Online Supplement (OS) (Not Intended for Publication)

This Online Supplement is intended purely to provide supplemental results to the paper that readers may find informative. Sections 1 and 2 of the Online Supplement discuss two Monte Carlo simulation analyses conducted to evaluate various estimators under alternative error structures but any conclusions drawn from the results should be done with care. OLS and PPML are conditional mean estimators while our suggested QR approach focuses on the conditional median, so the results cannot be directly compared. Any deviations from the set parameters in the simulations are reflective that the estimators are targeting different aspects of the conditional distribution (mean versus median), and should therefore be interpreted with caution. First, we use a Monte Carlo approached based upon the panel methodology in Poissonnier (2019). Among numerous results, we generally find that PPML with zeros has lower parameter deviations (from simulation coefficient value of 0.5 for EIA) relative to our three-step QR approach when the conditional variance of the trade flows is a constant (Case 1) or when the conditional variance of the flows is equal to its conditional mean (scaled by the index of dispersion) as in the Poisson distribution (Case 2), the latter in accord with our expectation. By contrast, if the conditional variance of the flows is equal to the (scaled) square of the conditional mean (Case 3) or is a quadratic function of the conditional mean (Case 4), our three-stage QR approach generally has lower parameter deviations. Second, we also provide a simulation analysis in the spirit of Santos Silva and Tenreyro (2011). An interesting outcome from this simulation is that PPML estimates have the lowest parameter deviations when firms are homogeneous (say, in productivities); however, with heterogeneous firms, our three-step QR estimator has lower parameter deviations at the median. Section 3 discusses a method for estimation of the partial effects of regressors on quantiles of the unconditional distribution of the dependent variable or "quantile treatment effect" developed by Firpo (2007). In the last section, we show evidence that our results are robust to including intra-national trade.

1 Monte Carlo Simulations Using the Poissonnier (2019) Approach

In the spirit of Santos Silva and Tenreyro (2006), Head and Mayer (2014), and Martin and Pham (2020), we conduct a large Monte Carlo simulation analysis of the sensitivity of EIA partial effect estimates across a wide array of "error structures." A novel difference of our simulation study with these three previous simulation studies is that we pay particular attention to the *panel* nature of our data. Most researchers, including the three studies cited above, use a cross-section approach and cross-section data in their simulations. To bear resemblance to the canonical gravity expression for panel data in our econometric work, we introduce the it, jt, and ij dimensions to our analysis and we use the same data as in section 4.4.

In this first simulation analysis, we consider ten alternative methods of estimation, three for positive flows only and seven for non-negative flows. For positive flows, we us OLS, PPML, and BVQCM. For non-negative flows, we use PPML, Logit(BVCM)-BVQCM, Logit(BV1FE)-BVQCM, Logit(3FE)-BVQCM, LPM(BVCM)-BVQCM, LPM(BV1FE)-BVQCM, and LPM(3FE)-BVCM (all to account for zeros). Also, we will separate results for positive trade flows only (Panel A in subsequent tables) from the results for non-negative trade flows (Panel B results).

1.1 Methodology

Our approach follows the structural gravity methodology described in Head and Mayer (2014) and adapted in Poissonnier (2019) for panel data. As in Head and Mayer (2014), we start with two RHS variables determining trade flows, a continuous time-invariant variable, $\ln DIST_{ij}$, and a time-varying variable, EIA_{ijt}.[44] In the simulations, the coefficients on $\ln DIST_{ij}$ and EIA_{ijt} are set to -1 and 0.5, respectively, and we use an iterative approach to solve for the multilateral resistance terms Φ_{jt} and Π_{it} using a structural gravity framework as in Poissonnier (2019); we assume positive trade flows. Also as in Poissonnier (2019), the convergence criterion is quadratic such that the matrix $\Pi_{it}\Phi^{'}_{jt}$ is continually updated until convergence, which differs from the approach for each multilateral resistance term in Head and Mayer (2014). The data generating process is defined such that variable trade costs (τ_{ijt}) raised to the trade elasticity ($-\theta$) are:

$$\tau_{ijt}^{-\theta} = exp(-\ln DIST_{ij} + 0.5EIA_{ijt}) * \eta_{ijt} \tag{18}$$

where η_{ijt} is a multiplicative error term such that $\eta_{ijt} = 1+\epsilon_{ijt}/\exp(\mathbf{Z}^{'}_{ijt}\beta)$ and $E(\eta_{ijt}|\mathbf{Z}_{ijt}) = 1$. We define the variance of η_{ijt} as σ^2_{ijt}. For notational convenience going forward, we define a term μ_{ijt} such that $\mu_{ijt} = exp(-\ln DIST_{ij} + 0.5EIA_{ijt})$. Hence, $\tau_{ijt}^{-\theta} = \mu_{ijt}\eta_{ijt}$, as in Santos Silva and Tenreyro (2006).

To see how estimates from our alternative estimators are sensitive to the structure of gravity's errors, we consider the same four cases for the error structure presented in Santos

[44] For the BV equivalent time-invariant and time-varying variables, each are constructed after the simulation data is created for each iteration.

Silva and Tenreyro (2006). Formally, the four cases we consider are:

1. Case 1: $\sigma^2_{ijt} = h \times \mu^{-2}_{ijt}$; $Var[X_{ijt}|EIA_{ijt}, \ln DIST_{ij}] = h$
2. Case 2: $\sigma^2_{ijt} = h \times \mu^{-1}_{ijt}$; $Var[X_{ijt}|EIA_{ijt}, \ln DIST_{ij}] = h \times \mu_{ijt}$
3. Case 3: $\sigma^2_{ijt} = h$; $Var[X_{ijt}|EIA_{ijt}, \ln DIST_{ij}] = h \times \mu^2_{ijt}$
4. Case 4: $\sigma^2_{ijt} = h \times (\mu^{-1}_{ijt} + \exp(x_{2ijt}))$; $Var[X_{ijt}|EIA_{ijt}, \ln DIST_{ij}] = h \times (\mu_{ijt} + \exp(x_{2ijt})\mu^2_{ijt})$

where the variable x_{2ijt} is a binary variable with mean 0.2.

We summarize each of the cases. In Case 1, the variance of X_{ijt} is a constant (h), which implies that the non-linear least squares estimator is optimal; as discussed below, we set $h = 4$ in the benchmark case. Santos Silva and Tenreyro (2006) argue this case is unrealistic for bilateral trade, but – as there – we include it for completeness. In Case 2, the conditional variance of X_{ijt} is equal to its conditional mean, scaled by the index of dispersion h, as in the Poisson distribution. In this case, PPML is the optimal estimator and $\lambda = 1$. In Case 3, $\lambda = 2$, so the conditional variance of X_{ijt} is equal to the square of its conditional mean, scaled by the index of dispersion h, as in the Gamma distribution. In this case, Gamma PML is the optimal estimator. In Case 4, the conditional variance of X_{ijt} is a quadratic function of the mean, but it is not proportional to the square of the mean.

Note that we use the same Constant Variance Mean Ratio (Case 2) and Constant Coefficient of Variation (Case 3) notation that is specified in Head and Mayer (2014). Similar to Head and Mayer (2014), we include an overdisperson parameter, h, that is set initially to 4 as in Head and Mayer (2014). PPML should still remain consistent and efficient. We also provide simulations where h was set to 1 and 10.

Unfortunately, the data generating process does not naturally generate zeros. Zeros can potentially be caused by economic considerations, such as export fixed costs relative to variable profits, or by institutional constraints that arbitrarily create cutoffs resulting in zeros in the data. We are agnostic here about their source. *However*, we need to create zeros for this simulation as a share of trade flows consistent with observed patterns. Consequently, we follow Head and Mayer (2014), p. 180, to generate zeros. As in a standard Melitz model, we assume that variable profits of a firm in country i for selling to country j (say, X_{ijt}/α) must exceed export fixed costs f_{ijt} to enter the market, where α is defined as the elasticity of substitution in consumption. Hence, trade can only occur if $X_{ijt} \geq \alpha f_{ijt}$. So we create a threshold (zero profit cutoff) such that trade is positive if:

$$X_{ijt} = \begin{cases} X_{ijt}, & \text{if } X_{ijt} \geq \alpha f_{ijt} \\ 0, & \text{if } X_{ijt} < \alpha f_{ijt} \end{cases} . \tag{19}$$

The mean and variance of the threshold are set to mimic the proportion of zeros observed in the current sample of countries from 1965 to 2010 at 5 year intervals. Total trade cost is defined as:

$$\phi_{ijt} = \tau_{ijt}^{-\theta} f_{it}^{-\left[\frac{\theta}{\alpha-1}-1\right]} \tag{20}$$

Following Head and Mayer (2014), we set $\theta = 5$ and $\theta/(\alpha - 1) = 2.5$ which implies that $\alpha = 3$. Head and Mayer (2014) note that $\theta/(\alpha - 1) = 2.5$ matches the estimates provided by Eaton et al. (2011) of this parameter.

As emphasized earlier, the OLS and PPML estimators are conditional mean estimators while our QR approach is focused on the conditional median (Q_{50}). We report the mean of the estimates and standard error for EIA in 250 iterations. In addition, the standard deviation of all 250 estimates is also provided and "Parm. Dev." is the percentage deviation of the mean of the estimates from 0.5.

1.2 Benchmark Simulation Results

For the benchmark simulations, we run all the models using our benchmark treatment of zeros in the data (adding ones only to the zeros). For all the results, we report the coefficient estimates from 250 iterations of the specifications. For OLS and PPML, we include it, jt, and ij fixed effects. For BVQCM, Logit-BVQCM, and LPM-BVQCM, we include our correlated random effects. For OLS and PPML specifications, we report the mean coefficient estimate for EIA_{ijt}, the mean standard error of the estimate, the standard deviation of all 250 coefficient estimates, and the parameter deviation (which is two times the deviation of the estimate from 0.5). For BVQCM, Logit-BVQCM, and LPM-BVQCM, we report the mean coefficient estimate for $EIABV_{ijt}$, the mean standard error of the estimate, the standard deviation of all 250 estimates, and the parameter deviation.

Table OS1 reports the benchmark results. Note that we consider all the estimators used earlier in the empirical analysis. As indicated above, the table (and subsequent tables) are divided into two panels. The top panel uses only positive trade flows. The bottom panel uses all non-negative trade flows. Going down the rows of the top panel in Table OS1, the three estimators are OLS, PPML, and BVQCM, using positive trade flows only for the four different error-structure Cases 1-4. In the bottom panel, we examine PPML, the three

Logit-BVQCM models, and the three LPM-BVQCM models using positive trade flows and zeros. For PPML, we use actual zeros, as standard.

For tractability, we discuss first the results for the top panel A for the first seven specifications, Tables OS1-OS7. In Table OS1 for Panel A using only positive trade flows, PPML has the lowest parameter deviations for all four cases, 1-4. OLS has the next lowest and BVQCM has the highest. The finding that PPML has the lowest parameter deviations across Cases 1-4 for positive flows is robust across *all seven specifications*, with the exception of the fifth specification (Table OS5) when OLS has slightly lower parameter deviations in Cases 3 and 4. For brevity (so that we can move on to Panel B for non-negative flows), we note that OLS generally has for the second lowest parameter deviations and BVQCM has the highest, except in two specifications.

We now move to the results in Panel B for non-negative trade for the remainder of this discussion. One important – and expected – main finding is that PPML has the lowest parameter deviations for Case 2, for which PPML should be optimal; recall that in Case 2, $\lambda = 1$ and the optimal estimator for the Poisson distribution is the PPML. However, in one alternative specification – Table OS4 with an increase in zeros by 25% – then Logit(BVCM)-BVQCM has a marginally lower parameter deviation in Case 2. Second, for Case 1 PPML has the lowest parameter deviations across five of the seven alternative specifications. Thus, despite PPML always having the lowest parameter deviations for all seven specifications for positive flows for Cases (error structures) 1 and 2, for non-negative flows some QR estimates have lower parameter deviations.

We now address the results for non-negative trade flows for all four Cases 1-4 across seven alternative specifications. Table OS1 reports the results across the four cases (error structures) for our benchmark specification (adding 1s only to 0s). For Case 1, Logit(BVCM)-BVQCM has a slightly lower parameter deviation than the other specifications, including PPML. However, for Case 2 PPML has the lowest parameter deviation, as expected. For Cases 3 and 4, however, Logit(3FE)-BVQCM have the lowest parameter deviations.

The remaining tables in section 1 of this Online Supplement report robustness analyses of the benchmark results in Table OS1. Table OS2 reports the results when we alter our benchmark treatment of zeros from adding 1's to only zeros to adding 1's to all observations (as done earlier in our empirical work). As shown in Panel B, for Cases 1 and 2, PPML has the lowest parameter deviations. For Case 3, LPM(BV1FE)-BVQCM has the lowest parameter deviation. For Case 4, Logit(BV1FE)-BVQCM has the lowest parameter deviation.

In the third simulation shown in Table OS3, we consider an alternative cutoff for zeros, as in the empirical analysis. Here we have a cutoff value for zeros at USD 500,000. For Cases 1 and 2 PPML again has the lowest parameter deviations for non-negative flows. For Cases 3 and 4, LPM(BV1FE)-BVQCM has the lowest parameter deviations.

In the fourth simulation, we alter the percentage of zeros in the data. Specifically, we increase the number of zeros by 25 percent. Table OS4 presents the results. In this specification, it is Logit(BV1FE)-BVQCM that has the lowest parameter deviations in Cases 1 and 2, not PPML. However, in Case 3 Logit(BVCM)-BVQCM has the lowest parameter deviation. In Case 4, LPM(BVCM)-BVQCM has the lowest parameter deviation.

In the fifth simulation, we decrease the number of zeros by 25 percent relative to the benchmark. Table OS5 presents the results. In Cases 1 and 2, PPML has the lowest parameter deviations. In Cases 3 and 4, Logit(3FE)-BVQCM has the lowest parameter deviations.

In the sixth simulation, we address overdispersion. We reduced the overdispersion index (h) from $h = 4$ to $h = 1$. Table OS6 presents the results. In Cases 1 and 2, PPML has the lowest parameter deviations. In Cases 3 and 4, LPM(BVCM)-BVQCM has the lowest parameter deviations.

In the seventh simulation, we address overdispersion again. Here we increased the overdispersion index (h) from the benchmark of $h = 4$ to $h = 10$. Table OS7 presents the results. For Case 1, we find that Logit(BV1FE)-BVQCM now has the lowest parameter deviation. However, as expected, PPML has the lowest parameter deviation for Case 2. For Case 3, Logit(3FE)-BVQCM has the lowest parameter deviation. For Case 4, Logit(BVCM)-BVQCM has the lowest parameter deviation.

Finally, for robustness we repeat simulations 1, 2, and 7 and report the results for Q_{25} and Q_{75} only reporting the three Logit-BVQCM models and the three LPM-BVQCM models *for non-negative trade flows only.* As one would expect, the parameter deviations are larger than the results for the median.

In summary of these simulations, our results indicate that the parameter deviations for each estimator depend materially on the underlying error structure. Though not reported in Tables OS8-OS10, we find unsurprisingly for the cases of constant variance (Case 1) and a Poisson distribution for errors (Case 2) that PPML has the lowest parameter deviations. However, as for the earlier simulations, for Cases 3 and 4, we find significant evidence that Logit-BVQCM or LPM-BVQCM yield the lowest parameter deviations in estimates for these non-negative flows.

Table OS1: Benchmark case of adding 1s to 0s only

Panel A:	EIA			
	Case 1	Case 2	Case 3	Case 4
	Trade>0			
(1)	(2)	(3)	(4)	(5)
OLS	0.5156	0.5399	0.5881	0.5947
Std. Dev.	0.0034	0.0066	0.0272	0.0275
S.E.	0.0029	0.0060	0.0264	0.0279
Parm. Dev.	0.0312	0.0798	0.1761	0.1894
PPML Pos	0.5003	0.5013	0.4447	0.4327
Std. Dev.	0.0001	0.0011	0.1171	0.1210
S.E.	0.0001	0.0009	0.0795	0.0839
Parm. Dev.	0.0005	0.0025	0.1106	0.1346
BVQCM Pos	0.5411	0.5578	0.5906	0.5928
Std. Dev.	0.0065	0.0090	0.0407	0.0437
S.E.	0.0108	0.0126	0.0372	0.0390
Parm. Dev.	0.0823	0.1156	0.1812	0.1856
Panel B:	Trade≥0			
(1)	(2)	(3)	(4)	(5)
PPML	0.4953	0.4959	0.4354	0.4235
Std. Dev.	0.0005	0.0012	0.1169	0.1208
S.E.	0.0005	0.0010	0.0794	0.0837
Parm. Dev.	0.0094	0.0081	0.1293	0.1531
Logit(BVCM)-BVQCM	0.5036	0.5132	0.5365	0.5350
Std. Dev.	0.0073	0.0095	0.0536	0.0591
S.E.	0.0130	0.0141	0.0491	0.0524
Parm. Dev.	0.0072	0.0264	0.0729	0.0701
Logit(BV1FE)-BVQCM	0.4815	0.5128	0.5260	0.5412
Std. Dev.	0.0193	0.0246	0.0697	0.0783
S.E.	0.0233	0.0272	0.0679	0.0701
Parm. Dev.	0.0371	0.0257	0.0521	0.0823
Logit(3FE)-BVQCM	0.4977	0.5353	0.5081	0.5204
Std. Dev.	0.0172	0.0217	0.0590	0.0601
S.E.	0.0199	0.0240	0.0565	0.0590
Parm. Dev.	0.0047	0.0706	0.0161	0.0407
LPM(BVCM)-BVQCM	0.3489	0.3888	0.5742	0.5914
Std. Dev.	0.0200	0.0186	0.0435	0.0465
S.E.	0.0241	0.0239	0.0376	0.0389
Parm. Dev.	0.3022	0.2224	0.1484	0.1828
LPM(BV1FE)-BVQCM	0.3317	0.3707	0.5456	0.5626
Std. Dev.	0.0209	0.0187	0.0432	0.0465
S.E.	0.0248	0.0245	0.0377	0.0389
Parm. Dev.	0.3367	0.2587	0.0913	0.1252
LPM(3FE)-BVQCM	0.3306	0.3695	0.5418	0.5584
Std. Dev.	0.0203	0.0187	0.0431	0.0462
S.E.	0.0246	0.0243	0.0378	0.0390
Parm. Dev.	0.3388	0.2609	0.0836	0.1167

Standard errors are clustered at the country-pair. The simulations used 250 iterations, Stata 18.0, and R 4.3.2. Adjustment to dependent only occurred when the log of trade values are used. PPML estimators use trade values in levels without any adjustment to dependent variable. Parm. Dev. is percentage difference (in decimal form) from 0.5 (or the EIA effect).

Table OS2: Adding 1s to all observations

Panel A:	EIA			
	Case 1	Case 2	Case 3	Case 4
	Trade>0			
(1)	(2)	(3)	(4)	(5)
OLS	0.5209	0.5446	0.5816	0.5845
Std. Dev.	0.0028	0.0062	0.0256	0.0262
S.E.	0.0028	0.0057	0.0257	0.0270
Parm. Dev.	0.0418	0.0891	0.1632	0.1689
PPML Pos	0.5007	0.5017	0.4488	0.4226
Std. Dev.	0.0001	0.0011	0.1259	0.1247
S.E.	0.0001	0.0009	0.0799	0.0845
Parm. Dev.	0.0014	0.0033	0.1025	0.1548
BVQCM Pos	0.5355	0.5584	0.5591	0.5582
Std. Dev.	0.0064	0.0100	0.0373	0.0397
S.E.	0.0113	0.0129	0.0354	0.0370
Parm. Dev.	0.0710	0.1169	0.1182	0.1164

Panel B:	Trade≥0			
(1)	(2)	(3)	(4)	(5)
PPML	0.4963	0.4968	0.4400	0.4144
Std. Dev.	0.0006	0.0012	0.1253	0.1243
S.E.	0.0004	0.0010	0.0797	0.0842
Parm. Dev.	0.0074	0.0064	0.1199	0.1713
Logit(BVCM)-BVQCM	0.4962	0.5087	0.5057	0.5128
Std. Dev.	0.0078	0.0096	0.0477	0.0537
S.E.	0.0128	0.0137	0.0468	0.0498
Parm. Dev.	0.0076	0.0175	0.0114	0.0255
Logit(BV1FE)-BVQCM	0.4652	0.4999	0.4851	0.4969
Std. Dev.	0.0192	0.0232	0.0694	0.0704
S.E.	0.0231	0.0266	0.0640	0.0664
Parm. Dev.	0.0697	0.0002	0.0299	0.0062
Logit(3FE)-BVQCM	0.4871	0.5254	0.4701	0.4759
Std. Dev.	0.0162	0.0206	0.0547	0.0597
S.E.	0.0200	0.0240	0.0515	0.0539
Parm. Dev.	0.0259	0.0507	0.0599	0.0481
LPM(BVCM)-BVQCM	0.3407	0.3719	0.5245	0.5441
Std. Dev.	0.0195	0.0185	0.0426	0.0430
S.E.	0.0243	0.0239	0.0367	0.0377
Parm. Dev.	0.3185	0.2562	0.0489	0.0881
LPM(BV1FE)-BVQCM	0.3238	0.3538	0.4978	0.5157
Std. Dev.	0.0201	0.0188	0.0425	0.0427
S.E.	0.0250	0.0245	0.0369	0.0379
Parm. Dev.	0.3524	0.2924	0.0045	0.0314
LPM(3FE)-BVQCM	0.3224	0.3534	0.4937	0.5119
Std. Dev.	0.0198	0.0187	0.0430	0.0429
S.E.	0.0248	0.0244	0.0371	0.0381
Parm. Dev.	0.3551	0.2931	0.0125	0.0237

Standard errors are clustered at the country-pair. The simulations used 250 iterations, Stata 18.0, and R 4.3.2. Adjustment to dependent only occurred when the log of trade values are used. PPML estimators use trade values in levels without any adjustment to dependent variable. Parm. Dev. is percentage difference (in decimal form) from 0.5 (or the EIA effect).

Table OS3: Cutoff Value for Trade to USD 500,000

Panel A:	EIA			
	Case 1	Case 2	Case 3	Case 4
	Trade>0			
(1)	(2)	(3)	(4)	(5)
OLS	0.5153	0.5405	0.5856	0.5932
Std. Dev.	0.0033	0.0066	0.0262	0.0256
S.E.	0.0029	0.0060	0.0263	0.0278
Parm. Dev.	0.0307	0.0809	0.1712	0.1863
PPML Pos	0.5002	0.5012	0.4443	0.4235
Std. Dev.	0.0001	0.0011	0.1206	0.1276
S.E.	0.0001	0.0009	0.0803	0.0840
Parm. Dev.	0.0005	0.0024	0.1113	0.1530
BVQCM Pos	0.5632	0.5596	0.6105	0.6252
Std. Dev.	0.0061	0.0077	0.0331	0.0351
S.E.	0.0106	0.0113	0.0352	0.0370
Parm. Dev.	0.1265	0.1193	0.2211	0.2504
Panel B:	Trade≥0			
(1)	(2)	(3)	(4)	(5)
PPML	0.4952	0.4959	0.4350	0.4140
Std. Dev.	0.0005	0.0012	0.1206	0.1272
S.E.	0.0005	0.0010	0.0801	0.0839
Parm. Dev.	0.0095	0.0082	0.1300	0.1720
Logit(BVCM)-BVQCM	0.5252	0.5286	0.5513	0.5610
Std. Dev.	0.0051	0.0071	0.0371	0.0380
S.E.	0.0101	0.0110	0.0387	0.0408
Parm. Dev.	0.0504	0.0573	0.1025	0.1220
Logit(BV1FE)-BVQCM	0.5112	0.5144	0.5593	0.5792
Std. Dev.	0.0156	0.0214	0.0499	0.0525
S.E.	0.0198	0.0222	0.0517	0.0534
Parm. Dev.	0.0225	0.0289	0.1187	0.1584
Logit(3FE)-BVQCM	0.5227	0.5150	0.5370	0.5554
Std. Dev.	0.0144	0.0198	0.0447	0.0479
S.E.	0.0181	0.0204	0.0480	0.0500
Parm. Dev.	0.0455	0.0299	0.0741	0.1108
LPM(BVCM)-BVQCM	0.4398	0.4463	0.5523	0.5684
Std. Dev.	0.0086	0.0097	0.0302	0.0335
S.E.	0.0106	0.0119	0.0317	0.0330
Parm. Dev.	0.1204	0.1073	0.1047	0.1367
LPM(BV1FE)-BVQCM	0.4266	0.4236	0.5189	0.5344
Std. Dev.	0.0092	0.0102	0.0303	0.0334
S.E.	0.0110	0.0123	0.0317	0.0330
Parm. Dev.	0.1469	0.1527	0.0378	0.0689
LPM(3FE)-BVQCM	0.4298	0.4287	0.5262	0.5416
Std. Dev.	0.0088	0.0102	0.0298	0.0333
S.E.	0.0109	0.0122	0.0321	0.0334
Parm. Dev.	0.1404	0.1427	0.0524	0.0832

Standard errors are clustered at the country-pair. The simulations used 250 iterations, Stata 18.0, and R 4.3.2. Adjustment to dependent only occurred when the log of trade values are used. PPML estimators use trade values in levels without any adjustment to dependent variable. Parm. Dev. is percentage difference (in decimal form) from 0.5 (or the EIA effect).

Table OS4: Increase in zeros by 25%

Panel A:	EIA			
	Case 1	Case 2	Case 3	Case 4
	Trade>0			
(1)	(2)	(3)	(4)	(5)
OLS	0.5047	0.5285	0.6217	0.6249
Std. Dev.	0.0019	0.0059	0.0302	0.0327
S.E.	0.0016	0.0050	0.0305	0.0320
Parm. Dev.	0.0094	0.0570	0.2435	0.2499
PPML Pos	0.5001	0.5014	0.4551	0.4328
Std. Dev.	0.0000	0.0011	0.1173	0.1282
S.E.	0.0000	0.0009	0.0803	0.0850
Parm. Dev.	0.0002	0.0028	0.0898	0.1344
BVQCM Pos	0.5413	0.5383	0.6109	0.6131
Std. Dev.	0.0084	0.0092	0.0397	0.0399
S.E.	0.0127	0.0130	0.0384	0.0401
Parm. Dev.	0.0826	0.0767	0.2218	0.2262
Panel B:	Trade≥0			
(1)	(2)	(3)	(4)	(5)
PPML	0.4809	0.4813	0.4281	0.4065
Std. Dev.	0.0019	0.0022	0.1161	0.1269
S.E.	0.0018	0.0020	0.0797	0.0844
Parm. Dev.	0.0381	0.0375	0.1438	0.1869
Logit(BVCM)-BVQCM	0.5104	0.5143	0.5359	0.5356
Std. Dev.	0.0062	0.0085	0.0445	0.0489
S.E.	0.0125	0.0133	0.0447	0.0472
Parm. Dev.	0.0207	0.0286	0.0717	0.0712
Logit(BV1FE)-BVQCM	0.5087	0.5118	0.5440	0.5504
Std. Dev.	0.0148	0.0160	0.0511	0.0543
S.E.	0.0189	0.0193	0.0501	0.0524
Parm. Dev.	0.0174	0.0236	0.0880	0.1007
Logit(3FE)-BVQCM	0.5236	0.5204	0.5513	0.5594
Std. Dev.	0.0147	0.0154	0.0466	0.0490
S.E.	0.0178	0.0180	0.0463	0.0485
Parm. Dev.	0.0472	0.0407	0.1025	0.1188
LPM(BVCM)-BVQCM	0.2447	0.2534	0.4521	0.4717
Std. Dev.	0.0177	0.0186	0.0384	0.0435
S.E.	0.0226	0.0230	0.0370	0.0380
Parm. Dev.	0.5107	0.4932	0.0959	0.0566
LPM(BV1FE)-BVQCM	0.1760	0.1808	0.3880	0.4073
Std. Dev.	0.0193	0.0203	0.0393	0.0441
S.E.	0.0243	0.0246	0.0375	0.0386
Parm. Dev.	0.6480	0.6385	0.2240	0.1854
LPM(3FE)-BVQCM	0.2054	0.2129	0.3980	0.4139
Std. Dev.	0.0190	0.0195	0.0386	0.0433
S.E.	0.0235	0.0239	0.0382	0.0393
Parm. Dev.	0.5893	0.5743	0.2039	0.1722

Standard errors are clustered at the country-pair. The simulations used 250 iterations, Stata 18.0, and R 4.3.2. Adjustment to dependent only occurred when the log of trade values are used. PPML estimators use trade values in levels without any adjustment to dependent variable. Parm. Dev. is percentage difference (in decimal form) from 0.5 (or the EIA effect).

Table OS5: Decrease in zeros by 25%

Panel A:	EIA			
	Case 1	Case 2	Case 3	Case 4
	Trade>0			
(1)	(2)	(3)	(4)	(5)
OLS	0.5173	0.5304	0.5571	0.5635
Std. Dev.	0.0047	0.0080	0.0244	0.0268
S.E.	0.0041	0.0070	0.0244	0.0259
Parm. Dev.	0.0346	0.0608	0.1141	0.1270
PPML Pos	0.5003	0.5007	0.4417	0.4140
Std. Dev.	0.0001	0.0010	0.1177	0.1264
S.E.	0.0001	0.0009	0.0805	0.0840
Parm. Dev.	0.0005	0.0013	0.1166	0.1719
BVQCM Pos	0.5618	0.5488	0.5984	0.6075
Std. Dev.	0.0051	0.0075	0.0328	0.0355
S.E.	0.0096	0.0103	0.0333	0.0352
Parm. Dev.	0.1236	0.0975	0.1967	0.2151
Panel B:	Trade≥0			
(1)	(2)	(3)	(4)	(5)
PPML	0.4990	0.4994	0.4388	0.4112
Std. Dev.	0.0002	0.0011	0.1176	0.1263
S.E.	0.0001	0.0009	0.0805	0.0839
Parm. Dev.	0.0020	0.0013	0.1224	0.1777
Logit(BVCM)-BVQCM	0.5300	0.5169	0.5533	0.5588
Std. Dev.	0.0043	0.0072	0.0348	0.0355
S.E.	0.0091	0.0102	0.0331	0.0343
Parm. Dev.	0.0600	0.0338	0.1065	0.1175
Logit(BV1FE)-BVQCM	0.5100	0.5336	0.5478	0.5619
Std. Dev.	0.0229	0.0274	0.0588	0.0583
S.E.	0.0242	0.0285	0.0519	0.0534
Parm. Dev.	0.0201	0.0672	0.0956	0.1237
Logit(3FE)-BVQCM	0.5041	0.5259	0.5222	0.5378
Std. Dev.	0.0187	0.0256	0.0563	0.0585
S.E.	0.0225	0.0267	0.0507	0.0525
Parm. Dev.	0.0081	0.0517	0.0443	0.0755
LPM(BVCM)-BVQCM	0.5207	0.5205	0.5986	0.6102
Std. Dev.	0.0053	0.0075	0.0304	0.0325
S.E.	0.0090	0.0099	0.0308	0.0324
Parm. Dev.	0.0413	0.0410	0.1972	0.2204
LPM(BV1FE)-BVQCM	0.5097	0.5084	0.5839	0.5948
Std. Dev.	0.0052	0.0073	0.0304	0.0336
S.E.	0.0088	0.0096	0.0307	0.0322
Parm. Dev.	0.0194	0.0168	0.1677	0.1897
LPM(3FE)-BVQCM	0.5087	0.5073	0.5817	0.5919
Std. Dev.	0.0051	0.0075	0.0311	0.0332
S.E.	0.0088	0.0096	0.0308	0.0323
Parm. Dev.	0.0174	0.0147	0.1633	0.1839

Standard errors are clustered at the country-pair. The simulations used 250 iterations, Stata 18.0, and R 4.3.2. Adjustment to dependent only occurred when the log of trade values are used. PPML estimators use trade values in levels without any adjustment to dependent variable. Parm. Dev. is percentage difference (in decimal form) from 0.5 (or the EIA effect).

Table OS6: Overdispersion Index Reduced from 4 to 1, i.e. $h = 1$

Panel A:	EIA			
	Case 1	Case 2	Case 3	Case 4
	Trade>0			
(1)	(2)	(3)	(4)	(5)
OLS	0.5068	0.5154	0.5603	0.5665
Std. Dev.	0.0021	0.0038	0.0178	0.0198
S.E.	0.0018	0.0034	0.0177	0.0194
Parm. Dev.	0.0136	0.0308	0.1205	0.1329
PPML Pos	0.5001	0.5003	0.4720	0.4534
Std. Dev.	0.0000	0.0005	0.0853	0.0975
S.E.	0.0000	0.0005	0.0549	0.0599
Parm. Dev.	0.0001	0.0007	0.0561	0.0931
BVQCM Pos	0.5647	0.5610	0.5978	0.5990
Std. Dev.	0.0061	0.0069	0.0233	0.0248
S.E.	0.0106	0.0106	0.0242	0.0258
Parm. Dev.	0.1294	0.1221	0.1956	0.1980
Panel B:	Trade≥0			
(1)	(2)	(3)	(4)	(5)
PPML	0.4953	0.4954	0.4650	0.4462
Std. Dev.	0.0005	0.0007	0.0851	0.0973
S.E.	0.0005	0.0006	0.0548	0.0598
Parm. Dev.	0.0094	0.0092	0.0700	0.1076
Logit(BVCM)-BVQCM	0.5294	0.5281	0.5451	0.5420
Std. Dev.	0.0051	0.0059	0.0233	0.0268
S.E.	0.0100	0.0102	0.0260	0.0279
Parm. Dev.	0.0587	0.0562	0.0902	0.0841
Logit(BV1FE)-BVQCM	0.5186	0.5132	0.5330	0.5350
Std. Dev.	0.0177	0.0196	0.0390	0.0403
S.E.	0.0204	0.0202	0.0385	0.0405
Parm. Dev.	0.0371	0.0264	0.0661	0.0700
Logit(3FE)-BVQCM	0.5300	0.5159	0.5216	0.5263
Std. Dev.	0.0157	0.0164	0.0331	0.0368
S.E.	0.0180	0.0184	0.0353	0.0373
Parm. Dev.	0.0600	0.0318	0.0431	0.0525
LPM(BVCM)-BVQCM	0.4426	0.4415	0.5014	0.5076
Std. Dev.	0.0086	0.0090	0.0220	0.0238
S.E.	0.0103	0.0109	0.0233	0.0246
Parm. Dev.	0.1147	0.1171	0.0029	0.0152
LPM(BV1FE)-BVQCM	0.4298	0.4259	0.4699	0.4767
Std. Dev.	0.0091	0.0093	0.0221	0.0242
S.E.	0.0107	0.0113	0.0234	0.0247
Parm. Dev.	0.1403	0.1483	0.0603	0.0465
LPM(3FE)-BVQCM	0.4331	0.4295	0.4759	0.4827
Std. Dev.	0.0086	0.0091	0.0221	0.0240
S.E.	0.0106	0.0112	0.0235	0.0248
Parm. Dev.	0.1338	0.1411	0.0483	0.0346

Standard errors are clustered at the country-pair. The simulations used 250 iterations, Stata 18.0, and R 4.3.2. Adjustment to dependent only occurred when the log of trade values are used. PPML estimators use trade values in levels without any adjustment to dependent variable. Parm. Dev. is percentage difference (in decimal form) from 0.5 (or the EIA effect).

Table OS7: Overdispersion Index Increased from 4 to 10, i.e. $h = 10$

Panel A:	EIA			
	Case 1	Case 2	Case 3	Case 4
	Trade>0			
(1)	(2)	(3)	(4)	(5)
OLS	0.5251	0.5655	0.5930	0.5934
Std. Dev.	0.0041	0.0090	0.0322	0.0341
S.E.	0.0039	0.0084	0.0317	0.0330
Parm. Dev.	0.0502	0.1311	0.1861	0.1868
PPML Pos	0.5005	0.5022	0.4370	0.3928
Std. Dev.	0.0001	0.0017	0.1415	0.1324
S.E.	0.0001	0.0015	0.0937	0.0982
Parm. Dev.	0.0010	0.0043	0.1260	0.2143
BVQCM Pos	0.5633	0.5650	0.6203	0.6261
Std. Dev.	0.0059	0.0092	0.0458	0.0444
S.E.	0.0108	0.0124	0.0421	0.0439
Parm. Dev.	0.1266	0.1301	0.2406	0.2523
Panel B:	Trade≥0			
(1)	(2)	(3)	(4)	(5)
PPML	0.4952	0.4965	0.4271	0.3826
Std. Dev.	0.0006	0.0017	0.1414	0.1316
S.E.	0.0005	0.0015	0.0935	0.0980
Parm. Dev.	0.0095	0.0070	0.1458	0.2349
Logit(BVCM)-BVQCM	0.5235	0.5337	0.5577	0.5605
Std. Dev.	0.0055	0.0088	0.0502	0.0465
S.E.	0.0102	0.0122	0.0463	0.0474
Parm. Dev.	0.0469	0.0673	0.1153	0.1211
Logit(BV1FE)-BVQCM	0.5030	0.5248	0.5951	0.5996
Std. Dev.	0.0161	0.0243	0.0602	0.0591
S.E.	0.0198	0.0254	0.0586	0.0601
Parm. Dev.	0.0060	0.0496	0.1902	0.1992
Logit(3FE)-BVQCM	0.5148	0.5266	0.5567	0.5657
Std. Dev.	0.0152	0.0214	0.0538	0.0540
S.E.	0.0184	0.0237	0.0550	0.0569
Parm. Dev.	0.0296	0.0532	0.1134	0.1314
LPM(BVCM)-BVQCM	0.4375	0.4573	0.5892	0.5986
Std. Dev.	0.0085	0.0098	0.0422	0.0382
S.E.	0.0110	0.0129	0.0366	0.0378
Parm. Dev.	0.1251	0.0854	0.1784	0.1972
LPM(BV1FE)-BVQCM	0.4225	0.4316	0.5591	0.5675
Std. Dev.	0.0089	0.0104	0.0418	0.0379
S.E.	0.0114	0.0132	0.0368	0.0380
Parm. Dev.	0.1550	0.1368	0.1183	0.1351
LPM(3FE)-BVQCM	0.4272	0.4354	0.5642	0.5709
Std. Dev.	0.0089	0.0103	0.0419	0.0373
S.E.	0.0112	0.0131	0.0372	0.0385
Parm. Dev.	0.1456	0.1293	0.1283	0.1418

Standard errors are clustered at the country-pair. The simulations used 250 iterations, Stata 18.0, and R 4.3.2. Adjustment to dependent only occurred when the log of trade values are used. PPML estimators use trade values in levels without any adjustment to dependent variable. Parm. Dev. is percentage difference (in decimal form) from 0.5 (or the EIA effect).

Table OS8: Benchmark case of adding 1s to 0s only

Panel A:	Q_{25} EIA			
	Case 1	Case 2	Case 3	Case 4
	Trade≥0			
(1)	(2)	(3)	(4)	(5)
Logit(BVCM)-BVQCM	0.5035	0.5129	0.5352	0.5360
Std. Dev.	0.0066	0.0095	0.0508	0.0571
S.E.	0.0130	0.0142	0.0490	0.0525
Parm. Dev.	0.0071	0.0258	0.0704	0.0720
Logit(BV1FE)-BVQCM	0.4808	0.5118	0.5285	0.5260
Std. Dev.	0.0201	0.0256	0.0702	0.0775
S.E.	0.0233	0.0274	0.0678	0.0706
Parm. Dev.	0.0383	0.0235	0.0571	0.0521
Logit(3FE)-BVQCM	0.4987	0.5327	0.5088	0.5105
Std. Dev.	0.0168	0.0217	0.0585	0.0620
S.E.	0.0199	0.0241	0.0565	0.0592
Parm. Dev.	0.0025	0.0653	0.0176	0.0210
LPM(BVCM)-BVQCM	0.3470	0.3855	0.5758	0.5925
Std. Dev.	0.0197	0.0180	0.0402	0.0410
S.E.	0.0240	0.0239	0.0376	0.0387
Parm. Dev.	0.3060	0.2290	0.1517	0.1851
LPM(BV1FE)-BVQCM	0.3293	0.3672	0.5487	0.5621
Std. Dev.	0.0204	0.0182	0.0388	0.0423
S.E.	0.0247	0.0245	0.0377	0.0387
Parm. Dev.	0.3413	0.2657	0.0974	0.1241
LPM(3FE)-BVQCM	0.3285	0.3665	0.5456	0.5583
Std. Dev.	0.0200	0.0179	0.0394	0.0418
S.E.	0.0245	0.0243	0.0378	0.0388
Parm. Dev.	0.3430	0.2671	0.0912	0.1165
Panel B:	Q_{75} EIA			
	Case 1	Case 2	Case 3	Case 4
	Trade≥0			
(1)	(2)	(3)	(4)	(5)
Logit(BVCM)-BVQCM	0.5895	0.5448	0.5614	0.5634
Std. Dev.	0.0076	0.0084	0.0351	0.0386
S.E.	0.0128	0.0127	0.0373	0.0392
Parm. Dev.	0.1790	0.0897	0.1229	0.1269
Logit(BV1FE)-BVQCM	0.5403	0.5075	0.5821	0.5834
Std. Dev.	0.0168	0.0205	0.0523	0.0512
S.E.	0.0184	0.0216	0.0491	0.0506
Parm. Dev.	0.0806	0.0150	0.1642	0.1669
Logit(3FE)-BVQCM	0.4844	0.4854	0.5557	0.5579
Std. Dev.	0.0167	0.0212	0.0525	0.0506
S.E.	0.0178	0.0211	0.0488	0.0508
Parm. Dev.	0.0312	0.0293	0.1114	0.1158
LPM(BVCM)-BVQCM	0.5789	0.5189	0.5549	0.5636
Std. Dev.	0.0079	0.0091	0.0322	0.0351
S.E.	0.0122	0.0120	0.0320	0.0336
Parm. Dev.	0.1578	0.0379	0.1098	0.1272
LPM(BV1FE)-BVQCM	0.5658	0.5060	0.5330	0.5409
Std. Dev.	0.0084	0.0098	0.0314	0.0357
S.E.	0.0131	0.0126	0.0325	0.0343
Parm. Dev.	0.1315	0.0119	0.0659	0.0818
LPM(3FE)-BVQCM	0.5716	0.5104	0.5475	0.5562
Std. Dev.	0.0077	0.0095	0.0323	0.0365
S.E.	0.0125	0.0122	0.0332	0.0349
Parm. Dev.	0.1433	0.0208	0.0950	0.1125

Standard errors are clustered at the country-pair. The simulations used 250 iterations, Stata 18.0, and R 4.3.2. Adjustment to dependent only occurred when the log of trade values are used. PPML estimators use trade values in levels without any adjustment to dependent variable. Parm. Dev. is percentage difference (in decimal form) from 0.5 (or the EIA effect).

Table OS9: Adding 1s to all observations

Panel A:	Q_{25} EIA			
	Case 1	Case 2	Case 3	Case 4
	Trade≥0			
(1)	(2)	(3)	(4)	(5)
Logit(BVCM)-BVQCM	0.4962	0.5083	0.5183	0.5111
Std. Dev.	0.0075	0.0089	0.0515	0.0556
S.E.	0.0128	0.0137	0.0468	0.0498
Parm. Dev.	0.0076	0.0165	0.0365	0.0221
Logit(BV1FE)-BVQCM	0.4660	0.5036	0.4914	0.4881
Std. Dev.	0.0199	0.0226	0.0612	0.0703
S.E.	0.0231	0.0266	0.0639	0.0665
Parm. Dev.	0.0681	0.0071	0.0172	0.0238
Logit(3FE)-BVQCM	0.4870	0.5270	0.4679	0.4724
Std. Dev.	0.0170	0.0217	0.0519	0.0564
S.E.	0.0199	0.0238	0.0511	0.0538
Parm. Dev.	0.0260	0.0540	0.0641	0.0551
LPM(BVCM)-BVQCM	0.3366	0.3715	0.5359	0.5441
Std. Dev.	0.0211	0.0194	0.0413	0.0433
S.E.	0.0243	0.0240	0.0367	0.0377
Parm. Dev.	0.3268	0.2569	0.0717	0.0881
LPM(BV1FE)-BVQCM	0.3195	0.3533	0.5095	0.5171
Std. Dev.	0.0213	0.0193	0.0421	0.0443
S.E.	0.0251	0.0247	0.0368	0.0379
Parm. Dev.	0.3609	0.2933	0.0190	0.0342
LPM(3FE)-BVQCM	0.3182	0.3525	0.5059	0.5138
Std. Dev.	0.0211	0.0190	0.0420	0.0436
S.E.	0.0249	0.0245	0.0370	0.0380
Parm. Dev.	0.3637	0.2949	0.0119	0.0275
Panel B:	Q_{75} EIA			
	Case 1	Case 2	Case 3	Case 4
	Trade≥0			
(1)	(2)	(3)	(4)	(5)
Logit(BVCM)-BVQCM	0.5962	0.5503	0.5533	0.5601
Std. Dev.	0.0071	0.0095	0.0352	0.0339
S.E.	0.0127	0.0124	0.0364	0.0378
Parm. Dev.	0.1924	0.1005	0.1066	0.1202
Logit(BV1FE)-BVQCM	0.5264	0.5007	0.5703	0.5811
Std. Dev.	0.0172	0.0186	0.0502	0.0475
S.E.	0.0174	0.0209	0.0478	0.0492
Parm. Dev.	0.0529	0.0014	0.1407	0.1622
Logit(3FE)-BVQCM	0.4753	0.4805	0.5470	0.5600
Std. Dev.	0.0171	0.0179	0.0493	0.0486
S.E.	0.0173	0.0206	0.0476	0.0489
Parm. Dev.	0.0493	0.0389	0.0940	0.1201
LPM(BVCM)-BVQCM	0.5756	0.5171	0.5455	0.5553
Std. Dev.	0.0080	0.0100	0.0331	0.0321
S.E.	0.0122	0.0118	0.0317	0.0329
Parm. Dev.	0.1511	0.0341	0.0910	0.1107
LPM(BV1FE)-BVQCM	0.5700	0.5089	0.5332	0.5426
Std. Dev.	0.0082	0.0105	0.0345	0.0312
S.E.	0.0130	0.0124	0.0322	0.0335
Parm. Dev.	0.1400	0.0178	0.0663	0.0853
LPM(3FE)-BVQCM	0.5735	0.5110	0.5457	0.5551
Std. Dev.	0.0077	0.0101	0.0339	0.0325
S.E.	0.0123	0.0119	0.0329	0.0342
Parm. Dev.	0.1471	0.0220	0.0915	0.1101

Standard errors are clustered at the country-pair. The simulations used 250 iterations, Stata 18.0, and R 4.3.2. Adjustment to dependent only occurred when the log of trade values are used. PPML estimators use trade values in levels without any adjustment to dependent variable. Parm. Dev. is percentage difference (in decimal form) from 0.5 (or the EIA effect).

Table OS10: Overdispersion Index Increased from 4 to 10, i.e. h = 10

Panel A:	Q_{25} EIA			
	Case 1	Case 2	Case 3	Case 4
	Trade≥0			
(1)	(2)	(3)	(4)	(5)
Logit(BVCM)-BVQCM	0.5048	0.5306	0.5348	0.5418
Std. Dev.	0.0073	0.0118	0.0636	0.0711
S.E.	0.0132	0.0166	0.0593	0.0615
Parm. Dev.	0.0097	0.0612	0.0697	0.0837
Logit(BV1FE)-BVQCM	0.4848	0.5387	0.5331	0.5489
Std. Dev.	0.0199	0.0270	0.0804	0.0938
S.E.	0.0236	0.0314	0.0775	0.0791
Parm. Dev.	0.0304	0.0774	0.0662	0.0978
Logit(3FE)-BVQCM	0.5009	0.5593	0.5086	0.5126
Std. Dev.	0.0174	0.0236	0.0690	0.0739
S.E.	0.0206	0.0282	0.0650	0.0675
Parm. Dev.	0.0018	0.1186	0.0172	0.0252
LPM(BVCM)-BVQCM	0.3478	0.4254	0.6323	0.6590
Std. Dev.	0.0188	0.0177	0.0474	0.0489
S.E.	0.0249	0.0236	0.0412	0.0417
Parm. Dev.	0.3044	0.1493	0.2645	0.3180
LPM(BV1FE)-BVQCM	0.3283	0.4061	0.6040	0.6298
Std. Dev.	0.0191	0.0184	0.0483	0.0488
S.E.	0.0257	0.0241	0.0413	0.0418
Parm. Dev.	0.3434	0.1878	0.2079	0.2597
LPM(3FE)-BVQCM	0.3282	0.4058	0.5990	0.6254
Std. Dev.	0.0190	0.0182	0.0478	0.0484
S.E.	0.0254	0.0240	0.0415	0.0420
Parm. Dev.	0.3436	0.1883	0.1979	0.2508
Panel B:	Q_{75} EIA			
	Case 1	Case 2	Case 3	Case 4
	Trade≥0			
(1)	(2)	(3)	(4)	(5)
Logit(BVCM)-BVQCM	0.5795	0.5241	0.5563	0.5668
Std. Dev.	0.0077	0.0101	0.0473	0.0461
S.E.	0.0125	0.0135	0.0444	0.0451
Parm. Dev.	0.1590	0.0482	0.1126	0.1336
Logit(BV1FE)-BVQCM	0.5326	0.4957	0.5978	0.6050
Std. Dev.	0.0172	0.0232	0.0607	0.0635
S.E.	0.0186	0.0248	0.0556	0.0567
Parm. Dev.	0.0651	0.0085	0.1957	0.2099
Logit(3FE)-BVQCM	0.4925	0.4817	0.5724	0.5804
Std. Dev.	0.0159	0.0246	0.0615	0.0611
S.E.	0.0182	0.0241	0.0561	0.0574
Parm. Dev.	0.0151	0.0365	0.1448	0.1607
LPM(BVCM)-BVQCM	0.5662	0.4975	0.5784	0.5892
Std. Dev.	0.0084	0.0105	0.0429	0.0410
S.E.	0.0122	0.0124	0.0380	0.0392
Parm. Dev.	0.1324	0.0050	0.1568	0.1783
LPM(BV1FE)-BVQCM	0.5537	0.4793	0.5581	0.5681
Std. Dev.	0.0087	0.0111	0.0425	0.0413
S.E.	0.0129	0.0129	0.0388	0.0400
Parm. Dev.	0.1075	0.0413	0.1161	0.1361
LPM(3FE)-BVQCM	0.5600	0.4848	0.5727	0.5834
Std. Dev.	0.0082	0.0109	0.0434	0.0412
S.E.	0.0123	0.0126	0.0396	0.0409
Parm. Dev.	0.1200	0.0303	0.1453	0.1668

Standard errors are clustered at the country-pair. The simulations used 250 iterations, Stata 18.0, and R 4.3.2. Adjustment to dependent only occurred when the log of trade values are used. PPML estimators use trade values in levels without any adjustment to dependent variable. Parm. Dev. is percentage difference (in decimal form) from 0.5 (or the EIA effect).

2 Monte Carlo Simulations Using the Santos Silva and Tenreyro (2011) and Breinlich et al. (2022) Approach

The simulation results presented in the previous section relied on a structural gravity methodology simulation approach described in Head and Mayer (2014) and adapted in Poissonnier (2019) for panel data and was implicitly a two-stage DGP. The key feature of this approach was the inclusion of fixed trade costs, f_{ijt}, in ϕ_{ijt} (overall trade costs). The simulations closely followed the theoretical gravity specification with export fixed costs developed by Melitz (2003), an extensive margin decision and an intensive margin decision. Given that the PPML estimator is not optimal for a two-stage modeling process, we will present an alternative one-step DGP here for non-negative outcomes.

2.1 Methodology

We now present simulation results as outlined in Santos Silva and Tenreyro (2011) and Breinlich et al. (2022). Similar to the simulations presented in Santos Silva and Tenreyro (2011), the dependent variable, y_i, has a significant proportion of zero outcomes while its expectation, conditional on its determinants, is expressed as:

$$E(y_i|x_i) = exp(x_i'\beta). \tag{21}$$

We note that this expectation can be represented as a finite mixture of two components: (1) m_i, which is a discrete random variable, and (2) a continuous random variable z_{ik}. In the context of international trade, for convenience let y_i be exports from country i (to some country j) and β a vector of coefficients on x_i. We let m_i denote the number of firms and let z_{ik} denote the exports of a particular firm, where k denotes a firm. Given the discussion, we can express the dependent variable as:

$$y_i = \sum_{k=1}^{m_i} z_{ik} \tag{22}$$

where m_i is a discrete non-negative integer (of possible export firms in country i).

If we assume that z_{ik} and m_i are independent, we can rewrite the expectation as:

$$E(y_i|x_i) = E(m_i|z_{ik})E(z_{ik}|x_i) \tag{23}$$

$$= exp(x_i'\gamma)exp(x_i'\delta) \tag{24}$$

so that $\beta = \gamma + \delta$. To simplify the simulation, both Santos Silva and Tenreyro (2011) and Breinlich et al. (2022) assume $\delta = 0$, which implies that:

$$E(y_i|x_i) = E(m_i|z_{ik}) = exp(x_i'\beta). \tag{25}$$

If z_{ik} is zero for all firms in country i, then the aggregate bilateral trade flow would be zero.

Following Santos Silva and Tenreyro (2011) and Breinlich et al. (2022), we specify that $E(m_i|z_{ik}) = exp(0.4 + \beta z_{ik})$ and $Var(m_i|z_{ik}) = aE(m_i|z_{ik}) + bE(m_i|z_{ik})^2$. Note that Santos Silva and Tenreyro (2011) used varying values of a and b that change the percentage of zeros in the simulated data. Following Breinlich et al. (2022), setting $a = 1$ and $b = 2$ implies that the simulated data will have roughly 50 percent zeros, which is similar to aggregate trade data used in the previous simulations and actual trade data. Both articles note that the Gamma PML is the optimal estimator given the simulation setup, which then should not favor the PPML *or* our three-stage QR approach. Similar to the previous simulation, the PPML estimator focuses on the conditional mean, while our proposed three-stage QR model targets the conditional median (though we will present results across various quantiles). This difference implies that the two estimators are not directly comparable, as they focus on different aspects of the conditional distribution. Therefore, any observed differences between them should be interpreted with caution, keeping in mind the distinct points of the distribution each method aims to estimate. We similarly exclude fixed effects, for simplicity. Additionally, we will report the results of the PPML conditional mean while we report $Q_{0.50}, Q_{0.60}, Q_{0.70}, Q_{0.80}, Q_{0.90}$ for our three-stage censored QR model.

We will consider two cases: (1) regressors and parameters are constant across k, and (2) parameters vary across k but each firm faces the same regressors. In Case 1, we consider homogeneous firms. In Case 2, we allow for heterogeneity across firms. In Case 2, we use σ_β to define variation in productivity across firms (i.e., heterogeneous firms or parameter heterogeneity):

- Case 1: Homogeneous firms
  - $\beta = -1$ and $z_i \sim \mathcal{N}(0, 1)$
- Case 2: Heterogeneous firms
  - $\beta \sim \mathcal{N}(-1, \sigma_\beta)$ and $z_i \sim \mathcal{N}(0, 1)$

The simulation will set $n \in \{100,000; 1,000,000\}$ and $k = 1$ (initially).[45] Additionally, we will allow the variation of the firm parameters such that $\sigma_\beta \in \{0.0, 0.25, 0.50, 0.75, 1.0\}$.

2.2 Results

The results are presented in two parts. The first part is the set of results in Table OS11. This table is composed of three panels. Panel A provides the results of estimating the model using PPML. We have two sets of samples: one uses a sample of $n = 100,000$ and the other uses a sample of $n = 1,000,000$. Across rows, we vary the degree of heterogeneity among firms (which is also representative of an index of heterogeneity across sectors). With homogeneous firms, PPML has virtually no parameter deviation. As we move down rows in Panel A, increasing heterogeneity among firms increases the parameter deviation of the

[45] The setup is analogous to Breinlich et al. (2022) and the choice of k will not be relevant given that our focus is on the aggregate y_i.

PPML conditional mean. The results are virtually identical across columns, that is, across the two sample sizes.

Panel B provides comparable results with $n = 100,000$ using Logit-BVQ. When there is no firm heterogeneity, Logit-BVQ has large parameter deviations at Q50, which is at the median of all flows, *including zeros* as described above. However, near the median of positive flows in the non-negative sample (Q80), Logit-BVQ has the least parameter deviations across rows. Moreover, as firm heterogeneity increases as we move down the rows of Panel B, the minimal parameter deviations at Q80 persists. Indeed, across every row for Logit-BVQ except for Q50, the parameter deviations stay *approximately the same* as heterogeneity increases. These results are confirmed with the larger sample in Panel C.

Table OS12 provides further results from the simulations analogous to those in Breinlich et al. (2022). The top panel presents the results allowing variation in the number of firms (k) in the sample. Clearly, the number of firms does not matter for the results. Moreover, the top panel shows clearly that – with firm homogeneity – PPML has the lowest parameter deviation.

However, the bottom panel of Table OS12 shows that – *with firm heterogeneity* – PPML, as found earlier, has considerable parameter deviations. Yet, Logit-BVQ has much smaller parameter deviations.

The bottom line is that which method has the least parameter deviations is a function of parameters of the model. Interestingly, under many scenarios, Logit-BVQ has smaller parameter deviations.

Table OS11: Simulation Results

Panel A: PPML	$n = 100,000$	$n = 1,000,000$
(1)	(2)	(3)
$\sigma_\beta = 0$	-1.000653	-0.9998776
	(0.0114556)	0.0034114
$\sigma_\beta = 0.25$	-1.066051	-1.066596
	(0.0205614)	0.0034114
$\sigma_\beta = 0.50$	-1.288403	-1.286394
	(0.0577706)	0.0173512
$\sigma_\beta = 0.75$	-1.542517	-1.544767
	(0.0599742)	0.0212977
$\sigma_\beta = 1.0$	-1.652002	-1.653248
	(0.0553482)	(0.0196304)

Panel B: Logit-BVQ	$n = 100,000$				
	$Q_{.50}$	$Q_{.60}$	$Q_{.70}$	$Q_{.80}$	$Q_{.90}$
(1)	(2)	(3)	(4)	(5)	(6)
$\sigma_\beta = 0$	-1.165343	-1.170695	-1.146849	-1.039147	-0.8966256
	(0.0817294)	(0.0323301)	(0.0166488)	(0.0071459)	(0.0059401)
$\sigma_\beta = 0.25$	-1.127125	-1.163318	-1.158612	-1.041043	-0.9031634
	(0.0845641)	(0.0349472)	(0.0189079)	(0.0074888)	(0.0062170)
$\sigma_\beta = 0.50$	-1.070306	-1.164445	-1.202391	-1.045299	-0.9171719
	(0.1055132)	(0.0400436)	(0.0185448)	(0.0083877)	(0.0070835)
$\sigma_\beta = 0.75$	-0.9884218	-1.178786	-1.24839	-1.045262	-0.9298841
	(0.1303428)	(0.0452036)	(0.0119897)	(0.0087541)	(0.0073268)
$\sigma_\beta = 1.0$	-0.7199572	-1.176696	-1.252054	-1.037991	-0.9407848
	(0.2033488)	(0.0494092)	(0.0116664)	(0.0098158)	(0.0081605)

Panel C: Logit-BVQ	$n = 1,000,000$				
	$Q_{.50}$	$Q_{.60}$	$Q_{.70}$	$Q_{.80}$	$Q_{.90}$
(1)	(2)	(3)	(4)	(5)	(6)
$\sigma_\beta = 0$	-1.159581	-1.168899	-1.146424	-1.038993	-0.8965574
	(0.0240261)	(0.0104824)	(0.0057150)	(0.0020882)	(0.0017831)
$\sigma_\beta = 0.25$	-1.127429	-1.162825	-1.159378	-1.041256	-0.9033191
	(0.0266184)	(0.0116302)	(0.0059946)	(0.0023218)	(0.0019798)
$\sigma_\beta = 0.50$	-1.074606	-1.165059	-1.201262	-1.045021	-0.9165881
	(0.0345299)	(0.0122133)	(0.0065913)	(0.0026530)	(0.0021705)
$\sigma_\beta = 0.75$	-0.9814089	-1.176771	-1.249682	-1.044930	-0.929834
	(0.0455976)	(0.0140130)	(0.0034241)	(0.0028412)	(0.0023617)
$\sigma_\beta = 1.0$	-0.7284294	-1.175973	-1.251161	-1.037126	-0.9396865
	(0.0607662)	(0.0155981)	(0.0037780)	(0.0031173)	(0.0027543)

The simulations used 500 repetitions in Stata 16.1. The table presents the mean and standard deviation (in parentheses) of results for the one-stage DGP simulation, as outlined in Santos Silva and Tenreyro (2011) and Breinlich et al. (2022). The variable σ_β represents the variation in productivity across firms. Panel A shows results for the PPML estimator, while Panels B and C display results for the Logit-QR at 100,000 and 1,000,000 observations, respectively.

Table OS12: Simulation Evidence that k is not relevant

	$\sigma_\beta = 0$		
(1)	(2)	(3)	(4)
	PPML	Logit-BVQ $Q_{.50}$	Logit-BVQ $Q_{.80}$
$k = 1$	-1.000653	-1.165343	-1.039147
	(0.0114556)	(0.0817294)	(0.0071459)
$k = 10$	-0.9999872	-1.167375)	-1.038725)
	(0.0105943)	(0.0792129)	(0.0069412)
$k = 50$	-0.9992448	-1.161576	-1.038804
	(0.0109702)	(0.0809577)	(0.0070298)

	$\sigma_\beta = 0.5$		
(1)	(2)	(3)	(4)
	PPML	Logit-BVQ $Q_{.50}$	Logit-BVQ $Q_{.80}$
$k = 1$	-1.288403	-1.070306	-1.045299
	(0.0577706)	(0.1055132)	(0.0083876)
$k = 10$	-1.285168	-1.072648	-1.044448
	(0.0550215)	(0.1077478)	(0.0083467)
$k = 50$	-1.287198	-1.07988	-1.04537
	(0.0590342)	(0.1003397)	(0.0080483)

Resutls for one-stage DGP as outlined in Santos Silva and Tenreyro (2011) and Breinlich et al. (2022). The simulations used 500 repetitions, Stata 16.1. The percentage of zeros of the dependent variable ranges .5029 to .5144 across repetitions. The variables k and σ_β represent the number of firms and the variation in productivity across firms, respectively. Mean values are reported with standard deviations in parentheses.

3 Quantile Treatment Effects

Firpo (2007) developed a method for analysis of the effects of binary regressors on quantiles of the unconditional distribution of the dependent variable. The method is a semiparametric two-stage approach where the first-stage logit model generates propensity scores and these propensity scores are used to re-weight the QR model suggested in Koenker and Bassett (1978). A key assumption is that the variable of interest is binary and exogenous, which causes complications with our current estimation strategy. First, as noted by Baier and Bergstrand (2004) it is quite difficult to predict EIAs between country-pairs due to the complexity of such agreements. Second, our estimation strategy requires the construction of bilateral regressors that have been demeaned (i.e., using the BV technique) such that the EIA variable is no longer binary. With these complications in mind, we estimated the model suggested by Firpo (2007) with one slight modification to our specification, namely, not demeaning EIA_{ijt}. We use a combination of the BV approximation approach and CREs in place of three-way fixed effects in our specification. Since the Firpo (2007) approach requires EIA_{ijt} to be binary, the second stage has re-weighted EIA_{ijt} values without the BV and CRE approximations in the second stage.

The first stage estimation of the Firpo (2007) methodology uses the (logit) EIA_{ijt} specification:

$$
\begin{aligned}
EIA_{ijt} = \beta_0 + \beta_1 \ln GDP_{it} + \beta_2 \ln GDP_{jt} + +\beta_3 DISTBV_{ij} + \beta_4 CONTIGBV_{ij} \\
+\beta_5 LANGBV_{ij} + \beta_6 LEGALBV_{ij} + \beta_7 RELIGBV_{ij} + \beta_8 COMCOLBV_{ij} \\
+\sum_{t=1}^{T} \alpha_t INTER_{ij} \times YEAR_t + \beta_9 \overline{\ln GDP}_i + \beta_{10} \overline{\ln GDP}_j + \eta_{ijt}.
\end{aligned} \tag{26}
$$

We know the quantile treatment effects (QTEs) for the unconditional outcomes are likely to differ from the benchmark conditional QR partial effect estimates. However, because we do not specify fully the determinants of EIA_{ijt} in this (first stage) logit specification (cf., Baier and Bergstrand (2004)), the second-stage QTE estimates may vary considerably from the benchmark conditional partial effect estimates in Table 5. Below in Table OS13, we report the QTEs using only positive flows (as in Table 5).

Table OS13: Quantile Treatment Effect: Positive Trade

(1)	(2) Q10	(3) Q20	(4) Q30	(5) Q40	(6) Q50	(7) Q60	(8) Q70	(9) Q80	(10) Q90
EIA	2.015***	1.922***	1.587***	1.371***	1.197***	0.968***	0.866***	0.738***	0.247***
	(0.234)	(0.161)	(0.163)	(0.146)	(0.120)	(0.108)	(0.107)	(0.083)	(0.084)
BV	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
INTER x Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CRE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	122999	122999	122999	122999	122999	122999	122999	122999	122999

The quantile treatment effect is estimated for a binary variable (EIA) using the methods described in Firpo (2007). The control variables include the BV terms, $DISTBV$, $CONTIGBV$, $LANGBV$, $LEGALBV$, $RELIGBV$, and $COMCOLBV$, GDPs and their means, and $INTER\,x\,Year\,FE$. The components of $EIABV$ are excluded and trim is set to (0.005 0.995).

4 Intra-national Trade

In principle, theoretical foundations for the gravity equation include a nation's trade "with itself." This is often referred to in the gravity-equation literature as "intra-national trade." Hence, in empirical work, several studies have addressed the robustness of results to including intra-national trade flows, which are effectively the output of a nation less its exports; alternatively, this can be termed the domestic expenditure of a country on its domestic output.

Table OS14 provides evidence that our results are robust to including intra-national trade. The specifications in columns (2) and (4) in *both* panels include only international trade flows. The specifications in columns (3) and (5) in both panels include international *and* intra-national trade flows. Intra-national trade is, on average, three times international trade. Because intra-national flows are very large relative to typical international trade flows, columns (3) and (5) in both panels include a dummy variable ($INTER \times YearFE$) that accounts for whether the flow is international or intra-national.[46]

Upon close examination, columns (2) and (3) in Panel A reveal that using OLS the additional inclusion of (relatively large) intra-national trade flows increases the EIA partial effect only a small amount, most likely due to skewing the trade distribution further to the right (and likely increasing the conditional mean EIA effect). This small increase in partial EIA effects is also found in Panel B using either BVQCM or Logit(BVCM)-BVQCM, for a likely similar reason. By contrast, the PPML estimate of the EIA partial effect is much more sensitive to the inclusion of intra-national trade, likely due to PPML's equal weighting of observations in levels.

[46]For brevity, we do not report the coefficient estimates for $(INTER \times Year)_{ijt}$.

Table OS14: Comparison of EIA Partial Effects With and Without Intra-national Trade

Panel A: OLS (Positive) and PPML (Non-negative)

(1)	(2) OLS	(3) OLS+INTRA	(4) PPML	(5) PPML+INTRA
EIA	0.383***	0.385***	0.109***	0.277***
	(0.034)	(0.034)	(0.029)	(0.050)
Exporter-Year FE	Yes	Yes	Yes	Yes
Importer-Year FE	Yes	Yes	Yes	Yes
Pair FE	Yes	Yes	Yes	Yes
INTER x Year FE	No	Yes	No	Yes
Adj. R2	0.857	0.862		
Pseudo R2			0.992	0.998
Obs	121417	122999	248123	249705

Panel B: BVQCM (Positive) and Logit(BVCM)-BVQCM (Non-negative)

(1)	(2) BVQCM	(3) BVQCM+INTRA	(4) Logit(BVCM)-BVQCM	(5) Logit(BVCM)-BVQCM+INTRA
EIABV	0.498***	0.537***	0.491***	0.542***
	(0.039)	(0.040)	(0.047)	(0.047)
BV	Yes	Yes	Yes	Yes
Year FE	Yes	No	Yes	No
CRE	Yes	Yes	Yes	Yes
INTER x Year FE	No	Yes	No	Yes
Pseudo R2	0.643	0.633	0.250	0.261
Obs	121417	122999	106959	108653

Clustered standard errors by country-pair are in parentheses * $p < .10$, ** $p < .05$, *** $p < .01$. OLS and PPML are average treatment effects. BVQCM is the median estimate of the positive trade flows. The prefix "Logit(BVCM)-" indicates that the three-stage estimation procedure described by Galvao et al. (2013) was implemented to account for zeros in quantile regressions; the first stage is a logit with BV methodology described in Baier and Bergstrand (2009a) and Baier and Bergstrand (2010) using all trade pairs (i.e. $T_{ij} \geq 0$). "Year FE" indicates whether year dummy variables were included or not in the second and third stages; "CRE" indicates whether correlated random effects were used or not in the second and third stages. The $INTER \times Year\ FE$ refers to the interaction of an indicator variable, 1 if an observed trade is across an international border (international) and 0 otherwise, and the year dummy variable. Following Machado et al. (2016), the squared correlation of $\ln X_{ijt}$ and the fitted values $Quant_q(\ln X_{ijt})$ are reported as the pseudo R^2 for panel B.