 1998 and increases with the introduction of notes in 2001. Finally, they use dynamic panel techniques and find the short- run Rose effect is 9% to 12%, with the long-run effect ranging from 21% to 34%.

Although it is useful to see the dynamic panel technique, the actual numbers should be ignored as far as policy is concerned. We know from the Melitz-Levy-Yeyati results that the trade effects monetary union (what happened in 1999) may be very different than the trade effects of currency union (what happened in 2002). The dynamic panel technique, however, views them as the same so the deepening of integration that came with physical euros is confounded with the delayed effect of the monetary union. The dynamic panel technique won't tell us anything sensible until we have at least a few years of post-2002 data.

3.3. Berger and Nitsch (2005)

MSO (2003) and Rose (2000) are seminal work; they used the best available data and econometrics to investigate an important policy-relevant issue. Even if subsequent data shows that all their conclusions were wrong, they will remain great papers.

To many economists, the MSO paper was important in that it raised the possibility that a Rose effect was happening in Europe's monetary union. What was particularly striking was that they found an effect on after just 4 years. The fact that the size of the effect was small just made there findings more believable. MSO, however, was never going to be the final word, we can get a reasonable estimate with just four year's of trade data.

MSO became an instant target for the shrink-the-Rose-effect brigade, and the brigade's informal captain – Volker Nitsch – was in the lead. Berger and Nitsch note four crazy things in MSO's findings, or as they put it politely, things that 'invite further study':

1) The trade effect of the euro is too large relative to the trade effect of EU membership. Their findings imply that the adoption of the euro has, in 4 years, had almost the same impact as the radical liberalisation of the Single Market Programme that has been gradual phased in since 1986, with much of it completed by 1992. Anyone who follows European integration knows this is just crazy.

2) Trade among EMU members seems to jump up in 1998, i.e., one year before the euro's launch as an electronic currency. Indeed, the Rose effect estimate seems to climb gradually over MSO's data period and this suggests that maybe something other than the euro was affect intra-Eurozone trade. More to the point, it suggests that MSO may be having trouble sorting out the effects of monetary integration among Eurozone nations and non-monetary integration.

3) The size of the Rose effect is quite sensitive to disaggregation by country. MSO find that the euro had the largest trade effect for DM-bloc countries, but it is negative for Finland, Greece and Portugal (although only significantly so for Greece). It is positive but insignificantly different from zero for Finland. That's means the 'magic wand' is not working correctly for 4 of the 11 Eurozone members (Belgium and Luxembourg are treated as one).

4) When the DM bloc is dropped from the sample the Rose effect disappears. This seems strange since – again referring to medical statistics – the dosage effect is all wrong (in showing that a drug helps, medical studies try to establish that the size of the benefit is sensitive to dosage as evidence that the result is not due to unobservable characteristics of the patient). The euro was a far, far bigger policy change for Greece than it was for Germany, yet Germany seems to get a significant positive Rose effect while Greece gets a significant negative effect whose magnitude is almost twice that of Germany's.

Shifting for critique to contribution, Berger and Nitch add a fifth year of data (namely 2003), and re-estimate MSO using recently revised trade data. Interestingly, both the data revision and extra year seem to greatly increase the Rose effect. (Below, I'll argue that this is a sign that, as Marcellus said so eloquently when Hamlet slipped off for a tête-à-tête with a ghost, "Something is rotten in the state of Denmark.")

They also put the adoption of the euro in historical perspective, viewing the Eurozone as "a continuation, or culmination, of a series of policy changes that have led over the last five decades to greater economic integration among the countries that now constitute the [Eurozone]" Specifically, they use data for MSO's developed country sample of the EU15 plus 8 reaching back to 1948! Their bottom line is that throwing in a time-trend-dummy for trade among the 11 Eurozone members wipes out the Rose effect completely. There is surely something to Eurozone-as-a-continuum idea – see Mongelli Dorrucci and Augur (2005) for a more elaborate formalisation of the idea that European trade and policy integration are a dialectic process – and this surely makes it hard to separate the Rose effect from the effects of other integration initiatives. However, I think it is too blunt to throw in a time trend for the EZ11 only. European integration has affected all EU members equally. In future drafts, I hope the authors will repeat more of the MSO robustness checks with their updated data, and redo the time trend exercise, but with a trend for EU membership as a whole. It would also be interesting to see if they could develop a data-based index of extraordinarily close integration among the DM bloc, rather than the EZ11. For example, one might take estimates of bilateral pass-through elastiticites as proxies for pair-specific trade integration, the notion being that pass through would be bigger between more tightly integrated partners.[37]

3.4. Flam and Nordstrom (2003)

This is probably the best paper in the field to date. It avoids the gold, silver and bronze medal mistakes that plague the rest of the papers in this literature. Moreover, they use a data set that probably has far fewer data issues than those used by MSO and its followers, namely they use bilateral exports rather than an average of bilateral exports and imports. The use of direction-specific bilateral trade flows – measured by exports reported by the exporting nation – is what the basic gravity theory suggests should be used. Moreover, it allows them to look at an issue that concerns all the non-Eurozone nations, whether the euro puts their exporters at a disadvantage in Euroland. Additionally, they also alert the reader to the problems with European trade data collection (much more on this below). Finally, they perform their regressions on sector data as well as aggregate data.

Their basic findings on the aggregate data are in line with MSO, both in terms of size and timing. Their preferred estimate uses the 3 non-Eurozone and eight extra rich nations as the control group and using this they find the Rose effect implies about 15% higher trade; Eurozone trade with other nations (in either direction) is boosted by about half that. When they use the cleanest definition of the control group – other EU nations – the Rose effect is only 8%. Their findings on the sectoral data suggest that the Rose effect is only present in sectors marked by differentiated products, confirming the earlier results of Taglioni (2002), and Baldwin, Skudelny and Taglioni (2003).

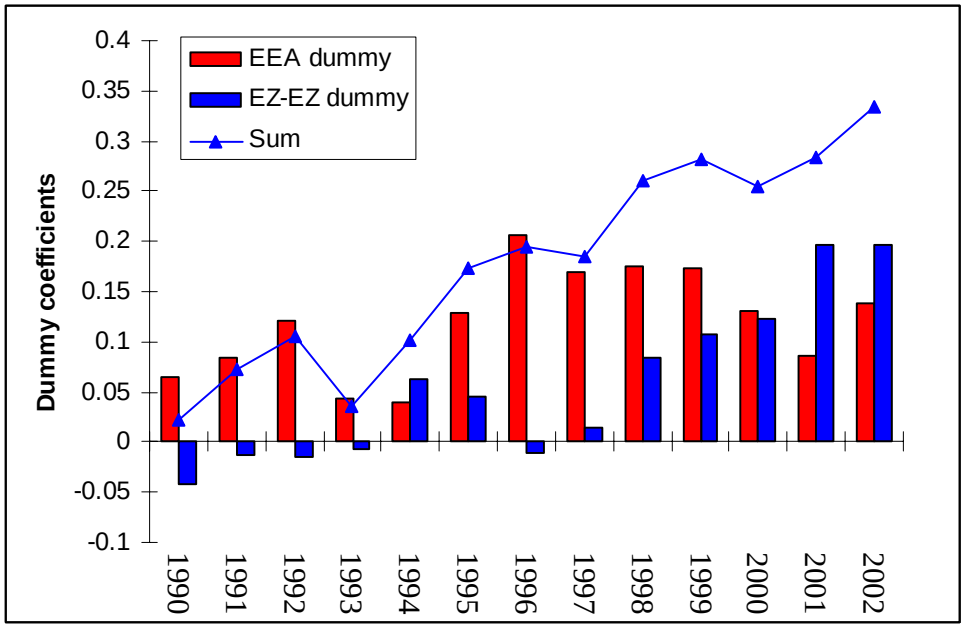
There are a few puzzling findings in Flam-Nordstrom. As in MSO, they find that the Single Market has about the same magnitude effect on trade as the euro. Also, they find that the Rose effect is larger in the broader sample of nations. When they run their preferred estimate on EU nations only, the Rose effect drops more than two standard deviations to about 8%. Moreover, the estimate for the exports of non-Eurozone nations to Eurozone nations is almost identical the intra-Eurozone effect, however, the Eurozone's exports to non-Eurozone nations seem to be unaffected by the new currency.

This finding is both intriguing and suspicious. The intriguing part is that if it is really coming from the euro's introduction, then the euro must making it easy, cheaper and/or safer to sell to Eurozone nations. Or to put it differently, the euro makes the Eurozone members extraordinarily good importers, rather than extraordinarily good exporters. It is suspicious since it suggests that it is not really the euro that is behind it all, but rather something that the Eurozone nations, or a subset of them, did around the time of the euro – something that made their markets more open to imports from all other EU nations.

[37] It is also strange that the Eurozone dummy falls when pair dummies are included. Almost everyone else finds that opposite since the omitted variables tend to be correlated with euro membership. This may be because they force the pair dummy to be the same for a half century. It would be interesting to see what happens if they allow, e.g. decadal pair dummies? My guess is that the Franco-Germany dummy will negative in early part of the sample but positive post 1958, or at least post 1968 (date of the customs union completion).

Another hint that it may not have been the euro causing the big trade effect comes from the authors' experiments with the sample. The estimate of the intra-Eurozone dummy jumps up by about one standard deviation when the sample includes Norway and Switzerland in addition to the EU14 (Flam and Nordstrom seem to exclude Greece from their aggregate regressions since they lacked Greek data for the sector regressions). Moreover, the dummy on Eurozone exports to non-EZ nations also jumps up, by two standard deviations. What could this mean?

Figure 12: Flam-Nordstrom estimates of Single Market and Eurozone dummies



Recall the difference-in-difference interpretation of the model estimated with pair dummies (a separate one for each direction-specific bilateral export). The treatment group in all cases are intra-EZ trade flows. What changes with the sample is the control group. In the EU sample, the control group is the 6 trade flows among the 3 'outs', i.e. non-EZ members, Britain, Denmark and Sweden, since Flam and Nordstrom include dummies for all trade flows between the ins and outs. Adding in Norway and Switzerland brings the total up to 20 control pairs. The fact that the EZ dummy estimate rises so much reflects the fact that trade between Norway and Switzerland and between these nations and the outs did not rise as much as trade among the outs did during this period. That in itself suggests that some strongly pro-trade factor was stimulating trade among EU members but omitted from the regression, regardless of the euro usage. Now the authors include an EEA dummy (that is the EU15 plus Norway in their sample), so this should have been controlled for, but Figure 12 shows that they may be missing something. The leftmost bar shows the year by year EZ dummy; the other bar in each pair of bars shows the estimated EU dummy (actually it is the EEA dummy even though they call it the EU dummy). What we see is that just as the Rose effect is estimated to be increasing sharply, in the 1999 to 2001 period, the effectiveness of the single market is estimated as diminishing by almost as much. The line shows the sum of the two. This also 'explains' why De Souza (2002) finds no Rose effect when he includes a time trend for EU integration, and why Berger and Nitsch (2003) are able to shrink the Rose effect to nothing by including a time trend for integration among the eventual Eurozone members.

All this, invites further study, as Berger and Nitsch would say. In particular, it would be interesting to see the Flam-Nordstrom robust procedures done on the EU sample alone, along the lines of MSO. It might also be interesting to interact an estimated EU integration trend with individual members' transposition deficits (i.e. the extent to which they are behind in implementing EU directives). It is worrying that the outs and the ins are so different when it comes to transposition in the face of rising overall integration. It would also be interesting to see the sensitivity to period with the EU sample alone.

Exchange rates and gravity. But does it work in theory?

One of the big methodological innovations in Flam and Nordstrom is their treatment of exchange rates. I believe this is extremely important and should become the norm in gravity studies, so it is worth taking some time to think through why and how exchange rates should matter. It also allows us to think more clearly about the biases remaining in Flam and Nordstrom and how we might correct them.

Figure 13: Euro against the dollar, 1999-2005.

The inclusion of exchange rates is important to control one of the potential sources of spurious Rose effects in the euro data. As the euro dropped sharply at birth, the intra-Eurozone goods came to look cheap compared to third nation goods, US goods in particular. In other words, the exchange rate altered relative prices inside the Eurozone in a way that would boost intra-Eurozone trade. If one fails to control for this properly, the coefficient on the EZ dummy will be biased by the omitted variable.

3.4.2. **The 'volume' version of the gravity equation and exchange rates**

Although the value version of the gravity equation is popular (since one does not need to find price deflators for trade flows), it is equally simple to specify a volume version. This is what Flam and Nordstrom (2004) do, implicitly (they don't include any theory in the paper). We start with the CES demand function rather than the expenditure function. Multiply the demand function for a single variety exported from nation-o to nation-d by the number of varieties produced in nation-o, n_o, the aggregate export volume is:

(13) $$X_{od} = n_o \left(\frac{p_{od}}{P_d} \right)^{-\sigma} \alpha_d \frac{E_d}{P_d}; \qquad P_d \equiv \left(\sum_{k=1}^{m} n_k \left(p_{kd} \right)^{1-\sigma} \right)^{1/(1-\sigma)}$$

where x is the quantity of the good exported from o to d, and α_d is nation-d's expenditure share on trade goods. Note that P_d here is not the GDP deflator; it is the price index for goods that compete with imports. I guess something like the importing nation's producer price index for goods would be a reasonable proxy.

Finally, we specify the exporting nation's general equilibrium condition as usual. The precise determinants of the allocation of a nation's productive factors in an open economy are extremely complex. It is called trade theory with trade cost; see Markusen and Venables (2000) and Benard, Redding and Schott (2003) for the latest evolutions. Indeed, most of the complexity in microfounding the gravity equation stems from these considerations (see Anderson 1979, Bergstand 1985, etc.). To make some headway without these complexities, I'll make a bold simplifying assumption. The number of varieties exported to nation-o is proportional to its real GDP:

(14) $$n_o = \chi_o \frac{E_o}{P_o^{GDP}}; \qquad \chi_o = f[\frac{E_o}{P_o^{GDP}}]$$

Here the price index should be nation-o's GDP deflator since we want a measure of the size of nation-o's stock of factors of production. In the simplest Helpman-Krugman trade model with no trade costs (and homothetic cost functions), n_o is exactly proportional to nation-o's supply of factors. However, in slightly more sophisticated model with trade costs, the famous Home Market Effect will be in operation so the number of varieties increases more than proportionally with nation-o's real GDP. To put that differently, χ may itself be a function of o's real GDP, so it is important to allow the exporter's real GDP to have a different coefficient in the regression.

Using (14) in the aggregated demand function (13), and assuming the 'f' in (14) is a power functinon:

(15) $$X_{od} = \left(\frac{p_{od}}{(\sum_{k=1}^{m} n_k (p_{kd})^{1-\sigma})^{1/(1-\sigma)}} \right)^{-\sigma} \alpha_d \frac{E_d}{P_d} \left(\frac{E_o}{P_o^{GDP}} \right)^{\eta_o}$$

where η_o is the exporter's variety elasticity. To get the volume of imports as a function of real GDP's and real exchange rates, we define nation-o's landed price in terms of its price-cost markup, bilateral trade costs, and marginal cost measured in nation-d's currency:

(16) $$p_{od} \equiv (\mu_{od})(\tau_{od})(e_{od}) mc_o$$

where μ is the pair specific mark-up (which may change with the exchange rate and thus result in less than perfect pass-through), τ is our usual all-inclusive bilateral trade cost and e is the nominal bilateral exchange rate that converts' nation-o's currency to nation-d's; mc_o is nation-o's marginal production cost (this too may change with the level of sales and thus dampen pass-through). Plugging this into the demand curve (15) and dividing top and bottom by P_d to get real exchange rates:

(17) $$X_{od} = \left(\frac{\tau_{od} \mu_{od} (e_{od} mc_o / P_d)}{(\sum_{k=1}^{m} n_k (\tau_{kd} (e_{kd} mc_k / P_d))^{1-\sigma})^{1/(1-\sigma)}} \right)^{-\sigma} \alpha_d \frac{E_d}{P_d} \left(\frac{E_o}{P_o^{GDP}} \right)^{\eta_o}$$

Assuming that α is a function the importing nation's per capita GDP. This can be written as:

(18)

$$X_{od} = \left(\frac{\tau_{od} \mu_{od} RER_{od}}{\left(\sum_k n_k \tau_{kd}^{1-\sigma} (RER_{kd})^{1-\sigma} \right)^{1/(1-\sigma)}} \right)^{-\sigma} \left(\frac{E_o}{P_o^{GDP} N_o} \right)^{\eta_d} \frac{E_d}{P_d} \left(\frac{E_o}{P_o^{GDP}} \right)^{\eta_o}; \quad RER_{kd} \equiv \frac{e_{kd} mc_k}{P_d}$$

Where η_d is the importing nation's income elasticity for imports. To estimate this, one would need a proxy for each exporting nation's marginal cost; its producer price index might serve well. Note that the bilateral real exchange rate and and nation-d's effective real exchange rate (i.e. the weighted sum of bilateral real exchange rates, where the right weights are approximately the importing nation's import shares).

There are a couple of important points here.

- First, the denominator is time-varying but the same for all exports to nation-d in a given year, thus a time-varying dummy for each importing nation could take care of this, thus alleviating the need to determine the appropriate weights on the RER_{kd}'s.
- Second, the price index for the importing nation in (19), i.e. P_d, is not the GDP deflator since not all goods are traded.

- Third, one rarely has perfect price deflators for bilateral trade. Flam and Nordstrom, for example, use nation-o's producer price index instead of the perfect export price index P_{od}.
- Fourth, population in most rich nations is flat over short periods, so the real GDP per capita and real GDP terms get conflated.

Thus one is estimating:

$$\textbf{(19)}\quad \frac{V_{od}}{P_o^P} = \left(\frac{\tau_{od}\,\mu_{od}\,RER_{od}}{\left(\sum_{k=1}^{m} n_k\,\tau_{kd}^{1-\sigma}\left(RER_{kd}\right)^{1-\sigma}\right)^{1/(1-\sigma)}} \right)^{-\sigma} \left(\frac{E_d}{P_d^{GDP}} \right)^{\eta_d} \left(\frac{E_o}{P_o^{GDP}} \right)^{\eta_o} \left\{ \frac{1}{N_o} \frac{P_{od}}{P_o^P} \frac{P_d^{GDP}}{P_d} \right\}$$

but omitting the terms in curly brackets.

Average bilateral trade instead of bilateral exports

If one adds the logs of exports from o to d together with the exports for d to o and divides by two – as MSO and most papers in the literature do – the bilateral real exchange rate term drops out, but then one must included the effective real exchange rate for both nations, or correct with time-varying country dummies. Note that MSO includes the silver-medal mistake and it would be useful to see what MSO's estimation techniques yield when the log of the sums and the sums of the logs are interchanged.

3.4.3. **Potential econometric problems in Flam and Nordstrom**

Note that the omitted terms in (19) are time-varying. The pair-specific dummies that Flam and Nordstrom use control for this residual on average, but leaves a time-varying residual that will be correlated with all time varying trade costs measures – including the EZ dummy (by construction, Pd includes τ_{od}). In short time samples, this problem may not be too bad, but in longer samples the bias may be important. Evidence for this can be found in Flam-Nordstrom sensitivity checks. They find that the size of the Rose effect increases with the length of the sample (a finding confirmed by others including Berger and Nitsch 2005). Of course, one can directly test for serial correlation in the errors and one should. It is particular important to check this for the Eurozone trade flows.

3.5. Other studies

3.5.1. **Barr, Breedon and Miles (2003)**

Barr, Breedon and Miles (2003) is another paper that was presented at the same Economic Policy Panel as MSO was. Estimating the Rose effect was not the central axis of investigation in this paper; they try to systematically compare EU members inside and outside the Eurozone. Their early drafts study a wide vary of issues, but comments from referees, Managing Editors and – above all the Panellists – led them to pare down the paper to a focus on trade. For comparative purposes, they also make preliminary estimates of the effect of monetary union on three other dimensions of economic performance: foreign direct investment, the development of financial markets and overall macroeconomic performance, though they recognise that their ability to control for other factors is more limited for these other indicators.[38] These authors make a much more thorough attempt than MSO to correct for reverse causality. Even after this correction, however, they that the Rose effect is positive and significant.

3.5.2. **Bun and Klaassen (2002)**

One of the other papers in line for the I-did-it-first prize is Bun and Klaassen (2002), since their first draft was circulated in 2001. This paper employs a dynamic fixed effects estimator. The results they obtain are quite similar to those reported for MSO using a similar estimator.

---

[38] This idea was to figure out how much of Britain's superior macro performance was due to their decision to stay out of the Eurozone, but with so few data points this proved elusive.

3.5.3. **De Souza (2002), Piscitelli (2003) and De Nardis and Vicarelli (2003)**

De Souza (2002) estimates the basic gravity model for the EU15 countries with the addition of a time trend. He finds no evidence for a significant Rose effect unless he removes the trend. This result is interesting, maybe even important, but throwing linear terms provides can do lots of things to a regression that gets most of its traction from the time-variation of the policy variable of interest. MSO's experiments with time trends (in the first draft) and a direct measure of EU integration in the published version do not line up with De Souza's findings.

Piscitelli (2003), following the 2001 draft of MSO, finds that lengthening the sample back to 1980 reduces the Rose effect estimates. The paper also find that the size of the Rose effect changes with the data used. OECD trade data uses the "cost, insurance and freight" (cif) methodology while the IMF trade data used in MSO (2003) takes the "free on board" (fob) approach. I'll have much more to say about this result, but I note here that fob is what you get when you rely on the exporter's data and cif when you rely on the importer's data (for most nations, the UN's ComTrade data base – the fount of all trade data – has four observations on bilateral trade, e.g. France to Germany as reported by the French and Germans, and Germany to France again by both nations; MSO average all of these to get their one estimate of bilateral trade).

De Nardis and Vicarelli (2003) was also one of the early papers on this. (One of the reasons Economic Policy decided to commission MSO was that others had found similar results.) They take a different tack at controlling for reverse causality, but get about the same answer as MSO; 10% in the short run and 20% in the long run.

3.5.4. **Anderton, Baltagi, Skudelny and De Sousa (2002)**

This paper uses more sophisticated econometrics -- three-stage least squares – to estimate import demand functions. They find no direct evidence of a Rose effect, but given the relatively small number of post-euro observations, it is hard to know what to make of this. Given the good data on Europe, however, this more direct approach to estimating the euro's impact should be tried again with the longer time series we have now.

4. COLLECTION OF CLUES[39]

I believe that we can be fairly sure that some form of Rose effect is occurring in the Eurozone. The cleanest test by a long shot is the Flam and Nordstrom (2003) estimate using only EU members on data from 1989 to 2002. Since they put in pair dummies using direction-specific exports, they have controlled for all time-invariant idiosyncratic relationships among the EU15, and reduced the risk of biases from the underreporting of imports. Because the time period is relatively short, the serial correlation that we know must be in their residuals should not pose too much of a problem in terms of biasing the point estimate of the Rose effect. And most importantly, because they only use EU members that have not joined the Eurozone, they have controlled for most of the bias that might emerge from unobserved pro- or anti-trade policies adopted by the EU in tandem with the euro's introduction. It would be useful to see a few more sensitive tests, but this results, combined with similar findings by MSO, Berger-Nitsch and many others, leads me to believe that the Rose effect is for real in Euroland.

If I had to provide 'the' number, I would – after plenty of provisos about the Rose effect not being a magic wand – say the number is between 5% and 10% to date. Most of the evidence suggests that this number may grow as time passes, maybe even doubling.

This section attempts to draw critical clues from the empirical literature, that is to say, to stylise the facts in a way that helps us think about the causes of the Rose effect. I organise the clues into spatial clues, timing clues and sector clues.

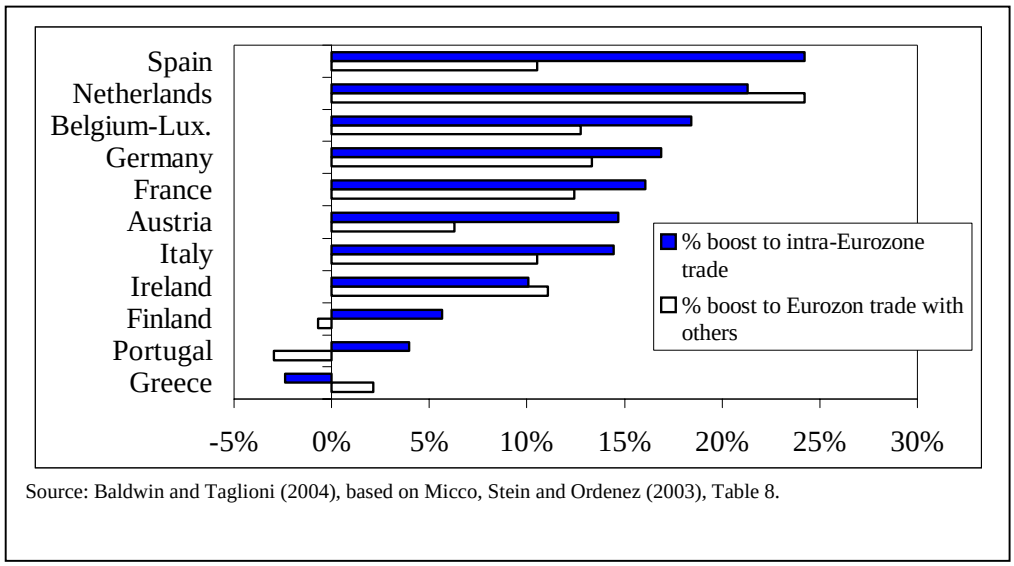
[39] This section draws heavily on Baldwin and Taglioni (2004).

4.1. Spatial variation of the Eurozone Rose effect

MSO (2003) and Flam and Nordstrom (2003) report extensive robustness checks. Deep within the MSO paper are nation-by-nation estimates the Rose effect for each Eurozone nations. Figure 14 (taken from Baldwin and Taglioni 2004) converts MSO's raw coefficients into percent increases in trade and plotted the results by nation. The nations are ordered by decreasing Rose effect. Three features are particularly relevant.[40]

1) Apart from Spain, the nations with the highest Rose effects are those that are already the most tightly integrated: the Benelux nations and Germany. Two aspects of this group may be relevant in our search for positive OCA criteria.

Figure 14: euro's trade effect by nation



Source: Baldwin and Taglioni (2004), based on Micco, Stein and Ordenez (2003), Table 8.

These nations have been in an informal, but very tight exchange rate arrangement called the DM-bloc for decades. As Figure 15 (from Anderton, Baldwin and Taglioni 2003) shows, intra-DM bloc volatility was very low, so the euro had only a very small impact on the bilateral exchange rate variability among these nations. This is a bit puzzling since one might have thought that the trade effects would have been largest among nations that had the largest, pre-euro bilateral volatility.

2) These nations are geographically proximate, so we suppose that the natural trade costs among these nations are quite low; gravity model estimates in Europe suggest that each doubling of the distance between capitals lowers trade by 70%. Moreover, these nations are among the most avid integrationists in the EU and thus have embraced the EU's deep trade integration even more tightly than other members.

For example, the Benelux nations formed a customs union even before the EU was founded in 1958, and Belgium and Luxembourg have shared a common currency since just after the war. As part of this distance-Rose-effect nexus, we note that the size of the euro's trade impact is lowest in the geographically peripheral Euroland nations: Greece, Portugal, Finland and Ireland. Again this suggests a negative relationship between trade costs and the Rose effect.

3) Berger and Nitsch (2003) point out that estimates of the Rose effect on an EU sample that excludes the DM bloc turn out to be insignificant. In other words, the effect is not just strong in these countries the aggregate numbers like 5% to 10% are driven by these nations.

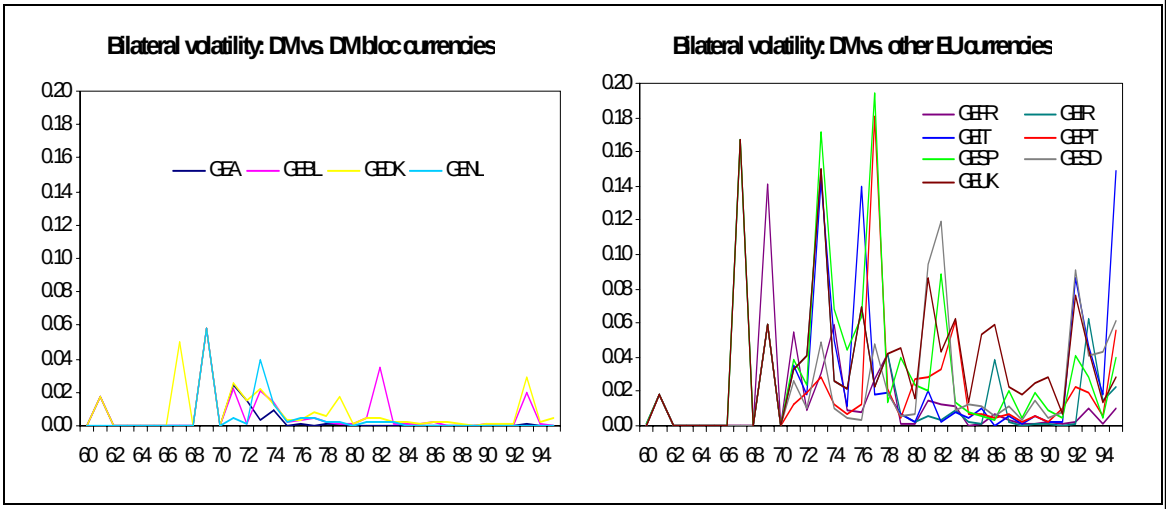
[40] The numbers for Greece, Portugal and Finland are not significantly different that zero, except Greece's EZ1 estimate which is significant at the 5% level of confidence.

The fact can be read in two ways. Pessimistically it says that it was not the euro, but some unobserved policy adopted by DM bloc nations that is driving the results. But what could it be; general product and labour market reforms that Britain, Denmark and Sweden had already undertaken? Optimistically, it could be that exactly because these nations had such low exchange rate volatility for so long, their firms were in a good position to profit from the removal of small costs. If this is right, we should see the Rose effect appearing in the non-DM Eurozone members, but more slowly. Glick and Rose (2001) present some evidence that the currency-trade link can take 30 years to fully work its magic.

4.1.1. **Trade with non-Eurozone nations**

Intriguingly, MSO (2003) find that trade between Eurozone nations and other nations rose with the euro's introduction, but not quite as much. Specifically, they estimate what might be called a one-sided euro dummy; its value is unity for any trading pair that involves only one Eurozone member (the regular euro dummy, or two-sided dummy, is one only for trading pairs where both nations are in the Eurozone). The results, again translated into percent increase in trade, are shown as the light bars in Figure 14. Roughly speaking, the one-sided impact is lower than the two-sided effect, but the nations with large two-sided effects also seem to have large one-sided effects.

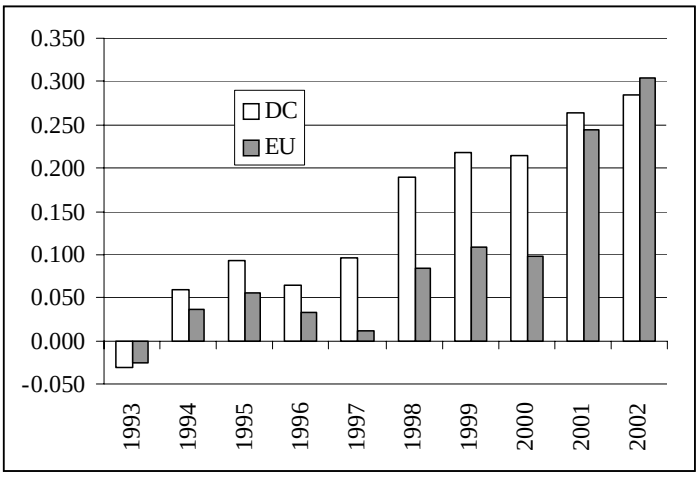
Figure 15: DM bloc exchange rate volatility, 1960-1994



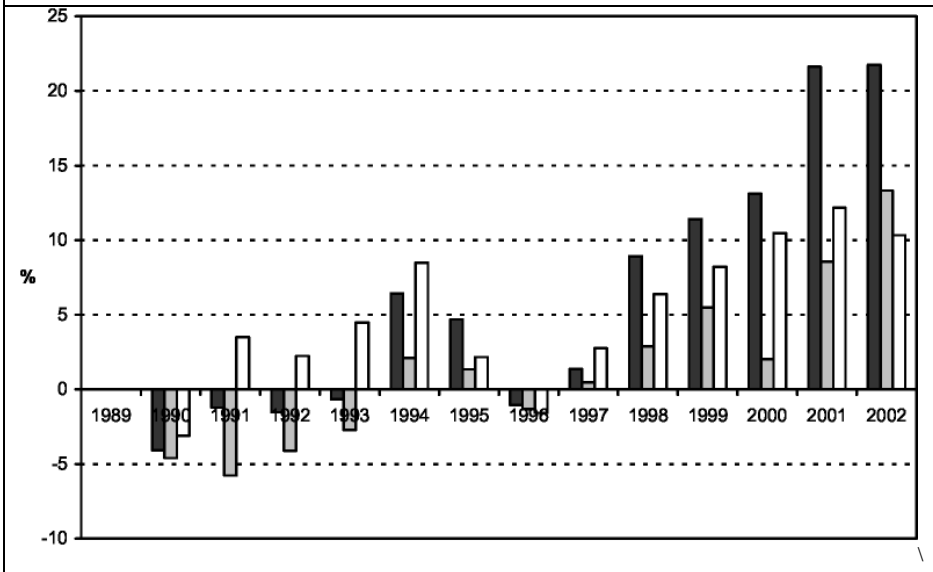
Source: Anderton, Baldwin and Taglioni (2003).

This result is intriguing. It provides a very significant hint as to the microeconomics of the Rose effect, or at least as to what it is not. Informal discussion of the trade effects of a monetary union typically refer to 'transaction costs' of having different currencies. In standard trade policy terminology, having a common currency is like reducing bilateral, non-tariff barriers. The evidence on the one-sided dummy tends to reject this view. If one could model the trade-reducing effects of volatility as a frictional trade barrier, the one-sided dummy should have been negative. The euro would have been akin to a discriminatory liberalisation and this should have reduced the exports of non-euro nations to Euroland.

Figure 16: euro's trade effect over time



Source: Micco, Stein and Ordenez (2003).



Notes: Top panel: The first bar is intra-Eurozone, the second is Eurozone exports to others, the third is Eurozone imports from others. Bottom panel: Source: Flam and Nordstrom (2004).

Flam and Nordstrom (2004) refine this clue by estimating direction-specific trade flows. In their cleanest regression – the one that only includes EU members – they find that EZ members have higher than expected imports from non-EZ members, but not higher exports. Indeed, the rise in exports from non-EZ members is statistically identical to the rise exports between EZ members. If one averaged the EZ imports with non-EZ members and EZ-exports to non-members, as MSO do, then it would seem that having one half of a trade pair inside the Eurozone increased trade by one half the amount that have both partners inside the Eurozone.

This is a powerful clue, if it is true. It suggests that the euro has acted more like a unilateral trade liberalisation than a preferential trade liberalisation.

If it is true, it also has some very important implications for the politics of Eurozone enlargement. I'll have a lot more to say about this below because it reverses some of the underpinnings of OCA theory. In basic OCA theory, you have to give up your monetary autonomy to get the benefits of reduced

transaction costs. If this result is right, it suggests that Britain, Denmark and Sweden were the clever ones from a mercantilist perspective – they got the better market access without sacrificing their main marco-policy tool.

4.2. Timing of the Eurozone Rose effect

Monetary union in Europe was never a sure thing until it actually happened. Although the treaty that laid out the path to the euro was signed in 1992, the Treaty had several major difficulties in becoming law. Moreover, the treaty laid down a series of conditions – the famous Maastricht conditions – for membership in the monetary union, and most European nations had trouble meeting these. Right up to the announcement of the names of the inaugural members in March 1998, sceptics doubted that the monetary union would ever become a reality.

4.2.1. **The effect appears in 1998**

Given this, the speed with which the euro's trade impact appeared is striking. Evidence for this comes from the MSO and Flam-Nordstrom estimate. The results are illustrated in Figure 16, which shows the estimated year-by-year dummies for intra-Eurozone trade; the dark bars show the estimates for the sample that includes only EU nations and the light bars show the estimates for the sample that includes all industrialised nations The main points are that the Rose effect jumps up and becomes statistically significant in 1998, the year before the monetary union was formed. It jumps up again in 2001, especially for the EU sample, the year before the monetary union became a currency union.

The rapid reaction of trade flows is quite remarkable since the MSO and Flam-Nordstrom controls for the main determinants of bilateral trade. The speed also provides us with an important hint as to what is *not* going on here. Such a rapid increase in trade would be very hard to explain if, for example, it was driven by the construction of new plants related to the unwinding of hedging-related foreign direct investment. The Flam-Nordstrom results show a longer time horizon.

4.2.2. **Sensitivity to the period of estimation**

Another clue relates to the nexus between the size of the estimated Rose effect and the sample period. The original draft MSO sent to Economic Policy six months before they presented the paper in Athens used two data sample, one from 1980 to 2002 and one from 1992 to 2002. Although the regressions have some serious problems that may vitiate the results, the early MSO seems to find that the Rose effect is bigger when a longer data set is used. Berger and Nitsch (2003) confirm this result when they extend the period back to 1948 and push it forward one more year to include 2003.

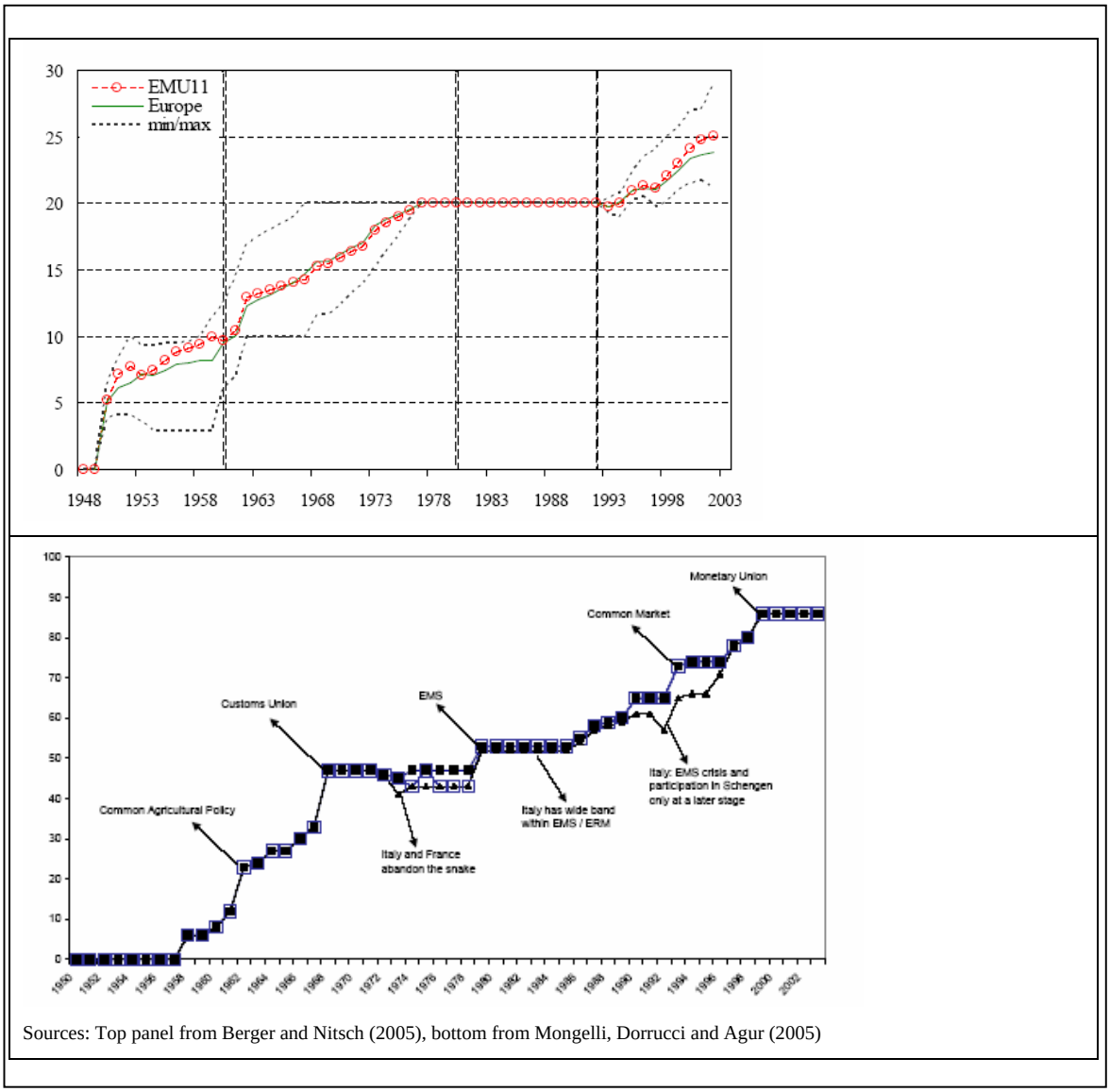
The finding that the Rose effect was bigger for sample starting further back is highly suspicious. It suggests that the various dummies for EU integration are not really removing everything-but-the-euro from the data. As Figure 17 shows, European economic integration has been an ongoing process for the last 50 years. In particular, economic integration was rising steeply just before the introduction of the euro. If pro-trade adjustments to pre-Eurozone integration take time, it could very well be that the lagged effects of Single Market measures are showing up in the post-1999 data and being confused with the trade effects of the euro.

At the heart of this suspicious is a misspecification of the lags and the role of European integration. In principle this is testable and correctable with proper econometrics. For example, Flam and Nordstrom (2003) allow for a time-varying EU dummy and they find that their Rose effect estimate is affected very little by a change in their sample period. In any case, the diagram showing the Flam-Nordstrom results for the EZ dummy and EU dummy suggest that it may be quite difficult to tease apart the effects of general European integration and the euro per se. Indeed, Berger and Nitsch (2005) go so far as to argue that since a time trend for integration among the Eurozone members wipes out the Rose effect, the MSO estimates are due to a misspecification.

4.2.3. **Sensitivity to data revisions**

Berger and Nitsch (2005) re-run the MSO regressions using updated data and adding an extra year. Surprisingly, running the same regression as MSO on the same years with the same nations, the Rose effect jumps up with the updated trade data. This, I believe, provides a clue as to the importance of data problems arising from VAT fraud. I'll discuss this more below, but the basic point is that VAT-cheats are systematically under-reporting imports. The data revisions are problem raising the level of intra-EU imports to more closely match the intra-EU export data. Since the MSO procedure, followed by Berger-Nitsch, involves averaging import and export data, such revisions would naturally boost the Rose effect.

Figure 17: Indices of European integration over time.



Sources: Top panel from Berger and Nitsch (2005), bottom from Mongelli, Dorrucci and Agur (2005)

4.3. Sectoral variable in the Eurozone Rose effect

While most studies of the euro's impact have focused on aggregate trade data, Taglioni (2002) and Baldwin, Skudelny and Taglioni (2003) run the standard gravity model using sectoral data. In addition to confirming the general findings of the aggregate studies when all the sectors are pooled, this paper

also provides sector-specific estimates of the Rose effect. The results are shown in Table 4. Flam and Nordstrom (2003) find a similar pattern.

Table 4: Rose effect and volatility impact by sector.

isic	industry	Rose effect	t-stat	Volatility	t-stat
40-41	electricity, gas and water supply	1.64	4.47	-15.78	-1.87
351	……building and repairing of ships and boats	0.57	2.00	-15.87	-2.42
15-16	food products, beverages and tobacco	0.40	2.64	-7.78	-2.23
25	….rubber and plastics products	0.35	2.25	-10.73	-3.04
35	….other transport equipment	0.34	1.84	-17.72	-4.23
30	……office, accounting and computing machinery	0.32	1.91	-5.77	-1.50
34	….motor vehicles, trailers and semi-trailers	0.31	1.81	-13.78	-3.53
32	……radio, television and communication equipment	0.27	1.68	-14.06	-3.74
36-37	manufacturing nec; recycling	0.27	1.76	-6.25	-1.76
353	……aircraft and spacecraft	0.27	1.09	-16.89	-2.98
33	……medical, precision and optical instruments	0.27	1.76	-7.75	-2.22
31	……electrical machinery and apparatus, nec	0.26	1.64	-14.13	-3.94
28	….fabricated metal products	0.25	1.66	-9.78	-2.85
17-19	textiles, textile products, leather and footwear	0.25	1.54	-12.00	-3.25
24	….chemicals and chemical products	0.25	1.52	-8.80	-2.38
20	wood and products of wood and cork	0.23	1.41	-7.78	-2.08
29	….machinery and equipment, n.e.c.	0.23	1.44	-9.29	-2.54
27	….basic metals	0.19	1.16	-14.23	-3.70
26	other non-metallic mineral products	0.19	1.24	-10.29	-2.91
271+2731	……iron and steel	0.14	0.74	-13.25	-3.08
2423	……pharmaceuticals	0.13	0.70	-8.04	-1.90
272+2732	……non-ferrous metals	0.12	0.63	-20.52	-4.72
01-05	agriculture, hunting, forestry and fishing	0.09	0.50	-7.59	-1.91
23	….coke, refined petroleum products and nuclear fuel	0.03	0.12	-7.83	-1.33
352+359	……railroad equipment and transport equipment n.e.c.	-0.05	-0.23	-14.09	-2.96
10-14	mining and quarrying	-0.21	-1.15	-9.84	-2.37

Source: Adapted from Baldwin, Skudelny and Taglioni (2003).

What these results show is a rough correlation between the size of the Rose effect and what we loosely call ICIR sectors (imperfect competition and increasing return sectors). At the bottom of the list, we have agriculture as well as mining and quarrying, while near the top, we have various types of machinery and highly differentiated consumer goods such as food products, beverages and tobacco. This finding opens the door to the possibility that ICIR like effects – for example, the impact of uncertainty on market structure – may be part of the story.

The Flam-Nordstrom paper also provides sector results, which are reproduced in Table 5. These are broadly in line with the earlier estimates in Table 4. The sectors without a Rose effect tend to be those marked by fairly homogenous products. Recall that trade inside Europe in agricultural goods is not free trade. Although there are no formal barriers, market intervention is pervasive.

Table 5: Flam-Nordstrom sectoral Rose effects.

Table 8. Results for sectors (one-digit SITC rev. 3), 1995-2002

	SITC 1-9 Aggregate export	SITC 0 Food & live animals	SITC 1 Beverages & tobacco	SITC 2 Crude materials, inedible, except fuels	SITC 3 Mineral fuels, lubricants & related material	SITC 4 Animal & vegetable oils, fats & waxes	SITC 5 Chemicals & related products, n.e.s.	SITC 6 Manufactured goods classified chiefly by materials	SITC 7 Machinery & transport equipment	SITC 8 Miscellaneous manufactured articles
EMU11	0.172***	0.014	0.352***	-0.033	-0.196	0.044	0.069*	0.124***	0.224***	0.071***
	(0.021)	(0.041)	(0.086)	(0.054)	(0.198)	(0.152)	(0.038)	(0.034)	(0.037)	(0.027)
EMU12	0.089***	0.047	0.129*	-0.063	-0.096	0.186	0.078**	0.002	0.087**	-0.002
	(0.018)	(0.037)	(0.072)	(0.052)	(0.172)	(0.125)	(0.033)	(0.032)	(0.035)	(0.023)
EMU21	0.089***	-0.088**	0.161*	-0.115***	0.075	0.139	0.059	0.088**	0.120***	0.009
	(0.019)	(0.039)	(0.087)	(0.044)	(0.167)	(0.133)	(0.036)	(0.036)	(0.036)	(0.025)
ln(RYi)	0.864***	-0.990***	-0.639**	0.928***	-0.652	1.926**	1.955***	-0.661***	1.142***	0.160
	(0.125)	(0.243)	(0.303)	(0.328)	(0.910)	(0.907)	(0.205)	(0.167)	(0.131)	(0.159)
ln(RYj)	1.077***	0.868***	1.370***	0.574**	0.580	0.645	1.105***	0.735***	1.044***	0.608***
	(0.102)	(0.225)	(0.442)	(0.248)	(0.722)	(0.547)	(0.164)	(0.126)	(0.160)	(0.144)
ln(REXRij)	-0.811***	-1.610***	-1.695***	-1.268***	-2.459***	-1.934***	-1.180***	-1.262***	-0.805***	-1.302***
	(0.056)	(0.115)	(0.224)	(0.131)	(0.371)	(0.332)	(0.095)	(0.091)	(0.105)	(0.072)
ln(REXRcj)	0.380***	1.341***	1.373***	0.821***	1.608***	1.979***	0.886***	0.912***	0.178	0.887***
	(0.076)	(0.167)	(0.322)	(0.206)	(0.560)	(0.504)	(0.143)	(0.125)	(0.154)	(0.110)
nomexr	0.373	4.622***	6.232**	-2.371	-2.005	0.804	-2.754**	2.279**	0.026	4.655***
	(0.633)	(1.327)	(2.763)	(1.615)	(4.926)	(3.977)	(1.198)	(1.059)	(1.112)	(0.867)
obs	2704	2702	2659	2702	2539	2460	2703	2704	2704	2704
R2	0.994	0.985	0.961	0.975	0.914	0.917	0.987	0.991	0.988	0.993

Robust standard errors in parentheses.
Significant at 10%; ** significant at 5%; *** significant at 1%.
Year, EU membership, "Rotterdam effect" and bilateral (fixed effects) dummies included but not reported.

5. WHAT COULD IT BE?

"This is indeed a mystery. What do you imagine that it means?" Watson remarked. "I have no data yet. It is a capital mistake to theorize before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts." ~ Sherlock Holmes, A Scandal in Bohemia

In this section I engage in a little Sherlock-ing about what might be going on. I'll entertain two broad sets of hypotheses.

1. The whole thing is a down to spurious results.
2. Something real has changed trade flows.

In the next section, we'll suggest a battery of diagnostic tests that should help us work out the true answer, or answers.

5.1. Spurious results

While the size of the estimated Eurozone Rose effects does not strain credibility the way the bigger estimates in the earlier literature did, many questions remain. The speed with which the effect appears is suspicious, as is the fact that it appeared before 1999. The lack of a trade diversion effect also raises questions, especially given that the cleanest estimates in the literature suggest that the ins and the outs gained equally in terms of selling more to Eurozone nations.

It is always exciting to find something new and real, but we must also consider the possibility that the effect is spurious. I believe there to be three hypotheses that need to be eliminated before we can be sure that the estimates are telling us anything real about the economy.

- The measured trade does not reflect true flows; VAT fraud creates a spurious Rose effect.

- Delayed effects of the euro depreciation create a spurious Rose effect.
- Effects of Eurozone implementation of Internal Market measures create a spurious Rose effect.

We discuss these in turn.

5.1.1. **Lies, damned lies and statistics**

This section consider that very real possibility that the Rose effect in the Eurozone is a statistical illusion stemming from the way trade figures are gathered. This is headache material, but I believe it is necessary since existing estimates of the effects are of the same order of magnitude as the estimated Rose effects in the Eurozone.

VAT fraud

One of the great coups of the 1986 Single European Act was to remove Europe's internal borders, at least as far as trade is concerned. This happened in 1993 and changed the way trade statistics were gathered on intra-EU trade. Data on intra-EU trade from 1993 onwards was collected from VAT authorities rather than customs.

The problem is that this creates a direct link between trade date and tax avoidance and evasion. Worse still, tax enforcement changes – and anticipation of the same – can create reactions that distort EU trade flows. These distortions can vary across time, across trade pairs and sectors.

Why would VAT authorities produce trade statistics? EU nations have VAT systems that are based on the so-called destination principle, i.e. a good pays the VAT rate of the nation where it is sold, not where it is made.[41] Practically speaking, this means that the exporting EU nation rebates its VAT to the exporting firm and the importing EU member imposes its own VAT rate on the importing firm. This is why VAT authorities have always kept track of imports and exports.

Although the VAT system was massively reformed in anticipation of the suppression of border controls – a major part of this being a narrowing of differences in VAT rates – the 1993 system was susceptible to fraud.

Box 2: Acquisition and Carousel and fraud

The easiest fraud is called 'acquisition fraud'. Criminals set up a company in, say, the UK and import goods from Germany at a zero-VAT price (the selling company gets the Germany VAT rebated). The importing firm sells the goods in Britain at something like the with-VAT price (since that is what honest importers have to charge) but they never pay the VAT; they go out of business before the VAT authorities can get them.

The Carousel fraud takes this one step further, but understanding this requires some background on why VAT is usually impervious to fraud. An EU firm that sells a good is liable for the VAT on the full sales price, unless they can prove that they bought inputs to make the good, in which case, they only pay the VAT on their value added. The point is that the VAT has already been paid on the inputs. How do we know the VAT's been paid on inputs? Well, whether it was made locally or imported, the local VAT rate was paid. How do we know that the firm won't exaggerate its purchases of inputs? Counteracting incentives is the answer. The input seller would like to under-report its sales to reduce its VAT bill, but the input buyer would like to over-report its purchases to reduce its VAT bill. The size of the gain and loss are identical so there is no reason to suspect an upward or downward bias. In short, buyers and sellers become informants on each other as far as VAT payments are concerned. Now, how does this work with exports? Suppose the export shipment is to Belgium and is worth 100,000 pounds. If the UK VAT rate is 20%, the British VAT authority pays the exporting firm 20,000 pounds – which is the amount of VAT that has been paid on the good in Britain.

But what if the VAT was never paid on the import into Britain because of an Acquisition fraud? In this case, the criminals pocket 20,000 pounds having paid very little, or maybe no VAT, in Britain. Since it

[41] This is a huge literature. See the Keen and Smith (1996) paper in Economic Policy for a accessible description – by the standards of the literature – of the basic issues.

worked once, they may be tempted to put the same goods through the same cycle again. The goods turn around and around like a carousel, each time showing up twice as an export and once or never as in an import.

This box draws heavily on Ruffles et al (2003).

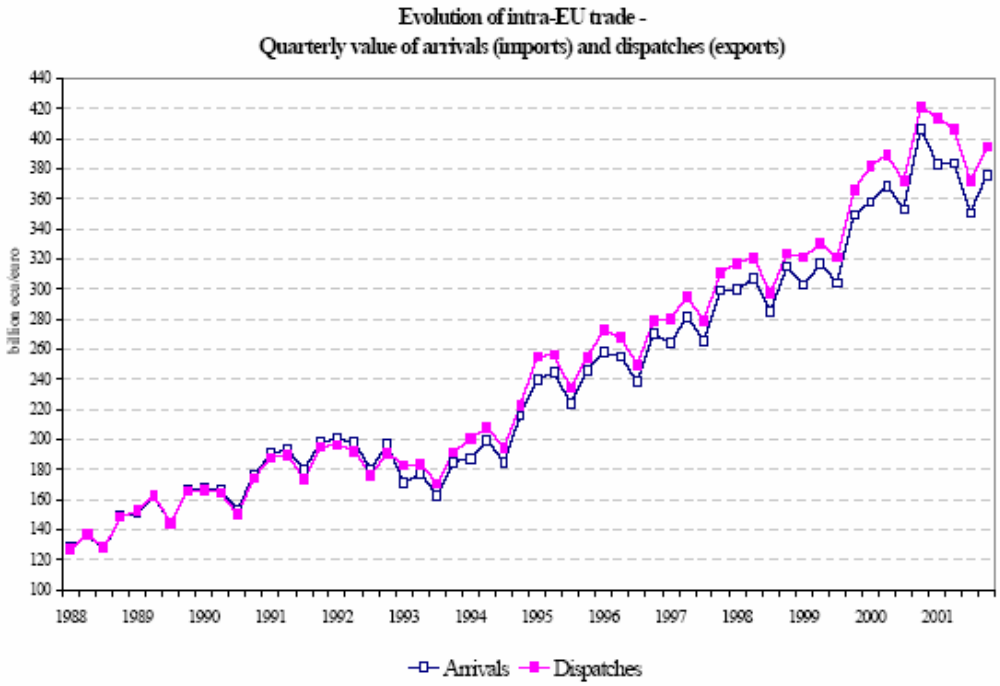
Fraud and intra-EU trade figures

As early as the mid-1990s, problems of VAT fraud were recognised, for example, by the European Parliament's first Temporary Enquiry Committee in 1997. Measures were taken to improve the system, but things were still problematic when the euro was launched. A European Commission report to the Council and European Parliament in 2000 used usually blunt language:

> "The transitional VAT arrangements have been in place for more than 6 years. During this period, one would have expected that the implementing problems should have been solved and that the system should be running smoothly. However, this does not appear to be the case. The 6 years appear to have given the fraudsters time to appreciate the possibilities offered by the transitional VAT arrangements to make money, while, generally speaking, Member States have not met the challenge posed by fraud. …There are indications that the level of serious fraud in intra-Community trade is growing."

The exact nature of such fraud is not easy to ascertain. Typically, it creates a gap between export statistics (every exporter wants the VAT rebate) and import statistics (some have an incentive to avoid paying the importing nation's VAT). But, it can also inflate the trade statistics as in the case of the so-called 'carousel fraud', see Box 2.

Figure 18: Difference between intra-EU exports and imports.



Source: Eurostat, June 2002

The effect is huge and anti-fraud activity differs across time and country pairs.

The effect of this fraud is so large that the UK has had to restate its national accounts (see Ruffles et al 2003). The revisions involve upward adjustments to imports of £1.7 billion in 1999, £2.8 billion in

2000, £7.1 billion in 2001 and £11.1 billion in 2002. Unadjusted imports in 2002 were £220 billion, so the effect is about 5%. The problem is not limited to the UK as Figure 18 show. The problem is large, on the order of 5%, and it varies over time. As inspection of the figure shows, the problem appears to increase substantially in the run up to the euro's introduction.

Policy reaction is correlated with the timing of the euro's introduction. The European Commission has coordinating the implementation of anti-fraud activities since the 1990s, and these are still not complete. The rate of implementation varies across nations and across types of trade. I presume the criminals involved in this fraud follow the process carefully, so it is entirely possible that they alter their behaviour in anticipation of changes.

If this fraud were simple – e.g. it consisted entirely of Acquisition Fraud, simple fixes might work. For example, researchers could use export data. But as the Carousel Fraud suggests, even these data may be exaggerated in ways that vary across time, country pairs and commodities.

The Rotterdam effect

A huge fraction of Germany's imports from nonEU nations arrive via Rotterdam. Some of this trade is record as a Dutch import from, say, New Zealand and then recorded as a German import from the Netherlands. Some of it, however, is recorded as a German import from New Zealand since it is subject to a 'transit' regime by which the tariff and VAT are not paid until the good arrives in Germany. One such system is called the TIR system ('Transports Internationaux Routiers') and it involves transferring sealed containers from ships to trucks (Rotterdam and Antwerp are the big centre for this, but it happens in other ports as well).

The removal of fiscal border checks within the EU, teamed with the rapid rise in the volume of trade, made fraud big problem in the EU. Anti-fraud measures lead to a reaction that resulted in a rise in the amount of third nation goods being declared twice, once and an import into Holland and once as an export from Holland to the true destination. There are a million stories, but here is one that gives an inkling of the problem: As a European Commission pamphlet on the EU's transit system tells it:

> "In the early 1990s, the TIR system also began to experience a significant increase in fraud leading to large losses of duties and charges. Much of the fraud concerned tobacco and alcohol, both subject to high rates of duties and charges. In those cases the USD 50 000 limit of the guarantee was often inadequate to meet claims made by customs. A special 'tobacco/alcohol' guarantee of USD 200 000 was therefore introduced on 1 January 1994. The situation was so bad that, with effect from 30 November 1994, the central pool of insurers were forced to withdraw their insurance cover for all guarantees for tobacco and alcohol. This meant it was no longer possible to move tobacco and alcohol under TIR. Furthermore, with effect from 1 April 1996, the national associations of some Community Member States withdrew their TIR guarantees for those sensitive goods that were banned from using the comprehensive guarantee in Community transit, for example beef, milk, cream and butter. As a result it is impossible for these goods to move under TIR into or out of the Community."

Of course, if these goods don't use the TIR, the intra-EU trade rises relative to the extra EU trade. The point is that before and after the good is counted as being imported once from the third nation, but in afterwards it is also counted as trade between Holland and Germany.

As with the VAT fraud, one could image simple fixes if the problem were simple, but unfortunately the magnitude of the problem varies over time and by member state. The worst part is that attempts to deal with these problems are phased in tandem with the euro as Figure 19 shows.

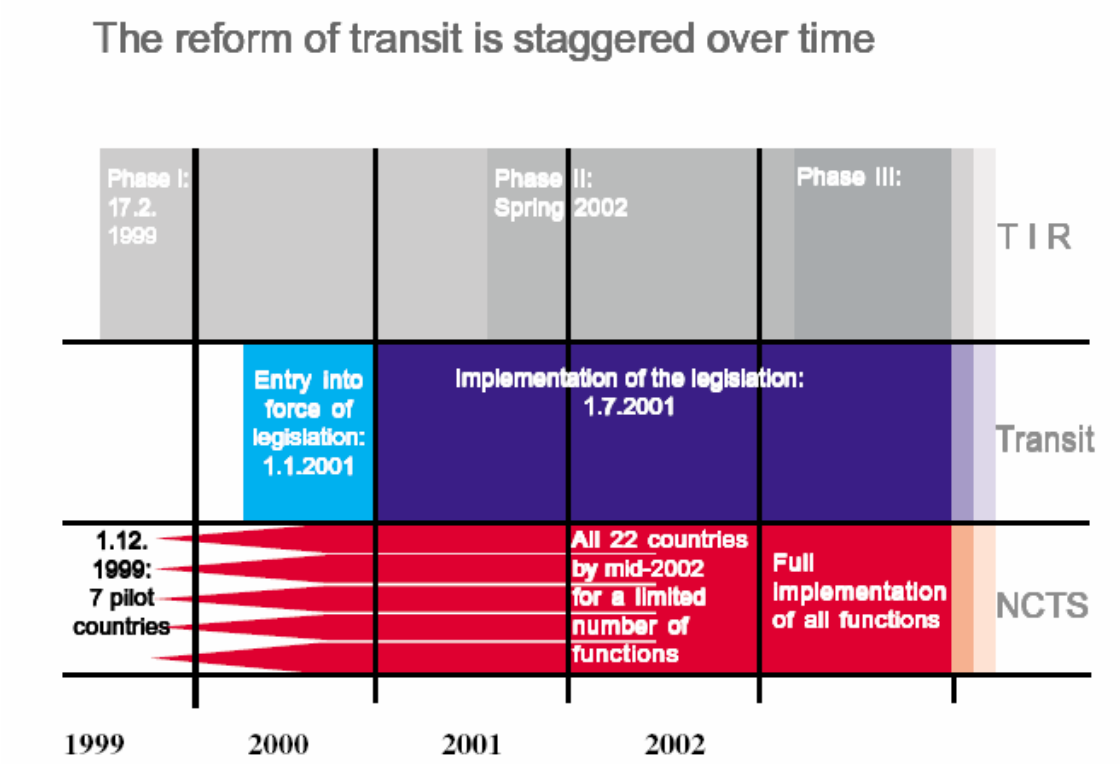
PECS: Woes with ROOs

The next problem with the statistics comes from another highly technical consideration – Rules of Origin, or ROOs as they are known to cognoscenti.

The EU is the world champion when it comes to preferential trade agreements. These cover not only intra-EU trade but also a very large share of EU imports from third nations ranging from Switzerland to Mexico. Preferential trade agreements, however, only cut tariffs on goods originating in nations that

have signed the agreement. To establish which goods get the tariff preference these agreements need 'rules of origin.'

Figure 19: The reform of EU transit regimes line up with the euro's introduction.



Source: European Commission (2001), "New customs transit systems for Europe."

Throughout my career as a trade economist, I've tried to ignore ROOs for two good reasons: they are dauntingly complex and mind-numbingly dull. My third reason for ignoring them – they don't matter much – turns out to be wrong. A string of recent papers demonstrates that they do affect trade flows, i.e. they are non-tariff barriers. A paper that will appear in the next issue of Economic Policy, Augier, Gasiorek and Lai Tong (2005), studies the impact of ROOs on European trade. In particular they study the impact of a change in which the EU applies its ROOs. This change, known as the Pan-European Cumulation System (PECS) was implemented in 1997.

The system is complex, but it was set up at the request of EU industry to reduce the existing complexity. Here's how. Staying competitive requires firms to set up a complex supply chain in which components were shipped among many nations. In the mid 1990s, there were something like 60 bilateral FTAs in Europe, each with its own complex set of origin rules. Such complexity made it difficult for firms to optimise manufacturing structures since could be difficult if not impossible for a firm to be absolutely sure how the outsourcing one of its intermediate goods would affect the origin status of its final-good exports.

The PECS simplified this in two ways: 1) it imposed uniform rules of origin in the EU15, EFTA nations and the ten nations that joined the EU in 2004, and 2) it allowed firms to count goods from any of these nations as originating in the EU.

Theoretically, the biggest impact on trade flows is between 'spoke' economies that had FTAs with the EU, but it could also encourage or discourage EU imports from nonEU nations both those that are part of PECS and those who aren't (see Augier, Gasiorek and Lai Tong 2005, or Krishna 2004).

The relevance here is that this could alter trade flows in the EU just about the time the euro was introduced. Augier, Gasiorek and Lai Tong (2005), for example, found that it had a statistically

significant impact on trade flows between the EU and nonEU PECS nations (both ways), and as well as boosting trade among the spokes. It is also not exactly clear whether PECS could warp the way in which EU imports are allocated across third nations. This is something that should be checked by ROO experts at the Commission.

5.1.2. **Euro depreciation and appreciation**

Another suspect for spurious results is the sharp depreciation of the euro at its birth. Here is the basic story. The gravity equation is a fancy demand curve. The sudden and sharp depreciation of the euro between 1999-2001 would make intra-Eurozone goods look cheaper relative to the extra-Eurozone goods. If this is not properly controlled for, the EZ dummy would pick of the expenditure shifting as a Rose effect.

Flam-Nordstrom attempt to control for this, but they have not addressed to problem of lags. Given the usual lags involved in trade, the impact on trade flows could last for a couple of years and thus still be biasing the results. Another point concerns the differential external exposures of Eurozone nations to external trade. For example, Greece does much less trade with the dollar zone than Ireland, so the euro depreciation could behind part of the differential effects that MSO find.

5.1.3. **Delayed Single Market effects**

As the papers by Berger and Nitsch (2005) and Mongelli, Dorrucci and Agur (2005) show, European integration is a work in progress. The doorstep to the euro, the 1992 to 1998 period, witnessed a particularly intense burst of deeper integration (see Figure 17). This would not be a problem for the Flam and Nordstrom or MSO estimates done on the EU sample, if only all EU members introduced these Single Market measures at the same time. But 'if only' is the bane of empirical economics. EU members differ widely on their pace of implementing EU Directives. Worse still, most of the tortoises are inside the Eurozone and most of the hares are outside. If you combine this with the likelihood that the pro-trade effects of many Directives may take a couple years of more to be fully realised, it is easy to see that there is a real problem. It is possible that the 'euro effect' is nothing more than the delayed and differential effect of pro-trade directives.

5.1.4. **Bottom line**

Here I end my 'Doubting Thomas' digression. All three of the sources for spurious results should be taken seriously by empirical researchers, but I believe a Rose effect happened. Just tell any group of European businesspeople – especially ones who own small or medium enterprises – that the euro had no impact on trade inside Europe. After a guffaw, you will almost surely hear 'of course, it did.' Now is its time to think about the microeconomics of why.

5.2. Microeconomic changes that might produce a Rose effect

In science, "fact" can only mean "confirmed to such a degree that it would be perverse to withhold provisional assent." I suppose that apples might start to rise tomorrow, but the possibility does not merit equal time in physics classrooms. -- Stephen Jay Gould

When I first started working on exchange rate changes and trade in the mid-1980s, the US deficit was the real puzzle. The dollar was falling like mad but the trade deficit wouldn't budge. Various economists had estimated the import demand equations and found nothing wrong, but others estimated the import pricing equations and found structural breaks. This puzzled many international economists since most had not heard of the 'new trade theory' with its increasing returns and imperfect competition – for them the J-curve was the nec-plus-ultra. Having taken my trade course from Helpman and Krugman (Helpman was visiting MIT that year), imperfect competition was the first thing I thought of, and I invented the Hysteresis in Trade – also known as the Beachhead Effect – to explain why we might

see a structural break with the falling dollar that we had not seen during the rising dollar.[42] In the euro's case, the facts seem to be reversed. We seem to have reasonably firm evidence of a structural break in the volume of trade, but not in the pricing. This makes me think that the real story will require the 'new, new' trade theory of Marc Melitz – but to start with, I want to consider a broader range of alternatives.

Theoretical literature on the Rose effect

I should have theory review to match the empirical review. That is easy. The only theoretical models of the Rose effect that I know of are the ones I have done in Baldwin Skuldelny and Taglioni (2004) and Baldwin and Taglioni (2004). Both turn on changes in the Beachhead costs and/or in uncertainty facing small firms using the heterogeneous-firms model of Melitz (2003). More on this below.

5.2.2. **An organising framework**

By my count, there four ways in which trade flows can change in the time framework discussed here, namely 5 years or so. If the Rose effect was for real, it is probably due on one or more of these. The best way to organise the discussion is to start with a bit of theory that just has to be right, the demand function. Taking the CES demand function from my discussion of the Flam-Nordstrom paper:

(20)
$$X_{od} = n_o \tau_{od} \mu_{od} \left(\frac{RER_{od}}{\left(\sum_k n_k \tau_{kd}^{1-\sigma} \left(RER_{kd} \right)^{1-\sigma} \right)^{1/(1-\sigma)}} \right)^{-\sigma} \frac{E_d}{P_d}; \qquad RER_{kd} \equiv \frac{e_{kd} mc_k}{P_d}$$

$$p_{od} = (\mu_{od})(\tau_{od})(e_{od}) mc_o$$

Recall that α_d is the share of expenditure on imports in the destination nation, μ_{od} is the price-marginal cost markup on sales from nation-o to nation-d, e_{od} is the bilateral exchange rate and mc_o is the marginal production cost in the exporting nation.

We have reasonably good data on are expenditures, bilateral exchange rates and aggregate price indices. We have very poor data on trade costs, markups, the number of exported varieties, n, and marginal cost, mc_o. Thus if we get a structure break when the euro in introduced (significance of the euro dummy is just a way of testing for structural breaks), then we can be reasonable sure that it is coming from a change in:

- The trade bilateral costs, τ
- the markup, μ
- the marginal costs, mc_o
- the number of varieties traded, n
- one of more of the above changes for third nation exports to the Eurozone.

5.2.3. **Trade costs**

This is the traditional optimal currency area story. If two nations share the same money, transaction costs are lower. The evidence on the Eurozone, however, suggests that this is a one-way trade cost reduction. In other words, it seems that the euro makes Eurozone nations exceptionally good at importing, rather than importing and exporting. Or, to put it differently, adoption of the euro was akin to unilateral trade liberalisation.

The problem with this one is that it is fairly hard to imagine the sorts of trade costs that could account for the magnitude of the Rose effect we seem to have seen. Even a 5% increase in aggregate trade, if it were due solely to lower trade costs, would require trade cost reduction of at least 5%. Such changes would be plainly obvious to practitioners. Since the sectoral evidence tells us that the Rose effect

[42] The paper I gave at the NBER's 1985 Summer Instittue when I was a third year grad student was eventually published in as Baldwin (1990), My papers with Krugman, Baldwin and Krugman (1988) and Krugman and Baldwin (1987) were written later but published earlier.

appeared only in some sectors, it would seem that we cannot search for marco changes, like a reduction in the cost of changing monies, or more intense banking competition driving down financial fees on international transactions.

A more subtle possibility

Flam and Nordstrom (2004) suggest that even small trade cost reductions could have large effects. It is based on work by Yi (2003) that shows how important the international fragmentation of trade is these days. The use of imported inputs in goods that are exported grew by almost 30 per cent in the OECD 1970 and 1990. In other words, goods – understanding a good to be the sum of its components – cross borders many times these days. Thus a small change in trade cost can, via a cumulative impact on costs – have big impact on trade volumes. Moreover, if the lower trade costs foster greater fragmentation, the process can be highly nonlinear.

5.2.4. **The markup**

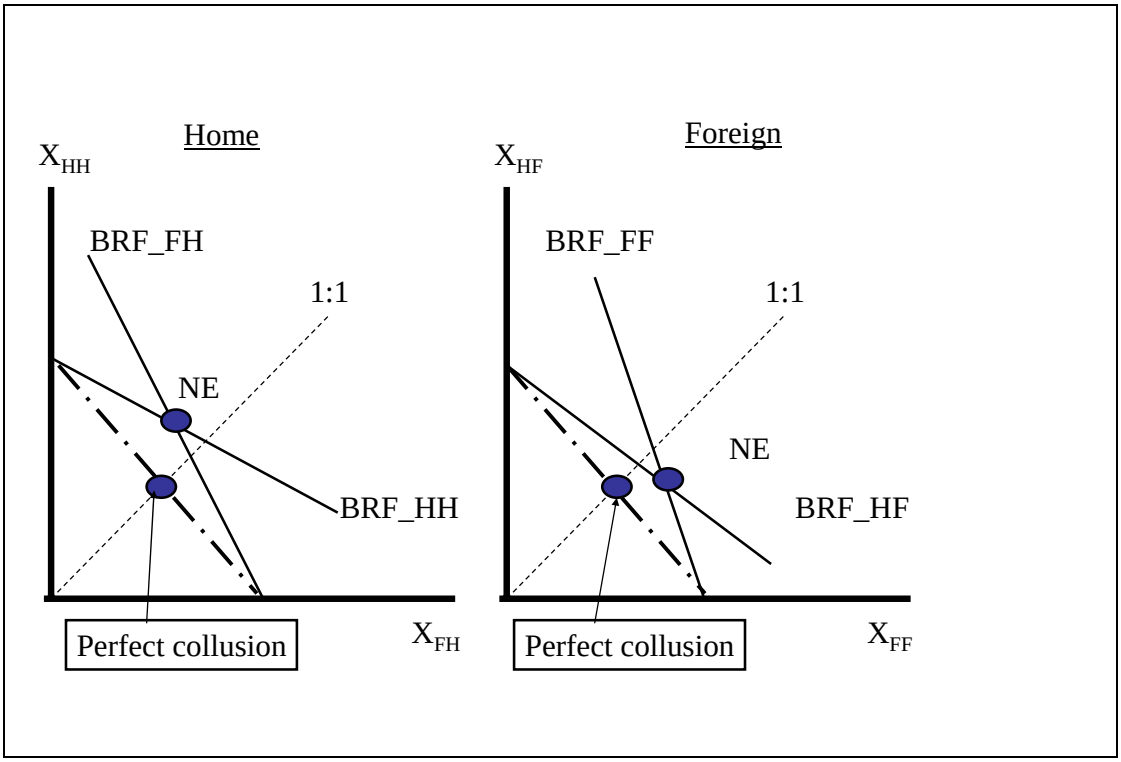
A good way of thinking about the markup is as a measure of the degree of competition. Changes in market structure can change the markup and thereby change the volume of trade. A very simple example is illustrated in Box 3. The critical point here is that if such a thing occurred, and it occurred due to the monetary union, it would show up in a gravity equation as a positive Rose effect.

One very attractive feature of the markup hypothesis is that it is sufficiently flexible to account for the national and sector diversity in the Rose effect estimates. One could expect that the euro would increase competition among Eurozone companies, and indeed this is what many trade practitioners will tell you. However, it is quite likely that the effect would interact in a complex manner with various nation-specific and sector-specific features. For example, if the euro lead to greater price transparency for big-ticket items like cars and trucks, the impact would be very different in, say, Greece, than it would be in Germany. The two nations differ greatly in terms of local producers and geographic distance to alternative suppliers. Likewise, the extent of the change in the markup could easily vary by sector. The obvious point is that the markup is pretty small already in some sectors and so unlikely to fall further. The more subtle point is that the pro-competitive effects of the euro could interact in complex ways with national regulations, a good example is banking. One would have thought that the elimination of currency risk would have allowed German homebuilders to get a mortgage in Luxembourg, but domestic regulations and practices effectively prevents this for now.

Box 3: Market structure and trade volumes

Consider the classic Brander-Krugman trade model of reciprocal dumping wherein there is one firm per nation and two nations. The standard presentation assumes the firms play Cournot in each market and so with trade costs, each sells some in the other's market, but has a bigger market share in its local market. One could, however, do the same assuming that the two firms act less competitively. For example, if they can collude perfectly, they arrange things such that the sum of their sales in each market equals the monopoly level, and somehow they agree to divide up the markets.

Both outcomes are shown in the diagram. The BRF lines are best response functions, one for each firm in each nation. The intersections show the Nash equilibrium (NE), the other dots show one example of perfect collusion. If we start from perfect collusion and somehow the euro creates more competition, say it breaks the collusion, then we get more trade. It is easy to see this since the NE points are further from the origin and thus involve a higher joint sales. Indeed in this example both firms sell more in both markets. As the exhaustive and exhausting new trade literature showed, there is absolutely nothing robust in this example. Nevertheless, it illustrates to point that a sudden change in the degree of competition can lead to a sudden increase in the volume of trade – i.e. it is a contender for the microfoundations of the Rose effect.



5.2.5. **Marginal costs**

If trading firms in the Eurozone experienced a drop in their marginal costs then there would appear to be a structural break in the import demand equation that did not use real import price data. That is to say, it would seem that these nations were trading more than they should be given the observables and this would make the Rose effect dummy positive and significant.

What could cause the marginal costs of Eurozone companies to drop with the euro introduction? I can only guess and as Sherlock Holmes said in *The Sign of Four*, "I never guess. It is a shocking habit -- destructive to the logical faculty." Well, perhaps I can make an exception. One guess would be that the euro led to a change in labour union behaviour in sectors where trade competition is particularly fierce. Another guess arises from the subtle point raised by Flam and Nordstrom. If goods are made up of parts that cross borders many times, then the total change in production cost can be many times the change in the once-across-the-border cost that most of have in mind.

The problem with the marginal cost explanation – regardless of what is driving it – lies in the trade diversion facts. If the euro is lowering manufacturing costs in the Eurozone more than that of nations outside the Eurozone, then the Eurozone exports to the rest of the EU should have flourished. Flam and Nordstrom find the opposite, at least on their cleanest set of results – those involving only the other EU members.

5.2.6. **Intensive and extensive margins – magic with the Melitz model**

A fascinating paper that I recommend as 'mind candy' to all international economists is Bernard and Jensen (2004). Using real data – and here I am talking about data from individual plants for the entire US manufacturing sector – they decompose sources of the US export boom in the late 1980s and early 1990s. They find that the preponderance of the increase in exports came from increasing export intensity at firms that were already exporting, but a non-negligible share came from firms that switched between only selling locally to selling locally and abroad. (Little known fact: most firms in most nations do not export even when they are in so-called traded goods sectors.)

The relevance here is that a change in 'n' – the number of varieties produced in nation-o and sold in nation-d – could produce a Rose effect since we do not have data on the n's. There are many things to recommend this idea, so many that I actually wrote of a simple theory model of how it might work (Baldwin and Taglioni 2004). Here it is in a nutshell.

The Baldwin-Taglioni story

The basic intuition is simple. Most European firms are not engaged in trade; they sell only in their local markets due to a variety of reasons – one of which is aversion to exchange rate uncertainty. Such uncertainly is a nuisance to giant companies like Nestle and Fiat, but to small and medium firms it is a very real barrier. Our story is that monetary union eliminated this uncertainty and thus increased the number of firms in the Eurozone that are engaged in exporting to other Eurozone markets. A sudden and permanent reduction of bilateral volatility within the Eurozone thus led to an increase in exports with little change in the basic production structure.[43] This story rests on the Melitz model – that's the young Melitz (Melitz 2003) rather than his father, my esteemed discussant – where the range of firms that export is endogenously determined and related to native firm-level productivity so that large firms export while small firms do not. (The firms are large and export since there are idiosyncratically more efficient than the non-exporters.)

Note that this story is flexible enough to account for sector variation and nation specific variation. In particular, the Melitz model works in industries marked by imperfect competition and increasing returns, but not in Walrasian sectors – just the pattern that the sector studies suggest. It is also flexible enough to allow for geographic variation if one goes into some of the craftier bits of the theory. To wit, there is a non-intuitive interaction between trade costs and the impact of exchange rate uncertainty in our model that predicts that the Rose effect should have been greatest in the nations that were already most tightly integrated.

Do the facts fit?

Inspector Gregory: "Is there any other point to which you would wish to draw my attention?"
Holmes: "To the curious incident of the dog in the night-time."
"The dog did nothing in the night time"
"That was the curious incident," remarked Sherlock Holmes.
From "The Adventure of Silver Blaze" by Arthur Conan Doyle

Consider the clues that Sherlock would have before him: The effect happened very quickly, far too quickly for the new trade to be explained by important changes in production structures. Moreover, given the very small size of the likely transactions cost reductions entailed in monetary union (even smaller than the one linked to currency union), the size of the effect seems to be too large to be explained by a standard export supply curve model using a reasonable export supply elasticity.

In poorly written murder mysteries, the critical clue remains obscure to the end of the novel. So it is with my paper. The keystone clue is that the scant evidence we have on the pricing equation in post-euro Europe – e.g. the Economic Policy article by Engles and Rogers (2003) – suggest that there has not been a break in the pricing relationships. Indeed, a more systemic study the ECB finds that although there are structural breaks in inflation relationships, they do not seem to be tied closely to the euro's introduction. The pricing equations are the dog that did not bark.

All of the other explanations – unobservable trade cost changes, unobservable change in markups and unobservable changes in marginal cost – should have also shown up as structural breaks in the pricing equation. But as even the most casual glance at (20) reveals, the number of exported varieties does not enter the pricing equation. A 'structural break' in 'n' would thus create a structural break in the trade volume equation without creating one in the trade price equation.

[43] The link is extremely non-linear, so throwing in a linear volatility term would not eliminate it. In Baldwin and Taglioni (2004), we find evidence the Rose effect is actually a highly nonlinear effect of exchange rate volatility.

6. BATTERY OF DIAGNOSTICS

This section discusses a battery of diagnostics that should give a better idea as to what is really going on.

6.1. Checking for spurious results

VAT fraud and the Rose effect

This is a very difficult one to check for. The approach by Flam-Nordstrom is a good start – they throw in various dummies – but as I pointed out above, the problem is likely to vary by trade pair, by time and by commodity. Surely the best place to start is with the extensive work done by Eurostat and member states' tax authorities to investigate this phenomenon. It is costing them billions, so they have surely put serious resources into working out how much fraud is occurring. In all likelihood, the problem is worse in for some trade pairs, e.g. cross channel trade with Britain, and in some goods, e.g. tobacco and alcohol. Indeed, the largest Rose effect in the Flam-Nordstrom paper is for the Tobacco and Beverage sector (since there are high excise taxes in addition to the VAT in some member states, fraud in this sector is particularly profitable). Direct evidence from anti-fraud studies would help us understand which data are particularly unreliable. It might also provide ways of adjusting the data for underreporting. Note however, that the VAT authorities are not interested in working out how much of the reported trade is 'for real' and how much is carosel trade.

One brave approach would be to estimate a model of the fraud – maybe based on VAT rate difference by commodity and by trade pair – and use it to instrument the trade data. Or at least to tell us where to put in fraud dummies akin the the Flam-Nordstrom Rotterdam dummies.

The Rotterdam effect

The Rotterdam effect whereby external trade gets mis-represented as being imported into the wrong EU nation and then generating false intra-EU trade is easier to deal with. Probably the best place to start is with data reported by third nations. For example, Australia has engaged in a correction exercise since many of its exports do not get reported in the nations it believes they are going to. Also, Switzerland participates in most of the European schemes and yet imports a lot of its goods via Rotterdam and Anterwp, so its trade may provide some sort of yardstick for separating the Rotterdam effect from the Eurozone effect.

PECS and ROOs

Here the challenge is to find out exactly how PECS affected intra-Eurozone trade by finding out if and how the system altered reporting of the country of origin. Also, since PECS affected all EU members and many other nations, there is hope that there is sufficient variation in the data to sort out the effects econometrically.

Euro depreciation

The obvious thing to do here is to use some real exchange rate indices that are specific to each importing nation. The gravity equation theory suggests the way forward, but I believe it is important to derive that proper control variable from the theory rather than just throwing in various real exchange rate indices. One will also have to throw in some lags to soak up delayed effects that can take years to show up.

Single Market initiatives

The first thing to do would be to use the Berger-Nitsch and Mangelli et al indices of European integration interacted with the Commission's implementation deficit scoreboard figures and see what this does in the Flam-Nordstrom set up. Ultimately, however, the only thing to do is wait until we have enough data to allow this sort of thing to average out. Alternatively, we could pursue a sectoral approach and introduce sector-specific indicators for Internal Market implementation by year and by nation.

6.2. Real changes

If I had a large research budget at my disposal to investigate the Rose effect in the Eurozone, one of the first things I'd do would be to survey business people, asking them how they think the euro has change the trading environment in their businesses. (The Commission did this for both the One Market One Money study and the Cecchini Report.) The next thing would be to shift the focus away from aggregate trade data studies and towards more specific effects and sectors.

The key, I believe, is to use an empirical methodology that combines the estimation of both price and quantity equations. If the unobservable trade costs, markup, or marginal cost stories are right, we should find breaks in the price equations that are congruous with the Rose effect by nation, by sector and by year. It can be hard to get bilateral price data (although, I believe Mike Knetter used German data in his 1980s papers on the pass-through puzzle), but even without this, one should be able to detect a break in EU nations for whom the Eurozone is the dominant trade partner. The examples of Belgium and Austria come to mind. Moreover, it seems that the ECB has access to some excellent price data for their inflation persistence project. From the public description, these data would be extremely useful in pinning down the source of the Rose effect.

The Flam-Nordstrom-Yi story could be tested directly by looking at trade in intermediates. There are by now several classifications of goods into final and components. If this story is right, much of the Rose effect should be coming through trade in components.

Changes in markups would also show up in the pricing equations. More directly, changes in market structure leave footprints in many data sets. For example, a rapid change in competitiveness on the scale necessary to produce a Rose effect should show up in the financial performance and thus stock market performance of the companies involved. Changes in the marginal cost of trading firms could also be tested directly by studying the financial performance of exporting firms.

7. CONCLUDING REMARKS

Fifteen years ago, Michael Emerson – the real author of the Cecchini report and high-ranking Eurocrat at the time – asked me to write a paper called the "On the Microeconomics of the European Monetary Union". In that paper I engaged in a practice that would have shocked Sherlock Holmes and Andy Rose – I theorised before I had the facts. I am grateful to the ECB for giving me this opportunity to reverse that error and to Andy Rose for stimulating an empirical debate that has provided many facts.

Although we are a long way from really knowing how the euro affected trade, it is now clear that the way forward needs to be guided by detailed theoretical hypothesis as to HOW the euro affects trade. There really is not enough information in the aggregate trade data to answer the question: "How much did the euro boost trade?" Indeed, the question itself is probably as unanswerable in the aggregate. Any conceivable theoretical accounting for the Rose effect would suggest that it should apply in different ways to different goods and different countries. What we need to ask is questions like: "If the euro boosted trade by sharpening competition, then in which dataset should we find the footprints?" And then go and check for footprints in the appropriate datasets, many of which will have nothing to do with trade.

In many ways, the whole Rose literature reminds me of the pass-through literature in the 1980s. It started with really bad econometrics on aggregate data. I myself estimated aggregate import pricing relationships for the US. Despite the manifest shortcomings of such an endeavour, it published in the AER because, like Samuel Johnson quipped about a dog walking on its hind legs, the interest lay not in the fact that it was done so well, but rather that it was done at all. Nowadays, researchers studying the pass through of exchange rate changes on prices use firm-level, product-specific data and take account of all the usual features of market under study. In the same way, it is now time to move beyond studies of 'how big is the magic.'

8. APPENDICES

8.1. Gravity for dummies, and dummies for gravity

A wide range of dummies have been thrown into the various regressions we have discussed, creating a bit of a muddled. Before turning the next branch of the climbing rose plant that this literature has become, it is worth sorting out the methodologies employed. Here I rely on a paper by I-Hui Cheng and Howard Wall. Cheng and Wall (2004) take a single dataset – a dataset that involves relatively homogenous nations (OECD) and observations every five years from 1982 to 1997 and estimates the same gravity equation using most combinations of dummies. Here is the list, working down from least restrictive to most restrictive:

- **Unidirectional pair dummies plus time dummies (FE)**. As I noted in my theory aside, the gravity model should be thought of as 'a demand equation with social ambitions'. Since the exports of Canada to the US face a different demand function than do US exports into Canada and US export prices are determined by different factor than Canadian export prices, one should include the two bilateral trade flows separately. If one is willing to make somewhat stronger theoretical restrictions, one may pool these flows but such assumptions are testable and in any case why bother? It just means you have to throw out half your data points. (Rose and most of the Rose Effect Hunters add up the two flows and divide by two before putting the values into the regression.) This specification is attributed by Cheng and Wall (2004) to Cheng (1999) and Wall (1999).
- **Symmetric pair dummies plus time dummies (SFE)**. In their data they force the unidirectional dummies to be the same, i.e. the Canada-to-US dummy to equal the US-to-Canada dummy. This is basically what Glick-Rose does.
- **Dual country dummies plus time dummies (XFE)**. As Anderson-Wincoop recommend (although Cheng and Wall ascribe the technique to Mátyás 1997), they put in a dummy for each nation as an importer and for each nation as an exporter. Of course, if one has reduced the information in the dataset by averaging a nation's bilateral exporters and imports, this is not possible in a balanced panel. In that case, one can only put in one dummy per nation. Unfortunately, Cheng and Wall (2004) don't compare theses two, reporting numbers only for the dual country pair dummies case.
- **Single country dummies plus time dummies**. This is like the previous one, but there is only one dummy per nation.
- **Time-varying Country dummies**. A balanced dataset has m time m-1 observations per year where m is the number of nations. If the directional bilateral flows are averaged first, the number is m(m-1)/2 times the number of years. This is a big number. The Anderson-Wincoop theory tells us that the country dummies SHOULD vary by year since they included, inter alia, GDP of various nations. One can include year-by-year country-specific fixed effect, i.e. T times 2 times m dummies. This is the correct way to control for the real-prices-matter terms, but it excludes pair dummies, so the researcher faces a trade-off. Unfortunately, Cheng and Wall (2004) don't look at this, but I hope they will in a future draft. **Pooled Cross Section with time dummies (PCS)**. This is the preferred technique in Rose (2000). One estimates the model on the whole panel allowing only the intercepts to vary across years.

8.1.1. **Comparing dummies**

What do they find? Well, unfortunately they haven't done this exercise on the Rose effect question, but they do estimate a stripped-down gravity model involving only GDPs and populations of origin and destination nations (the GDPs are real for unexplained reasons and they deflate the value of trade flows by the US CPI). They also estimate a gravity model with a few regional trade agreement dummies.

Table 6: Comparing various FE estimators for the gravity model (Cheng and Wall 2004).

	Pooled Cross-Section	*Unrestricted FE Model*	*Restricted FE Models*	
	PCS	FE	SFE	XFE
intercept	6.852*	-	-	-
	(0.546)			
origin GDP	0.617*	0.122*	0.213*	0.122*
	(0.038)	(0.023)	(0.025)	(0.055)
destination GDP	0.511*	0.208*	0.117*	0.208*
	(0.035)	(0.027)	(0.024)	(0.054)
origin population	0.141*	-0.390	0.935*	-0.390
	(0.038)	(0.298)	(0.268)	(0.565)
destination population	0.214*	2.313*	0.989*	2.313*
	(0.038)	(0.319)	(0.268)	(0.584)
distance	-1.025*			
	(0.023)			
contiguity	-0.125			
	(0.085)			
common language	1.075*			
	(0.072)			
1987	0.077	0.199*	0.199*	0.199*
	(0.067)	(0.029)	(0.038)	(0.063)
1992	0.014	0.357*	0.357*	0.357*
	(0.068)	(0.043)	(0.053)	(0.093)
1997	0.051	0.482*	0.481*	0.482*
	(0.064)	(0.058)	(0.070)	(0.122)
observations	3188	3188	3188	3188
parameters	11	804	408	63
log-likelihood	-5163.27	-1663.07	-2863.46	-4704.08
R^2	0.690	0.954	0.916	0.768

Notes: White-errors in parentheses; * denotes 5% significance level. FE is Cheng-Wall fixed effects.

In terms of priors on point estimates, the pooled cross section does best, with the distance elasticity about -1.0 as usual, the GDP elastiticites closest to the expected unit (but still more than ten standard deviations away from unity), and the population variables positive but less than the GDP variables (so GDP/population is pro-trade). The other estimators yield GDP elastiticites that are hard to believe give that trade has expanded faster than output in almost every year since the war. By the way, the hard-to-believe GDP elastiticites tend to go away with one uses the product of GDPs, although the authors don't look at this.[44]

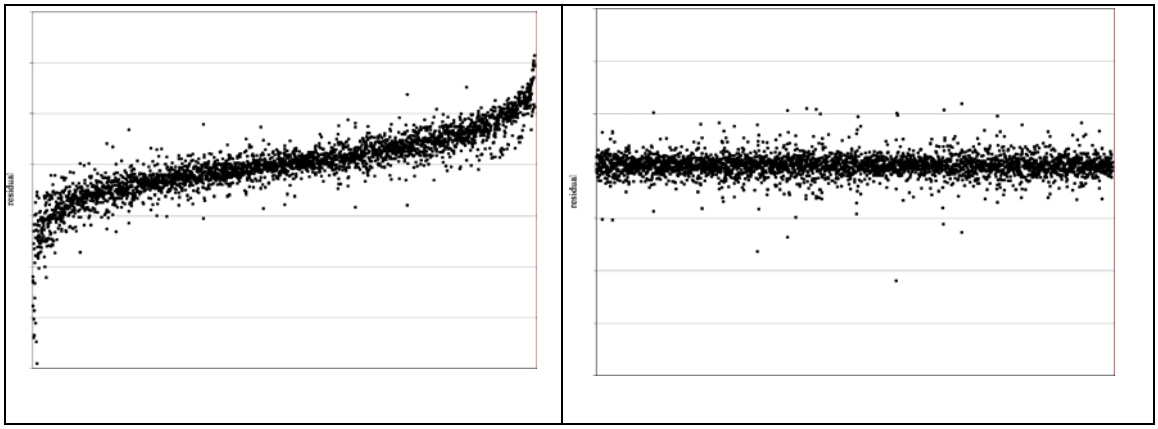
Seeing this, it is somewhat easier to understand why so few people have used sophisticated fixed effects estimator. But which method is better econometrically?

8.1.2. **Checking residuals**

One thing that Cheng-Wall do, which I think should be standard practice for all Rose-effect hunters, is to plot the residuals. The left panel of Figure 22 the residuals for the pooled cross section model for the 797 unidirectional country pairs in their data set. These are ordered by the average of pairs residuals. Remember we have to assume that these residuals are white noise if the estimator is valid, but this is plainly not the case. Thus although one might prefer the PCS model based on the numbers in Table 7, a quick eyeballing of the residual tells us that not everything is coming up roses. The right panel is for the Cheng-Wall direction-specific pair dummies. These look much better.

[44] One thing I don't understand is why the point estimates for the FE and XFE models are exactly the same bare-bones gravity model, but different when one throws in other proxies for trade cost (see their Table 2).

Figure 20: Residuals from PCS and FE models.



Residual claimants

When a model is mis-specified, there is a residual variation in the trade data that needs a home. If this residual variation and some included explanatory variable find that they have something in common in a correlation sense, an unholy alliance can emerge. Thus it is useful to run a simple model and inspect the residuals. In the old days, when 'lots of data' meant more than one hundred observations, a researcher could undertaken this sort of inspection pretty easily – most of the old econometric packages let you plot actual and fitted values for this task. Things are harder today with big datasets, but the need has not disappeared.

Howard Wall kindly sent me the residuals from the pooled cross-section regression he ran on the OECD. This is not directly related to the Rose effect since no one in his data had a currency union, but it is useful in thinking about the sorts of nonlinearities that could arise even among OECD nations. What I did was to order the residuals according to various criteria and look for patterns. Of course, since these are regression residuals, the are supposed to be white noise and we know they are orthogonal to the regressors. But what about nonlinear combinations of the regressors? Especially ones that theory suggests might matter.

The new trade theory, which is a quarter of a century old this year, tells us to expect that north-north trade and south-south trade should be different than north-south trade. There aren't any really poor nations in the dataset, but I labelled Argentina, Brazil and Mexico as developing nations (the south) leaving Australia, Austria, Belgium-Luxembourg, Canada, Denmark, Finland, France, Germany, Greece, Hong Kong, Ireland, Israel, Italy, Japan, Korean Republic, the Netherlands, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, Switzerland, the United Kingdom, Uruguay, and the United States as developed nations (the north). Then I lined up all the residuals with those corresponding to north-north and south-south trade flows first and north-south second (the order with in these groups is pretty randon – alphabetic by destination nation).

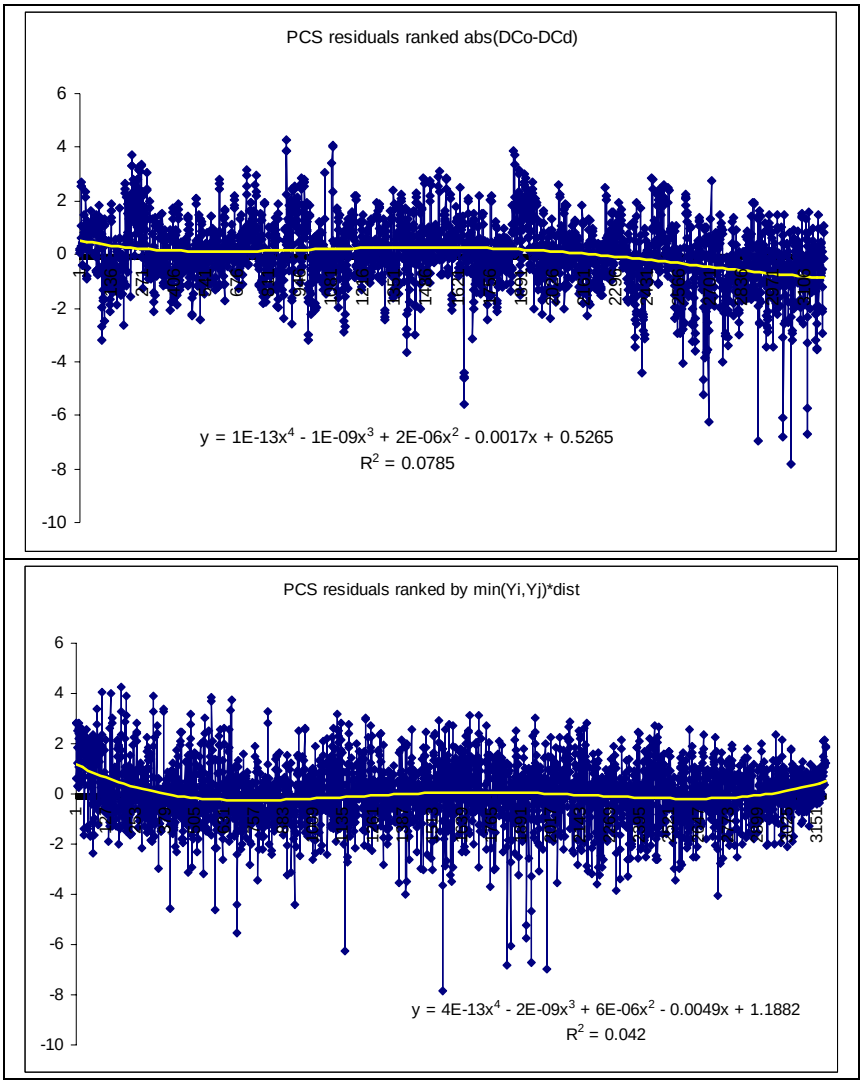


Figure 21: Residuals from PCS and FE models.

The result is shown in the top panel of Figure 22, with an Excel trendline (4th order polynomial) added for emphasis. The North-South flows are at the right and they are systematically negative. In other words the gravity model works systematically worse for North-South trade, a point that has been know to gravit-istas for decades.

The second ranking is a bit more subtle. It ranks the residuals by the minimum of the real GDP of the smaller partner times the bilateral distance. The idea is that made trade with small distant nations is more influenced by unusual, unquantifiable factors. The result, plotted in the bottom panel suggests deviations on both ends of the spectrum. The observation further to the right is US-Japan (the smallest is big and they are far apart), the leftmost observation is Netherlands-Belgium.

Honestly, I can't say exactly what all this means apart from the fact that we should be looking at residuals instead of relying purely on summary statistics.

8.1.3. **Mistaken policy inferences**

Cheng and Wall do the same exercise throwing in fairly standard dummies for preferential trade agreements. I won't go through their policy results in detail since there are not directly relevant, but one point jumps out. They provide an useful example of how one can be misled by the pooled cross section

(PCS) estimator employed by Rose (2000). Cheng-Wall estimate the impact of the Israel-U.S. FTA using the various estimators. Plainly there is something special about this trade relationship that would be impossible to quantify accurately a variable that was valid for a panel of 100+ nations. Allowing for fixed effects, this je-ne-sais-quoi is swept away and the Israel-U.S. FTA is found to have no significant impact on trade, a sensible result in my opinion. With the PCS method however, the FTA picks up all the je-ne-sais-quoi; the estimates suggest that the FTA boosts bilateral trade by an unbelievable 5 times. That's larger than all but the largest estimate of the Rose effect! Interestingly, the Rose-Wincoop method also finds a significant but much smaller effect, while the other methods find the FTA to be insignificant.

REFERENCES

Alho, K. (2002) "The Impact of Regionalism on Trade in Europe," Royal Economic Society Annual Conference 2003 3, Royal Economic Society.

Anderson, James and Eric van Wincoop, 2001, "Gravity with gravitas: a solution to the border puzzle," NBER Working Paper 8079, forthcoming in American Economic Review.

Anderson, James, 1979, "The theoretical foundation for the gravity equation," American Economic Review 69, 106-116.

Anderton, Robert, B. Baltagi, F. Skudelny and N. Sousa (2002) "Intra- and Extra-Euro Area Import Demand for Manufactures" European Trade Study Group Annual Conference 2002 4, European Trade Study Group.

Baldwin, R. and D. Taglioni, (2004) "OCA".

Baldwin, Richard E. (1991) "On the Microeconomics of the European Monetary Union" European Economy, 21-35.

Barr, David, Francis Breedon and David Miles (2003) "Life on the Outside" Economic Policy, 573 - 613.

Barr, David, Francis Breedon and David Miles, 2003, "Life on the outside: economic conditions and prospects outside Euroland," Imperial College, University of London, manuscript

Bernard, A. and J. Jensen (2004), "Entry, Expansion, and Intensity in the US Export Boom," Review of International Economics, Volume 12 Issue 4 Page 662 - September 2004.

Bun, Maurice and Franc Klaassen, 2002, "Has the euro increased trade?," http://www1.fee.uva.nl/pp/klaassen/

Bun, Maurice J. G. and Franc J. G. M. Klaassen (2002) "Has the Euro Increased Trade?", Tinbergen Institute Discussion Paper, TI 2002-108/2.

De Nardis, Sergio and Claudio Vicarelli (2003) "Currency Unions and Trade: The Special Case of EMU", World Review of Economics, 139 (4): 625-649.

De Sousa, Lucio V. (2002) "Trade Effects of Monetary Integration in Large, Mature Economies", A Primer on European Monetary Union, Kiel Working Paper No. 1137.

Deardorff, Alan, 1998, "Determinants of bilateral trade: does gravity work in a neoclassical world?," in Jeffrey Frankel (ed.), The regionalization of the world economy, University of Chicago Press.

Engel, Charles and John H. Rogers (2004) "European Product Market Integration after the Euro", forthcoming European Policy.

European Commission, 2000, REPORT FROM THE COMMISSION TO THE COUNCIL AND THE EUROPEAN PARLIAMENT, COM(2000) 28 final.

European Commission, 2000, REPORT FROM THE COMMISSION TO THE COUNCIL AND THE EUROPEAN PARLIAMENT, COM(2000) 28 final.

European Commission, 2000, REPORT FROM THE COMMISSION TO THE COUNCIL AND THE EUROPEAN PARLIAMENT, COM(2000) 28 final.

Feenstra, Robert. (2003) Advanced International Trade. Princeton University Press, Princeton, New Jersey.

Flam, Harry, and Hakan Nordstrom (2003) "Trade Volume Effects of the Euro: Aggregate and Sector Estimates", Institute for International Economic Studies unpublished.

Frankel, Jeffrey A. (1997). *Regional Trading Blocs in the World Economic System.* Institute for Internal Economics, Washington, D.C.

Glick, Reuven and Andrew Rose (2002) "Does a Currency Union Affect Trade? The Time Series Evidence", *European Economic Review* 46-6, 1125-1151.

Gomes, T, Chris Graham, John Helliwell, Takashi Kano, John Murray, Larry Schembri (2004). "The Euro and Trade: Is there a Positive Effect?" pdf file of a preliminary and incomplete draft, not for quotation without permission.

Helpman, Elhanan and Paul Krugman, 1985, Market structure and foreign trade, MIT Press.

Hummels, David, Jun Ishii and Kei-Mu Yi, 2001, "The nature and growth of vertical specialization in world trade," Journal of International Economics 54, 75-96.

Journal of Political Economy 111, 52-102.

Kenen, Peter B. (2002) "Currency Unions and Trade : Variations on Themes by Rose and Persson" Reserve Bank of New Zealand Discussion Paper 2002/08.

Linnemmann, H. (1966). *An Econometric Study of International Trade Flows*, North-Holland, Amsterdam.

Melitz, Jacques (2001) "Geography, Trade and Currency Union", CEPR Discussion Paper, No. 2987.

Micco, Alejandro, Ernesto Stein and Guillermo Ordoñez, 2003, "The currency union effect on trade: early evidence from EMU," http://www.economic¬policy.org/paneldrafts.asp

Micco, Alejandro, Ernesto Stein, Guillermo Ordonez (2003) "The Currency Union Effect on Trade: Early Evidence From EMU", *Economic Policy*, 316-356.

Mongelli, F. P., Ettore Dorrucci and Itai Agur (2005).What does European Institutional Integration tell us about Trade Integration?, 9 March 2005. PDF file.

Nitsch Volker (2004). “Have a break, have a …national currency: when do monetary unions fall apart?” CESIFO WORKING PAPER NO. 1113.
Nitshc, V. (2005). "Currency Union Entries and Trade", forthcoming in "Globalization of Capital Markets and Monetary Policy", edited by Jens Hölscher and Horst Tomann, Palgrave Macmillan, 2005.
Pakko, Michael R. and Howard J. Wall (2001) “Reconsidering the Trade-Creating Effects of a Currency Union”, *Federal Reserve Board of St. Louis Review*, 83-5, 37-45.
Persson, Torsten (2001) “Currency Unions and Trade: How Large is the Treatment Effect?”, *Economic Policy* 33, 435-448.
Persson, Torsten, “Currency unions and trade: how large is the treatment effect?,” Economic Policy 33, 435-448.
Piscitelli, L. (2003) mimeo, available from UK Treasury.
Poyhonen, Pentti. 1963a. “A Tentative Model for the Volume of Trade Between Countries”, *Welwirtschaftliches Archiv*, 90 (1): 93-99.
Poyhonen, Pentti. 1963b. “Toward a General Theory of International Trade”, *Ekonomiska Samfundets Tidskrift*, 16(2): 69-78.
Rose, Andrew K. (2000) “One Money, One Market: Estimating the Effect of Common Currencies on Trade”, *Economic Policy* 30, 9-45.
Rose, Andrew K. and E. van Wincoop (2001) “National Money as a Barrier to Trade: The Real Case for Monetary Union”, *American Economic Review* 91-2, 386-390.
Rose, Andrew K. and Eric van Wincoop, 2000, “National money as a barrier to international trade: the real case for currency union,” http://faculty.haas.berkeley.edu/arose/
Rose, Andrew K., 2000, “One money, one market: the effect of currency unions on trade,” Economic Policy 30, 7-46.
Rose, Andrew K., 2001, “Currency unions and trade: the effect is large,” Economic Policy 33, 449-461.
Tenreyro, S. (2001) “On the Causes and Consequences of Currency Unions” Harvard University mimeo.
Tenreyro, Silvana and Robert J. Barro, 2003, “Economic effects of currency unions,” NBER Working Paper 9435.
Yi, Kei-Mu, 2003, “Can vertical specialization explain the growth of world trade?,”

9. APPENDIX ON METHODS

9.1. A direct comparison of gravity estimators

Here I go through a very interesting paper that looks at various estimators side-by-side. I believe that the Rose effect literature needs much more of this sort of methodical, methodological studies.

A wide range of dummies have been thrown into the various regressions we have discussed, creating a bit of a muddled. It is worth sorting out the methodologies employed. Here I rely on a paper by I-Hui Cheng and Howard Wall. Cheng and Wall (2004) take a single dataset – a dataset that involves relatively homogenous nations (OECD) and observations every five years from 1982 to 1997 and estimates the same gravity equation using most combinations of dummies. Here is the list, working down from least restrictive to most restrictive:

- **Unidirectional pair dummies plus time dummies (FE)**. As I noted in my theory aside, the gravity model should be thought of as 'a demand equation with social ambitions'. Since the exports of Canada to the US face a different demand function than do US exports into Canada and US export prices are determined by different factor than Canadian export prices, one should include the two bilateral trade flows separately. If one is willing to make somewhat stronger theoretical restrictions, one may pool these flows but such assumptions are testable and in any case why bother? It just means you have to throw out half your data points. (Rose and most of the Rose Effect Hunters add up the two flows and divide by two before putting the values into the regression.) This specification is attributed by Cheng and Wall (2004) to Cheng (1999) and Wall (1999).
- **Symmetric pair dummies plus time dummies (SFE)**. In their data they force the unidirectional dummies to be the same, i.e. the Canada-to-US dummy to equal the US-to-Canada dummy. This is basically what Glick-Rose does.
- **Dual country dummies plus time dummies (XFE)**. As Anderson-Wincoop recommend (although Cheng and Wall ascribe the technique to Mátyás 1997), they put in a dummy for each nation as an importer and for each nation as an exporter. Of course, if one has reduced the information in the dataset by averaging a nation's bilateral exporters and imports, this is not possible in a balanced panel. In that case, one can only put in one dummy per nation. Unfortunately, Cheng and Wall (2004) don't compare theses two, reporting numbers only for the dual country pair dummies case.
- **Single country dummies plus time dummies**. This is like the previous one, but there is only one dummy per nation.
- **Time-varying Country dummies**. A balanced dataset has m time m-1 observations per year where m is the number of nations. If the directional bilateral flows are averaged first, the number is m(m-1)/2 times the number of years. This is a big number. The Anderson-Wincoop theory tells us that the country dummies SHOULD vary by year since they included, inter alia, GDP of various nations. One can include year-by-year country-specific fixed effect, i.e. T times 2 times m dummies. This is the correct way to control for the real-prices-matter terms, but it excludes pair dummies, so the researcher faces a trade-off. Unfortunately, Cheng and Wall (2004) don't look at this, but I hope they will in a future draft. **Pooled Cross Section with time dummies (PCS)**. This is the preferred technique in Rose (2000). One estimates the model on the whole panel allowing only the intercepts to vary across years.

9.1.1. **Comparing dummies**

What do they find? Well, unfortunately they haven't done this exercise on the Rose effect question, but they do estimate a stripped-down gravity model involving only GDPs and populations of origin and

destination nations (the GDPs are real for unexplained reasons and they deflate the value of trade flows by the US CPI). They also estimate a gravity model with a few regional trade agreement dummies.

Table 7: Comparing various FE estimators for the gravity model (Cheng and Wall 2004).

	Pooled Cross-Section	*Unrestricted FE Model*	*Restricted FE Models*	
	PCS	FE	SFE	XFE
intercept	6.852*	-	-	-
	(0.546)			
origin GDP	0.617*	0.122*	0.213*	0.122*
	(0.038)	(0.023)	(0.025)	(0.055)
destination GDP	0.511*	0.208*	0.117*	0.208*
	(0.035)	(0.027)	(0.024)	(0.054)
origin population	0.141*	-0.390	0.935*	-0.390
	(0.038)	(0.298)	(0.268)	(0.565)
destination population	0.214*	2.313*	0.989*	2.313*
	(0.038)	(0.319)	(0.268)	(0.584)
distance	-1.025*			
	(0.023)			
contiguity	-0.125			
	(0.085)			
common language	1.075*			
	(0.072)			
1987	0.077	0.199*	0.199*	0.199*
	(0.067)	(0.029)	(0.038)	(0.063)
1992	0.014	0.357*	0.357*	0.357*
	(0.068)	(0.043)	(0.053)	(0.093)
1997	0.051	0.482*	0.481*	0.482*
	(0.064)	(0.058)	(0.070)	(0.122)
observations	3188	3188	3188	3188
parameters	11	804	408	63
log-likelihood	-5163.27	-1663.07	-2863.46	-4704.08
R^2	0.690	0.954	0.916	0.768

Notes: White-errors in parentheses; * denotes 5% significance level. FE is Cheng-Wall fixed effects.

In terms of priors on point estimates, the pooled cross section does best, with the distance elasticity about -1.0 as usual, the GDP elastiticites closest to the expected unit (but still more than ten standard deviations away from unity), and the population variables positive but less than the GDP variables (so GDP/population is pro-trade). The other estimators yield GDP elastiticites that are hard to believe give that trade has expanded faster than output in almost every year since the war. By the way, the hard-to-believe GDP elastiticites tend to go away with one uses the product of GDPs, although the authors don't look at this.[45]

I should also note that population in these regressions is probably acting a bit like a time-invariant country dummy since population in the OECD nations has varied little in this time frame. Note how much the population point estimates change with country dummies (XFE) and without (SFE), for example. The theory suggests population should have a negative coefficient if import income elasticise exceed unity and nations tend to produce more traded goods as they develop (or vice versa).

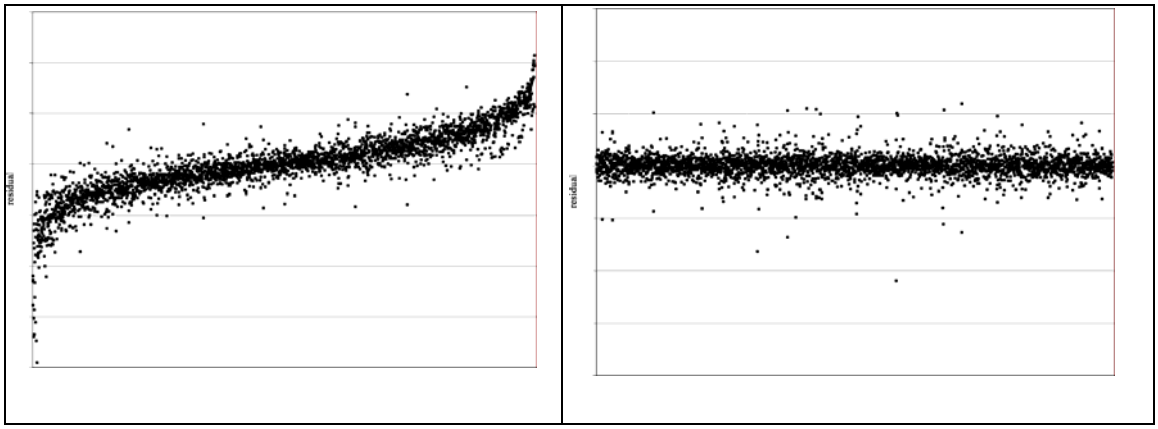
Seeing this, it is somewhat easier to understand why so few people have used sophisticated fixed effects estimator. But which method is better econometrically?

[45] One thing I don't understand is why the point estimates for the FE and XFE models are exactly the same bare-bones gravity model, but different when one throws in other proxies for trade cost (see their Table 2).

9.1.2. **Checking residuals**

One thing that Cheng-Wall do, which I think should be standard practice for all Rose-effect hunters, is to plot the residuals. The left panel of Figure 22 the residuals for the pooled cross section model for the 797 unidirectional country pairs in their data set. These are ordered by the average of pairs residuals. Remember we have to assume that these residuals are white noise if the estimator is valid, but this is plainly not the case. Thus although one might prefer the PCS model based on the numbers in Table 7, a quick eyeballing of the residual tells us that not everything is coming up roses. The right panel is for the Cheng-Wall direction-specific pair dummies. These look much better.

Figure 22: Residuals from PCS and FE models.



Residual claimants

When a model is mis-specified, there is a residual variation in the trade data that needs a home. If this residual variation and some included explanatory variable find that they have something in common in a correlation sense, an unholy alliance can emerge. Thus it is useful to run a simple model and inspect the residuals. In the old days, when 'lots of data' meant more than one hundred observations, a researcher could undertaken this sort of inspection pretty easily – most of the old econometric packages let you plot actual and fitted values for this task. Things are harder today with big datasets, but the need has not disappeared.

Howard Wall kindly sent me the residuals from the pooled cross-section regression he ran on the OECD. This is not directly related to the Rose effect since no one in his data had a currency union, but it is useful in thinking about the sorts of nonlinearities that could arise even among OECD nations. What I did was to order the residuals according to various criteria and look for patterns. Of course, since these are regression residuals, the are supposed to be white noise and we know they are orthogonal to the regressors. But what about nonlinear combinations of the regressors? Especially ones that theory suggests might matter.

The new trade theory, which is a quarter of a century old this year, tells us to expect that north-north trade and south-south trade should be different than north-south trade. There aren't any really poor nations in the dataset, but I labelled Argentina, Brazil and Mexico as developing nations (the south) leaving Australia, Austria, Belgium-Luxembourg, Canada, Denmark, Finland, France, Germany, Greece, Hong Kong, Ireland, Israel, Italy, Japan, Korean Republic, the Netherlands, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, Switzerland, the United Kingdom, Uruguay, and the United States as developed nations (the north). Then I lined up all the residuals with those corresponding to north-north and south-south trade flows first and north-south second (the order with in these groups is pretty randon – alphabetic by destination nation).

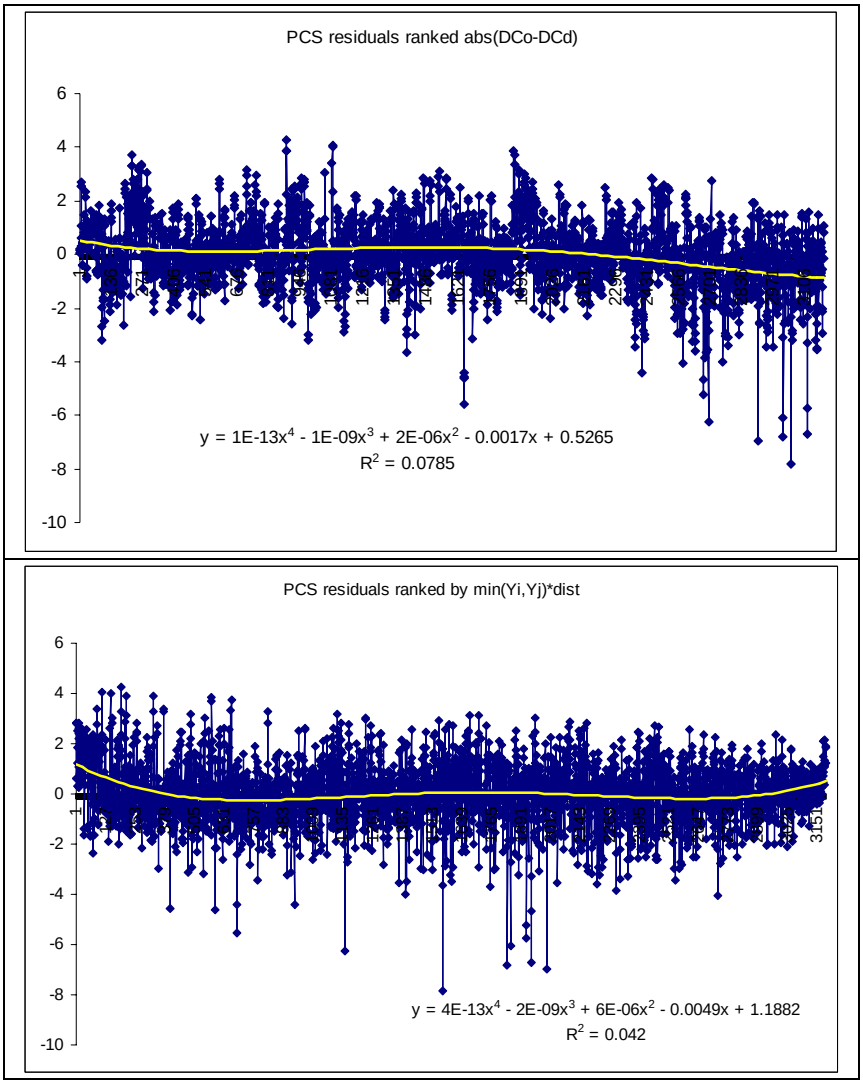


Figure 23: Residuals from PCS and FE models.

The result is shown in the top panel of Figure 22, with an Excel trendline (4th order polynomial) added for emphasis. The North-South flows are at the right and they are systematically negative. In other words the gravity model works systematically worse for North-South trade, a point that has been know to gravit-istas for decades.

The second ranking is a bit more subtle. It ranks the residuals by the minimum of the real GDP of the smaller partner times the bilateral distance. The idea is that made trade with small distant nations is more influenced by unusual, unquantifiable factors. The result, plotted in the bottom panel suggests deviations on both ends of the spectrum. The observation further to the right is US-Japan (the smallest is big and they are far apart), the leftmost observation is Netherlands-Belgium.

Honestly, I can't say exactly what all this means apart from the fact that we should be looking at residuals instead of relying purely on summary statistics.

9.1.3. **Mistaken policy inferences**

Cheng and Wall do the same exercise throwing in fairly standard dummies for preferential trade agreements. I won't go through their policy results in detail since there are not directly relevant, but one point jumps out. They provide an useful example of how one can be misled by the pooled cross section

(PCS) estimator employed by Rose (2000). Cheng-Wall estimate the impact of the Israel-U.S. FTA using the various estimators. Plainly there is something special about this trade relationship that would be impossible to quantify accurately a variable that was valid for a panel of 100+ nations. Allowing for fixed effects, this je-ne-sais-quoi is swept away and the Israel-U.S. FTA is found to have no significant impact on trade, a sensible result in my opinion. With the PCS method however, the FTA picks up all the je-ne-sais-quoi; the estimates suggest that the FTA boosts bilateral trade by an unbelievable 5 times. That's larger than all but the largest estimate of the Rose effect! Interestingly, the Rose-Wincoop method also finds a significant but much smaller effect, while the other methods find the FTA to be insignificant.