Many Unions, One Estimate? Disaggregating the Currency Union Effect on Trade

Joseph Kopecky

Published online: 15 May 2023.

Submit your article to this journal ⬀

View related articles ⬀

View Crossmark data ⬀

Routledge
Taylor & Francis Group

OPEN ACCESS   Check for updates

Many Unions, One Estimate? Disaggregating the Currency Union Effect on Trade

Joseph Kopecky

Department of Economics, Trinity College Dublin, Dublin, Ireland

ABSTRACT

A large literature estimates the impact of currency unions on trade. Often ignored in these estimates are the dramatic differences in the characteristics of countries adopting common currencies, hidden by aggregation into a single currency union effect. I show that currency unions have substantial differences in their observable characteristics, relative to non-unions, making them a poor comparison group for estimation of policy treatment. Further, these differences are heterogeneous across individual currency unions, making one aggregate estimate likely inappropriate. Using inverse propensity score methods, I find that adjusting these gravity equation estimates to account these differences, both via weighting and via sample adjustment, meaningfully impacts the estimated policy effects. I find a wide range of currency union effects across individual, disaggregated, currency unions. My results suggest that future work on currency unions, and other macroeconomic policies, should be careful to check for such underlying heterogeneity when estimating policy effects.

1. Introduction

Happy families are all alike; every unhappy family is unhappy in its own way.
 - Leo Tolstoy, Anna Karenina

The currency union effect on trade is a contentious topic, with revised attention in recent years as many have studied the impact of the European Monetary Union (EMU). Far less work has explored the disaggregated effect of other currency unions, which largely consist of emerging market economies. In this article I make two contributions to this literature. The first shows that estimates of currency union trade impact generally fail to account for the imbalance that exists between treatment and control samples on observable characteristics. Nonunion observations have substantial observable differences from countries in a currency union, making them likely a poor comparison group. My approach corrects for this and suggests that this may account for the variability in existing currency union estimates. Second, I show that estimating effects individual currency unions reveals substantial heterogeneity, not only in the estimates of currency union effects across disaggregated unions, but also in the nature of this imbalance between members of these groups and non-currency union observations. Most currency unions studied in large trade datasets are developing countries using either a colonizer's currency, that of another large, developed nation, or who are in multilateral arrangements with other developing economies with their common currency pegged to a large country. I find evidence that modeling the choice of entry into a disaggregated currency union, using propensity score methods to create a closer comparison group to members within that union, substantially changes the estimates of the trade impacts of these policies. My findings suggest that one-size-fits-all measures of these policies are

likely quite poor and provide misleading evidence of their impact on bilateral trade. While trade improvement is not the only factor in determining whether a currency union is an optimal[1] policy choice, it reflects one of the large potential benefits. Improving and better contextualizing these estimates will better inform these monumental policy decisions.

Since the seminal work of Rose (2000), a great deal of effort has gone into the estimation of the currency union effect on trade. Work such as Nitsch (2002) purported to have shrunk the eye-popping estimates from the original work which were by the author's own admission "*embarrassingly and implausibly large*" (Rose 2002). These estimates were smaller, but still substantial when utilizing the time-series variation among currency unions in Glick and Rose (2002). With a growing time-series for one of the largest currency union policy changes in history, a large body of research has studied the EMU, generally finding smaller effects than those of Glick and Rose (2002). Rose (2017) surveys these estimates and suggests that the reason behind the diminished EMU effect in other research comes from limiting the sample to include only large, rich, economies, preferring the Glick and Rose (2016) estimate that implies a 50% increase in trade among members. In a meta-analysis of these estimates, Polák (2019) suggests a range somewhere between 2% and 6%, though this work focuses on publication bias in choosing estimates (Rose suggests the literature average is more like 12% focusing on only author preferred estimates), and not on appropriateness of sample choice. While my focus is largely on emerging market economies, my results also speak to this literature, showing that properly weighting and truncating the sample to account for comparability between currency union and nonunion members does change estimates in a downward direction, likely improving them.

The primary focus of this work, and much less well established, are the effects of non-EMU currency unions. It is somewhat understandable that the creation of such a large multilateral currency union arrangement would create such interest, but much can be learned from the other currency unions in the Glick and Rose (2016) sample, which for the most part are either small developing countries using a larger trading partner's currency, or a group of relatively small countries adopting a common currency. Glick and Rose (2016) does provide an estimate for disaggregated unions, suggesting dramatic differences among these groups. Saiki (2005) shows that for developing countries who use a larger country currency that the currency union effect is asymmetric. Their work finds that exports from the USA and France to dollarized and CFA Franc countries increase in response to the policy, but that the effects are neutral on imports into those larger countries from their developing counterparts. Campbell and Chentsov (2017) consider the disaggregated currency unions used in my sample (and others), showing that the historical context matters for these policies. In similar work, Campbell (2013) showed that much of the aggregate currency union effect on trade is driven by the major geopolitical events causing currency union dis-allusion, largely decolonization and war. Other works such as Head, Mayer, and Ries (2010) and Glick and Taylor (2010) suggest that specific historical context may matter a great deal. While my work provides a simple, and data-driven, way of trying to adjust for context, the large impact I find on estimates supports work that takes such context seriously.

Adopting the language of the causal inference literature, consider currency unions as the *treated* group. Non-currency union observations are thus a *control*. If country pairs were randomly assigned into treated and control groups, then one could simply take a weighted average of the estimates to arrive at an average treatment effect. Of course, this is not how these relationships are formed. Controlling for various observable characteristics and fixed effects can help improve estimates given the *selection on observable* assumption, but still falls far short of causal inference in cases where selection into the policy treatment is endogenously determined. I use two different estimators which weight the estimates from gravity equations with probabilities of treatment derived from estimations of selection into currency unions. The inverse probability-weighted regression adjustment (IPWRA) estimator, described in detail in Imbens (2004),[2] Lunceford and Davidian (2004), and Wooldridge (2007), simply incorporates these weights into a standard gravity equation. The augmented inverse probability weights (AIPW), described in detail in Glynn and Quinn (2010), augments an inverse probability weighted (IPW) estimator of group means with regression adjustments. These estimators seek to resolve to problems with previous estimates of the currency union effect on trade. I show that

weighting alone often fails to bring treatment and control groups to match on observables, but that doing so along with truncating the estimation sample to exclude worst fitting controls gets fairly close.

One of the many critiques leveled against existing aggregate gravity equation estimates[2] of the currency union effect on trade is that they fail to account for the endogeneity of choice of currency unions. Countries enter into these arrangements as a result of factors that will bias estimates of their trade relationship. Some existing work has attempted to address these selection problems. Persson (2001) does this on the original aggregate currency union estimate from Rose (2000), finding smaller implied estimates. Chintrakarn (2008) carries out a similar exercise on the EMU, using matching estimators to show a reduced trade effect. In both cases using estimates of probability of entering the union to create matched treatment and control groups weakens the currency union estimate. Rebuttals point out that these estimators throw out the majority of data, often choosing one or a few observations to use as comparisons for each observation of a country-pair in a currency union. Matching estimators such as those of Persson (2001) and Chintrakarn (2008) use similar logic to my approach. However, the IPWRA and AIPW estimators I use are better able to balance the potential usefulness of the well-studied gravity equation approach with these probability weighted estimators.

My estimation method also shares something in common with estimates from Barro and Tenreyro (2007), who use an IV estimation motivated by the model of Alesina and Barro (2002), whereby probit estimations are used to estimate the probability that a country adopts the currency of one of six potential anchor countries. Their approach is particularly concerned with exactly the selection issue that mine seeks to address. They calculate the probability that two countries adopt a common anchor as an instrument in a bilateral trade regression, finding quite large estimates in line with Rose (2000). My estimation strategy is a more reduced form approach that makes fewer assumptions about the reasoning behind currency union membership. Unlike theirs, I focus on the disaggregated currency union effect, while they estimate one effect across many currency unions. It is hard to directly compare to their results, as they do not incorporate the time-varying country fixed effects suggested by Head and Mayer (2014), nor the dyadic pair effects which substantially decrease the trade effect in Glick and Rose (2002) and are shown by Baldwin and Taglioni (2007) to be important empirically in estimation of currency union effects. However, my estimates suggest that an update of their work, which focuses specifically on the motivations of unions who adopt common currencies to peg against a larger country,[3] would be useful on a disaggregated basis.

In Section 2, I describe the data and methods used. Here I provide further motivation by showing that union and nonunion pairs have large differences in observable characteristics and define the estimators I use to resolve these differences. I present estimates for standard unweighted gravity equations, the IPWRA on a full and truncated sample, and the AIPW estimators in Section 3. Section 4 repeats these estimates using the more robust PPML estimator. Section 5 contextualizes these findings, provides some discussion for paths of future work, and concludes.

2. Data and Methodology

My estimations use the International Monetary Fund's *Direction of Trade* data. For ease of comparison, I will use the same coverage as Glick and Rose (2016), whose study contains bilateral trade data from 200 countries from 1948 to 2013. Before discussing estimation methods, I establish some motivating evidence that the observable characteristics of countries in currency unions are quite different from non-currency unions. Moreover, these differences change depending on the currency union relationship at hand, suggesting that not only is it important to carry out some adjustment to rebalance these groups, but also that a one-size-fits-all approach to doing so is impossible in this context.

I consider the largest seven currency unions in this sample, as measured by the number of country-pair-year observations. I limit to these as restricting my outcome model of trade to *"theory consistent"* method suggested in Head and Mayer (2014), which includes the use of a large number importer-year and exporter-year fixed effects, which at times makes estimation of unions with few observations

infeasible. The unions I consider are: the European Monetary Union (EMU), the CFA Franc, the East Caribbean Currency Union (ECCU), the Indian Rupee zone, and countries using Australian dollars, US dollars, and British pounds. The purpose of this study is to explore the extent to which directly modeling the individual currency relationship affects estimates of trade. Restricting my estimates to the Glick and Rose (2016) data improves comparisons to well-known estimates, but it is likely that this limits the ability of my propensity score model to fully contextualize the many reasons why currency unions form and break. However, it will become clear that even a stylized model of selection is able to substantially improve the comparability of currency union and nonunion members, with large implications for their estimated effects on trade.

Table 1 shows the mean and standard deviation for four variables under policy treatment (currency union) and control (non-currency union) groups. The first two of these are standard gravity equation variables, the log product of the bilateral pair GDP, and GDP per capita. The second two reflect the absolute value of the difference between these two variables in exporter and importer countries. These latter values are of particular interest as they distinguish between trade partners of relatively similar size and income (such as in developing-to-developing pairs), or those with large imbalances in these characteristics. I do not report means for the full sample as they would be in general indistinguishable from the control sets, which can be seen to vary only trivially across these different currency union groups. This is because even the aggregate currency union only consists about 4.2% of the overall sample, so the non-treated observations make up the vast majority of total observations. The bottom of the table reports the difference in treated and control sub-populations. Large positive or negative deviations here mean that the currency union in question has large observable differences in characteristics from the country-pairs that are being used as a comparison in standard estimates. Random assignment of policy would make these differences zero. While macroeconomic observational studies will never match this experimental ideal, research in these areas cannot ignore the problems this poses for estimation of an average treatment effect.

Consider the two multilateral arrangements of developing countries: the CFA Franc and East Caribbean Currency Union (ECCU). These are monetary unions of small developing countries who not only share a common currency, but also peg that currency to a large, developed country (the French Franc/Euro and United States dollar). These pairs have substantially smaller output relative to the non-currency union sample, with the ECCU countries having roughly average living standards, as

Table 1. Summary statistics: disaggregated currency unions.

	(1) All CUs	(2) EMU	(3) CFA	(4) ECCU	(5) Aus. Dollar	(6) British Pound	(7) Rupee	(8) US Dollar
Control (CU=0)								
$log(GDP_1 GDP_2)$	49.49	49.44	49.49	49.47	49.45	49.45	49.45	49.45
	(2.641)	(2.681)	(2.667)	(2.667)	(2.686)	(2.687)	(2.687)	(2.686)
$log(GDPpc_1 GDPpc_2)$	17.47	17.44	17.48	17.45	17.45	17.45	17.45	17.45
	(1.780)	(1.799)	(1.788)	(1.807)	(1.806)	(1.806)	(1.805)	(1.806)
$log(|GDP_1 - GDP_2|)$	25.71	25.68	25.71	25.69	25.68	25.68	25.68	25.68
	(1.903)	(1.941)	(1.918)	(1.928)	(1.941)	(1.941)	(1.940)	(1.938)
$log(|GDPpc_1 - GDPpc_2|)$	8.92	8.90	8.92	8.90	8.90	8.90	8.90	8.90
	(1.42)	(1.44)	(1.42)	(1.44)	(1.44)	(1.44)	(1.44)	(1.44)
Treated (CU=1)								
$log(GDP_1 GDP_2)$	47.15	52.79	45.15	40.76	45.84	49.68	49.30	52.26
	(4.149)	(2.213)	(1.236)	(1.418)	(1.018)	(1.415)	(1.242)	(1.652)
$log(GDPpc_1 GDPpc_2)$	16.56	20.57	14.28	17.79	18.37	17.30	14.11	18.89
	(2.875)	(0.706)	(1.105)	(0.878)	(0.480)	(1.140)	(1.210)	(1.452)
$log(|GDP_1 - GDP_2|)$	23.80	26.84	22.10	19.82	26.88	27.09	27.06	29.46
	(3.029)	(1.431)	(1.328)	(1.386)	(0.431)	(0.165)	(0.992)	(0.542)
$log(|GDPpc_1 - GDPpc_2|)$	7.58	9.11	6.49	7.97	10.07	8.82	5.27	9.43
	(1.90)	(1.32)	(1.53)	(1.09)	(0.35)	(0.52)	(1.78)	(1.07)
$Treat - Con$: lGDP	−2.34	3.34	−4.33	−8.70	−3.62	0.23	−0.15	2.81
$Treat - Con$: lGDPpc	−0.91	3.13	−3.19	0.34	0.92	−0.15	−3.34	1.44
$Treat - Con$: lDiff. GDP	−1.91	1.16	−3.61	−5.87	1.20	1.41	1.38	3.79
$Treat - Con$: lDiff. GDPpc	−1.34	0.21	−2.43	−0.93	1.17	−0.08	−3.63	0.53
N	738876	738876	738876	738876	738876	738876	738876	738876

measured by GDP per capita, while the CFA Franc union is extremely low income. The CFA in particular is the only disaggregated union whose sign on for all four characteristics matches that of the aggregate union.

Notably the most widely studied individual union, the EMU, has seemingly nothing in common with other currency union groups, further motivating a disaggregated approach. The Eurozone represents pairs of countries that are much larger and richer than the average in the sample, reflected by the large positive deviations from their product GDP and GDP per-capita. While they are relatively diverse in terms of size, with a larger difference within the pair in terms of GDP than in the control set, they have somewhat close to average differences in GDP per-capita. So, while the EMU sub-sample is quite rich relative to the controls the differences among trading partners within the EMU are closer to the non-CU average.

The last four unions in this table are a group that all share the characteristic of being pairs consisting of a very large and very small economy. This is reflected in all four having a positive difference across treated and controls in the absolute value of the difference between trading partner GDP. They have little else in common, with the sheer size of US and UK GDP making their product of GDPs larger than the control groups, despite most partner countries being extremely small, while the Australian dollar reflects much smaller joint size. The difference in living standards in both Australian and US dollar arrangements is substantial, reflecting much larger gaps in living standards between the large central country and those adopting their currency than among the average trading partners in the sample, while the opposite is true for the Indian Rupee zone, where per-capita output in India was quite low relative to its smaller trading partners.

There should be two broad takeaways from Table 1. The first is that currency union countries are very different from the average nonunion trading partner in this sample. The second is that every currency union member is different in its own way, with some common themes that can be identified, but no systematic pattern that holds across groups. It is clear that context matters a great deal, and looking at this table one would not expect an aggregate estimate of currency union to do a good job describing the experiences across these groups.

2.1. Currency Union Estimator: The Gravity Equation of Trade

I specify the theory consistent representation of the gravity equation of trade, as described in Head and Mayer (2014). This involves including a full set of exporter-year and importer-year fixed effects to my estimation. These time varying fixed effects should pick up some of the spurious correlations revealed to bias prior estimates in Campbell (2013) but are unlikely to solve all problems of identification and endogeneity. In addition to this, I will follow Glick and Rose (2016), whose preferred estimates include pair specific fixed effects. I wish to keep my gravity specification close to theirs so that my estimates will be easily compared. As a baseline I estimate the following:

$$ln(X_{ijt}) = \gamma CU_{ijt} + \beta Z_{ijt} + \lambda_{it} + \psi_{jt} + \varphi_{ij} + \in_{ijt} \qquad (1)$$

Where X_{ijt} are exports from country i to country j at time t, CU_{ijt} is a dummy variable representing a currency union relationship between the country-pair in year t, Z_{ijt} is an arbitrary set of time-varying controls, λ_{it} exporter-time fixed effects, ψ_{jt} importer-time fixed effects, and φ_{ij} time-invariant country-pair fixed effects. This specification of the gravity equation is identical to the preferred estimates in the section of Glick and Rose (2016) that uses this "newer" theory consistent export model. I will present baseline estimates that are quite close to theirs. Additionally, I will include their regional trade agreement variable in Z_{ijt}, though I will omit other common gravity controls as these are generally captured by the rich set of fixed effects.

One difference between my baseline specification and that of Glick and Rose (2016) is that I use currency unions that are *non-transitive*. This is the definition in their earlier work (Glick and Rose 2002). Transitivity suggests that if countries x and y are in a currency union, while countries x and

z are also in a currency union, then countries y and z are also in one. While it is true that this transitive property accurately reflects a shared currency among countries, it has implications for how I estimate my first stage selection into currency unions. Multilateral currency union arrangements, such as the EMU, will inherently have such transitive properties in the data. This transitive property comes into play for countries who unilaterally adopt a particular currency, such as developing countries using dollars, who then find such *sibling* relationships with other countries who make the same choice. Since the choice to join a currency union *together* is not necessarily made in these cases I choose to ignore these observations, as my later propensity score methods will attempt to model that selection directly. They are dropped from analysis entirely so as to not introduce bias through the control set, though there is little impact on my baseline estimates relative to those of Glick and Rose (2016) who include them and including them in my matching estimates appears to not have a substantial impact.

A common problem associated with the OLS estimates of the log-gravity equation specified in Equation (1) relates to bias that may be introduced in log-linearized estimates in the presence of heteroskedasticity. Silva and Tenreyro (2006) and Silva and Tenreyro (2011) show that this bias is potentially large in the context of traditional gravity equation estimates, proposing a Poisson pseudo-maximum likelihood (PPML) estimator. This has the added benefit of admitting zeros that are common in trade data. As such I will also estimate my the following equation:

$$X_{ijt} = e^{\gamma CU_{ijt} + \beta Z_{ijt} + \lambda_{it} + \psi_{jt} + \varphi_{ij}} + \in_{ijt} \tag{2}$$

using their PPML approach. Here I compare estimates to Larch et al. (2019) who have a comparable PPML estimate to that of Glick and Rose (2016) as a baseline for disaggregated currency union estimates.

2.2. Propensity Score Weighting: Two Doubly Robust Estimators

While it is possible that the multilateral resistance terms, along with my control for regional trade agreements may properly account for problems of selection in Equation (1), the large gap in Table 1 of currency union and nonunion populations warrants some caution. I use inverse propensity score weighting in the estimation of Equation (1) to adjust the regression estimates on the currency union effect of trade as a way to properly balance the characteristics of currency union and non-currency union estimates. I leverage two forms of regression adjustment that have the *doubly robust* property.[4] These estimators model both the outcome variable of interest, here exports, and the selection into treatment groups (currency union membership). The doubly robust property refers to the fact that these estimators should provide consistent estimates of the average treatment effect if *either* of these two models are correctly specified. For more detailed derivation of these propensity score estimators in a macroeconomic context see Jordà and Taylor (2016).[5]

There is a long history of propensity scores in the literature on currency unions and trade, where Persson (2001) used matching methods to determine comparable groups among broad currency union membership to show that under matching techniques the estimates of currency union impacts on trade are much lower. Chintrakarn (2008) uses similar methods on the Euro area, again finding much smaller estimates than those using an unweighted gravity equation. Matching methods inherently throw out a large amount of data by choosing to either use only the best matches, or a small sub-sample of close matches. This is because they are designed to seek identification by attempting to convert observational data to something closer to a randomized control trial, through re-randomization. To my knowledge there is limited work using these regression adjustments with propensity score data in the trade literature, with Millimet and Tchernis (2009) providing one example of using the IPWRA to estimate an effect of the EMU on a much smaller sample. Given that the gravity equation of trade in Equations (1) and (2) are well-studied and theoretically grounded objects, it seems reasonable to assume that the information from this regression adjustment is useful, even if there are gains to be made through the traditional propensity score adjustments.

To the extent that treatment and control samples are significantly different, it is possible to improve non-weighted estimates of traditional gravity estimates by implementing an inverse propensity score weighting with regression adjustment (IPWRA). I follow similar notation to Jordà and Taylor (2016), who provide a full discussion of how these doubly robust estimators are derived in a potential outcomes framework in a macroeconomic context. The IPWRA estimator is:

$$\widehat{ATE}_{IPWRA} = \frac{1}{n_1^*}\sum\left[\frac{CU_{ijt}\left(m_1\left(Z_{ijt},\hat{\gamma}\right)\right)}{\hat{p}_{ijt}}\right] - \frac{1}{n_0^*}\sum\left[\frac{\left(1-CU_{ijt}\right)\left(m_0\left(Z_{ijt},\hat{\gamma}\right)\right)}{1-\hat{p}_{ijt}}\right] \tag{3}$$

Where CU_{ijt} is a country-pair dummy, $\hat{p}_{ijt}$ is the predicted probability of treatment, estimated from a first stage model for the likelihood of currency union membership. The term $m_{0/1}\left(Z_{ijt},\hat{\gamma}\right)$ represents an estimate of the mean for both control and treated (0/1) groups conditional on controls[6] given by: Z_{ijt}. This conditional mean in my case the gravity equation estimation coming from either Equation (1) or Equation (2), with $\hat{\gamma}$ the vector of estimated regression coefficients. In principle, it is possible to allow for separate estimates of $\hat{\gamma}$ among treated and control populations, but because the number of disaggregated currency union observations is relatively small this is not possible in practice due to the rich set of multilateral resistance terms and dyadic fixed effects. Allowing different conditional means on treated and control would require estimating a weighted version of Equation (1) on just the treated subpopulation, and doing so as described would generally imply more parameters to estimate than observations. As such, in practice I must assume that the coefficients estimating these conditional means are the same across these groups. Finally, I normalize the inverse probability weights with $n_1^* = \sum\frac{CU}{\hat{p}}$ and $n_0^* = \sum\frac{1-CU}{1-\hat{p}}$ as suggested in Hirano and Imbens (2001) and Imbens (2004), which simply requires the weights to sum to one within each group. The AIPW is given by:

$$\widehat{ATE}_{AIPW} = \frac{1}{n}\sum\left\{\left[\frac{CU_{ijt}\left(m_1\left(Z_{ijt},\hat{\gamma}\right)\right)}{\hat{p}_{ijt}} - \frac{\left(1-CU_{ijt}\right)\left(m_0\left(Z_{ijt},\hat{\gamma}\right)\right)}{1-\hat{p}_{ijt}}\right]\right.$$
$$\left.- \frac{CU_{ijt}(\ln\left(X_{ijt}\right)}{\hat{p}_{ijt}\left(1-\hat{p}_{ijt}\right)}\left[\left(1-\hat{p}_{ijt}\right)m_1\left(Z_{ijt},\hat{\gamma}\right) + \hat{p}_{ijt}m_0\left(Z_{ijt},\hat{\gamma}\right)\right]\right\} \tag{4}$$

where the first term in this expression is a standard inverse probability weights (IPW) estimator, which simply takes a weighted difference of the group means across treated and control populations. The second term is an adjustment which is mean zero in large samples. A particularly useful property of this adjustment term is that it should provide stability of the estimator when probability weights approach the zero/one extremes, a problem that has been shown to introduce instability in estimates of the IPWRA estimator. These properties make the AIPW estimator potentially useful in the case of trade data, where sample sizes are quite large and can take advantage of some of these benefits, but large fractions of the data in the control group are assigned weights near zero. However, they also put more weight on the IPW first stage model, which in this case is less well theoretically justified.

Work such as Glynn and Quinn (2010) show using a Monte Carlo design that this estimator preforms dramatically better than standard IPW estimator. They do not compare it directly to the properties of the IPWRA. Because some of my disaggregated currency unions have extremely poor fit of their first stage selection equation, I generally prefer estimates using the IPWRA estimator, but report both below. As I will show, the AIPW appears to be quite volatile in cases where the propensity score is poorly modeled.

2.3. Modeling Selection into Currency Unions

Both the IPWRA and AIPW estimators require a first stage estimation of the likelihood of entering currency unions arrangements for each bilateral pair over time. My first-stage specification is to

predict these probabilities of treatment with a logit model that uses some standard gravity estimators (the log product of GDP and GDP per capita). In addition to this, I include terms that aim to fit differences among trading partners.

$$CU_{ijt} = \theta_0 + \theta_1 ln\left(Y_{it} \times Y_{jt}\right) + \theta_2 ln\left(y_{it} \times y_{jt}\right) + \theta_3 ln\left(Dist_{ijt}\right) + \theta_4 ln\left|Y_{it} - Y_{jt}\right| + \theta_5 ln\left|y_{it} - y_{jt}\right|$$
$$+ \theta_6\left|g_{it} - g_{jt}\right| \tag{5}$$

In Equation (5), in addition to the first three terms, which are standard gravity equation estimates of size of output ($Y_{i/j,t}$), incomes ($y_{i/j,t}$), and distances ($Dist_{ij,t}$) between the two countries, I also include the differences in output, per-capita GDP, and GDP growth rates ($g_{i/j}$). These are important to my motivation because the standard gravity terms may do a poor job of capturing differences between a multilateral union of rich countries (like the EMU) and matches that take place between wealthy countries and their developing trade partners. I also include two terms for the square of the log product of GDP and GDP per-capita. This improves fit of the model, and Millimet and Tchernis (2009) suggest that there are benefits of *over-specifying* the propensity score estimator. Because I wish to keep my results as comparable to the existing literature on currency unions and trade, I work only with variables readily available in standard bilateral trade datasets. Given that work such as Rose (2017) shows that choice of data has large implications on estimates, it seems prudent to work from the same starting point to demonstrate the estimates obtained from the IPWRA and AIPW estimators.

I specify the model in this way to allow that some currency union pairs will be among countries of similar economic size, income, or growth trajectories. The difference terms allow that unions of similar income (like the EMU) have different selection motivation than unions of small emerging markets in unions with large economies. As is seen in Table 1, there are many important differences along these dimensions. The more traditional gravity estimates capture the relative size of unions, which may be small for developing-developing pairs, larger for developing-developed, and very large for developed-developed.

2.4. Dealing with Limited Overlap

While my first-stage estimates from Equation (5) generally provide good fit, with large amounts of overlap between the treated and control sub-populations, many countries-pair observations are assigned with incredibly small probabilities of entering into a currency union. This has been shown to be a source of bias in propensity score estimates. It is common in the literature on propensity scores to trim the sample, with somewhat arbitrary rule of thumb conventions quite common. Imbens (2004) suggests a number of solutions to lack of proper overlap. The first is to trim the sample to exclude outliers, whose potential influence increases in small samples. Export trade datasets are much larger than many of these studies where conventional cutoffs of [0.05, 0.95] or [0.1, 0.9] are common. However, the number of extreme outliers in many of my estimations is still large, and potentially a cause for concern. As such, I will present estimates of the IPWRA estimator on both the full sample and those where any estimated probabilities of treatment fall outside of the [0.001, 0.999] range. In most cases no currency union observations are dropped, but the majority of data which falls below this range is removed. This is because for some disaggregated measures of currency unions, very few trading partners in the control set are close comparisons.

In Appendix A, I discuss the implications of these outliers in my truncated samples. These may have a large impact for the ECCU and Rupee unions as there are so few observations that the sample size becomes small enough where such outliers may influence results. In general, the large sample size should reduce the potential for treatment observations with very low (or control observations with very high) probability of treatment from having an outsized impact. In the PPML estimates larger samples reduce the potential for these issues, but the ECCU still remains potentially problematic.

An important contribution of this article is to show that this truncation actually may improve estimates of the currency union effect. Rose (2017) conducts a meta-analysis, investigating *why* such

a large amount of variation in estimates for the EMU exist and concludes that there are systematic differences based on sample choice. His conclusion is that using larger samples tends to increase the estimates of the currency union on trade, while smaller samples shrink them. He summarizes his findings, saying: *"truncating the sample by omitting countries that are small or poor biases downward the estimates of the country-time fixed effects … this leads to downward bias in the estimated partial effect of EMU on exports"* (Rose 2017). He refers to dropping countries from the sample as *inappropriate*, without providing clear justification why keeping the data is inherently better.

I will show that the non-currency union pairs have substantial differences with those in a currency union, a problem that can become even larger when disaggregating the currency unions. This is in line with similar evidence from Kopecky (2023), who shows that poor comparability of control groups explain why estimates for the EMU using large samples have vastly different estimates than those that restrict the sample to similar, high-income, countries. Here I show that this is even more critical in the context of other disaggregated currency unions, which represent a diverse set of often emerging market economies. Simply weighting on propensity scores, via the IPWRA estimator, without changing the estimation sample can slightly improve this metric, but large differences remain. Truncating the sample to remove extreme outliers, while keeping propensity score weights, appears do a better job creating an appropriate control set.

The claim that more data is better need not be true for its own sake, particularly when the data is expanded by introducing observations with poor comparison to those who are receiving the policy treatment. Consider the analogy of an observational medical study where the population receiving a drug is systematically more likely to be sick than the general population, as is usually the case. Including large amounts of *untreated* observations can easily make the drug look like it's causing poor health outcomes. An obvious choice might be to construct a control sample of those individuals with similar medical history, preexisting conditions, and diet. This is precisely what propensity score estimators attempt to do in situations where it is not possible to pre-randomize and one must try to do so ex-post. While these re-randomizations do not improve credibility of estimation results to the extent that a randomized control trial would,[7] they certainly are a step in the right direction from the perspective of rebalancing the sample to avoid these more obviously poor comparison groups.

3. Baseline Results: Log Gravity

In this section I present my results using the OLS estimates of the log-linear gravity specification from Equation (1).

3.1. Currency Union Selection: Constructing an Appropriate Comparison Group

I first show estimation of inverse probability-weights. These estimated probabilities are used in all the IPWRA and AIPW estimators that follow to create the inverse propensity score weights for each disaggregated currency union. I estimate Equation (5) using a logistic regression.[8] I separately estimate these probabilities using identical models for each of the disaggregated currency unions. Acquiring separate estimates of these currency unions, rather than fitting a single first-stage model for aggregate currency union treatment, is important to my approach. The wide range in magnitude and signs in the differences between treatment and control across disaggregated unions in Table 1 suggests that it is not possible for a single model to jointly capture this heterogeneity. In Table 2, I show the coefficients from this logistic estimation for the full aggregated currency union as well as for each of the seven disaggregated unions I study.

Though the coefficients themselves provide little information on my estimators of interest, there are a few remarks worth making about Table 2. The first is that using a pseudo-R^2 measure of fit there is a wide range of performance in terms of variability in selection into these currency union groups picked up by this estimator. Using the aggregate currency union is quite poor, but so too is the fit for countries using the British pound. This is likely a measure of the variability of countries within these

Table 2. First stage logistic regression, log gravity.

	All CUs	EMU	CFA	ECCU	Aussie Dollar	British Pound	Rupee	US Dollar
$Log(|GDP_1-GDP_2|)$	−0.05***	0.14***	−0.40***	−1.10***	1.61***	0.92***	2.26***	2.81***
	(0.01)	(0.02)	(0.01)	(0.07)	(0.08)	(0.04)	(0.13)	(0.07)
$Log(|GDPpc_1-GDPpc_2|)$	−0.11***	−0.13***	−0.03*	−0.17*	0.77***	−0.28***	−0.59***	−0.22***
	(0.01)	(0.02)	(0.01)	(0.07)	(0.12)	(0.03)	(0.05)	(0.04)
Log(Dist)	−1.00***	−1.11***	−0.97***	−2.70***	−1.50***	−0.80***	−2.38***	−1.81***
	(0.01)	(0.02)	(0.02)	(0.12)	(0.11)	(0.05)	(0.13)	(0.06)
$Log(|GDP_1 GDP_2|)$	−3.04***	1.16***	23.29***	−6.16***	42.60***	15.72***	23.61***	8.37***
	(0.06)	(0.31)	(0.62)	(1.19)	(5.02)	(1.58)	(3.92)	(1.21)
$Log(|GDPpc_1 GDPpc_2|)$	−3.51***	18.44***	9.22***	24.52***	53.93***	8.04***	−2.86**	2.02**
	(0.06)	(1.20)	(0.28)	(2.78)	(7.82)	(0.83)	(1.05)	(0.70)
$|growth_1-growth_2|$	−2.14***	−9.37***	−1.28***	−0.54	3.90***	−0.50	−2.99	1.74**
	(0.18)	(0.83)	(0.21)	(1.58)	(0.92)	(0.72)	(1.67)	(0.61)
$Log(|GDP_1 GDP_2|)^2$	0.03***	−0.01***	−0.26***	0.06***	−0.48***	−0.16***	−0.25***	−0.08***
	(0.00)	(0.00)	(0.01)	(0.01)	(0.05)	(0.02)	(0.04)	(0.01)
$Log(|GDPpc_1 GDPpc_2|)^2$	0.10***	−0.42***	−0.32***	−0.65***	−1.45***	−0.23***	0.05	−0.05**
	(0.00)	(0.03)	(0.01)	(0.08)	(0.21)	(0.02)	(0.03)	(0.02)
Observations	733356	724511	727816	723059	722170	722486	722134	722596
Pseudo R-squared	0.294	0.447	0.502	0.920	0.531	0.128	0.555	0.472

Note: Standard errors clustered by country-pairs, *p < .10, &**p < .05, ***_p < .01.

samples with the British pound union, which covers a wide range of diverse countries from the start of the sample until the late 1960s, and now (other than the UK itself) consists of just a few small territories who still use the currency. The East Caribbean Currency Union, on the other hand, is a small group of quite similar countries, and the model does a good job fitting treatment, though, as will be apparent when truncating the sample below, it finds few non-ECCU trade partners that make suitable comparisons in the data. Another notable feature of Table 2 is that the coefficients differ significantly across sample, further driving home the need to estimate selection separately for these various currency unions.

I motivated my approach in Table 1 by showing that there was both a substantial difference in the mean of currency unions and nonunion pairs, and large heterogeneity in the size and sign of these differences across disaggregated unions. In Figure 1, I show that these differences can be substantially improved by modeling selection into each union as outlined in Section 2. I plot the difference in means
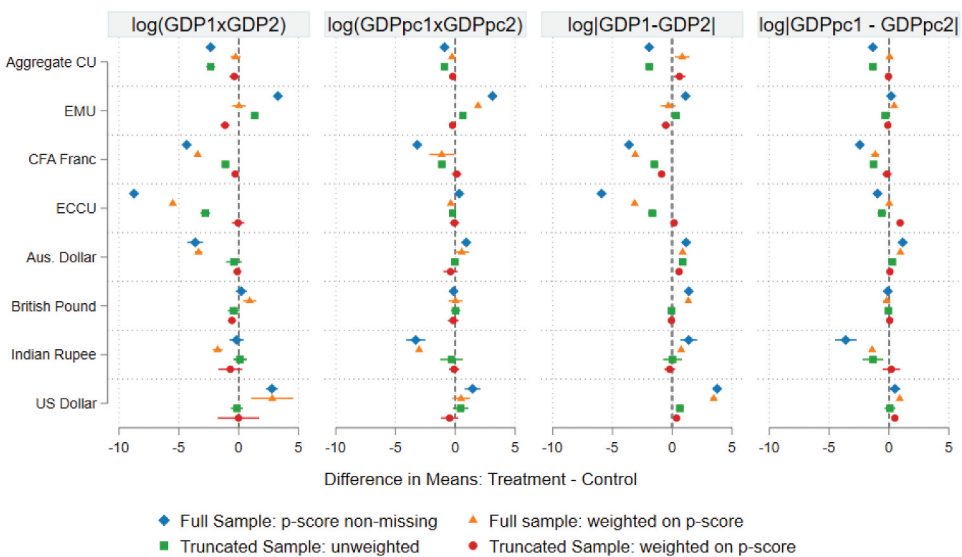
Figure 1. Differences in means (Treatment – Control), various samples.

between treatment and control groups for these same four covariates across each currency union under four different models. The first is the full sample, which excludes other currency unions, conditional on the first stage providing a non-missing estimate of probability of selection into currency unions. The second uses this same sample, but weights observations using the inverse probability weights as in Equation (3). These provide a comparison for the pure effect of weights on improving similarity across treated and control groups. I then repeat this exercise on the truncated sample, keeping only observations where the probability of treatment is between [0.001,0.999]. For the aggregate CU I note that very little data is dropped when doing so, but as can be readily seen in my results (Table 3), this will drop most of the data when estimating each individual currency union.

Figure 1 highlights the primary contribution of my approach. The diamonds reflect the differences in means in from Table 1, which show large deviations in the comparability of treatment and control observations, without a consistent pattern across these heterogeneous unions. Weighting on propensity scores in most cases improves on this fit, bringing the difference in means closer to zero in all but a few cases. However, it is fairly clear that keeping as much data as possible and relying on the weights alone cannot bring these values to zero for all covariates and all currency unions. Samples truncating only on the ability to estimate a probability of selection are still extremely large, over 720,000 observations, so even allowing for extremely low probability weights it is difficult for any weighting to fully offset impact keeping a relatively big sample of poor comparisons for a small number of treated observations. The red circles are, in general, the closest to zero in this figure, suggesting that both truncation and weighting provide the best comparability between the individual currency unions and the comparison groups used in estimating their treatment effects. Importantly, truncation without weights (squares) often gets closer to zero than weights without truncation (triangles), though differences can be useful when interpreting the results of estimation presented in the next section.

3.2. Estimates of Currency Union Effect in Log Gravity Equation of Trade

I now present estimates for the effect of currency unions on trade that make use of traditional log-gravity specifications shown in Equation (1). Table 3 reports these results for both the OLS and IPWRA specifications. The first column provides the unweighted and untruncated sample and shows estimates that are quite similar to those estimated with the same *theory consistent* method of estimating the gravity relationship via exports as in Glick and Rose (2016), whose outcome of interest is the EMU but report jointly estimated disaggregated specification. My sample is slightly smaller than that of Glick and Rose (2016), due to missing data for some of the controls used in the logistic model of currency union selection and because all other currency unions are dropped when estimating each disaggregated union.[9] This sample reduction ensures proper comparability between my estimates using propensity score weights with these unweighted currency union effects. The second column of Table 2 reports estimates limiting estimation to the truncated sample without using the IPWRA weights. This has a substantial impact on all of the disaggregated estimates. While the aggregate currency union sample does not change much, each individual currency union only a fraction of observations are within the [0.001,0.999] range for estimated probabilities. For most, this sample selection effect decreases the estimated coefficient. The Australian dollar and Indian Rupee unions have larger point estimates, though both are insignificant when clustering of country-pairs. Reducing the estimation sample in this way substantially reduces the power, making these estimates less significant. Column 3 of Table 3 shows the IPWRA estimates on the full sample, while the IPWRA on the truncated sample are shown in Column 4. I report both conventional heteroskedastic robust standard errors and errors clustered by country-pairs. Unless otherwise noted in what follows I refer to these more restrictive standard errors.

As is well established, these baseline estimates reflect a substantial estimated currency union trade impact, with the average effect of 0.39 reflecting a 48% ($e^{0.39}-1$) improvement in exports upon entering into a currency union. These of course vary a great deal with estimates slightly larger, but quantitatively similar estimates to this baseline in the EMU, Aussie Dollar, and British pound currency unions,

Table 3. Log gravity equation: OLS estimates, various models.

	Baseline OLS	Baseline OLS (Truncated)	IPWRA	IPWRA (Truncated)
All CUs	0.39	0.39	0.43	0.43
clustered s.e.	(0.05) ***	(0.05)***	(0.05) ***	(0.05) ***
Het. Robust s.e.	(0.03) ***	(0.03)***	(0.02) ***	(0.02) ***
R^2	0.861	0.863	0.921	0.9219
N	732,143	721,500	732,143	721,500
EMU	0.45	0.15	0.40	0.10
clustered s.e.	(0.05) ***	(0.05)***	(0.05) ***	(0.05) **
Het. Robust s.e.	(0.04) ***	(0.03)***	(0.02) ***	(0.02) ***
R^2	0.862	0.937	0.924	0.9688
N	723,300	118,250	723,300	118,250
CFA Franc	0.75	0.53	0.50	0.29
clustered s.e.	(0.22) ***	(0.25)**	(0.23) **	(0.29)
Het. Robust s.e.	(0.08) ***	(0.12)***	(0.10) ***	(0.15)*
R^2	0.860	0.781	0.873	0.8798
N	726,606	125,936	726,606	125,936
ECCU	−1.61	−1.90	−1.41	−2.12
clustered s.e.	(0.21) ***	(0.59)***	(0.17) ***	(0.61) ***
Het. Robust s.e.	(0.19) ***	(0.66)***	(0.08) ***	(0.50) ***
R^2	0.860	0.907	0.889	0.958
N	721,847	3,643	721,847	3,643
Aussie $	0.42	0.67	0.32	1.00
clustered s.e.	(0.31)	(0.74)	(0.36)	(0.75)
Het. Robust s.e.	(0.26)	(0.67)*	(0.13)**	(0.37)***
R^2	0.860	0.860	0.887	0.992
N	720,271	13,455	720,271	13,455
British Pound	0.47	0.24	0.43	0.36
clustered s.e.	(0.14) ***	(0.13) *	(0.10) ***	(0.11) ***
Het. Robust s.e.	(0.08) ***	(0.10)**	(0.03) ***	(0.05) ***
R^2	0.860	0.857	0.938	0.947
N	721,277	129,104	721,277	129,104
Indian Rupee	0.55	1.02	0.51	1.11
clustered s.e.	(0.50)	(1.06)	(0.38)	(1.06)
Het. Robust s.e.	(0.17) ***	(0.58)*	(0.09)	(0.64)*
R^2	0.860	0.887	0.881	0.9542
N	720,923	6,038	720,923	6,038
US $	0.13	−0.15	0.05	−0.28
clustered s.e.	(0.11)	(0.23)	(0.09)	(0.30)
Het. Robust s.e.	(0.11)	(0.21)	(0.03)	(0.24)
R^2	0.860	0.951	0.934	0.9693
N	721,385	39,194	721,385	39,194

Note: Each currency union estimate reflects a separate estimation of Equation (1) either by standard OLS or using the IPWRA adjustment in Equation (3). Results are presented on the full sample (excluding other currency unions) and truncated such that estimated probabilities must fall within a [0.001,0.999] range. All estimates include exporter-year, importer-year, and pairwise fixed effects. Standard errors are reported for both heteroskedastic robust (Huber/White) errors, and with clustering by country-pairs, *$p < .10$, &**$p < .05$, ***_$p < .01$.

whose estimates reflect a 56%, 52%, and 60% increase respectively. Much larger are the estimates for the CFA Franc and Indian Rupee with implied increases of 112% and 73%. Puzzlingly low, but in line with the results from Glick and Rose (2016), is the large negative value for the ECCU, with a point estimate reflecting an 80% reduction in trade.

At first glance the IPWRA estimates look similar to those from the unweighted regression. All estimates of individual currency unions are smaller (in absolute value), with the aggregate currency union estimate slightly larger. Truncating the sample has effects that are less clear-cut increasing some estimates and decreasing others. Looking first at the EMU estimate the parameter is slightly larger than the range reported in the literature survey by Polák (2019). Notably my estimates in Table 3 for the EMU are extremely close to those of Rose (2017) when limiting the sample to only rich economies. Here I estimate a coefficient of 0.10, while Rose (2017) finds a coefficient of 0.11 when estimating the Euro effect among a sample of

high-income countries. While I did not choose rich countries specifically, the estimation of Equation (4) for selection into the EMU is ultimately trying to match on the bilateral trade relationship of relatively large and relatively rich economies. This can be readily seen from the EMU's differences in means in Figure 1, where the EMU is a large outlier to the right of zero for log product of GDP per capita in the unweighted sample (higher incomes than the average control. It would seem my data-driven method of selecting a proper control more or less converged on a similar sample to the ad-hoc method used in their work. Kopecky (2023) further explores the dependency of underlying sample to determine estimates of the EMU trade effect.

The CFA Franc union is roughly one standard error lower when using the full sample and IPWRA than the baseline estimate, reflecting a trade effect that is still sizable (64.9%), but nearly fifty percentage points lower than estimates that fail use these first stage weights. A similar result comes when using the truncated sample without weighting. This is perhaps because each method of adjusting captures slightly different variation, as seen in Figure 1, where the full sample with weights substantially improves the comparability for the second and fourth covariates, while having only a small impact on the comparability of the first and third, which come close to zero with truncation even without weighting. If these two approaches shrink the estimator, but for different reasons, then it is intuitive that in column 4, the weighted and truncated estimates are smaller still with a point estimate of 0.29, reflecting a roughly 33% improvement in trade. This is only marginally significant and only when using the traditional heteroskedastic robust standard errors.

Estimates for countries using the British pound suggest potentially large reductions, but most appears to be driven by sample selection. Notably, this is the one estimate where the unweighted full sample and IPWRA on the full sample appear to perform similarly poorly in adjusting comparability in Figure 1 (the IPWRA is perhaps worse). The two truncated estimates are smaller, reflecting an increase in trade of 27% on for the unweighted OLS estimator and 43% on the IPWRA truncated estimate.

Reduction of the magnitude of the negative ECCU coefficient is similar to the CFA Franc in terms of significance (a roughly one standard error reduction), but the implied reduction in trade is smaller with a 75.5% ($e^{-1.41}-1$) reduction in trade relative to the prior estimate of roughly 80%. Notably truncation appears to increase the magnitude for these estimates, as is reflected in both the unweighted and weighted estimates, though I note that the model is quite good at predicting selection into the ECCU (as reflected in the strong fit in Table 2), but there are very few non-ECCU observations with which to compare these and so outliers may strong influence on these truncated estimates. The comparability in Figure 1 of the differences in GDP per capita are among the poorest of these estimates. The Aussie dollar is another union where the effects of weighting and truncation do not obviously move in the same direction. Compatibility for this union appears to be affected very little by weighting alone in Figure 1, with both truncated samples having similar differences in means. The larger (though only significant without clustering standard errors) estimates in columns 2 and 4 are those that perform best by this measure.

I note that Estimates for the ECCU, Aussie Dollar, and Indian Rupee now rely on small sample sizes. Using the arithmetic from Imbens (2004), these now allow for some extreme effects as reported in the table. To test if these probability weights were greatly affecting these estimates, I re-estimated this model on the same sample, while censoring maximum and minimum probabilities to the [0.01,0.99] range. The resulting estimates were not statistically distinguishable, suggesting that the sample reduction choice has driven changes in these estimators, not large weights on outliers. In addition to having such small samples, the estimates for these three have jumped substantially, though the implausibly large estimate for the ECCU is common across all of my estimates, and indeed all of the disaggregated results presented in Glick and Rose (2016). This increase for the Australian dollar and Indian Rupee unions remains insignificant for cluster robust errors.

3.3. *Augmented Inverse Probability Weighting Estimator*

The AIPW estimator should provide more efficient estimates than the IPWRA, and while both estimators have the doubly robust property it is not clear which should perform better under misspecification in both models. These estimates use the probabilities from Table 2 along with the conditional means from the weighted regression estimations in Table 3 to construct the AIPW estimator as defined in Equation (3). This combines a simpler IPW estimator (the first term in Equation (3)) with an augmented regression adjustment term, which comes from a weighting of the estimates of the same gravity equation from Equation (1), specified in the same way as for the IPWRA estimator. These are reported in Table 4.

There are two large outliers in these estimates, with positive coefficients that are implausible. These are the aggregate currency union and the British pound. As discussed above, these are the two treatments with the worst first stage fit in Table 2, with the British pound currency union having extremely low probabilities of treatment across even the treated sub-sample. Given that these are estimated on the full sample, one might expect the estimates to be universally closer to those estimated on the full sample in Table 3. This is not the case. Two of them, the EMU and CFA, have estimates quite close to their truncated IPWRA estimates, which were substantially smaller than those estimated on the full sample. In the case of the EMU this is now imprecisely estimated, while the CFA estimates are here significant. The Aussie Dollar and Indian Rupee, on the other hand, have estimates that are extremely close to their non-truncated IPWRA in Table 3, in the case of the Rupee zone almost the same as the estimate from the baseline gravity in Table 3. Notably these are two cases where the limited sample in the truncated case created some cause for concern. The ECCU's persistently large and negative coefficient remains but is less extreme here.

It is interesting to note that the aggregate currency union effect, British pound, and US dollar were fairly consistent across the full and truncated samples in Table 3, but are dramatically changed in Table 4. This estimator puts substantially more weight on the traditional IPW estimator of weighted group means than the conditional mean estimates from the gravity equations, with these implausible jumps in estimates likely pointing to failure of the modeling of selection into currency unions for these particular disaggregated unions, something that is clearly true for the aggregate and British pound groups. It is not obvious that the fit for the US dollar unions is substantially worse than others. Because my modeling choices for selection is an atheoretical reduced-form model, I am cautious of reading too closely into results that rely more heavily on the ability of these first stage probabilities to re-weight group means, and while the regression adjustment should stabilize these estimates it clearly struggles to do so in the cases of extremely poor fit in the aggregate CU and British pound. With that said, the EMU, CFA, Australian, and Indian currency unions all have estimates from the AIPW that fit in sensibly with those estimated from the IPWRA estimates.

4. PPML Estimates of Gravity Equation

In this section I repeat the results of Section 3 with the PPML estimator of Equation (2). This approach has been shown in the literature to remove potential bias introduced due to the presence of heteroskedasticity in log-linearized estimators. This also expands the dataset as missing trade flows, often assumed to be zeros, can now be included. This may improve the

Table 4. AIPW estimator.

	All CUs	EMU	CFA	ECCU	Aussie $	British £	Rupee	US $
ATE$_{AIPW}$	2.45***	0.12	0.22**	−0.81***	0.32***	3.88***	0.58***	0.75***
	(0.26)	(0.14)	(0.08)	(0.20)	(0.02)	(0.55)	(0.09)	(0.12)
N	733,143	723,300	726,606	721,847	720,271	721,277	720,923	721,385

Note: Standard errors clustered by country-pairs, †$p<.15$, *$p < .10$, &**$p < .05$, ***_$p < .01$. In this table I do not report estimates on the truncated sample as the AIPW estimator appears extremely sensitive to outlier probability weights in small samples.

fit of my first stage model, especially for smaller developing countries that may have more in common with pairs with missing trade data. As such, I re-estimate the selection equations allowing for inclusion of a larger dataset. I treat missing trade flows as zero if export data is missing but GDP data is available for a given country-pair-year. To conserve space, I report the first stage logistic results in Appendix B. Figure 2 shows the same set of differences in means when using the same weighting and truncation measures as in Section 3 on this larger data and with the new probability weights.

My results for these PPML estimations are presented in Table 5. I note that while the full sample available for estimation of these results is roughly 1.48 million observations, many are dropped as they are singleton groups or "separated" by a fixed effect (see Correia, Guimaraes, and Zylkin (2019) and Correia, Guimaraes, and Zylkin (2020)). Larch et al. (2019) provide a comparison for the baseline estimates in this section. Their work generate similar results to that of Glick and Rose (2016) while using the more robust PPML method. This includes both an aggregate currency union estimate and a jointly estimated disaggregated estimation. The Larch et al. (2019) for the aggregate currency union is a weakly significant 0.15, similar to what is found here. This actually increases when using the IPWRA on the full sample, though truncation reduces both the weighted and unweighted estimates slightly. While there is clearly a large reduction in aggregate currency union estimates due to removing the bias induced by log-linearization, weighting actually strengthens the size of these estimates predicting a 16% rise in trade in the weighted and truncated sample.

The EMU baseline estimate is similar in size and significance in Larch et al. (2019) to the baseline that I find here, and consistent with similar estimates found in Kopecky (2023), who focuses on selection effects into the EMU specifically. This result is weakly estimated, but becomes negative and significant in the truncated IPWRA model. This small negative estimate is not completely out of line with this larger literature on the Eurozone and may be affected by using a sample that contains the global financial crisis and European sovereign debt crisis.

Estimates for the CFA Franc are now extremely small when using the full sample and quite large when truncating. As with the log-gravity specifications the comparability of treatment and control are unambiguously closer to zero in Figure 2 when both weighting and truncation are used, but here the estimated increase in trade is an implausibly large 177% increase. This is in stark contrast to the PPML estimate in Larch et al. (2019), which is −0.126, and like my baseline PPML estimate not statistically
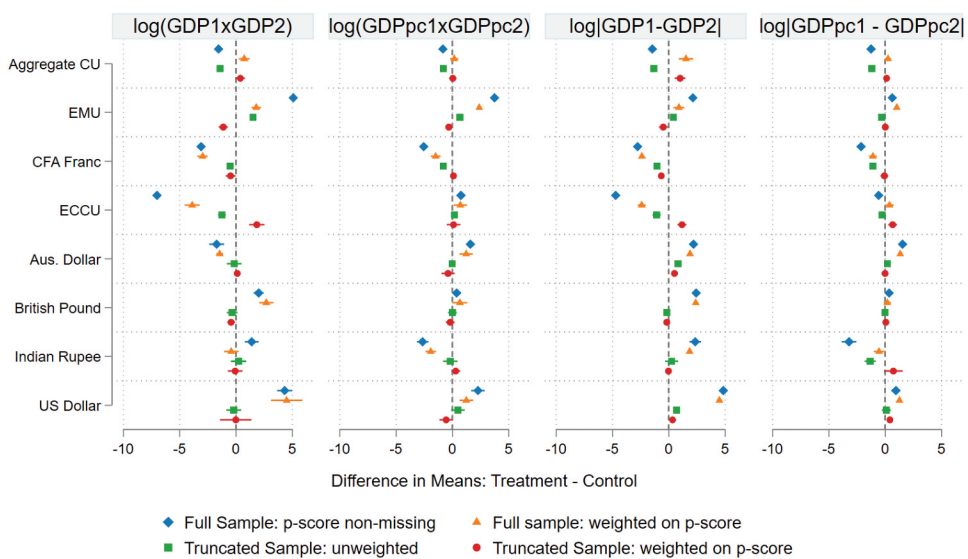
Figure 2. Difference in means (Treatment -Control), various models, PPML sample.

distinguishable from zero. While the weights used in regression seem to have little impact here the method of creating a more comparable sample using propensity scores has a dramatic impact on estimates.

The strong negative point estimates for the ECCU survive the use of the PPML, but for three of these estimates shrink substantially and remain significant using cluster robust standard errors[10] on the full samples. Here the use of IPWRA estimator to weight the full sample results shrinks the negative estimated impact of the union from a 66% reduction in trade to a 55% decline.

The only other union with significant estimates is the British pound, which as before is very strongly significant for all models. Here the large positive baseline of 1.09 is nearly identical to the results reported using similar estimators in Larch et al. (2019); however, the truncation once again substantially reduces this estimate, putting it closer in line with those reported in the log-gravity case above. Once again using the PPML estimator appears to minimize impact from the weights directly; however, the use of the first stage model to select an appropriate comparison group still has a substantial impact.

While the Aussie dollar estimates are only significantly positive using the baseline PPML estimates, it is worth noting that here truncation and weighting both work to reduce the estimated impact, suggesting that improving the comparison group brings this seemingly strong positive estimate (again in line with the significant result reported in Larch et al. (2019)) to be indistinguishable from zero, or weakly positive if using heteroskedastic robust errors without clustering.

4.1. *Augmented Inverse Probability Weighting Estimator, PPML Estimates*

I report estimates for the AIPW in Table 6. As outlined above, this augments the IPWRA estimates with more traditional propensity score estimators. Estimates for the AIPW on the truncated sample in Section 3.3 were omitted as the IPW component of the estimator used to construct these estimates appeared to be sensitive to outlier probability weights,[11] particularly when samples are small. Because the PPML retains more data, and the results appear to provide useful context for Table 5 I include both here, but caution over-interpretation of point estimates of the truncated results for this reason.

The AIPW estimator gives a broadly similar picture to what is found in Table 5. Estimates are slightly more significant, but the point estimate for the CFA, ECCU, and British pound are very much in line with the results discussed above, including differences between full sample estimates and truncated. Here a notable difference is that the Rupee union, which wasn't significant in any prior estimates other than the previous AIPW is now strongly significant, with the estimated impact switching sign when moving from the full sample to the truncated one. Using once again Figure 2 as a guide the truncated sample does appear to substantially improve comparability between treated and control units and the coefficient of 0.45 is similar in magnitude to that of the log-OLS AIPW of 0.58. The Aussie dollar estimate for the full sample is nearly the same as those for the truncated and weighted estimates presented above, but are now significant. Truncating the sample this disappears completely.

5. Conclusions

What can be taken away from this exercise? It is not the goal of this article to advocate for any particular estimates above as the *true* currency union effect on trade, nor even the *true* average treatment effect for a particular disaggregated currency union. Rather I hope to demonstrate how important the choice and weighting of the comparison group used in such analyses are in determining the estimated effects. These results suggest that there should be a change in thinking about how currency union or any macroeconomic policy estimates are estimated. I have intentionally limited myself to the well-studied data of Glick and Rose (2016) as a means of connecting with the large existing literature on these estimates to highlight the potential flaws associated with failure to correct for the endogeneity of selection into a particular policy choice. I see three main takeaways.

Table 5. Gravity equation: PPML estimates, various models.

	Baseline PPML	Baseline PPML (Truncated)	IPWRA	IPWRA (Truncated)
All CUs	0.11	0.09	0.18	0.15
clustered s.e.	(0.04)***	(0.04) **	(0.04) ***	(0.04)***
Het. Robust s.e.	(0.01)***	(0.01)***	(0.01)***	(0.01)***
R^2	0.984	0.985	0.991	0.991
N	1,197,812	1,043,696	1,197,812	1,043,696
EMU	0.03	−0.06	0.01	−0.08
clustered s.e.	(0.04)	(0.04)	(0.04)	(0.04)*
Het. Robust s.e.	(0.01)***	(0.01)***	(0.01)***	(0.01)***
R^2	0.985	0.996	0.994	0.996
N	1,183,040	136,026	1,183,040	136,026
CFA Franc	0.002	0.97	0.02	1.02
clustered s.e.	(0.37)	(0.34) ***	(0.34)	(0.35) ***
Het. Robust s.e.	(0.11)	(0.17)***	(0.11)	(0.19)***
R^2	0.984	0.922	0.984	0.912
N	1,190,143	297,397	1,190,143	297,397
ECCU	−1.08	−0.69	−0.82	−2.68
clustered s.e.	(0.28) ***	(0.47)	(0.25)***	(1.61)
Het. Robust s.e.	(0.10)***		(0.46)	
R^2	0.860	0.907	0.889	0.958
N	1,182,707	5,366	1,182,707	5,366
Aussie $	0.43	0.21	0.27	0.18
clustered s.e.	(0.25)*	(0.16)	(0.25)	(0.15)
Het. Robust s.e.	(0.12)***	(0.09)**	(0.14)*	(0.09)*
R^2	0.980	0.900	0.980	0.798
N	1,179,269	12,863	1,179,269	12,863
British £	1.09	0.22	1.17	0.22
clustered s.e.	(0.15) ***	(0.09)**	(0.12) ***	(0.08) **
Het. Robust s.e.	(0.04) ***	(0.03)***	(0.04)***	(0.04)***
R^2	0.860	0.857	0.938	0.947
N	1,180,986	128,686	1,180,686	128,686
Indian Rupee	−0.08	0.32	−0.04	0.45
clustered s.e.	(0.42)	(0.49)	(0.27)	(0.48)
Het. Robust s.e.	(0.14)	(0.32)	(0.09)	(0.32)
R^2	0.984	0.957	0.984	0.958
N	1,180,659	11,053	1,180,659	11,053
US $	0.004	0.11	−0.03	−0.11
clustered s.e.	(0.07)	(0.14)	(0.06)	(0.24)
Het. Robust s.e.	(0.02)	(0.13)	(0.02)	(0.15)
R^2	0.983	0.995	0.992	0.996
N	1,181,095	43,476	1,181,095	43,476

Note: Each currency union estimate reflects a separate estimation of Equation (2) either by standard PPML or by using the IPWRA adjustment in Equation (3). Results are presented on the full sample (excluding other currency unions) and truncated such that estimated probabilities must fall within a [0.001,0.999] range. All estimates include exporter-year, importer-year, and pairwise fixed effects. Standard errors are reported for both heteroskedastic robust (Huber/White) errors, and with clustering by country-pairs, *p < .10, &**p < .05, ***_p < .01.

Table 6. AIPW estimator.

	All CUs	EMU	CFA	ECCU	Aussie $	British £	Rupee	US $
Full Sample								
ATE_{AIPW}	0.18	0.01	−0.02***	−1.01***	0.27***	1.17***	−0.11***	−0.02
	(0.63)	(0.49)	(0.00)	(0.00)	(0.00)	(0.12)	(0.09)	(0.03)
N	1,197,812	1,1830400	1,190,143	1,182,707	1,179,269	1,180,659	1,180,659	1,181,095
Truncated Sample								
ATE_{AIPW}	0.15	−0.07	1.09***	−0.72***	0.00	0.22***	0.45***	−0.07**
	(0.69)	(3.59)	(0.01)	(0.00)	(0.00)	(0.05)	(0.11)	(0.03)
N	1,043,696	136,026	297,397	5,366	12,863	128,686	11,053	43,476

Note: Standard errors clustered by country-pairs, † p<.15, *p < .10, &**p < .05, ***_p < .01.

First, macroeconomic estimates of trade relationships should focus more on individual policy choices, acknowledging that the choice of Liberia to accept dollars as legal tender is taken under a completely unrelated set of circumstances than the choice of Luxembourg to enter into a multilateral currency regime with Germany. These are likely both different from the decision of Mali and Senegal to band together with neighboring African countries in a multilateral currency union of developing countries with a fixed exchange rate to the Franc/Euro. These differences almost certainly impact the effectiveness that these currency unions have on trade and comparing them to each other could lead to misleading understanding of the potential gains and losses from currency union formation and destruction.

Another key takeaway is that it is critical to establish a proper counterfactual control group for these policy experiments. The wealth of literature around the currency union and EMU impact on trade has made clear that adding huge amounts of data of developing country bilateral trade partners inflates both of these estimates. As shown in Figures 1 and 2, these samples make for poor comparisons to the countries actually receiving policy treatment, making them a poor comparison group. I have shown above that this largely comes from constructing a sample where the trade partners share few observable characteristics. This is likely equally true for emerging market country unions, as explored in Campbell and Chentsov (2017), though there is unfortunately much less existing work studying their individual impact to make such connections with. While weighting using an estimator such as the IPWRA appears less critical when using the PPML approach, selection of sample using the same propensity scores can have a large impact even when using these more robust estimators.

Finally, future work must do better to empirically model the choice of entering currency arrangements themselves. It will be useful to replicate this analysis with a more sophisticated dataset on the relevant institutional, demographic, and political factors that might affect first stage selection. This can not only improve the fit of the first stage probability model over the somewhat crude one used above, but more importantly will better identify those countries who serve as adequate control groups, and likely provide better understanding of *why* a currency union might have better or worse trade impact. Also useful would be to motivate this choice of data with a theoretically consistent model such as that of Alesina and Barro (2002). Their model provides key insight for small developing economies that adopt a currency union to anchor to larger trading partners, such as the US dollar or Pound unions. A different model might be appropriate when considering the EMU, or another still for the CFA or ECCU. It is clear from the heterogeneity in Figures 1 and 2 that future work needs be flexible in the type of pairs being studied and how that may inform such decisions, or rather limit its scope to just one type. Even if the majority of data are still poor matches with smaller developing country currency unions, as they were above, improvement of the comparison group will provide more overlap, and therefore should reduce the bias of these estimates. Having a richer dataset to understand the common factors leading to selection into these arrangements will also facilitate a deeper understanding of the trade estimates, by providing better indications of the non-trade related motivations behind the decision to adopt a common currency.

Matching estimators, particularly flexible ones such as the IPWRA and AIPW, should be useful in macroeconomic analyses of these estimates going forward. Even in situations where their weighting does not significantly impact estimates, they provide a useful framework for thinking about finding the proper comparison group for which to estimate the effects of nonrandomly assigned policies. So too might synthetic control estimates that seek to construct appropriate comparison groups from the data when no adequate match is available. This article shows that even using standard trade models such estimators can make for large differences, but it is difficult to fully understand the historical context without empirically modeling the salient factors for each individual union separately.

Notes

1. See Alesina, Barro, and Tenreyro (2002), Bayoumi and Eichengreen (1993), and more recently Aizenman (2016) and Chari, Dovis, and Kehoe (2020), for a discussion on the many costs and benefits of entering into a monetary union.
2. For a full discussion of this debate in the context of the Eurozone estimates see Frankel (2010).
3. The CFA Franc and East Caribbean Currency Union are both of this type.
4. For detailed discussion on these and other propensity score estimators see: Imbens (2004), Lunceford and Davidian (2004), and Wooldridge (2007).
5. Though their estimators use dynamic local projections framework, these can easily be converted to a static context as I do here.
6. To condense notation I refer to the control set as simply $Z_{ij,t}$, but this includes the rich set of fixed effects in Equation (1).
7. There are still likely unobserved confounders that are not being addressed.
8. Using a probit model instead has little impact on the results, but appears to have slightly worse fit in most cases as judged by pseudo-R^2.
9. I am able to perfectly replicate the Glick and Rose (2016) results when using the full data. See Table 5 of their paper for the equivalent specification.
10. For the EECU in the truncated and weighted case heteroskedastic standard errors cannot be estimated as the variance matrix is singular. This is, in part, due to the small sample and large number of fixed effects.
11. Nearly all became greater than one in absolute value.
12. Fewer are dropped for the aggregate currency union, in part because the treatment model has a hard time fitting the diverse subset of countries and so fewer countries are assigned the near-zero probabilities.[13] Which often looks at smaller observational studies in fields such as medicine.

Acknowledgments

Disclosure Statement

ORCID

Joseph Kopecky 🆔 http://orcid.org/0000-0003-1118-9219

References

Aizenman, J. 2016. Optimal currency area: A 20th century idea for the 21st century? Technical report, National Bureau of Economic Research.

Alesina, A., and R. J. Barro. 2002. Currency unions. *The Quarterly Journal of Economics* 117 (2):409–36. doi:10.1162/003355302753650283.

Alesina, A., R. J. Barro, and S. Tenreyro. 2002. Optimal currency areas. *NBER Macroeconomics Annual* 17:301–45. doi:10.1086/ma.17.3585292.

Baldwin, R., and D. Taglioni. 2007. Trade effects of the euro: A comparison of estimators. *Journal of Economic Integration* 22:780–818. doi:10.11130/jei.2007.22.4.780.

Barro, R., and S. Tenreyro. 2007. Economic effects of currency unions. *Economic Inquiry* 45 (1):1–23. doi:10.1111/j.1465-7295.2006.00001.x.

Bayoumi, T., and B. Eichengreen. 1993. One money or many? On analyzing the prospects for monetary unification in various parts of the world. Technical report.

Campbell, D. L. 2013. Estimating the impact of currency unions on trade: Solving the Glick and rose puzzle. *The World Economy* 36 (10):1278–93. doi:10.1111/twec.12062.

Campbell, D. L., and A. Chentsov. 2017. Breaking badly: The currency union effect on trade.

Chari, V. V., A. Dovis, and P. J. Kehoe. 2020. Rethinking optimal currency areas. *Journal of Monetary Economics* 111:80–94. doi:10.1016/j.jmoneco.2019.01.023.

Chintrakarn, P. 2008. Estimating the euro effects on trade with propensity score matching. *Review of International Economics* 16 (1):186–98. doi:10.1111/j.1467-9396.2007.00725.x.

Correia, S., P. Guimaraes, and T. Zylkin. 2019. Verifying the existence of maximum likelihood estimates for generalized linear models. *arXiv Preprint arXiv* 1903.01633. doi:10.48550/arXiv.1903.01633.

Correia, S., P. Guimaraes, and T. Zylkin. 2020. Fast poisson estimation with high dimensional fixed effects. *The Stata Journal* 20 (1):95–115. doi:10.1177/1536867X20909691.

Frankel, J. 2010. The estimated trade effects of the euro: Why are they below those from historical monetary unions among smaller countries? In *Europe and the Euro*, ed. A. Alberto and G. Francesco, 169–212. University of Chicago Press.

Glick, R., and A. K. Rose. 2002. Does a currency union affect trade? The time-series evidence. *European Economic Review* 46 (6):1125–51. doi:10.1016/S0014-2921(01)00202-1.

Glick, R., and A. K. Rose. 2016. Currency unions and trade: A post-emu reassessment. *European Economic Review* 87:78–91. doi:10.1016/j.euroecorev.2016.03.010.

Glick, R., and A. M. Taylor. 2010. Collateral damage: Trade disruption and the economic impact of war. *The Review of Economics and Statistics* 92 (1):102–27. doi:10.1162/rest.2009.12023.

Glynn, A. N., and K. M. Quinn. 2010. An introduction to the augmented inverse propensity weighted estimator. *Political Analysis* 18 (1):36–56. doi:10.1093/pan/mpp036.

Head, K., and T. Mayer. 2014. Gravity equations: Workhorse, toolkit, and cookbook. In *Handbook of international economics*, ed. Gita. G, H. Elhanan and K. Rogoff, vol. 4, 131–95. Elsevier.

Head, K., T. Mayer, and J. Ries. 2010. The erosion of colonial trade linkages after independence. *Journal of International Economics* 81 (1):1–14. doi:10.1016/j.jinteco.2010.01.002.

Hirano, K., and G. W. Imbens. 2001. Estimation of causal effects using propensity score weighting: An application to data on right heart catheterization. *Health Services & Outcomes Research Methodology* 2 (3):259–78. doi:10.1023/A:1020371312283.

Imbens, G. W. 2004. Nonparametric estimation of average treatment effects under exogeneity: A review. *The Review of Economics and Statistics* 86 (1):4–29. doi:10.1162/003465304323023651.

Jordà, O., and A. M. Taylor. 2016. The time for austerity: Estimating the average treatment effect of fiscal policy. *The Economic Journal* 126 (590):219–55. doi:10.1111/ecoj.12332.

Kopecky, J. 2023. A match made in Maastricht: Estimating the treatment effect of the euro on trade. *SSRN Working Paper*. doi:10.2139/ssrn.4271796.

Larch, M., J. Wanner, Y. V. Yotov, and T. Zylkin. 2019. Currency unions and trade: A ppml re-assessment with high-dimensional fixed effects. *Oxford Bulletin of Economics and Statistics* 81 (3):487–510. doi:10.1111/obes.12283.

Lunceford, J. K., and M. Davidian. 2004. Stratification and weighting via the propensity score in estimation of causal treatment effects: A comparative study. *Statistics in Medicine* 23 (19):2937–60. doi:10.1002/sim.1903.

Millimet, D. L., and R. Tchernis. 2009. On the specification of propensity scores, with applications to the analysis of trade policies. *Journal of Business & Economic Statistics* 27 (3):397–415. doi:10.1198/jbes.2009.06045.

Nitsch, V. 2002. Honey, i shrunk the currency union effect on trade. *The World Economy* 25 (4):457–457. doi:10.1111/1467-9701.00443.

Persson, T. 2001. Currency unions and trade: How large is the treatment effect? *Economic Policy* 16 (33):434–48. doi:10.1111/1468-0327.00081.

Polák, P. 2019. The euro's trade effect: A meta-analysis. *Journal of Economic Surveys* 33 (1):101–24. doi:10.1111/joes.12264.

Rose, A. 2002. Honey, the currency union effect on trade hasn't blown up. *The World Economy* 25 (4):475–79. doi:10.1111/1467-9701.00444.

Rose, A. K. 2000. One money, one market: The effect of common currencies on trade. *Economic Policy* 15 (30):08–45. doi:10.1111/1468-0327.00056.

Rose, A. K. 2017. Why do estimates of the emu effect on trade vary so much? *Open Economies Review* 28 (1):1–18. doi:10.1007/s11079-016-9420-1.

Rose, A. K., and T. D. Stanley. 2005. A meta-analysis of the effect of common currencies on international trade. *Journal of Economic Surveys* 19 (3):347–65. doi:10.1111/j.0950-0804.2005.00251.x.

Saiki, A. 2005. Asymmetric effect of currency union for developing countries. *Open Economies Review* 16 (3):227–47. doi:10.1007/s11079-005-1023-1.

Silva, J. S., and S. Tenreyro. 2006. The log of gravity. *The Review of Economics and Statistics* 88 (4):641–58. doi:10.1162/rest.88.4.641.

Silva, J. S., and S. Tenreyro. 2011. Further simulation evidence on the performance of the poisson pseudo-maximum likelihood estimator. *Economics Letters* 112 (2):220–22. doi:10.1016/j.econlet.2011.05.008.

Wooldridge, J. M. 2007. Inverse probability weighted estimation for general missing data problems. *Journal of Econometrics* 141 (2):1281–301. doi:10.1016/j.jeconom.2007.02.002.

Appendices

Appendix A

One note of caution when considering IPWRA estimates is that extreme outliers in probability of treatment could exercise influence. The large sample size should minimize the impact of any one observation. Although the majority of the data is assigned a vanishingly small probability weight for each of the individual estimations of disaggregated currency unions.[12] Indeed this appears to be a problem that increases with the quality of the first stage fit as improvements in the estimates of Table 3 to identify characteristics associated with treatment tend to further discount country-pairs that do not fit the description. Usual rule of thumb estimates consider dropping observations outside of 1%/10% thresholds. Imbens (2004) suggests choosing based on sample size, noting that for a weighted group means estimator that if the maximum allowable weight of any unit is: $1/[N \times (1 - p_{max})]$ for units at the top of the probability spectrum and $1/[N * p_{min}]$ at the bottom. He uses the example that with a sample of 1000, then limiting the impact of any unit to less than 5% then implies a range of [0.02, 0.98]. Thus, because of the large sample size I will allow for a much larger range than is used in much of the propensity score matching literature[13] and drop only observations with a probability below 0.001 and above 0.999. Even this wide range drops an incredible amount of data for all individual currency unions, as most trading partners are poor matches. The challenge is that using a smaller range drops still more data, allowing these outliers more influence. Only in the case of the ECCU are observations above this threshold dropped, and there only a small number. While I do not model this trade-off to find an optimal truncation, I report in this truncated sample the value of the maximum influence, as measured by $1/[N * p_{min}]$. The results for my IPWRA estimates on this truncated sample are reported in Table A1.

Table A1. Maximum influence in truncated samples.

	ALL	EMU	CFA	ECCU	Aus. Dollar	Pound	Rupee	US Dollar
OLS								
N	721,385	118,250	125,936	3,643	13,455	129,104	6,038	39,194
$1/N*p_{min}$	0.0014	0.0085	0.0079	0.275	0.0743	0.0077	0.165	0.0255
PPML								
N	1,043,696	136,026	297,397	5,366	12,863	128,686	11,053	43,476
$1/N*p_{min}$	0.00096	0.0074	0.0034	0.1864	0.0778	0.0077	0.0904	0.023

Appendix B Estimates of Logistic Regression on PPML Sample

Here I report first stage estimates of the logistic regression on the PPML sample. While there are a few small differences, this is quite similar to the result from the smaller log-gravity first stage.

Table B1. First stage logistic regression, PPML estimates.

	All CUs	EMU	CFA	ECCU	Aussie Dollar	British Pound	Rupee	US Dollar
$Log(|GDP_1-GDP_2|)$	−0.08***	0.17***	−0.33***	−0.60***	1.65***	0.98***	2.10***	2.81***
	(0.01)	(0.02)	(0.01)	(0.05)	(0.07)	(0.04)	(0.09)	(0.06)
$Log(|GDPpc_1-GDPpc_2|)$	−0.12***	−0.12***	−0.13***	−0.30***	0.83***	−0.19***	−0.58***	−0.22***
	(0.01)	(0.02)	(0.01)	(0.05)	(0.10)	(0.03)	(0.04)	(0.04)
Log(Dist)	−1.34***	−1.15***	−1.41***	−3.18***	−1.76***	−0.86***	−2.25***	−1.59***
	(0.01)	(0.02)	(0.01)	(0.08)	(0.10)	(0.05)	(0.09)	(0.06)
$Log(|GDP_1GDP_2|)$	−1.61***	1.14***	10.76***	6.06***	42.80***	17.31***	13.03***	4.48***
	(0.05)	(0.28)	(0.27)	(0.98)	(5.01)	(1.45)	(1.96)	(0.82)
$Log(|GDPpc_1GDPpc_2|)$	−2.69***	13.87***	8.21***	16.18***	57.58***	5.97***	2.12	1.43*
	(0.05)	(0.97)	(0.20)	(1.16)	(7.52)	(0.69)	(1.22)	(0.68)
$|growth_1-growth_2|$	−1.68***	−8.08***	−0.93***	−3.35**	2.24**	−1.51*	−3.60**	1.32*
	(0.12)	(0.74)	(0.14)	(1.02)	(1.02)	(0.72)	(1.31)	(0.63)
$Log(|GDP_1GDP_2|)^2$	0.02***	−0.01***	−0.12***	0.09***	−0.09***	−0.18***	−0.14***	−0.05***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.05)	(0.01)	(0.02)	(0.01)
$Log(|GDPpc_1GDPpc_2|)^2$	0.08***	−0.31***	−0.29***	−0.44***	−1.55***	−0.18***	−0.11**	−0.03
	(0.00)	(0.02)	(0.01)	(0.03)	(0.20)	(0.02)	(0.04)	(0.02)
Observations	1480694	1466302	1473768	1465194	1463856	1464220	1463900	1464358
Pseudo R-squared	0. 280	0.477	0.394	0.897	0.503	0.164	0.556	0.513

Note: Standard errors clustered by country-pairs, * $p < .10$, &** $p < .05$,***_p < .01.

Note that in the PPML estimates of **Table A.3.2**. a number of estimates in the IPWRA truncated sample are missing their standard error. In these cases the variance matrix was non-singular. These standard errors are clustered by pair, and perhaps a different method of clustering may be useful if exploring these results further.

Table C2. Gravity equation: PPML estimates, no pair fixed effects.

	Baseline OLS	Baseline OLS (Truncated)	IPWRA	IPWRA (Truncated)
All CUs	−0.08	−0.08	0.16	0.13
clustered s.e.	(0.07)	(0.07)	(0.05) ***	(0.05) ***
R^2	0.943	0.945	0.967	0.966
N	1,397,203	1,238,501	1,397,203	1,238,501
EMU	−0.12	−0.09	−0.09	−0.10
clustered s.e.	(0.07)	(0.08)	(0.07)	(0.08)
R^2	0.944	0.936	0.970	0.969
N	1,381,343	139,370	1,381,343	139,370
CFA Franc	0.50	1.11	0.51	1.71
clustered s.e.	(0.25) *	(0.22) ***	(0.23) **	(0.23)***
R^2	0.941	0.832	0.944	0.834
N	1,388,687	444,930	1,388,687	444,930
ECCU	1.17	1.00	1.51	1.12
clustered s.e.	(0.31) ***	(0.23) ***	(0.30) ***	(.)
R^2	0.941	0.882	0.943	0.822
N	1,381,511	5,709	1,381,511	5,709
Aussie $	1.96	1.01	1.94	0.93
clustered s.e.	(0.27)***	(0.32)***	(0.24)***	(.)
R^2	0.930	0.856	0.936	0.798
N	1,377,454	14,937	1,377,454	14,937
British Pound	1.61	0.36	1.82	0.59
clustered s.e.	(0.17) ***	(0.13)***	(0.14) ***	(.)
R^2	0.941	0.832	0.976	0.984
N	1,379,292	132,860	1,379,292	132,860
Indian Rupee	−0.79	0.65	−0.20	1.02
clustered s.e.	(0.36)**	(0.48)	(0.22)	(0.49)**
R^2	0.941	0.873	0.946	0.928
N	1,378,965	11,465	1,378,965	11,465
US $	−0.29	0.11	−0.31	0.07
clustered s.e.	(0.16)*	(0.37)	(0.10)***	(.)
R^2	0.941	0.959	0.977	0.978
N	1,379,435	43,783	1,379,435	43,783

Note: Each currency union estimate reflects a separate estimation of Equation (1) either by standard OLS or using the IPWRA adjustment in Equation (3). Results are presented on the full sample (excluding other currency unions) and truncated such that estimated probabilities must fall within a [0.001,0.999] range. All estimates include exporter-year, importer-year fixed effects and traditional gravity variable controls. Standard errors are reported for both heteroskedastic robust (Huber/White) errors, and with clustering by country-pairs, *$p < .10$, &**$p < .05$, ***_$p < .01$.