Currency Unions and Trade:
The Effect is Large

Andrew K. Rose[1]

June 6, 2001

Et Tu, Torsten?

This critique of my *Economic Policy* paper generates many emotions. It is certainly an honor for my research to be taken seriously by such an eminent scholar as my friend Torsten Persson. And while any publicity is good, a negative citation dominates a positive one.[2] There is irony; I came up with the original idea in a hallway conversation with Harry Flam a few feet from Persson's office. There was indubitably surprise, especially at the AEA meetings where Persson's research was first presented. There is fear; neither *Economic Policy* nor I want a reputation for publishing sloppy work. Persson is a ruthlessly efficient scholar with an almost fanatical devotion to the truth. But even if I wasn't expecting a Spanish Inquisition, honor demands a satisfactory response.[3]

Defense

My original estimate of the effect of currency union on trade was large, at least compared to my intuition (there were no extant estimates in the literature). The surprising thing was how difficult it was to reduce this effect. More precisely, the coefficient (denoted γ) on a currency union (CU) dummy in an empirical "gravity" model of bilateral trade seemed doggedly positive, and significant in both economic and statistical terms. Its value rarely fell below 1.2, implying an effect of currency union on trade of at least $e^{1.2} \approx 300\%$.

I tried hard – plenty hard – to reduce the size of the effect. I went to pains to perform sensitivity analysis, providing over fifty estimates of the effect of currency union on trade. I cut my sample in different ways, added many factors, took simultaneity seriously, measured

[1] University of California, Berkeley, CA 94720-1900. Tel: +1 (510) 642-6609, Fax: +1 (510) 642-4700. E-mail: arose@haas.berkeley.edu, URL: haas.berkeley.edu/~arose

[2] Negative citations are a more serious indication of interest and sometimes generate further work … and citations.

[3] The honor of my son, who inspired the paper. Asher was born a few weeks after my return from Stockholm; nearly all the work was done within a month of his arrival. He is thanked in the introductory footnote, and his website contains pictures of us working together on the manuscript.

variables in different ways, and so forth. I can honestly say that none of the suggestions given to me have been ignored; many appear in the published paper.[4] Others have replicated my results. For instance, Ken Rogoff assigned his Harvard graduate students last Spring a "search and destroy" mission; they were assigned to replicate my results and find the silver bullet. Such is the risk associated with posting your data sets on the web. To my knowledge, all accomplished the former, and none the latter. Such is the joy.

I made a number of caveats in my original paper. Some of these might seem relevant to Persson's paper, but are not. For instance, I cautioned about applying my results to EMU since pre-EMU currency union members tend to be smaller and/or poorer than those in EMU:

> "In 330 observations two countries trade and use the same currency. Many (though not all) of the countries involved are small, poor or both, unlike most of the EMU-11. Thus, any extrapolation of my results to EMU may be inappropriate since most currency union observations are for countries unlike those inside Euroland. (p.15)
>
> "Of course, the effects may be overstated for modern industrialized countries like those in EMU. Still, if my estimate of γ is over-stated by a factor of five, the growth of trade inside EMU would still be large." (Fn 31)

Persson's critique is not the applicability of my results to countries like Argentina, Sweden and the UK, which are considering currency unions, but the effect of currency unions on trade *for countries comparable to those that are actually in (or have been in) currency unions*.[5]

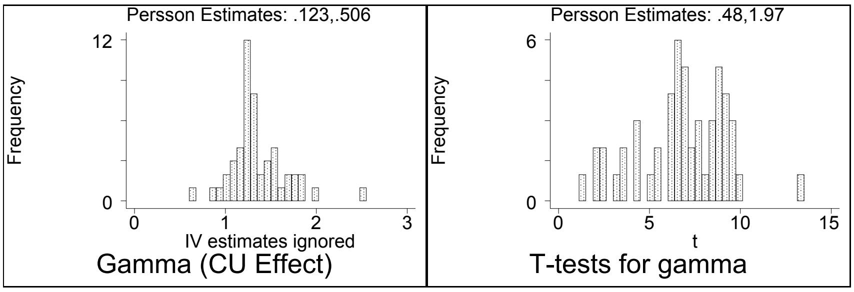
I also urged people not to take the exact size of the key estimate too literally, given my doubts about its size:

> "Without taking the precise γ estimates too literally, it seems clear that trade is substantially higher for countries that use the same currency, holding other things equal." (p.17)
>
> "It is clear that a common currency should encourage trade. The puzzle in this paper is that the effect seems to be so enormous. Why does sharing a currency have such a big effect on trade? I don't know. … we simply don't know why a common currency seems to facilitate trade so much." (pp 31-32)

[4] For the record, most of my sensitivity analysis was done *before* the paper was circulated to *Economic Policy* and was intended to pre-empt criticism, not respond to it.

[5] I handle the applicability issue in my work with van Wincoop.

Yet despite all the robustness checks, Persson's (non-regression based) results are miles from mine in both economic and statistical terms; too far for the caveats to be of relevance. Consider histograms of my estimates for γ and its associated t-statistics:



What's going on?

The Eye of the Beholder

Persson argues that my estimate of γ is high for two reasons: non-linearities and non-random selection. Say it ain't so.

Persson's motivation for the importance of non-linearities is uncompelling. He cites the literature without providing empirical evidence. But evidence there is. In my original paper, I provided: a) sample sensitivity analysis; b) non-parametric estimation; and c) direct inclusion of a variety of non-linearities (e.g., interactions between currency union and gravity regressors, and quadratic terms). None of these results support his case; all support mine; all are ignored.[6]

Persson's motivation for the importance of non-random selection is also less than compelling. Persson relies on country-pairs inside currency unions being dissimilar from those not in currency unions. Persson argues "As Table 1 shows, pairs with a common currency are smaller, poorer, and geographically closer together; they more often share a common language, common borders, …" Yet, Danny Quah, one of the discussants of my original paper, followed up

[6] Again: Persson states "For instance, sharing a common colonial history … and a common language as well as geographical proximity, might reduce trade costs between two countries much more than the sum of the partial effects of the three features." It *might*, but it *doesn't*. Interactions between the three colonial dummies and the common language dummy are indeed jointly significant when added to my default specification. But γ is robust at 1.15 (se=.14). Interactions between the colonial factors and distance deliver γ=1.10 (.15); interactions with common border imply 1.19 (.14). Adding all three sets of interactions result in γ =1.08 (.16).

Parenthetically, I use the same data set that Persson uses in my results below and can replicate his results exactly. There are a few minor coding errors in this data set associated with official Belgian and Swiss languages; these are discussed on my website.

a suggestion made by me and tabulated (in his Table 9, essentially the same as Persson's Table 1) the means of trade and the key gravity regressors for both currency union and non-union observations.[7] He wrote (p. 38): "Glancing down the main columns, the two groups seem surprisingly alike…" I look at Persson's Table 1 (or, equivalently, Quah's Table 9) and agree with … Quah. The two sets of observations do not seem far apart in economic terms, especially given the variation in the data.[8]

At the outset: I am uncomfortable with Persson's unfamiliar techniques. I am unhappy with the fact that the techniques substitute matching for conditioning. We know that currency union countries tend to trade less, simply because they are poor and/or small. That is, *unconditionally there is less trade between currency union members than between non-members*, a fact I reported in my original paper (p. 15).[9] I tried to take these facts into account through conditioning trade on country income, etc. through the gravity model. Yet this conditioning is only done implicitly (if at all) through the propensity score. People are more homogeneous than pairs of countries, so this technique may be better suited to medicine or labor economics (where it originated) than to modeling bilateral trade flows. There are also other, more technical issues.[10]

An Equation Too Far

Persson's technique critically relies on an empirically successful (probit) model of (bilateral) currency union membership. He uses this to generate his propensity score, which is in turn used to stratify and match the data. Yet when I had to model the same thing to generate my instrumental variable estimates for γ, I stated:

---

[7] Footnote 6 which appears in the working paper version (e.g., CEPR DP No. 2329) but was cut by the editorial staff, states "The average values of the key gravity regressors for currency union observations are below but close to those for the rest of the sample."

[8] Again: Persson needs systematic links between currency unions and gravity regressors. His "first cursory look … suggests that such a correlation is present in Rose's data set. Indeed the correlation is 0.2 or higher (in absolute value) for six out of the nine variables …" He could have alternatively written "a correlation is not present in Rose's data set. The correlation is 0.3 or lower (in absolute value) for nine out of the nine variables."

[9] Still, the ratio of *aggregate* trade to output is economically and statistically significantly higher for currency union countries compared to other countries, as shown in Table 1 of Rose and Engel (2000) and Table A2 of Frankel and Rose (2000).

[10] For instance, I know nothing of its small sample properties (and Persson's technique reduces my sample considerably, sometimes by 98%). Propensity score uncertainty isn't included in the standard errors of the currency union effect. How important is the 1:1 nature of the matching? Would 2:1 result in dramatically different results?

> "I treat these results with a grain of salt, given the difficulty of finding appropriate instrumental variables for the incidence of currency unions … it is difficult to find good instrumental variables for common currency arrangements." (p. 31)

Persson's strategy for modeling currency union membership is the opposite of an identifying scheme; he includes everything as a conditioning variable (rather than excluding variables to achieve identification). How does he do?

Persson's probit model of currency union membership lacks theoretical underpinnings. There is an almost complete absence of the political economy considerations relevant to the decision to join a currency union.[11] Mundell identified business cycle synchronization and labor mobility; others have added pass-through considerations, risk sharing, central bank credibility, and political considerations. All of these are absent in Persson's model. Their absence raises the question of whether it is legitimate to condition on the observables chosen.

The nerd within me has technical objections to the model. I am disturbed by the absence of time effects, which are usually significant in such applications.[12] There are also no country effects.[13] More generally, Persson does not seem to take into account the panel nature of the data set. But that is critical, since country-pairs enter the data set repeatedly and are decidedly non-independent. For instance, using robust covariance estimators (as I did in my gravity work) reduces his t-statistics considerably.[14] Adding other plausible gravity regressors such as landmass, landlocked, and island status might matter to Persson's results.[15]

I am being unfair. Robustness is relevant only if the model performs well. But Persson's model performs poorly. The pseudo-R^2 statistic of .48 is misleading. As Persson and I both stress, 99% of my observations are for pairs of countries *not* in currency unions; thus the model will fit well simply by predicting that *no* country pair is in a currency union. And indeed Persson's probit model (which generates the propensity score) mis-predicts 84% of the

[11] And some of the evidence on trading costs – e.g., the effects of regional trade agreements – goes the wrong way, as Persson acknowledges.

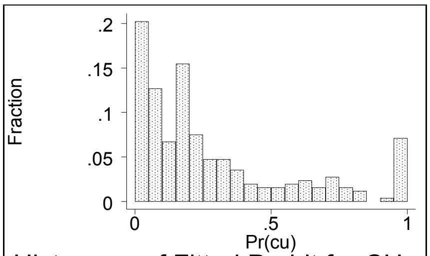
[12] And they're indeed significant at the .000 level when added to Persson's model.

[13] They make a difference. Adding country-effects dramatically reduces the sample size and changes Persson's results considerably. For instance, instead of output having a negative and significant impact on currency union membership, with country-effects it has a positive and insignificant one.

[14] E.g., the t-statistic on GDP falls from 7.5 to 3.8. While only the propensity score is used, the coefficients of the probit are still auxiliary checks on the model. Also, Persson's estimates seem to wander over time, when estimated on individual years.

[15] The hypothesis that they're insignificant is rejected at the .0000 level.

observations for actual currency unions, but only .02% of the observations for non-unions. A histogram of the propensity score (which predicts currency union membership) *for actual currency union members* is revealing:[16]



Histogram of Fitted Probit for CUs

Nevertheless, a well-functioning probit model for currency union membership seems critical to Persson's techniques; it drives his matching and stratification. If the model fits poorly, this jeopardizes the validity of stratifying into bins. In particular, most observations are then bundled into the lowest probability bin, making this essentially an unconditional test. That seems to be exactly what has happened: the bins with p<.25 have over 60% of the actual currency unions (and 98% of the non-unions).[17][18] Alternatively, if his probit model fits poorly, then countries that are actually in but are not predicted (by the model) to be in currency unions can be matched implausibly.[19]

Consider an out-of-sample test (so long as I am being unfair). Persson's model of currency union predicts: bigger, richer countries are *less* likely to join currency unions; a land border has *no* effect; a common language *adds*, and a free trade agreement *detracts* from the likelihood of countries belonging to a currency union. Oops! Persson's model predicts the likelihood of El Salvador dollarizing at less than .0005; it did so in 2001. Ecuador dollarized in 2000; Persson's model predicted this event would occur with probability .0001. The model works better at predicting EMU. But not much: the median likelihood of a country-pair joining EMU is .0009 and the highest prediction for any country-pair is only .02.

16 The analogue for non-unions consists essentially of a spike of probability at the extreme left (i.e., 0).
17 That's why Persson only discards a small fraction of his observations in the stratification procedure; the common support is big, since so much of the mass is close to zero.
18 Even then, his bin with p<.1 rejects equal trade resoundingly (in the direction of my finding); it is overturned in the aggregate by instances where his model mis-predicts (i.e., predicts a currency union which doesn't exist).
19 For instance, the 1990 US-Panama currency union observation is matched to 1990 trade between Zimbabwe and Norway; the 1985 observation is matched to 1985 trade between Pakistan and Denmark.

Succinctly, Persson's probit model is problematic. It lacks theory, fits and predicts poorly, and seems sensitive to small reasonable empirical perturbations (in exactly the way I took pains to show that my original estimates were not). As a Californian living near a fault line, I am nervous about building an elaborate structure on shaky ground.

Offense

In a choice between using more data or more technique, I choose the former. Others may choose differently. But one can always criticize the exact properties of a particular estimator. And a good data set is hard to beat.

My original (United Nations) data set was large: it covered 186 countries from 1970 through 1990. In ongoing work with Reuven Glick (FRBSF), I have developed a new data set is very large. It covers over 230 IMF country codes between 1948 and 1997, and includes all bilateral trade covered in the IMF's *Direction of Trade*. Only about 1% of the sample covers currency unions (as with my earlier data set). But a number of the currency unions that existed at the end of WWII dissolved during the sample (almost always before 1970), a fact I shall exploit below.[20]

Some descriptive statistics for the new data set follow. Sample means for the key gravity regressors are similar for currency unions and non-unions (at least to me), the only exception being the common language variable:

Table 1: Descriptive Statistics for IMF Data Set

	Non-Unions	**Currency Unions**
Observations	422,987	4,255
Log Real Trade	10.7 (3.7)	10.5 (3.1)
Log Distance	8.2 (.8)	7.1 (1.0)
Log product GDP	47.9 (2.6)	44.7 (3.0)
Log product GDP/capita	16.1 (1.4)	14.4 (1.6)
Common Language Dummy	.15 (.35)	.85 (.36)
Land Border Dummy	.02 (.14)	.18 (.38)

Means, with standard deviations reported in parentheses

[20] As with all my data sets, this one will be web-accessible when the working paper is released.

The standard gravity equation works well on the new data set to estimate the effect of currency union on bilateral trade. Using Persson's notation, I estimate:

$$t_{it} = \gamma C_{it} + \varphi X_{it} + \varepsilon_{it}$$

OLS estimates of γ and a few of the more important gravity coefficients are reported below. The model delivers a γ estimate that is comparable to and slightly *higher* (in both economic and statistical significance) than my original (UN-based) estimates:

Table 2: Pooled OLS Gravity Estimates from IMF Data Set

Currency Union	1.41 (11)
Log Distance	-1.11 (47)
Log Product Real GDPs	.93 (93)
Log Product Real GDP/capita	.45 (29)
Common Language	.37 (8.7)
Common Land Border	.40 (3.4)
Observations	219,558
R^2	.64
RMSE	2.02

Intercept and year controls not recorded. Other controls not recorded: a) regional FTA membership, b) # landlocked; c) # islands; d) area; e) common colonizer; f) colony/colonizer; g) common country.
Absolute t-statistics (computed with standard errors robust to country-pair clustering) in parentheses.
Annual data for 231 countries, 1948-1997.

The gravity model also works well in a purely cross-sectional sense.[21]

Low Tech, High Ec

Persson's work violates Frank Fisher's Iron Law of Non-Linear Econometrics: "Don't do it." While I did some high-tech estimation in my original paper, I tried to keep things simple (plausible). The only obvious technique I did not pursue in the paper was a fixed-effects estimator. I did this for the reason given in footnote 25: "The paucity of countries that either

---

[21] Estimates of γ for individual years include (t-statistics in parentheses): 1950: 1.14 (3.3); 1955: 1.17(4.2); 1960: .85 (4.4); 1965: 1.07 (6.9); 1970: 1.44 (7.0); 1975: 1.44 (6.1); 1980: 1.34 (5.7); 1985: 1.88 (8.1); 1990: 2.48 (9.9); and 1995: 1.61 (7.0).

join or leave currency unions means that a time-series "within" estimator (i.e., one that exploits only country-pair fixed effects) is untenable." Persson agrees: he writes (p. 2): "As there are very few regime changes – countries entering or leaving common currencies – in the data, the treatment effect of a common currency on trade has to be identified from the cross-sectional variation."

My original reason for eschewing fixed-effects estimation seems warranted. For the sake of science I tabulate the actual "within" estimates from the original data set below, along with a random effects estimate that is even more demanding in terms of econometric courage. The point estimate for γ is negative but insignificantly different from zero given the very large standard error (shades of …). The latter is inevitable: there were only eight switches in currency union status during my sample.[22]

Table 3: Pooled Panel Gravity Estimates from UN Data Set

	Fixed-effects ("within")	**Random-effects GLS**
Currency Union	-.38 (0.6)	1.23 (6.0)
Log Distance		-1.16 (34.8)
Log Product Real GDPs	1.35 (15.0)	.80 (84.3)
Log Product Real GDP/capita	-.16 (1.9)	.60 (35.8)
Common Language		.39 (5.4)
Common Land Border		.69 (4.1)
R^2: Within	.09	.09
R^2: Between	.44	.65
R^2: Overall	.43	.63
Hausman Test (p-value)		.00

22,948 observations in 6,707 country-pair groups. Obs per group within [1,5], mean=3.4.
Intercepts and year controls not recorded. Other controls not recorded: a) regional FTA membership, b) nominal exchange rate volatility; c) common colonizer; d) colony/colonizer; e) common country.
Absolute t-statistics in parentheses.
Annual data for 186 countries at 5-year intervals, 1970-1990.

[22] Ireland left the pound; Mali joined the CFA franc zone, and so forth.

This is a shame, since the fixed effect estimator is the most appropriate way to exploit the panel nature of the data set without making heroic assumptions.[23] The fixed-effect estimator exploits variation over time, *which answers the policy question of interest*, namely the (time series) question "What is the trade effect of a country joining (or leaving) a currency union?"[24] It estimates γ by comparing trade for a pair of countries before CU creation/dissolution to trade for the sample pair of countries after CU creation/dissolution. This comparison of "like" to "like" removes the need for a matching technique. It also counters the important Anderson-van Wincoop critique that is ignored by Persson.

Fortunately, the new IMF data set has enough time-series variation to estimate γ precisely using a fixed effects estimator. Instead of 8 switches in bilateral currency union status (for the UN data set), there are 146. Thus a within estimator is eminently feasible. It delivers:[25]

Table 4: Pooled Panel Gravity Estimates from IMF Data Set

	Fixed-effects ("within")	**Random-effects GLS**	**Between Estimator**	**Maximum Likelihood**
Currency Union	.74 (14)	.82 (16)	1.57 (6.7)	.80 (15)
Log Distance		-1.35 (43)	-1.42 (44)	-1.36 (37)
Log Product Real GDPs	.05 (5.8)	.27 (38)	.98 (74)	.23 (30)
Log Product Real GDP/capita	.79 (56)	.52 (45)	.45 (22)	.57 (47)
Common Language		.29 (4.6)	.43 (6.9)	.27 (3.8)
Common Land Border		.52 (3.2)	.48 (2.9)	.53 (2.8)
R^2: Within	.12	.12	.11	
R^2: Between	.23	.52	.63	
R^2: Overall	.23	.47	.58	

219,558 observations in 11,178 country-pair groups. Obs per group within [1,50], mean=19.6.
Intercepts not recorded. Other controls not recorded: a) regional FTA membership, b) # landlocked; c) # islands; d) area; e) common colonizer; f) colony/colonizer; g) common country.
Absolute t-statistics in parentheses.
Annual data for 231 countries, 1948-1997.

[23] There are only two drawbacks to the "within" estimator: the impossibility of estimating time-invariant factors, and a potential lack of efficiency. While I am prepared to treat second moments as being of second order, the first issue effectively precluded estimating γ since currency union status changes so infrequently.

[24] As opposed to the cross-sectional question of "How much more do countries within currency unions trade than non-members?"

[25] Adding time-effects delivers similar results.

The bad news is that an estimate of $\gamma \approx .74$ is certainly smaller than my original estimate of $\gamma \approx 1.2$. Since $e^{.74} \approx 2.1$, the fixed effect estimator implies that currency union doubles trade. The good news is that the effect is still economically and statistically significant, and entirely consistent with the spirit of my original paper.[26] It is also interesting to note that Persson's matching estimator delivers a positive and statistically significant effect of currency union on trade when applied to the new data set.[27]

Before the Curtain Falls

To isolate the effects of currency unions on trade, one needs to model trade or currency unions (or both). Modeling bilateral trade is easy. Modeling currency union membership is not. As I took pains to point out originally, the gravity model is one of the most – if not *the* most – successful models of applied economics. *But gravity is a model of trade, not currency unions.* Modeling the latter is fraught with peril. I could avoid that issue in my original paper (absent endogeneity considerations); Persson cannot. The validity of his technique hinges on an implausible model of currency union membership; hence my distrust of it.

I think that Persson is precisely correct that γ seems implausibly big (as I have always maintained). I also think he is exactly wrong in his diagnosis of the reason and the solution. Defending his technique requires considerable bravery; it is poorly suited to this problem and consequently unreliable. More importantly, a new larger data set confirms my results, using his technique, my technique, or the preferred fixed-effects technique that neither of us could use on the original data set.

As the Fat Lady Sings

Persson concludes: "It is worthwhile to study – theoretically and empirically – the selection into common currencies jointly with the effect of those currencies on trade or other

---

[26] Adding quadratic terms for GDP and GDP/capita delivers an estimate of γ=.66 with a t-statistic>12.
[27] More precisely, I used the propensity score from Persson's model to match currency union observations with non-union observations that had propensity scores within .000001 of the score for the union observations. This is Persson's "radius" estimator. When I apply this technique to the UN data set, I get, like Persson, a currency union effect on trade that is negative and verges on significance. When I repeat this on the IMF data set, I get a *positive* currency union effect on trade with a t-statistic of 4.3. This effect is essentially unchanged if I use a radius of 0000001.

variables." As I stated originally: "…Research on the determinants of currency unions remains an interesting research issue." (p. 31). It seems like we should agree to agree.

But perhaps not. In my original paper, I provide north of fifty estimates of the γ, the effect of currency union on trade; the smallest effect was $\gamma \approx .67$ (with a t-statistic exceeding 4). Since then, I've provided twelve more estimates (in my work with Engel, Frankel and van Wincoop), and a further five in this work; Persson has contributed another six. *Almost all these estimates are large*. Persson's dubious stratification estimator delivers his lower-bound estimate; it implies trade will increase 13% as a result of currency union, though Persson implies that the effect is probably around 40%. All the other estimates are higher, most of them substantially so.

To conclude, two conclusions. First, we seem to be converging to an estimate of the effect of currency union on trade that is economically very large indeed. Second, we seem to be converging to an estimate that is much higher than seemed plausible to me when I started this work. It's almost unpatriotic to suggest that further work in the area isn't needed; but it seems either we need a radically new approach, or … to adjust our priors.

References

Anderson, James E. and Eric van Wincoop (2001) "Gravity with Gravitas: A Solution to the Border Puzzle" NBER WP No. 8079.

Frankel, Jeffrey A. and Andrew K. Rose (2000) "Estimating the Effect of Currency Unions on Trade and Growth" CEPR DP 2631.

Rose, Andrew K. (2000) "One Money, One Market: Estimating the Effect of Common Currencies on Trade" *Economic Policy*.

Rose, Andrew K. and Charles Engel (2000) "Currency Unions and International Integration" CEPR DP 2659.

Rose, Andrew K. and Eric van Wincoop (2001) "National Money as a Barrier to Trade: The Real Case for Monetary Union" forthcoming *American Economic Review*.