Hierarchical Bayesian Method for the Gravity Equations*

Meixin Guo†

Department of Economics, School of Economics and Management, Tsinghua University

First draft May 2010; revised August 2010

Abstract

Using panel data on country pairs, there are many ways in which fixed effects may be included in gravity models to control for countries' business cycle properties; estimates of key parameters can vary considerably, as noted in the literature. This paper tries to answer whether the gravity model estimations should include time-varying or time-invariant dummies, nation or separate importer and exporter dummies, and symmetric or asymmetric pair dummies. These model selection questions cannot be easily answered with conventional likelihood ratio tests and Wald tests because the high dimensionality of the parameter space due to many fixed effects and hundreds of constraints associated with these hypothesis tests lead to a large size distortion (the type I error). Here the hierarchical Bayesian method is applied to reduce the over-parameterization problem and to help choose models more credibly. This study estimates the gravity equations for the Euro Zone effect on trade for 22 developed countries during 1980-2004, a case used previously by Baldwin and Taglioni (2006). The Bayesian results show that the model with time-invariant importer and exporter dummies is preferred.

1 Introduction

Since the 1960s, the gravity equation has been used to explain trade flows between pairs of countries in terms of the countries' incomes, distances and other factors. There are four categories of variables involved in estimation: trade flows, economic mass variables (expenditures, output and income per capita), fixed effects, and paired trade costs.[1] Several issues

*The author is greatly indebted to Paul Bergin, Guido Kuersteiner, and Katheryn Russ for their valuable advice and constant encouragements, and also very grateful to Colin Cameron and Òscar Jordà for valuable suggestions. The author also appreciates the comments from macro Brown bag participants at UC Davis.

†Department of Economics, School of Economics and Management, Tsinghua University, Beijing, China, 100084. guomx@sem.tsinghua.edu.cn. Phone: 86-10-62795839.

[1]The trade cost factors have three main categories: a. geography variables, including distance, land-locked and island status, land area, country adjacency, latitude and longitude etc.; b. culture and history

are debated in this empirical gravity equation literature, including the choice of dependent variables and economic mass variables.[2]

A more severe issue in the gravity estimations is the choice of dummy variables, to include as fixed effects to control for business cycles and bilateral relations between countries. There is an increasing debate over how to choose dummies; conclusions on the key parameters vary with the inclusion of different groups of dummies. The essential questions are whether to include time-varying or time-invariant dummies, nation or separate importer and exporter dummies, and symmetric or asymmetric country pair dummies. The model with time-varying importer and exporter dummies has a theoretical foundation from Anderson and Van Wincoop (2003) and econometric support in Feenstra and Drive (2002). Evidence of heterogeneous preferences across countries provides support for the asymmetric pair dummies as in Guo (2010). These three papers taken together favor a model with both time-varying importer and exporter dummies, and asymmetric country pair dummies in panel regressions. However, researchers adopt different specifications of fixed effects in empirical studies.

For example, the two literatures on the WTO and currency union effects show the greatest disagreements on the choice of dummy variables. Rose (2004) considers time-invariant country fixed effects and finds little evidence that WTO/GATT promotes the total trade flow (sum of imports and exports) between a pair of countries. However, Subramanian and Wei (2007) take into account time-varying importer and exporter fixed effects, and find a strong positive and asymmetric effect of WTO on bilateral imports across countries and sectors, increasing world imports around 120% on average.

In the rich literature on currency union effect, Rose and Stanley (2005) evaluate a long list of estimates using meta regression analysis under two cases: random and fixed effect models. They find that the positive effect is 47% with a 95% confidence interval of 20%-80%, and claim that fixed effect estimates provide modest numbers, but suffer biases because of the heterogeneity. A more recent notable study on Euro Zone effect (EZ) is given by Baldwin and Taglioni (2006). They find EZ has no significant positive impact on trade controlling for both time-varying nation and symmetric country pair dummies, in contrast with 58% in Rose and Van Wincoop (2001) using country fixed effects and 5% - 20% in Micco et al. (2003) by pair and year dummies. Baldwin and Taglioni (2006) provide great summary and many comments on the gravity equation literature, but leave open the choice of dummy variables.

---

variables, including language, legal system, colony history, current colony status, independence date, development status etc.; c. special agreement dummies, such as common currency and common regional trade agreement. See Rose (2004) for details.

[2]Two main choices have been considered by researchers, trade levels and ratios, including unidirectional/total trade flow and imports over countries' outputs. See appendix A for details.

There is no paper in the literature proposing formal empirical criteria to select different groups of dummy variables. This paper tries to answer the question on choosing dummies to control for country fixed effects, the multilateral resistance terms in Anderson and Van Wincoop (2003). The study uses panel data with 22 countries during 1980-2004 (appendix A) and focuses on the Euro Zone (EZ) and European Union (EU) effects following Baldwin and Taglioni (2006) and Micco et al. (2003). The variations of these countries (EU vs non-EU, EZ vs non-EZ) enable discrimination between the EZ and EU effects on trade. I specify nine models with different choices of dummy variables, which are most commonly used in the literature, and do the hypothesis testing. The baseline model, one with time-varying importer and exporter fixed effects and asymmetric country pair fixed effects, is the most general specification and is implied by the theoretical models. The simplest model only controls for year and nation fixed effects.

The ordinary panel regressions show that the EZ increases the bilateral import ratio 18% on average, and the effect varies from a negative 0.3% (insignificant) to 49% (significant) (in table 5). The Wald test using Newey-West standard errors, the traditional likelihood ratio test (LR test) and F test based on spherical errors all support the baseline model, where EZ has no significant effect on imports. However, the LR test and Wald test suffer a really large size distortion (the type I error) due to the high dimensionality of the parameter space. The size distortions are also sensitive to the choice of estimates on spherical errors due to the hundreds of constraints associated with these hypothesis tests. In another words, dimention adjustment methods change the size distortion in different ways for the LR and Wald tests.

By comparing coefficients and testing hypotheses across models with different fixed effects, we find symmetric and asymmetric pair fixed effects have a similar influence in estimating EZ and EU effects. Results also show that individual importer/exporter fixed effects play a similar role to the nation fixed effects. But the debate on time-varying/time-invariant fixed effects cannot be resolved by the traditional method. Here the hierarchical Bayesian method is used to alleviate the small-sample size distortion, to reduce the number of (key) constraints across models, to estimate the distributions of the fixed effects, and to help select the model which fits data best. The Bayesian results support the model with time-invariant importer and exporter fixed effects, where the EZ increases 43% import ratios. With relatively limited data sets, the different conclusions on the model selection reveals a trade-off between the small sample problem and omitted variable issue in estimating the high dimenstional gravity equation.

The structure of the paper is as follows. Section 2 introduces the gravity equation and different specifications for the dummy variables. Section 3 provides results from least squares estimation (LS) and maximum likelihood estimation (MLE); and Bayesian results are presented in Section 4. The last section concludes.

2 Euro Effect or European Union Effect?

This section first presents the theoretical gravity model and nine popular empirical specifications for fixed effects (section 2.1) in the gravity equation literature. The subsequent subsections provide the results on EZ/EU effects using LS and MLE, and test hypotheses on different groups of dummy variables.

2.1 Gravity Model

The gravity equation in Anderson and Van Wincoop (2003) augmented with heterogeneous preferences in consumption baskets for each country can be shown to yield[3]

$$\begin{aligned} IM_t^{ik} &= \frac{EXP_t^i}{WOUT_t} * \frac{OUT_t^k}{WOUT_t} * WOUT_t * \alpha^i(k) \left(\frac{P_t^i \Pi_t^k}{\tau_t^{ik}}\right)^{\eta-1} \\ &= \frac{EXP_t^i * OUT_t^k}{WOUT_t} * \alpha^i(k) \left(\frac{P_t^i \Pi_t^k}{\tau_t^{ik}}\right)^{\eta-1}, \end{aligned} \tag{1}$$

where the multilateral resistance for good k is defined as

$$\left(\Pi_t^k\right)^{1-\eta} = \sum_{j=1}^{N} \left(\frac{P_t^j}{\tau_t^{jk}}\right)^{\eta-1} \alpha^j(k) \frac{EXP_t^j}{WOUT_t},$$

and the aggregate price level (the multilateral resistance for country i) is defined as

$$P_t^i = \left(\sum_{k=1}^{N} \alpha^i(k) \left[\tau_t^{ik} Q_t(k)\right]^{1-\eta}\right)^{\frac{1}{1-\eta}}.$$

The gravity equation specifies that the bilateral imports IM_t^{ik} are positively influenced by world output $WOUT_t$, importers' expenditure shares of world output "$EXP_t^i/WOUT_t$", exporters' output shares of world output "$OUT_t^k/WOUT_t$", and preference on good k "α^{ik}", but are impeded by trade costs "τ_t^{ik}".[4] The choice of dummy variables enters in the specification of two "multilateral (gravitational inconstant) trade resistance terms" (MLR), i.e. $MLR = (P_t^i \Pi_t^k)^{\eta-1}$.

[3]See section 2 of Guo (2010) for details.

[4]An alternative form of the gravity equation can be found in Head and Mayer (2004) to measure access to markets, and in Jacks et al. (2008) to measure trade costs.

The most general regression equation to estimate EZ and EU effects, the baseline model, is shown below after we take logs on equation (1)

$$w_t^{ik} = cons + EZ + EU + \delta^{ik} + \theta_t^i + \phi_t^k + \sum_{j=1}^{3} \gamma_{jt} g_j^{ik} + \varepsilon_t^{ik}. \tag{2}$$

The dependant variable bilateral import ratio is defined as $w_t^{ik} = log(\frac{IM_t^{ik} * WOUT_t}{EXP_t^i * OUT_t^k})$,[5] and variables δ^{ik} are asymmetric country pair dummies to capture the preferences α^{ik}. Both time-varying importer and exporter dummies are used to control for the MLR, $log(MLR) = \theta_t^i + \phi_t^k$.

The form of bilateral trade costs is assumed as follows

$$\left(\tau_t^{ik}\right)^{1-\eta} \equiv \prod_{j=1}^{J} \left(G_j^{ik}\right)^{\gamma_{jt}}. \tag{3}$$

Taking logs of both sides of the equation obtains the log of trade costs, $\sum_{j=1}^{J} \gamma_{jt} g_j^{ik}$, where variables $g_j^{ik} (\equiv log\left(G_j^{ik}\right))$, include logs of bilateral distances (log(dist.)) and six dummies for border (contig.), common official language (comlang.), land lock status for both exporter and importer countries (locked_im and locked_ex), Euro Zone membership (EZ), and European Union membership (EU). The variable "EZ" is equal to 1 if both countries belong to the Euro Zone; and the variable "EU" is equal to 1 if the two countries belong to the European Union.

Tables 1 and 2 list another eight variations of equation (2) based on different assumptions on MLR, which are all nested in the baseline model. The two tables respectively use two versions of dependent variables— log of bilateral import level: $lim_t^{ik} = log\left(IM_t^{ik}\right)$, and log of bilateral import ratio: $w_t^{ik} = log\left(\frac{IM_t^{ik} * GOUT_t}{EXP_t^i * OUT_t^k}\right)$. The eight models arise from combinations of four key assumptions regarding fixed effects. The first assumption is on country pair dummies, excluding asymmetric pair dummies ($\delta^{ik} = 0$) or imposing symmetric restrictions on pair dummies ($\delta^{ik} = \delta^{ki}$) in the regression. The second one assumes the constant multilateral resistance; for example, the time-invariant fixed effects take the forms $log(MLR) = \theta^i + \phi^k$ or $log(MLR) = \mu_t + \theta^i + \phi^k$. The third ignores the different roles for importer and exporter, and contains nation dummies only, such as $log(MLR) = \theta^i + \theta^k$. Lastly, if the multilateral resistance is country-pair specific, the model will have $log(MLR) = \mu_t + \widetilde{\zeta^{ik}}$ and so that estimated country pair dummies ζ^{ik} is the sum of two parts, i.e. $\zeta^{ik} = \delta^{ik} + \widetilde{\zeta^{ik}}$.

[5] See appendix A for another two versions of estimation equations commonly used in the literature.

2.2 Standard Panel Regressions on EZ and EU Effects

The annual data for 22 OECD countries during 1980-2004 are collected mainly following Baldwin and Taglioni (2006) (appendix A). There are 14 countries in EU by year 1995 and eight non-EU countries. In the EU group, four countries, Denmark, Greece, Sweden, and United Kingdom did not join in the EZ by year 2000. The EZ and EU effects on trade can be distinguished by the variation of these countries. All models are given in tables 1 and 2, and analyses focus on the table 2 using import ratios to avoid potential endogenous economic mass variables and non-stationary issue (see appendix A). The baseline model includes time-varying importer and exporter fixed effects, and asymmetric country pair dummies. Note that the variable "lyy" in table 1, the product of current GDPs of the importers and exporters, is collinear with pair dummies and time-varying nation dummy in table 1, and hence we could not use the estimated coefficient on this variable to judge which estimation equation is more supported by intuitions and theories as in Baldwin and Taglioni (2006).

Table 3, using just eleven years of panel data, replicates the key results given in table 4 of Baldwin and Taglioni (2006). The results from the baseline model show that both EZ and EU have no significant effect on bilateral imports. Tables 4 and 5 respectively use the whole 25-year data on import levels and ratios.[6] Like in table 3, EZ has no significant effect on trade in the baseline model of tables 4 and 5 while EU has significant and positive effect on imports, 25 % for import level in tableolslevel and 14% for import ratios in table 5. The column "YNAP" with time-varying nation and pair dummies also shows an insignificant EZ effect on trade, consistent to the result in table 5 of Baldwin and Taglioni (2006). Table 4 for import levels yields qualitatively similar conclusion to table 5. From here on, I focus on the results for bilateral import ratios to avoid possible endogenous problem.

In table 5, models with symmetric or asymmetric pair dummies and time-varying country fixed effects (columns of Base, YIMEXP and YNAP) shows that EZ has no significant effect on imports. However, estimations with time-varying country fixed effects (columns of YIMEX and YNA) show that using Euros as domestic currency for the trade parterners can increasethe import ratio by 27%, compared to non-EZ members.[7] Estimations with time-invariant country fixed effects (columns of IMEXYear and NAYear) show that both EZ and EU have a large effect in promoting import, 23% and 9% respectively. The average EZ

[6] The trade cost variables are redundant because of multi-collearity with the asymmetric country pair dummies; time-varying importer and exporter dummies also drop one for each country due to the same reason. In this paper with 22-country and 25-year panel data, the base line model drops total 72 dummies, including, 3 trade costs variables, 25 time-varying importer dummies for USA (θ_t^{USA}), 1 time-varying exporter dummy in 2004 for USA, and another 43 asymmetric country pair dummies.

[7] The number 27% is equal to $e^{(0.24)} - 1$ due to the logs. So do other percentages calculated based on the estimated coefficients.

effect increasing import ratios is 18% based on the 11 models and has a range from -0.3% to 49%. The EU effect on import is more stable compared to EZ effect, and varies from 9% to 37% (16% on average). These results illustrate that the EZ and EU effects vary prominantly with the choice of fixed effects though the standard errors for EZ and EU are similar across different specifications.

Compared with the above LS results, tables 6 and 7 provide results estimated by MLE for bilateral import levels and ratios, assuming i.i.d. normally distributed country pair random effects. MLE results provide robustness check on the variations of EZ and EU effects with the choice of fixed effects. On average, the MLE results show EZ effect in increasing import ratio is 11% in table 7, compared to 18% using LS in table 5. The extent of EZ effects depends on the choice of time-varying or time-invariant country fixed effects. Estimations with time-varying fixed effects in columns of Base, YIMEXP, YIMEX, YNAP, and YNA, do not provide evidence to support the EZ effect, but show significant effect from EU membership (15% more in import ratios). The YearPair model used in Micco et al. (2003) show a 12% increase in import ratio because of the EZ effect. In contrast, models with time-invariant country fixed effects in columns of IMEXYear, IMEX, and NAYear, shows both EZ and EU have significant effects on imports, around 21% and 25 % respectively. If we use simple difference-in-difference method (DID) within 14 EU member group with four non-EZ countries and take year 1999 as the breaking point, estimations with time-invariant country fixed effect show that using euros can increase 15% import ratios. However, after controlling the time-varying country fixed effect, the EZ effect drops to 5%.

These results from LS, MLE and DID imply three conclusions in estimating the EZ and EU effects: 1) the difference between symmetric and asymmetric country pair fixed effect does not matter so much for this sample; 2) isolating the role of importing or exporting country yields similar results as using nation fixed effects only; 3) furthermore, the choice on time-varying or constant fixed effects matters magnificently for the estimations on EZ and EU effects.

2.3 Hypothesis Tests

Based on estimations in table 5, four hypothesis tests on models with different fixed effects are given in table 8. The classical LR test results are in the LR1 column while the dimension adjusted LR test is in the LR2 column suggested by Italianer (1985).[8] The LR test and F

[8]Italianer (1985) finds that the LR test statistic is chi-squared distributed with the correction factor m/N, where N is the number of observation and m is equal to $(N - r - 0.5 * d_n)$ with the number of restrictions r and the dimension d_n of the restricted model (the null hypothesis). See append B.1 and table 10 for calculations on LR2.

test assumes i.i.d. error terms. The "Wald_NW" column takes the heteroskedasticity into account, provides the Wald statistics using Newey-West standard errors with 2 lags,[9] which are robust to HAC (Newey and West (1987, 1994)). All four tests reject the null hypothesis in the first eight combinations of the null and alternative models. That is, the baseline model cannot be rejected. If only considering choices on time-varying and time-invariant importer and exporter dummies, the null hypothesis— time-invarying importer and exporter fixed effect in combination (9) is supported by all four tests.

These tests suffer from a size distortion due to many fixed effects, pursued in section 3. Before considering this weakness, we use an alternative model selection procedure: Bayesian Information Criterion (BIC)

$$BIC = -2 * LL\,(y \mid \Theta) + d * log\,(N)\,, \tag{4}$$

where $LL\,(y \mid \Theta)$ is the log-likelihood, d is the number of the estimated coefficients for a model, and N is the number of observations. This criterion puts a large penalty on over-parameterization and is known to favor smaller models compared to hypothesis tests. BIC statistics for the eleven models,[10] are listed in table 9, based on the LS and MLE results in tables 5, 4, 7, and 6. In the first two LS columns, the baseline model is supported by BIC. However, in the next two MLE columns, the model with time-invariant nation and year dummies is preferred by BIC. These two different results on model selection using BIC come from different specifications of the error terms. The LS assumes spherical errors only, whereas the MLE additionally considers the i.i.d. country-pair random effects.

3 Size Distortion

The preceding regressions include a large number of dummy variables, especially for the baseline model; and the hypothesis tests have hundreds of constraints (the degree of freedom (DF) for the test, the column of "DF" in table 8). Though the hypothesis tests all support the baseline model, statistical inference may suffer from a small sample problem due to the unusually high dimensionality of the parameter space and the impressively large group of constraints associated with the hypothesis tests. As we know, the statistic from the LR test is asymptotically chi-square distributed with the number of different parameters in the null and alternative models (DF), when sample size becomes larger. However, with a large number of parameters and limited data, Evans and Savin (1982) and Italianer (1985) show that the

[9]The conclusion remains up to 5-lags.

[10]the eleven models include a simple OLS, model with year dummies only and the nine models in tables 1 and 2.

finite sample distribution of the statistic is biased towards the conventional large sample asymptotic chi-square distribution. Therefore, the critical values for the 5% significant level should be adjusted. But the LR test uses the reported critical values without corrections on the high dimensionality.

Monte Carlo simulations with i.i.d errors show a large size distortion for the classical LR test (the LR1 column with spherical errors) in table 10,[11] more than 50% on average (the type I error, the rate of rejecting the null when the null is true). The Wald test (the Wald1 column) has a even larger size distortion, 77% on average, using the same consistent but biased estimated variance as in LR1 column. In contrast, the second Wald test (the Wald2 column) with consistent and unbiased estimated variance and F tests have normal size. Particularly, the size for the Wald2 test is around 6.7% with relatively larger number of the constraints (those null models have smaller dimensions compared to the alternative models). In other words, with a larger group of constraints, the tests have relative greater size distortions. We reject the models with time-invariant dummies slightly more often (1.7 % more than 5 %) using Wald2 test, and too often for the LR1 and Wald1. The conclusion is robust to different dependent variables (level or ratio).

Supplementary detailed information for rejection of the null model is provided in figure 1, which lists the power curves for four combinations of null and alternative models (see appendix B.1). For example, the first top graph is for the combination of null model "YNAP"— symmetric country pair and time-varying nation fixed effects, and the alternative model "Base" — asymmetric country pair and time-varying importer and exporter fixed effect. The YNAP model is recommended by Baldwin and Taglioni (2006), which can be obtained by imposing symmetric conditions on the baseline model. Comparison between the YNAP and the baseline model is to test the null hypotheses $H0$ are $\delta^{ik} = \delta^{ki}$ and $\phi_t^k = \theta_t^k$. By defining $\widetilde{\delta^{ik}}$ $(= \delta^{ik} - \zeta^{ik})$ with $\zeta^{ik} = \zeta^{ki}$, and $\kappa_t^k (= \phi_t^k - \theta_t^k)$, the null hypotheses $H0$ are equivalent to test $\widetilde{\delta^{ik}} = 0$ and $\kappa_t^k = 0$. Accordingly, the null model is nested in the baseline model. The more volatile $\widetilde{\delta^{ik}}$ and κ_t^k, the more frequently the null model will be rejected. Provided that group of parameters $\widetilde{\delta^{ik}}$ and κ_t^k follow normal distributions with a zero mean and a common standard deviation σ_Δ, increasing the σ_Δ from zero to 0.1, such as 0.01 each time, leads to a higher chance of rejection on the null model. As a result, the power curve of the test is determined by the levels of the σ_Δ. All four tests in figure 1 achieve almost 100% rejection on the null models when $\sigma_\Delta = 0.04$.

The size distortions shown in table 10, are also sensitive to the choice of estimated variances (LR1 vs LR2, and Wald1 vs Wald2). In another words, adjusting dimensions on the statistics reduces the size distortions magnificently. We take the combination (4), IMEX

[11] See appendix B.1 for simulation details for columns of LR1, LR2, Wald1, Wald2 and F.

vs Base, as an example to illustrate this fact using three figures. Figures 2, 3 and 4 provide the chi-square densities with degree of freedom (DF) 1448 for three cases (appendix B.1): 1) figures 2 is the ideal case, drawing 1 million observations from the Chi-square distribution directly; 2) figures 3 plot the LR1 and LR2 statistics using 1000 simulations; 3) figures 4 plot the Wald1 and Wald2 statistics using 1000 simulations. The solid lines use original values and dash lines are adjusted by dimensions. The shaded areas in all three figures are 5% rejection regions based on the observed densities and the theoretical critical (CV) for chi-square distribution with DF 1448 is 1537.639.

With DF 1448, the Chi-squared distribution for LR and Wald tests (figures 3 and 4) are biased compared to the ideal distribution figure 2, where the means are 1555 and 1665 respectively, much larger than true mean 1448. A small adjustment on the statistics (the dash densities) reduces the size distortion prominently based on the theoretical CV 1537.639. For example, in figures 2, a value, such as 1539, in the 5% shaded area under the solid line is changed into 1342.86 with an adjustment ($0.873 = \frac{11550-1448-0.5*48}{11550}$); this value is no longer significant compared with the CV 1537.639. In figure 3, LR1 (the solid line) has a 62.1 % rejection rate on the null model "IMEX" (the size, probability larger than the CV 1537.7). However, the size becomes zero using LR2 (10). Similar phenomenon occurs in figure 4 for Wald tests comparing Wald1 with Wald2. The size is 97.6% in column Wald1, but 6.9% in column Wald2 after the Wald1 statistics are adjusted by $0.87(=\frac{11550-1496}{11550})$. The three graphs support the first two conclusions drawn from table 10: the high demensions in the models and large number of constraints associated with the hypotheis tests lead to a biased symptotical chi square distribution/large size distortion; futhermore, size distortions could be sensitive to dimenstion adjustments, the simple adjustment proposed by Italianer (1985) does not work well for the LR test.

The third conclusion we can draw from table 10 is that misspecified errors evoke close to 100% size distortion except the two cases: combinations (9) and (10). Considering heteroskedasticity and autocorrelation, we consider three HAC forms in the paper (appendix B.2) and have qualitatively similar results. We mainly focus on "HAC1", taking the form as below,

$$\epsilon_t^{ik} = g^{ik} + \nu_t^{ik} \qquad \nu_t^{ik} = b_\nu \nu_{t-1}^{ik} + \mu_t^{ik} \qquad var(g^{ik}) = \sigma^2_{g^{ik}} \qquad var(\mu_t^{ik}) = \sigma^2_\mu$$

There is no contemporaneous correlation across county pairs.[12] This parametric assumption considers the heterogeneous fixed effect in the variance covariance matrix Ξ $(= var(\epsilon))$, which is a block diagonal matrix with Ω^{ik} (462 pairs) for one specific country pair (importer

[12]Models with contemporaneous correlation across county pairs can be estimated by spacial regression.

i and exporter k) and Ω^{ik} has the following form,

$$\Omega^{ik} = \sigma^2_{g^{ik}} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \vdots & \vdots \\ \vdots & \vdots & 1 & 1 \\ 1 & \cdots & 1 & 1 \end{bmatrix} + \frac{\sigma^2_{\mu}}{1-b^2_{\nu}} \begin{bmatrix} 1 & b_{\nu} & \cdots & b_{\nu}^{T-1} \\ b_{\nu} & 1 & \vdots & \vdots \\ \vdots & \vdots & 1 & b_{\nu} \\ b_{\nu}^{T-1} & \cdots & b_{\nu} & 1 \end{bmatrix}.$$

In simulations, the true errors have HAC1; and we consider two cases: considering HAC (the column HAC1) and no controlling for HAC1 (assuming spherical errors) noted as HAC1(M). The interesting point is that the combinations (9) and (10) have zero rejection for four tests, in contrast with almost 100% size distortion for other combinations. This result seems odd at first glance. However, table 11 provides evidence on the large bias in estimating the coefficients because of the misspecification on the errors. The country-pair specific variance structure remarkably influences the estimation on the coefficients, particularly for the country fixed effects. Without controlling for true HAC, both null and alternative models cannot estimate the coefficients consistently.

Table 11 and the following three tables present the size of the Wald test on the artificial coefficients and Wald statistics[13] for the estimates over 1000 simulations (see appendix B.3). The null hypothesis in table 11 is $H0 : \hat{B}_m = B_0$ for $m \in \{a, n\}$, and the Wald statistic is calculated by following formula

$$Wald = (\hat{B}_m - B_0) * \left[\widehat{var(\hat{B}_m)}\right]^{-1} * (\hat{B}_m - B_0). \tag{5}$$

where $\widehat{var(\hat{B}_m)} = \hat{\sigma^2_m} * (X_m' X_m)^{-1}$, $\hat{\sigma^2_m} = RSS_m/(N - K_m)$, and $m = a/n$ refers to the alternative/null model. Then the rejection rates (size) for the test is the frequency of rejection over 1000 simulations. Beside the previous over-rejection fact in the HAC1(M) columns, table 11 also shows that when the null model has a larger dimension (for example, YIMEXP and YNAP), the size is relative larger, more than 5%. In particular, table 12 focuses on the EZ and EU effects and gives a similar conclusion on size disortion due to the misspecified error structure.

Tables 13 and 14 (for EZ and EU effects only) calculate the Wald statistics for $H0 : \bar{\hat{B}}_m = B_0$ for $m \in \{a, c\}$

$$Wald = (\bar{\hat{B}}_m - B_0) * \left[\widehat{var(\bar{\hat{B}}_m)}\right]^{-1} * (\bar{\hat{B}}_m - B_0), \tag{6}$$

[13] From here on, we use Wald2 as the Wald test with a normal size.

where the variance covariance matrix is

$$var(\bar{\hat{B}}_m) = \bar{\hat{\sigma}}^2_m * (X'_m X_m)^{-1}/1000,$$

and $\bar{\hat{\sigma}}^2_m = \overline{RSS}_m/(N - K_m)$ based on the average sum of squared residuals. The Wald statistic follows chi-squared distribution given degrees of freedom/number of constraints. Both tables show a more rejection in HAC1(M) columns than HOMO and HAC1 columns since the regressions in HAC1(M) do not consider the HAC. Particularly, in table 14, Wald statistics for the null H0 models (YIMEX and YNA in HOMO column) with yearly importer and exporter or nation fixed effects, reject the true (artificial) EZ and EU coefficients, indicating that high dimensionality with a small sample leads to biased estimates on average. A weird fact is that the alternative models (the H1 column) perform better than the null model (the H0 column), and they can estimate consistent EZ and EU effects on average for both HOMO and HAC1 cases, except the two combinations (9) and (10). In these two combinations (9) and (10), over-parameterization in the alternative model with time-varying fixed effects (H1: YIMEX) leads to bias on EZ and EU effects in contrast with the null models with constant fixed effects (H0: IMEX or IMEXY).

Exercises on the size distortion fairly show models with either asymmetric or symmetric pair fixed effect (Base and YIMEXP) have similar performance in estimating the EZ and EU effects, and separating importer fixed effect from the exporter fixed effect or using nation fixed effect only (IMEXYear vs NAYear; and YNA vs YIMEX) leads to similar results too. However, the issue on time-varying or time-invariant fixed effects (YIMEXP/YIMEX vs IMEX, YNA vs NAYear) has not been resolved. We need other method to provide the clues on the problem.

4 Hierarchical Bayesian Method

Previous estimation results and simulations provide relatively consistent support on symmetric pair and nation fixed effects for this specific data sample. But there is no clear-cut evidence against or for time-varying fixed effects. The Bayesian method can estimate the distributions of all parameters, including the distributions of the variance across years, which provides direct visual evidence on the volatility of these fixed effect. The hierarchical structure has the advantage to reduce the dimension of the key parameter space and number of constraints to test different models, so that alleviate the large dimensionality problem in LS and MLE. Section 4.1 provides the hierarchical Bayesian model and estimation methodology; and the next section shows the Bayesian results on model selection.

4.1 Estimation Models and Priors

The general two-level hierarchical linear Bayesian model with normal prior distributions for regression coefficients and Gamma/Wishart prior distributions for variance coefficients is given as follows (Gamerman and Lopes (2006), Gelman et al. (2004), and Koop et al. (2007))

$$Y = X * \beta_1 + \varepsilon_t^{ik} \tag{7}$$

$$\begin{aligned} \text{where } \overline{Y}| \quad & \beta_1, \Sigma_1 && \sim N(X * \beta_1, \Sigma_1) \\ \beta_1| \quad & \beta_2, \Sigma_2 && \sim N(W * \beta_2, \Sigma_2) \\ \beta_2| \quad & \beta_3, \Sigma_3 && \sim N(Z * \beta_3, \Sigma_3) \\ \Sigma_{1,2,3}| \quad & \beta_{1,2,3} && \sim Gamma/Wishart, \end{aligned}$$

where the β_2 is the key parameter vector. This Normal-Wishart prior distribution gives analytical posterior (conditional) distributions for all parameters and has an advantage for estimation. In this paper, the hierarchical structure of the baseline model can be specified as follows,

$$\begin{aligned} w_t^{ik} &= cons + EZ + EU + \delta^{ik} + \theta_t^i + \phi_t^k + \varepsilon_t^{ik} \qquad (8) \\ w_t^{ik}|\Theta &\sim N\left(\overline{w_t^{ik}}, \left(\sigma^{ik}\right)^2\right) \\ \text{where } \overline{w_t^{ik}} &= cons + EZ + EU + \delta^{ik} + \theta_t^i + \phi_t^k \\ \frac{1}{\left(\sigma^{ik}\right)^2} &\sim G(a_1, a_2), \end{aligned}$$

where the error term has a simple heteroskedastic form without serial correlations. The choice on this type of error term mainly relies on the fact that heterogeneity of cross-sectional country pair significantly affects the estimation of the EZ and EU effects based on previous sections 3. Furthermore, with a heteroskedastic and serially correlated variance and covariance structure, the time-varying component of the data would be absorbed in the error ε_t^{ik} instead of the country fixed effects θ_t^i and ϕ_t^k; consequently, the model with time-invariant fixed effect is more easily accepted and preferred.

The specific values and prior distributions for the parameters are listed below:[14]

$(cons, EZ, EU)' \sim N(0, a_3 * I(3))$;

$\delta^{ik} \sim N\left(p, \sigma_p^2\right)$ with mean $p \sim N(0, a_3)$ and variance $\frac{1}{\sigma_p^2} \sim G(a_1, a_4)$;

[14] This hierarchical linear model is a special case of mixed linear models in Colin and Trivedi (2005) with randomly varying intercepts (p774).

$$\theta_t^i \sim N\left(\theta^i, (\sigma_{\theta t})^2\right) \text{ with mean } \theta^i \sim N(0, a_3) \text{ and variance } \frac{1}{(\sigma_{\theta t})^2} \sim G(a_1, a_4);$$

$$\phi_t^k \sim N\left(\phi^k, (\sigma_{\phi t})^2\right) \text{ with mean } \phi^k \sim N(0, a_3) \text{ and variance } \frac{1}{(\sigma_{\phi t})^2} \sim G(a_1, a_4);$$

where $a1 = 3$, $a2 = 0.5$, $a3 = 10$ and $a4 = 1$, close to the values picked in Ranjan and Tobias (2007), who uses hierarchical Bayesian method to estimate a non-parametric threshold Tobit gravity model because of too many zeros in the trade data.[15] The time-invariant country pair fixed effects δ^{ik} are assumed to have a common mean and variance to capture the common features of these 22 OECD countries. The time-varying importer and exporter fixed effect, θ_t^i and ϕ_t^i are expected to have heterogeneous means across countries and variances across years. They are used to control for the business cycle properties and shocks for each country and each year, denoted as the multilateral resistant terms.

Depending on the specifications of the dummies, the model and priors vary. For example, the IMEX model with time-invariant fixed effects assumes $\widetilde{\theta}^i \sim N(0, a_3)$ and $\widetilde{\phi}^i \sim N(0, a_3)$ only; and the two variance terms $\sigma_{\theta t}$ and $\sigma_{\phi t}$ are zeros. Volatile $\sigma_{\theta t}$ and $\sigma_{\phi t}$ across years in the models of Base, YIMEX, YNAP, and YNA, provide supportive evidence for the time-varying country fixed effects.

The estimation algorithm is the Markov Chain Monte Carlo Simulations via Gibbs Sampler, which has the following five steps:

Step 1. Give initial values for the variances, $\Sigma_{1,2,3}$ ($(\sigma^{ik})^2$, σ_p^2, $\sigma_{\theta t}^2$, $\sigma_{\phi t}^2$, and $a3$) and the second level parameters β_2 (p, θ^i, ϕ^k);

Step 2. Draw values from the posterior distributions for the first level parameters β_1 ($cons, EZ, EU$, θ_t^i, ϕ_t^k, and δ^{ik}), given $\Sigma_{1,2,3}$ and β_2;

Step 3. Draw values from the posterior distributions for the variances $\Sigma_{1,2}$ ($(\sigma^{ik})^2$, σ_p^2, $\sigma_{\theta t}^2$, and $\sigma_{\phi t}^2$), given Σ_3, β_1 and β_2.

Step 4. Draw values from the posterior distributions for the parameters β_2, given β_1 and $\Sigma_{1,2,3}$.

Step 5. Repeat from the second step until the chain converges.

Two convergence tests are used here to determine the burning and draw times: 1) Gelman-Rubin statistic (BGR) with $|R - 1| < 0.05$ for single parameter and multiple pa-

[15]The import data in this paper has rare zeros. Also diffuse priors, such as $a1 = a3 = 1000$ and $a2 = a4 = 0.001$, select the model "IMEX" the best specification, same as in section 4.2.

rameters (Gelman (2006));[16] 2) Geweke chi-squared test (Geweke (1992)).[17] Most of the parameters in the small models like IMEXYear, IMEX, NAYear, and YearPair converge after 5000 burning times based on the two tests. However, the large models have more than one thousand parameters and converge slowly. After 100,000 burning times, parameters in all models have converged based on BGR statistics for 50000 draws (thin 10) from either single chain or multiple chains. Coefficients in small models and the second level (key) parameters (β_2) in large models converged based on Geweke Chi-squared statistics.

4.2 Bayesian Results

Table 15 shows the estimated coefficients and standard deviations on variables EZ and EU for nine models using Bayesian method, which are close to those in table 5. The average EZ effect on promoting import ratios is 15%. The EZ effects are significantly positive in six models: YIMEX, IMEXYear, IMEX, NAYear, YNA, and YearPair, but the magnitude is a little smaller than that using LS in table 5. In another three models: Base, YIMEXP, and YNAP, EZ has no significant effect on trade. In comtrast, the EU membership is consistently estimated to have a significantly positive effect on import, 17% more in import ratios on average. The variation of EZ and EU effects due to the different choices on dummy variables remain as LS and MLE estimations.

Since Bayes factor/odds ratio is sensitive to the dimensionality of the parameter space,[18] three information criteria: DIC, BIC, and AIC, are used to select models. The three criteria are developed based on likelihood with a penalty on the number of dimensions.[19] Spiegelhalter et al. (2002), Congdon (2005), Gelman et al. (2004), and Iliopoulos et al. (2007) recommend Deviance Information Criterion (DIC) as a main criterion,[20] which is calculated using log-likelihood,

$$\begin{aligned} DIC &= \overline{\widehat{D\left(y,\Theta\right)}} + p_D \\ \text{where } p_D &= \overline{\widehat{D\left(y,\Theta\right)}} - D\left(y, \overline{\widehat{\Theta}}\right) \\ \text{with } \overline{\widehat{D\left(y,\Theta\right)}} &= -2 * \overline{\widehat{LL\left(y \mid \Theta\right)}} \text{ and } D\left(y, \overline{\widehat{\Theta}}\right) = -2 * LL\left(y \mid \overline{\widehat{\Theta}}\right). \end{aligned}$$

The $LL\left(y \mid \Theta\right)$ is the log-likelihood and $D\left(y,\Theta\right)$ is the deviance, a measure of how well the

[16]Matlab code reference: the GNU General Public License.

[17]Matlab code reference: James P. LeSage, Dept of Economics Texas State University-San Marcos, jlesage@spatial-econometrics.com.

[18]See Gelman et al. (2004) for an extensive discussion, page 185-186.

[19]The BIC is an asymptoticly consistent approximation to Bayes factors between models (Berger et al. (2003)).

[20]Iliopoulos et al. (2007) estimate two-way contingency table with Bayesian method using MCMC.

model fits the data. The symbol "— " refers to mean of the variables. The posterior mean deviance $\widehat{D(y,\Theta)}$ is equal to -2 times the mean of posterior log-likelihood $\widehat{LL(y \mid \Theta)}$, and the deviance $D\left(y,\overline{\widehat{\Theta}}\right)$ uses the log-likelihood calculated by the mean of posterior parameters $\overline{\widehat{\Theta}}$. The p_D, represents the penalty with a larger number of parameters. The smaller number of DIC implies a better fit of the model. Another two commonly used criteria are Akaike Information Criterion(AIC) and Bayesian Information Criterion (BIC) (Congdon (2005) and Iliopoulos et al. (2007)) shown as below,

$$AIC = D\left(y,\overline{\widehat{\Theta}}\right) + 2 * d$$
$$BIC = D\left(y,\overline{\widehat{\Theta}}\right) + d * log\left(N\right)$$

where d is the (effective) number of estimated parameters and N is the number of observations. Smaller values in AIC and BIC indicate better fitness of the model. In table 15, all three criteria give the lowest values for the model "IMEX" with time-invariant importer and exporter fixed effects, and hence favor this specification to other eight models. If we calculate the traditional LR statistics using the posterior log-likelihood $LL\left(y \mid \overline{\widehat{\Theta}}\right)$, the model IMEX cannot be rejected by any other model who has a higher posterior log-likelihood (Base, YIMEX, and YNA).

The supportive evidence for the model IMEX can also be found in figure 5 with fairly constant posterior means for the variances of time-varying fixed effects. Figure 5 shows 95% posterior credible intervals for the variances $\sigma^2_{\theta t}$ and $\sigma^2_{\phi t}$ in three models with time-varying fixed effects. The top two graphs are for the variances of time-varying importer and exporter fixed effects $\sigma^2_{\theta t}$ and $\sigma^2_{\phi t}$ respectively in the baseline model;[21] and the bottom two are for the variances of time-varying nation fixed effects in models YNA and YNAP. The non-volatile variances give another evidence to favor the model IMEX with time-invariant dummies.

5 Conclusion

This paper studies how the choice of dummy variables affects the magnitude of the Euro Zone effect on increasing bilateral import when we estimate gravity models. Three groups of dummies, which are included in the gravity equations to control for individual country and country-pair fixed effects, are compared, asymmetric or symmetric country pair dummies, time-varying or time-invariant country dummies, separating importer/exporter or nation dummies. Depending on the choice of dummies, EZ and EU effects on trade during

[21]The model YIMEX has similar posterior distributions of $\sigma^2_{\theta t}$ and $\sigma^2_{\phi t}$.

the period 1980–2004 vary greatly using the LS and MLE methods, from -0.3% to 49%. Based on LS, conventional Wald test, LR test and F test are used to assess the necessity of the different dummies; but these tests have a large size distortion/biased Chi-square distributions. Two factors contribute this large size distortion. First the high dimensionality of the parameter space leads to a biased asympotical chi-square distribution for LR test and Wald tests. Second, hundreds of constraints associated with the hypothesis tests drive the size distortion sensitive to the dimension adjustment method applied to the test statistics.

Among the issue on selecting three groups of dummies, it is the most difficult to make a choice on time-varying or constant country dummies. Results from LS regressions and Monte Carlo simulations on size distortions for the tests more or less show that symmetric pair dummies influence the estimated coefficients similarly to the asymmetric pair effects, and that separating the role of importer and exporter in the estimations does not significantly change the coefficients estimated from the model with symmetric nation dummies only. However, there is no deterministic evidence against or for the time-varying country dummies. Hence, hierarchical Bayesian method is adopted to provide the distributions of all parameters and give a close investigation on the variances of the time-varying country fixed effects. The non-volatile posterior distributions of the variance parameters in the models with time-varying country dummies provide evidence on the constant country effects. Three information criteria based on Bayesian results also show that the model "IMEX" with constant importer and exporter effects is favored among all nine models.

References

J.E. Anderson and E. Van Wincoop. Gravity with gravitas: a solution to the border puzzle. *American Economic Review*, 93(1):170–192, 2003.

A. Aviat and N. Coeurdacier. The geography of trade in goods and asset holdings. *Journal of International Economics*, 71(1):22–51, 2007.

R. Baldwin and D. Taglioni. Gravity for dummies and dummies for gravity equations. *C.E.P.R. Discussion Papers*, Sep 2006.

J.O. Berger, J.K. Ghosh, and N. Mukhopadhyay. Approximations and consistency of Bayes factors as model dimension grows. *Journal of Statistical Planning and Inference*, 112(1-2):241–258, 2003.

C.A. Colin and P.K. Trivedi. Microeconometrics: Methods and applications, 2005.

P. Congdon. *Bayesian Models for Categorical Data*. Chichester, United Kingdom John Wiley and Sons Ltd, 2005.

G.B.A Evans and N.E. Savin. Conflict among the criteria revisited: The w, lr and lm tests. *Econometrica*, 50(3):737–48, May 1982.

R.C. Feenstra and O.S. Drive. Border effects and the gravity equation: consistent methods for estimation. *Scottish Journal of Political Economy*, 49:491–506, 2002.

A. Gelman. Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*, 1(3): 515–533, 2006.

A. Gelman, J. Carlin, H.S. Stern, and D.B. Rubin. *Bayesian Data Analysis*. New York, NY: Chapman and Hall, 2004.

J. Geweke. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. *Bayesian statistics*, 4(2):169–193, 1992.

M. Guo. Frictions, heterogeneous preferences and international consumption risk sharing. *University of California, Davis. Manuscript*, August 2010.

K. Head and T. Mayer. The Empirics of Agglomeration and Trade, in Handbook of Regional and Urban Economics Vol.. *Henderson, V. and JF Thisse*, 2004.

G. Iliopoulos, M. Kateri, and I. Ntzoufras. Bayesian estimation of unrestricted and order-restricted association models for a two-way contingency table. *Computational Statistics and Data Analysis*, 51(9): 4643–4655, 2007.

A. Italianer. A small-sample correction for the likelihood ratio test. *Economics Letters*, 19(4):315–317, 1985.

D.S. Jacks, C.M. Meissner, and D. Novy. Trade Costs, 1870-2000. *American Economic Review*, 98(2): 529–534, 2008.

A. Micco, E. Stein, G. Ordoñez, K.H. Midelfart, and J.M. Viaene. The currency union effect on trade: early evidence from EMU. *Economic Policy*, 18(37):317–356, 2003.

W.K. Newey and K.D. West. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–08, May 1987.

W.K. Newey and K.D. West. Automatic lag selection in covariance matrix estimation. *Review of Economic Studies*, 61(4):631–53, October 1994.

P. Ranjan and L.J. Tobias. Bayesian inference for the gravity model. *Journal of Applied Econometrics*, 22 (4):817–838, 2007.

A.K. Rose. Do we really know that the wto increases trade? *American Economic Review*, 94(1):98–114, 2004.

A.K. Rose and T.D. Stanley. A meta-analysis of the effect of common currencies on international trade. *Meta-regression analysis: issues of publication bias in economics*, page 53, 2005.

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

D.J. Spiegelhalter, N.G. Best, B.P. Carlin, and A. van der Linde. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, pages 583–639, 2002.

A. Subramanian and S.J. Wei. The wto promotes trade, strongly but unevenly. *Journal of International Economics*, 72(1):151–175, 2007.

A Gravity Models and Data

Two main choices on dependent variables have been considered by researchers in estimating the gravity equation, trade levels and ratios. The level dependent variable can be the log of bilateral (unidirectional) imports/exports, or the average/sum of imports and exports between countries, whereas the later one suffers from the silver medal error proposed by Baldwin and Taglioni (2006). These trade flow data can be measured by current dollar or deflated by price index (US CPI). However, estimations with deflated trade values suffer from the bronze medal error shown in Baldwin and Taglioni (2006). The model with log of bilateral import levels is shown as below,

$$lim_t^{ik} = cons + lyy + EZ + EU + \delta^{ik} + \theta_t^i + \phi_t^k + \sum_{j=1}^{J} \gamma_{jt} g_{jt}^{ik} + \varepsilon_t^{ik}. \tag{9}$$

The dependent variable, lim_t^{ik}, is the log of bilateral import levels, $log(IM_t^{ik})$, which is determined by heterogeneous preferences ($\delta^{ik} \equiv log(\alpha^{ik})$) , the product of importers' expenditures and exporters' outputs ($lyy \equiv log(EXP_t^i * OUT_t^k)$), trade costs ($g_{jt}^{ik} \equiv log(\tau_t^{ik})$) and fixed effects. The trade costs g_{jt}^{ik} include log of distance, dummies for border, common language, land-lock, Euro Zone and European Union. With symmetric conditions, $\tau_t^{ik} = \tau_t^{ki}$ and $\alpha^i(k) = \alpha^k(i)$, trade balance for each country is zero and total output is equal to total expenditure, $EXP_t^i = OUT_t^i$, so that $lyy = log(OUT_t^i * OUT_t^k)$ if replacing expenditure EXP_t^i using output OUT_t^i.

This type of estimation has a potential endogeniety problem because the economic mass data "*lyy*" are included in the explanatory variables. Therefore, researchers use the second choice: the log of the bilateral import ratio— imports divided by the product of the importer's expenditure and exporter's output as in Anderson and Van Wincoop (2003) with cross-section data. This estimation restricts the unit effect of economic mass variables on bilateral trade. The model using bilateral import ratios in Anderson and Van Wincoop (2003) and Aviat and Coeurdacier (2007) is given below,

$$limr_t^{ik} = cons + +EZ + EU + \delta^{ik} + \theta_t^i + \phi_t^k + \sum_{j=1}^{J} \gamma_{jt} g_{jt}^{ik} + \varepsilon_t^{ik}, \tag{10}$$

where the dependent variable is defined as

$$limr_t^{ik} = log(\frac{IM_t^{ik}}{OUT_t^i * OUT_t^k}) \text{ or } log(\frac{IM_t^{ik}}{EXP_t^i * OUT_t^k}).$$

The version of bilateral import ratios may be non-stationary with long panel data. Hence, as in section 3 of Guo (2010), this paper uses the import ratios in equation (2) of section 2.1

$$limr_t^{ik} = log(\frac{IM_t^{ik} * WOUT_t}{OUT_t^i * OUT_t^k}).$$

The data contain 22 OECD countries. There are fourteen countries in EU by 1995—AUT, BEL-LUX, DEU, DNK, ESP, FIN, FRA, GBR, GRC, IRL, ITA, NLD, PRT, and SWE, among which four countries did not join in the EZ in 2000—DNK, GBR, GRC, and SWE. Another eight countries—AUS, CAN, CHE, JPN, USA, ISL, NOR, and NZL— do not belong to EU. The data source is listed as below,

1) Current dollar value of bilateral import/export data: IMF DOTS.

2) Current dollar value of GDP: WDI and IMF DOTS (robustness check).

3) Current dollar value of private consumption expenditure: World Bank's World Development Indicators (WDI).

4) Bilateral trade costs variables: distance, dummies for border connection, landlock, and common language are taken from CEPII. Geodesic (great circle) distances are measured as kilometers between capital cities.[22]

5) EU and EZ dummies: constructed by author following the dates of countries' participation in European Union and Euro Zone.

B Monte Carlo Simulations

B.1 Size under Homoscedastic Errors

All models with different groups of dummy variables are nested in the baseline model. I use the baseline model (the alternative) and the model "IMEX" (the null) as an example to illustrate the Monte Carlo simulations for size distortion presented in table 10

$$w_t^{ik} = cons + EZ + EU + \delta^{ik} + (\widetilde{\theta}_t^i + \widetilde{\theta}^i) + (\widetilde{\phi_t^k} + \widetilde{\phi^k}) + \sum_{j=1}^{3} \gamma_{jt} g_j^{ik} + \varepsilon_t^{ik},$$

where $\theta_t^i = \widetilde{\theta}_t^i + \widetilde{\theta}^i$ and $\phi_t^k = \widetilde{\phi_t^k} + \widetilde{\phi^k}$. In order to obtain the model "IMEX", we need to impose the following 1448 restrictions on the baseline model: $\delta^{ik} = 0$, $\widetilde{\theta}_t^i = 0$, and $\widetilde{\phi_t^k} = 0$. All simulations are performed 1000 times.

[22] http://www.cepii.franglaisgraph /bdd/distances.htm

1) Obtain the coefficients B_0 ($\widetilde{\theta^i}$, $\widetilde{\phi^k}$, EU, EZ, γ_{jt} and the constant intercept) and variance σ_0^2 (=$var(\epsilon_0)$) based on the model $y = XB_0 + \epsilon_0$. I estimate the coefficients from the model "IMEX" using the real data shown in appendix A. The dependent variable is import ratio $log\left(\frac{(1+IM_t^{ik})*WOUT_t}{EXP_t^i*OUT_t^k}\right)$. The coefficients on trade costs γ_{jt} are listed in table 5. The variance is the mean of the squared residual.

2) Simulate the dependent variables $\hat{y}$ for 1000 times given B_0, σ_0^2, and the covariates X from the model "IMEX". The random sample comes from the random draws of the error term.

3) Fit the simulated $\hat{y}$ using both the null and alternative models ($y = XB_m + \epsilon_j$ and $m \in \{n, a\}$, the subscript"n" and "a" is represented the the null and alternative models respectively.) and obtain the estimated coefficients ($\hat{B_n}$ and $\hat{B_a}$) and variance ($\hat{\sigma_n^2}$ and $\hat{\sigma_a^2}$) for 1000 times assuming i.i.d.

4) Calculate the statistics for the LR test and rejection rate (size). I use the formula

$$LR1 = N * \left[log(\hat{\sigma_n^2}) - log(\hat{\sigma_a^2})\right]$$

to calculate the statistic for the LR test,"LR1", where $\hat{\sigma_j^2} = RSS_j/N$ and RSS_j is the residual sum of squares of model m. Following Italianer (1985), the "LR2" adjusts the dimensions of the models (footnote 8); that is

$$LR2 = LR1 * (N - r - K_n/2) = LR1 * (11550 - 1448 - 0.5 * 48)/11550.$$

The statistic is chi-squared distributed with 1448 degrees of freedom and the critical value at 5% significant level is 1537.639. The size is the percentage of rejecting the null model "IMEX" with 1000 simulations when the null model "IMEX" is true.

5) The power curve. The difference between model "IMEX" and the baseline model is the $\sigma_{\widetilde{\theta}}$ and $\sigma_{\widetilde{\phi}}$, and σ_δ. For example, $\widetilde{\theta_t^i} = \theta_t^i - \widetilde{\theta^i}$ with zero mean and $var\left(\widetilde{\theta_t^i}\right) = \sigma_{\widetilde{\theta}}^2$. The standard deviation $\sigma_{\widetilde{\theta}}$ is equal to zero in model "IMEX"; so do $\sigma_{\widetilde{\phi}}$ and σ_δ. By increasing the $\sigma_{\widetilde{\theta}}$, $\sigma_{\widetilde{\phi}}$, and σ_δ by the same scale, i.e. 0.01, the rejection rate of the null model "IMEX" goes up. The power curve plots the rejection rate along with the increasing standard deviations.

6) Calculate the statistic for the F test. The F test can be used to compare models with homoscedastic error terms. In table 10 the F test statistic for null and alternative models is calculated as

$$F = \frac{(RSS_n - RSS_a)/(K_a - K_n)}{RSS_a/(N - K_a)},$$

where RSS is the residual sum of squares and K is the number of estimated coefficients. The F statistic has the degrees of freedom ($K_a - K_n = 1496 - 48 = 1488$ and $N - K_a = 11550 - 1496 = 10054$) and the critical value for the significant level 5% is 1.067.

7) Calculate the statistic for the $Wald$ test. The constraint matrix R_n can be constructed using the conditions $\delta^{ik} = 0$, $\widetilde{\theta_t^i} = 0$, and $\widetilde{\phi_t^k} = 0$. The Wald statistic is calculated as follows

$$Wald = (R_n \hat{B}_a)' * [R_n var(\hat{B}_a) R_n']^{-1} * (R_n \hat{B}_a),$$

where $var(\hat{B}_a) = \hat{\sigma_a^2} * (X_a' * X_a)^{-1}$. We use the consistent and biased estimate $\hat{\sigma_a^2} = RSS_a/N$ to calculate the Wald1 statistics, and use consistent and unbiased estimate $\hat{\sigma_a^2} = RSS_a/(N - K_a)$ to calculate (dimension adjusted) Wald2 statistics. The statistic is chi-squared distributed with degrees of freedom 1448 and the critical value at 5% significant level is 1537.639. The size is the percentage to reject the null model "IMEX" with 1000 simulations when the null model "IMEX" is true.

8) Kernel density graphs for chi-square distribution with DF 1448 in figures 2, 3, and 4. The areas with shades in three figures are the 5% rejection regions. The kernel density for chi-square distribution with DF 1448 in figure 2 is plotted with 1 million of simulations— the solid line; adjustment on dimensions (times 0.873(=11550-1448-0.5*48)/11550)) gives the dash line. Figures 3 and 4 plot the LR and Wald statistics with 1000 simulations. The solid lines are for the LR1 and Wald1 cases respectively, and the dash lines are adjusted with demensions for the LR2 and Wald2 cases.

B.2 Errors with Heteroskedasticity and Autocorrelations

In table 8, I calculate the Wald statistics using Newey-West standard errors with 2 lags, robust to the heteroskedasticity and autocorrelation (HAC). The Monte Carlo simulations assume HAC error terms for a specific importer i and exporter k pair (462 pairs) and specify three parametric forms for the HAC. The conclusions on three hypothesis tests and size distortions (with misspecification or not) are robust to the choices on HAC.

The first HAC, "HAC1", in tables (10), (11), (13), (12), and (14) takes the form as below,

$$\epsilon_t^{ik} = g^{ik} + \nu_t^{ik} \qquad \nu_t^{ik} = b_\nu \nu_{t-1}^{ik} + \mu_t^{ik} \qquad var(g^{ik}) = \sigma_{g^{ik}}^2 \qquad var(\mu_t^{ik}) = \sigma_\mu^2$$

There is no contemporaneous correlation across county pairs. This parametric assumption considers the role of fixed effect in the variance covariance matrix Ξ $(= var(\epsilon))$. The matrix Ξ $(= var(\epsilon))$ is a block diagonal matrix with Ω^{ik} (462 pairs) for one specific importer i and exporter k pair and Ω^{ik} has the following form,

$$\Omega^{ik} = \sigma^2_{g^{ik}} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \vdots & \vdots \\ \vdots & \vdots & 1 & 1 \\ 1 & \cdots & 1 & 1 \end{bmatrix} + \frac{\sigma^2_{\mu}}{1-b^2_{\nu}} \begin{bmatrix} 1 & b_{\nu} & \cdots & b_{\nu}^{T-1} \\ b_{\nu} & 1 & \vdots & \vdots \\ \vdots & \vdots & 1 & b_{\nu} \\ b_{\nu}^{T-1} & \cdots & b_{\nu} & 1 \end{bmatrix}.$$

The second HAC, "HAC2", does not take the fixed affect into account and takes the form,

$$\epsilon_t^{ik} = b\epsilon_{t-1}^{ik} + \upsilon_t^{ik}, \text{ and } var(\upsilon_t^{ik}) = \sigma^2_{\upsilon^{ik}}$$

The variance Ξ is a block diagonal matrix with Ω^{ik}, where

$$\Omega^{ik} = \frac{\sigma^2_{\upsilon_t^{ik}}}{1-b^2} \begin{bmatrix} 1 & b & \cdots & b^{T-1} \\ b & 1 & \vdots & \vdots \\ \vdots & \vdots & 1 & b \\ b^{T-1} & \cdots & b & 1 \end{bmatrix}.$$

Considering both cases in HAC1 and HAC2 leads to the third HAC form, "HAC3", which takes the form,

$$\epsilon_t^{ik} = g^{ik} + \nu_t^{ik} \qquad \nu_t^{ik} = b_{\nu}\nu_{t-1}^{ik} + \mu_t^{ik} \qquad var(g^{ik}) = \sigma^2_{g^{ik}} \qquad var(\mu_t^{ik}) = \sigma^2_{\mu^{ik}}$$

The variance Ξ is a block diagonal matrix with Ω^{ik}, where

$$\Omega^{ik} = \sigma^2_{g^{ik}} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \vdots & \vdots \\ \vdots & \vdots & 1 & 1 \\ 1 & \cdots & 1 & 1 \end{bmatrix} + \frac{\sigma^2_{\mu^{ik}}}{1-b^2_{\nu}} \begin{bmatrix} 1 & b_{\nu} & \cdots & b_{\nu}^{T-1} \\ b_{\nu} & 1 & \vdots & \vdots \\ \vdots & \vdots & 1 & b_{\nu} \\ b_{\nu}^{T-1} & \cdots & b_{\nu} & 1 \end{bmatrix}.$$

With HAC2, the size distortions are larger than those with HAC1 and conclusions remain. With HAC3, the size distortions are close to either HAC1 or HAC2 depending on the null hypothesis models. Simulations with only heteroskedastic errors without serial correlation gives similar results too. The results with HAC1 only are reported in the paper to save space.

B.3 Monte Carlo Simulations for the Misspecified Case

I continue to use the combination of the null model "IMEX" and the alternative baseline model as an example to illustrate the Monte Carlo simulations on the misspecified case (HAC1(M)) in tables 10, 11, 13, 12, and 14. The misspecification refers (no controlling for HAC) to the fact that the simulated data have HAC in the error term, but the regressions ignore the HAC and assume homoscedastic error terms to estimate the variance-covariance matrix of the coefficients.

1) Obtain the coefficients B_0 and variance Ξ_0 $(= var(\epsilon_0))$ based on $(y = XB_0 + \epsilon_0)$. The (estimated) variance covariance matrix Ω^{ik} has the form either HAC1 or HAC2 or HAC3 in appendix B.2.

2) Simulate the dependent variables $\hat{y}$ for 1000 times given B_0, Ξ_0, and covariates in the null model "IMEX". The random sample comes from the random draws of the error term.

3) Fit the simulated data into models, same as in the appendix B.1 assuming homoscedasticity.

4) Calculate the statistics for three tests, including "LR1" and "LR2" for the LR test, "F" for F test and "Wald1" and "Wald2" for Wald test. Then obtain the rejection rates (size) for each test, which follows the appendix B.1 assuming homoscedasticity.

B.4 Monte Carlo Simulations for the Wald Test on B_0

Tables (11), (12), (13) and (14) show the Wald hypothesis tests (Wald2) on the estimated coefficients from both the null (H0) and alternative (H1) with respect to the artificial B_0. The "HAC1(M)" refers to the misspecification case discussed in append (B.3) without controlling for the HAC1. Particularly, tables (12) and (14) provide details for EZ and EU effects, a subset of the B_0. I continue using the same example to illustrate the simulation.

1) Obtain the coefficients B_0 and variance, either homoscedasticity σ_0^2 or heteroskedasticity Ξ_0 based on $(y = XB_0 + \epsilon_0)$ as in append B.1 and B.3.

2) Simulate the dependent variables $\hat{y}$ for 1000 times given B_0,σ_0^2 or Ξ_0, and covariates in small model (c). The random sample comes from the random draws of the error term.

3) Fit the simulated data into models, same as in the appendix B.1 if with homoscedasticity. With HAC, I transform the $\hat{y}$ by multiply the cholesky decomposition of the variance matrix Ξ_0, which has no misspecification. The case with HAC and misspecification is the fact that the simulated data have HAC in the error term, but the regressions assume homoscedastic error terms.

4) Calculate the statistic for the Wald test in table (11) for both null and alternative models. The null hypothesis in the Wald test is $H0 : \hat{B}_m = B_0$ for $m \in \{a, n\}$, and Wald statistic (Wald2) is calculated by following formula

$$Wald = (\hat{B}_m - B_0) * \left[\widehat{var(\hat{B}_m)}\right]^{-1} * (\hat{B}_m - B_0),$$

where $\widehat{var(\hat{B}_m)} = \hat{\sigma^2_m} * (X'_m X_m)^{-1}$ and $\hat{\sigma^2_m} = RSS_m/(N - K_m)$. Then obtain the rejection rates (size) for the test under different assumptions of the error terms.

5) Obtain the rejection rates (size) in table (12) for both null and alternative models based on the choice of the subset of the coefficients B_0.

6) Calculate the Wald statistics in table 13. The null hypothesis is $H0 : \bar{\hat{B}}_m = B_0$ for $m \in \{a, c\}$, the statistic is calculated as

$$Wald = (\bar{\hat{B}}_m - B_0) * \left[\widehat{var(\bar{\hat{B}}_m)}\right]^{-1} * (\bar{\hat{B}}_m - B_0),$$

where the variance covariance matrix is

$$\widehat{var(\bar{\hat{B}}_m)} = \bar{\hat{\sigma^2_m}} * (X'_m X_m)^{-1}/1000.$$

with mean of the estimated variance $\bar{\hat{\sigma^2_m}} = \overline{RSS}_m/(N - K_m)$ and mean of the sum of squared residual $\overline{RSS}_m$. The Wald statistic follows chi-squared distribution given degrees of freedom $K = 48$ and the critical value for the significant level 5% are 65.171.).

Table 1: List of equations: bilateral import levels

Models	Equations
Base	$lim_t^{ik} = cons + EZ + EU + lyy + \delta^{ik} + \theta_t^i + \phi_t^k + \sum_{j=1}^{3} \gamma_{jt} g_j^{ik} + \varepsilon_t^{ik}$
YIMEXP	$lim_t^{ik} = cons + EZ + EU + lyy + \zeta^{ik} + \theta_t^i + \phi_t^k + \sum_{j=1}^{3} \gamma_{jt} g_j^{ik} + \varepsilon_t^{ik}$
YIMEX	$lim_t^{ik} = cons + EZ + EU + lyy + \theta_t^i + \phi_t^k + \sum_{j=1}^{3} \gamma_{jt} g_j^{ik} + \varepsilon_t^{ik}$
IMEXYear	$lim_t^{ik} = cons + EZ + EU + lyy + \mu_t + \theta^i + \phi^k + \sum_{j=1}^{3} \gamma_{jt} g_j^{ik} + \varepsilon_t^{ik}$
IMEX	$lim_t^{ik} = cons + EZ + EU + lyy + \theta^i + \phi^k + \sum_{j=1}^{3} \gamma_{jt} g_j^{ik} + \varepsilon_t^{ik}$
NAYear	$lim_t^{ik} = cons + EZ + EU + lyy + \mu_t + \theta^i + \theta^k + \sum_{j=1}^{3} \gamma_{jt} g_j^{ik} + \varepsilon_t^{ik}$
YNAPair	$lim_t^{ik} = cons + EZ + EU + lyy + \zeta^{ik} + \theta_t^i + \theta_t^k + \sum_{j=1}^{3} \gamma_{jt} g_j^{ik} + \varepsilon_t^{ik}$
YNA	$lim_t^{ik} = cons + EZ + EU + lyy + \theta_t^i + \theta_t^k + \sum_{j=1}^{3} \gamma_{jt} g_j^{ik} + \varepsilon_t^{ik}$
YearPair	$lim_t^{ik} = cons + EZ + EU + lyy + \zeta^{ik} + \mu_t + \sum_{j=1}^{3} \gamma_{jt} g_j^{ik} + \varepsilon_t^{ik}$

Table 2: List of equations: bilateral import ratios

Models	Equations
Base	$w_t^{ik} = cons + EZ + EU + \delta^{ik} + \theta_t^i + \phi_t^k + \sum_{j=1}^{3} \gamma_{jt} g_j^{ik} + \varepsilon_t^{ik}$
YIMEXP	$w_t^{ik} = cons + EZ + EU + \zeta^{ik} + \theta_t^i + \phi_t^k + \sum_{j=1}^{3} \gamma_{jt} g_j^{ik} + \varepsilon_t^{ik}$
YIMEX	$w_t^{ik} = cons + EZ + EU + \theta_t^i + \phi_t^k + \sum_{j=1}^{3} \gamma_{jt} g_j^{ik} + \varepsilon_t^{ik}$
IMEXYear	$w_t^{ik} = cons + EZ + EU + \mu_t + \theta^i + \phi^k + \sum_{j=1}^{3} \gamma_{jt} g_j^{ik} + \varepsilon_t^{ik}$
IMEX	$w_t^{ik} = cons + EZ + EU + \theta^i + \phi^k + \sum_{j=1}^{3} \gamma_{jt} g_j^{ik} + \varepsilon_t^{ik}$
NAYear	$w_t^{ik} = cons + EZ + EU + \mu_t + \theta^i + \theta^k + \sum_{j=1}^{3} \gamma_{jt} g_j^{ik} + \varepsilon_t^{ik}$
YNAPair	$w_t^{ik} = cons + EZ + EU + \zeta^{ik} + \theta_t^i + \theta_t^k + \sum_{j=1}^{3} \gamma_{jt} g_j^{ik} + \varepsilon_t^{ik}$
YNA	$w_t^{ik} = cons + EZ + EU + \theta_t^i + \theta_t^k + \sum_{j=1}^{3} \gamma_{jt} g_j^{ik} + \varepsilon_t^{ik}$
YearPair	$w_t^{ik} = cons + EZ + EU + \zeta^{ik} + \mu_t + \sum_{j=1}^{3} \gamma_{jt} g_j^{ik} + \varepsilon_t^{ik}$

Table 3: Euro effect and European Union effect on (log) bilateral import level by LS: 1992-2002

Dummies	Base	1) YIMEXP	2) YIMEX	3) IMEXYear	4) IMEX	5) NAYear	6) YNAP	7) YNA	8) YearPair	9) OLS
Time-varying	Yes	Yes	Yes	No	No	No	Yes	Yes	No	No
Imp. & Exp.	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No
Nation	No	No	No	No	No	Yes	Yes	Yes	No	No
Year	No	No	No	Yes	No	Yes	No	No	Yes	No
Pair	Asym.	Sym.	No	No	No	No	Sym.	No	Sym.	No
/Var.										
EZ	-0.008	-0.008	0.349**	0.266**	0.267**	0.266**	-0.008	0.349**	0.120**	0.233**
	0.039	0.038	0.107	0.052	0.044	0.052	0.037	0.105	0.028	0.048
EU	0.067	0.067	0.284*	0.205°	0.216*	0.205°	0.067	0.284*	-0.006	0.052
	0.047	0.046	0.144	0.105	0.101	0.105	0.045	0.140	0.032	0.088
lyy	1.172**	0.923**	0.841	0.558**	0.609**	0.558**	0.850**	0.786**	0.523**	0.803**
	0.013	0.022	0.036	0.057	0.043	0.057	0.009	0.043	0.055	0.020
locked_im									-1.493**	-0.328**
									0.063	0.101
locked_ex									-1.661**	-0.496**
									0.064	0.115
log(dist.)			-0.880**	-0.897**	-0.895**	-0.897**		-0.880**		-0.763**
			0.075	0.069	0.069	0.069		0.073		0.044
contig.			0.223	0.217	0.218	0.217		0.223		0.329*
			0.142	0.136	0.136	0.135		0.138		0.128
comlang.			0.391**	0.384**	0.385**	0.384**		0.391**		0.624**
			0.112	0.108	0.108	0.108		0.110		0.126
Obs.	5082	5082	5082	5082	5082	5082	5082	5082	5082	5082
RSS	150.3	550	1450	1512	1517	1932	995	1895	1019	2646
Adj.R2	0.993	0.975	0.935	0.932	0.932	0.913	0.955	0.915	0.954	0.881
LL	1735	-1561	-4024	-4131	-4140	-4754	-3067	-4704	-3128	-5553

This table replicates table 4 in Baldwin and Taglioni (2006). ** significant at 1%; * significant at 5 %; ∘ significant at 10%. The standard errors reported below the coefficient estimates for all models are clustered on time-invariant country pairs. "LL" is the log-likelihood based on i.i.d. errors. "EZ" variable is the dummy for Euro Zone, taking one when both countries are in the euro zone and zero when at least one country is not in the euro zone. "EU" variable is the dummy for European Union, taking one if country pairs are members of European Union (tracing back to European Coal and Steel Community and Treaty of Rome).

Table 4: Euro effect and European Union effect on (log) bilateral import level by LS: 1980-2004

Dummies	Base	1) YIMEXP	2) YIMEX	3) IMEXYear	4) IMEX	5) NAYear	6) YNAP	7) YNA	8) YearPair	9) OLS
Time-varying	Yes	Yes	Yes	No	No	No	Yes	Yes	No	No
Imp. & Exp.	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No
Nation	No	No	No	No	No	Yes	Yes	Yes	No	No
Year	No	No	No	Yes	No	Yes	No	No	Yes	No
Pair	Asym.	Sym.	No	No	No	No	Sym.	No	Sym.	No
/Var.										
EZ	-0.004	-0.004	0.413**	0.330**	0.314**	0.330**	-0.004	0.413**	0.166**	0.009
	0.055	0.055	0.112	0.057	0.051	0.057	0.053	0.110	0.037	0.054
EU	0.227**	0.227**	0.171	0.156*	0.167*	0.156*	0.227**	0.171	0.175**	0.008
	0.056	0.055	0.112	0.074	0.072	0.074	0.054	0.109	0.039	0.078
lyy	1.062**	0.867**	0.766**	0.584**	0.508**	0.584**	0.904**	0.773**	0.602**	0.767**
	0.021	0.027	0.037	0.058	0.013	0.057	0.015	0.045	0.057	0.019
locked_im									-1.350**	-0.320**
									0.055	0.095
locked_ex									-1.522**	-0.492**
									0.058	0.107
log(dist.)			-0.890**	-0.894**	-0.893**	-0.894**		-0.890**		-0.765**
			0.069	0.064	0.064	0.064		0.067		0.041
contig.			0.212	0.212°	0.212°	0.212°		0.212°		0.410**
			0.130	0.125	0.125	0.125		0.127		0.122
comlang.			0.424**	0.422**	0.423**	0.422**		0.424**		0.614**
			0.109	0.105	0.105	0.105		0.107		0.124
Obs.	11550	11550	11550	11550	11550	11550	11550	11550	11550	11550
RSS	751.6	1615	3463	3775	3801	4505	2504	4351	2646	6782
Adj.R2	0.984	0.966	0.928	0.928	0.928	0.915	0.949	0.914	0.949	0.872
LL	-610	-5029	-9433	-9931	-9970	-10951	-7559	-10751	-7879	-13314

This table replicates the table 5 in Baldwin and Taglioni (2006). See table notes in table 3.

Table 5: Euro effect and European Union effect on (log) bilateral import ratio by LS: 1980-2004

Dummies	Base	1) YIMEXP	2) YIMEX	3) IMEXYear	4) IMEX	5) NAYear	6) YNAP	7) YNA	8) YearPair	9) OLS
Time-varying	Yes	Yes	Yes	No	No	No	Yes	Yes	No	No
Imp. & Exp.	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No
Nation	No	No	No	No	No	Yes	Yes	Yes	No	No
Year	No	No	No	Yes	No	Yes	No	No	Yes	No
Pair	Asym.	Sym.	No	No	No	No	Sym.	No	Sym.	No
/Var.										
EZ	-0.003	-0.003	0.240**	0.203**	0.402**	0.203**	-0.003	0.240**	0.109**	0.277**
	0.033	0.032	0.068	0.035	0.032	0.035	0.031	0.066	0.023	0.053
EU	0.135**	0.135**	0.088	0.083∘	0.287**	0.083∘	0.135**	0.088	0.103**	0.313**
	0.033	0.033	0.069	0.046	0.045	0.046	0.032	0.067	0.024	0.084
locked_im									-2.087**	-0.187*
									0.041	0.090
locked_ex									-2.188**	-0.289**
									0.038	0.100
log(dist.)			-0.518**	-0.520**	-0.483**	-0.520**		-0.518**		-0.324**
			0.043	0.040	0.042	0.040		0.042		0.043
contig.			0.099	0.099	0.096	0.099		0.099		0.487**
			0.079	0.076	0.077	0.076		0.078		0.110
comlang.			0.259**	0.258**	0.284**	0.258**		0.259**		0.495**
			0.066	0.063	0.064	0.063		0.064		0.109
Obs.	11550	11550	11550	11550	11550	11550	11550	11550	11550	11550
RSS	274.4	588.2	1269	1380	1939	1633	896	1577	949.7	5746.6
Adj.R2	0.968	0.932	0.857	0.858	0.801	0.832	0.902	0.830	0.901	0.411
LL	5208	805.6	-3636	-4120	-6084	-5093	-1625	-4890	-1961	-12357

See table notes in table 3.

Table 6: Euro effect and European Union effect on (log) bilateral import level by MLE: 1980-2004

Dummies	Base	1) YIMEXP	2) YIMEX	3) IMEXYear	4) IMEX	5) NAYear	6) YNAP	7) YNA	8) YearPair	9) OLS
Time-varying	Yes	Yes	Yes	No	No	No	Yes	Yes	No	No
Imp. & Exp.	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No
Nation	No	No	No	No	No	Yes	Yes	Yes	No	No
Year	No	No	No	Yes	No	Yes	No	No	Yes	No
Pair	Asym.	Sym.	No	No	No	No	Sym.	No	Sym.	No
/Var.										
EZ	-0.004	-0.004	0.002	0.169**	0.177**	0.168**	-0.004	0.001	0.166**	0.168**
	0.022	0.023	0.023	0.017	0.015	0.017	0.025	0.025	0.017	0.015
EU	0.227**	0.227**	0.226**	0.175**	0.188**	0.175**	0.227**	0.227**	0.175**	0.179**
	0.018	0.018	0.018	0.014	0.014	0.014	0.020	0.020	0.014	0.014
lyy	1.120**	0.914**	0.806**	0.602**	0.514**	0.602**	0.904**	0.783**	0.602**	0.523**
	0.002	0.004	0.009	0.018	0.004	0.018	0.108	0.044	0.018	0.004
locked_im									-1.350**	-0.379
									0.372	0.145
locked_ex									-1.522**	-0.552
									0.372	0.145
log(dist.)			-0.890**	-0.895**	-0.892**	-0.895**		-0.890**		-0.668**
			0.047	0.047	0.046	0.052		0.053		0.040
contig.			0.210∘	0.211∘	0.211∘	0.211∘		0.210∘		0.718**
			0.109	0.108	0.108	0.122		0.123		0.179
comlang.			0.430**	0.424**	0.426**	0.424**		0.430**		0.734**
			0.092	0.092	0.092	0.104		0.104		0.137
sigma_u	0.000	0.268	0.484	0.481	0.481	0.543	0.367	0.545	0.364	0.881
	0.002	0.009	0.016	0.016	0.016	0.018	0.012	0.018	0.012	0.030
sigma_e	0.255	0.260	0.260	0.310	0.314	0.310	0.286	0.286	0.310	0.314
	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
rho	0.000	0.515	0.776	0.706	0.701	0.754	0.621	0.784	0.580	0.887
Obs.	11550	11550	11550	11550	11550	11550	11550	11550	11550	11550
R2-P	0.976	0.864	0.841	0.677	0.666	0.673	0.762	0.747	0.688	0.643
LL	-610	-1612	-1879	-3821	-3953	-3876	-2814	-2993	-3696	-4233

** significant at 1%; * significant at 5%; ∘ significant at 10%. MLE assumes the random effects, i.e. $\epsilon_t^{ik} = u^{ik} + e_t^{ik}$, $var(u^{ik}) = \sigma_u^2$ and $var(e_t^{ik}) = \sigma_e^2$.

Table 7: Euro effect and European Union effect on (log) bilateral import ratio by MLE: 1980-2004

Dummies	Base	1) YIMEXP	2) YIMEX	3) IMEXYear	4) IMEX	5) NAYear	6) YNAP	7) YNA	8) YearPair	9) OLS
Time-varying	Yes	Yes	Yes	No	No	No	Yes	Yes	No	No
Imp. & Exp.	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No
Nation	No	No	No	No	No	Yes	Yes	Yes	No	No
Year	No	No	No	Yes	No	Yes	No	No	Yes	No
Pair	Asym.	Sym.	No	No	No	No	Sym.	No	Sym.	No
/Var.										
EZ	-0.003	-0.003	0.000	0.110**	0.312**	0.110**	-0.003	0.000	0.109**	0.308**
	0.015	0.014	0.014	0.010	0.014	0.010	0.015	0.015	0.010	0.014
EU	0.135	0.135**	0.134**	0.103**	0.442**	0.103**	0.135**	0.134**	0.103**	0.449**
	0.012	0.011	0.011	0.009	0.012	0.009	0.012	0.012	0.009	0.012
locked_im									-0.821**	-0.149
									0.222	0.106
locked_ex									-0.922**	-0.251*
									0.222	0.106
log(dist.)			-0.516**	-0.519**	-0.461**	-0.519**		-0.516**		-0.293**
			0.028	0.028	0.029	0.032		0.032		0.030
contig.			0.098	0.098	0.093	0.098		0.098		0.489
			0.066	0.066	0.067	0.074		0.074		0.131
comlang.			0.264**	0.261**	0.303**	0.261**		0.264**		0.510
			0.056	0.056	0.057	0.062		0.063		0.101
σ_u	0.000	0.162	0.293	0.291	0.294	0.327	0.219	0.328	0.217	0.646
	0.001	0.006	0.010	0.010	0.010	0.011	0.007	0.011	0.007	0.021
σ_e	0.154	0.157	0.157	0.187	0.289	0.187	0.172	0.172	0.187	0.289
	0.001		0.001	0.001	0.002	0.001	0.001	0.001	0.001	0.002
rho		0.514	0.776	0.708	0.508	0.753	0.618	0.784	0.575	0.83299
Obs.	11550	11550	11550	11550	11550	11550	11550	11550	11550	11550
R2-P	1.338	1.938	1.878	1.451	0.373	1.439	1.683	1.642	1.480	0.294
LL	5208	4208	3939	2024	-2814	1972	3068	2884	2156	-3170.25

** significant at 1%; * significant at 5%; ○ significant at 10%. MLE assumes the random effects, i.e. $\epsilon_t^{ik} = u^{ik} + e_t^{ik}$, $var(u^{ik}) = \sigma_u^2$ and $var(e_t^{ik}) = \sigma_e^2$.

Table 8: Hypotheses testing: 1980-2004

Comb.	H0	H1	Wald_NW	LR1	LR2	CV(chi2)	F-Test	CV(F)	DF
(1)	YIMEXP	Base	7626	8805	8154.541	244.808	54.751	1.168	210
(2)	YIMEX	Base	48458	17688	16223.954	464.554	87.601	1.119	416
(3)	IMEXYear	Base	64788	18656	16297.752	1512.903	28.447	1.067	1424
(4)	IMEX	Base	91504	22584	19705.762	1537.639	42.121	1.067	1448
(5)	NAYear	Base	84735	20602	17979.035	1534.548	34.449	1.067	1445
(6)	YNAP	Base	20752	13666	12346.137	799.181	30.987	1.091	735
(7)	YNA	Base	72206	20196	18065.366	985.444	50.72	1.081	941
(8)	YearPair	Base	25994	14338	12641.026	1320.968	19.986	1.071	1238
(9)	IMEXyear	YIMEX	903	968	880.503	1082.973	0.909	1.078	1008
(10)	IMEX	YIMEX	4907	4896	4448.366	1107.847	5.356	1.077	1032
(11)	YNAP	YIMEXP	2839	4861	4480.090	579.412	10.231	1.107	525
(12)	YNA	YIMEX	1365	2508	2333.743	579.412	4.84	1.107	525

The LR statistics (LR1 and LR2), log of likelihood ratio for the null and the alternative models are asymptotically distributed as chi squared with the degrees of freedom (DF) as given assuming i.i.d. in the error terms. LR1 is traditionally calculated, but the LR2 uses the method in Italianer (1985) to adjust the dimensions (see appendix B.1). Assuming HAC in the error term, the chi square distributed Wald test statistics are shown in the column "Wald_NW", using the Newey-West kernel for panel data with 2 lags. The "DF" column is the difference of the dimensions in two models (the number of constraints imposed on the alternative models). The "CV(Chi2)" and "CV(F)" columns show the critical value at the 5% significance level for Chi-square distribution and F distribution respectively. All null hypotheses are rejected at the 5% significance level except the comb.(9).

Table 9: BIC for MLE and LS estimations: 1980-2004

Models	LS_ratio			LS_level		MLE_ratio		MLE_level	
	d	LL	BIC	LL	BIC	LL	BIC	LL	BIC
Base	1497	5208	3587.598	-610	15223.6	5208	3596.952	-610	15232.95
YIMEXP	1287	805.6	10427.97	-5029	22097.17	4208	3632.52	-1612	15272.52
YIMEX	1081	-3636	17384.15	-9433	28978.15	3939	2243.505	-1879	13879.51
IMEXYear	73	-4120	8922.874	-9931	20554.23	2024	-3355.77	-3821	8343.583
IMEX	49	-6084	12626.37	-9970	20407.72	-2814	6095.722	-3953	8383.076
NaYear	52	-5093	10672.43	-10951	22397.79	1972	-3448.22	-3876	8257.14
YNAP	762	-1625	10378.08	-7559	22246.08	3068	1001.438	-2814	12765.44
YNA	556	-4890	14981.07	-10751	26703.07	2884	-557.577	-2993	11196.42
YearPair	259	-1961	6344.8	-7879	18180.8	2156	-1879.85	-3696	9824.155
OLS	9	-12357	24798.19	-13314	26721.54	-3170	6433.544	-4233	8568.899
Year	33	-11763	23834.7	-12734	25786.05	1661	-3003.95	-4004	8335.405

$BIC = -2 * LL + d * log(N)$ where N is the number of observations. "d" is the number of estimated parameters in the LS_ratio, and adjusted for other cases. MLE assumes the country-pair random effects and the LS estimations assume i.i.d. error terms. See table notes in 3 and 6.

Table 10: Actual size of the tests with homoskedasticity and HAC1: 1980-2004

Comb.	H0	H1	Homoskedastic Errors					HAC1									
								Control for HAC1: HAC1					No Control for HAC1: HAC1(M)				
			LR1	LR2	Wald1	Wald2	F	LR1	LR2	Wald1	Wald2	F	LR1	LR2	Wald1	Wald2	F
(1)	YIMEXP	Baseline	37.1	12.5	41.5	4.4	4.3	37.5	13.6	40.5	4.4	4.0	100	100	100	100	100
(2)	YIMEX	Baseline	53.4	10.7	65	4.7	4.5	53.6	11.4	65.3	4.8	4.5	100	100	100	100	100
(3)	IMEXYear	Baseline	62.7	0	97.5	6.9	5.3	59	0	97.7	6.4	5.3	100	100	100	100	100
(4)	IMEX	Baseline	62.1	0	97.6	6.9	5.5	58.4	0	97.9	6.5	5.6	100	100	100	100	100
(5)	NAYear	Baseline	62.7	0	97.5	7.1	5.6	58.3	0	97.4	6.3	5.4	100	100	100	100	100
(6)	YNAP	Baseline	63.3	5.2	83.9	5.5	4.7	62.2	5.2	82.6	5.4	5.0	100	100	100	100	100
(7)	YNA	Baseline	66.0	1.9	90.1	5.4	4.3	65.2	2.5	89.9	6.4	5.4	100	100	100	100	100
(8)	YearPair	Baseline	65.7	0.2	95.5	6.2	5.1	61.4	0.4	95.5	5.9	4.6	100	100	100	100	100
(9)	IMEXyear	YIMEX	33.5	0.6	71.2	6.0	5.3	30.1	0.4	68.1	6.1	5.2	0	0	0	0	0
(10)	IMEX	YIMEX	32.9	0.4	71.8	5.7	4.8	30.6	0.3	69.7	6	4.6	0	0	0	0	0
(11)	YNAP	YIMEXP	47.6	7.3	62.5	5.7	4.9	43.4	8.1	59.8	6.3	5.3	90.8	84.6	91.7	83.6	83.3
(12)	YNA	YIMEX	34.0	5.3	50.7	5.7	4.9	32.9	5.3	48.3	5.6	5.3	39.9	32.5	42.8	32.6	32.3

See appendix B.1 for simulations of Wald1, Wald2, F, LR1 and LR2 tests for the homoskedastic errors; see appendix B.2 for Controlling for HAC1; see appendix B.3 without controlling for HAC1, the misspecified error case, which is briefly noted as HAC1(M) in later tables. See DF for each combination of null and alternative models in table 8.

Table 11: Size of the Wald test (Wald2), H0: $\hat{\beta}_m = B_0$ ($m \in \{n, a\}$): 1980-2004

Comb.	H0	H1	H0			H1		
			HOMO	HAC1	HAC1(M)	HOMO	HAC1	HAC1(M)
(1)	YIMEXP	Base	7.2	6.3	100	7.3	6.3	100
(2)	YIMEX	Base	5.7	5.2	44.2	6.2	6.2	100
(3)	IMEXYear	Base	4.7	5	100	4.1	5.5	100
(4)	IMEX	Base	4.5	5.6	100	5.3	5.1	100
(5)	NAYear	Base	4.8	4.8	100	4.3	5.7	100
(6)	YNAP	Base	6	6.1	100	5.9	5.4	100
(7)	YNA	Base	4.7	5.8	47.9	6	5.5	100
(8)	YearPair	Base	5	6.1	100	4.2	5.5	100
(9)	IMEXyear	YIMEX	4.7	5	100	5	5.1	77
(10)	IMEX	YIMEX	4.5	5.6	100	4.9	5.3	90.6
(11)	YNAP	YIMEXP	6	6.1	100	6.3	5.8	100
(12)	YNA	YIMEX	4.7	5.8	47.9	6.2	5.1	44.5

See appendix B.1 for Monte Carlo simulation for HOMO columns; see appendix B.2 for HAC1 columns; see appendix B.3 and table notes in table (10) for HAC1(M) columns. The subscripts n and a represents the null and alternative models respectively. Also see appendix B.4 for the calculations on the size distortions.

Table 12: Actual size of the Wald test (Wald2), H0: $\hat{EZ}_m = EZ_0$ and $\hat{EU}_m = EU_0$.

Comb.	H0	H1	H0			H1		
			HOMO	HAC1	HAC1(M)	HOMO	HAC1	HAC1(M)
(1)	YIMEXP	Base	4.7	5.4	18	4.7	5.4	44.4
(2)	YIMEX	Base	5.3	5.8	74.8	4.7	5.1	46.3
(3)	IMEXYear	Base	4.2	5.4	62.8	4.7	5.5	49.3
(4)	IMEX	Base	4.1	5	61.6	4.7	5.1	59.2
(5)	NAYear	Base	4.2	5.3	60.8	4.7	5.1	49.3
(6)	YNAP	Base	4.7	5.2	12	4.7	5.2	46.4
(7)	YNA	Base	5.3	5.2	74.1	4.7	5	47.2
(8)	YearPair	Base	4.5	5.3	17.8	4.7	5.1	48.9
(9)	IMEXyear	YIMEX	4.2	5.4	62.8	5.3	6	74.3
(10)	IMEX	YIMEX	4.1	5	61.6	4.7	4.9	71.9
(11)	YNAP	YIMEXP	4.7	5.2	12	5.3	5.2	12.6
(12)	YNA	YIMEX	5.3	5.2	74.1	4.7	5.2	74.1

This table focuses on the EZ and EU coefficients particularly. See table notes in table 11 and appendix B.4.

Table 13: Wald statistics (Wald2), H0: $\bar{\hat{\beta}}_m = B_0$

Comb.	H0	H1	DF	CV	H0			H1		
					HOMO	HAC1	HAC1(M)	HOMO	HAC1	HAC1(M)
(1)	YIMEXP	Base	1286	1370.54	1215.2	1228.7	7342.8	1244	1259	20132
(2)	YIMEX	Base	1080	1157.566	1016.4	1029.1	784.6	1017	1045	9369
(3)	IMEXYear	Base	72	92.80827	73.6	67.2	649.3	75	78	2778
(4)	IMEX	Base	48	65.17077	43.1	44.8	564.6	54	57	2565
(5)	NAYear	Base	51	68.66929	55.9	62.6	281.8	36	52	2600
(6)	YNAP	Base	761	826.2871	706.4	791.2	7621.9	745	771	24327
(7)	YNA	Base	555	610.9143	507.6	588	353.5	518	545	5523
(8)	YearPair	Base	258	296.4659	256.8	265	7442.5	268	279	16586
(9)	IMEXyear	YIMEX	72	92.80827	73.6	67.2	649.3	70	78	47
(10)	IMEX	YIMEX	48	65.17077	43.1	44.8	564.6	52	50	55
(11)	YNAP	YIMEXP	761	826.2871	706.4	791.2	7621.9	715	730	7830
(12)	YNA	YIMEX	555	610.9143	507.6	588	353.5	516	520	431

This table focuses on the mean of the estimated coefficients. See table notes in table 11 and appendix B.4.

Table 14: Wald statistics (Wald2), H0: $\bar{\hat{EZ}}_m = EZ_0$ and $\bar{\hat{EU}}_m = EZ_0$.

Comb.	H0	H1	CV	H0			H1		
				Homo	HAC1	HAC1(M)	HOMO	HAC1	HAC1(M)
(1)	YIMEXP	Base	5.991	1.1	1.6	2.6	1.1	1.6	6
(2)	YIMEX	Base	5.991	7.8	3.6	6.5	1.1	3	9.6
(3)	IMEXYear	Base	5.991	5	5.3	2	1.1	2.9	10.8
(4)	IMEX	Base	5.991	5.7	7.1	1.4	1.1	3.3	19.2
(5)	NAYear	Base	5.991	5	4.6	4.2	1.1	2.8	8.7
(6)	YNAP	Base	5.991	1.1	2.2	2.7	1.1	2.2	7.7
(7)	YNA	Base	5.991	7.8	2.4	9.9	1.1	2.8	7.4
(8)	YearPair	Base	5.991	0.4	4.3	1.4	1.1	2.3	9.9
(9)	IMEXyear	YIMEX	5.991	5	5.3	2	7.8	3.4	2.6
(10)	IMEX	YIMEX	5.991	5.7	7.1	1.4	7.8	3.3	3.2
(11)	YNAP	YIMEXP	5.991	1.1	2.2	2.7	1.1	2.2	2.8
(12)	YNA	YIMEX	5.991	7.8	2.4	9.9	7.8	2.4	9.9

This table focuses on the mean of the estimated EZ and EU coefficients particularly. See table notes in table 11 and appendix B.4.

Table 15: Model selections by three information criteria and estimates for EZ/EU effects

Models	$\overline{\widehat{LL}}$	$\overline{\widehat{SSR}}$	$\widehat{SSR}$	$\widehat{LL}$	DIC	AIC	BIC	DF	EZ	EZ-SE	EU	EU-SE
Base	-22202	935	699	-21723	45361	47552	62651	2053	0.043	0.034	0.197**	0.026
YIMEXP	-22173	896	674	-21830	45031	47345	60899	1843	0.038	0.034	0.171**	0.028
YIMEX	-21889	1544	1349	-21613	44328	46496	58521	1635	0.175**	0.033	0.115**	0.020
IMEXYear	-21889	1464	1447	-22059	43437	45185	49113	534	0.161**	0.022	0.095**	0.016
IMEX	-21709	2014	1999	-21890	43057	44800	48551	510	0.355**	0.021	0.336**	0.017
NAYear	-21793	1697	1684	-21968	43235	44961	48734	513	0.156**	0.022	0.090**	0.017
YNAP	-22033	1176	995	-22005	44121	46553	55908	1272	0.029	0.035	0.169**	0.029
YNA	-21787	1741	1633	-21727	43693	45582	53407	1064	0.180**	0.034	0.106**	0.021
YearPair	-22027	1144	1043	-22250	43606	45943	51253	722	0.100**	0.022	0.118**	0.019

This table presents model selection results using the Bayesian method along with the estimated EZ and EU effects and their standard deviation (SE). Burning time: 100,000; thin: 10; simulation: 50,000. ** significant at 1%. See section 4.1 for details.

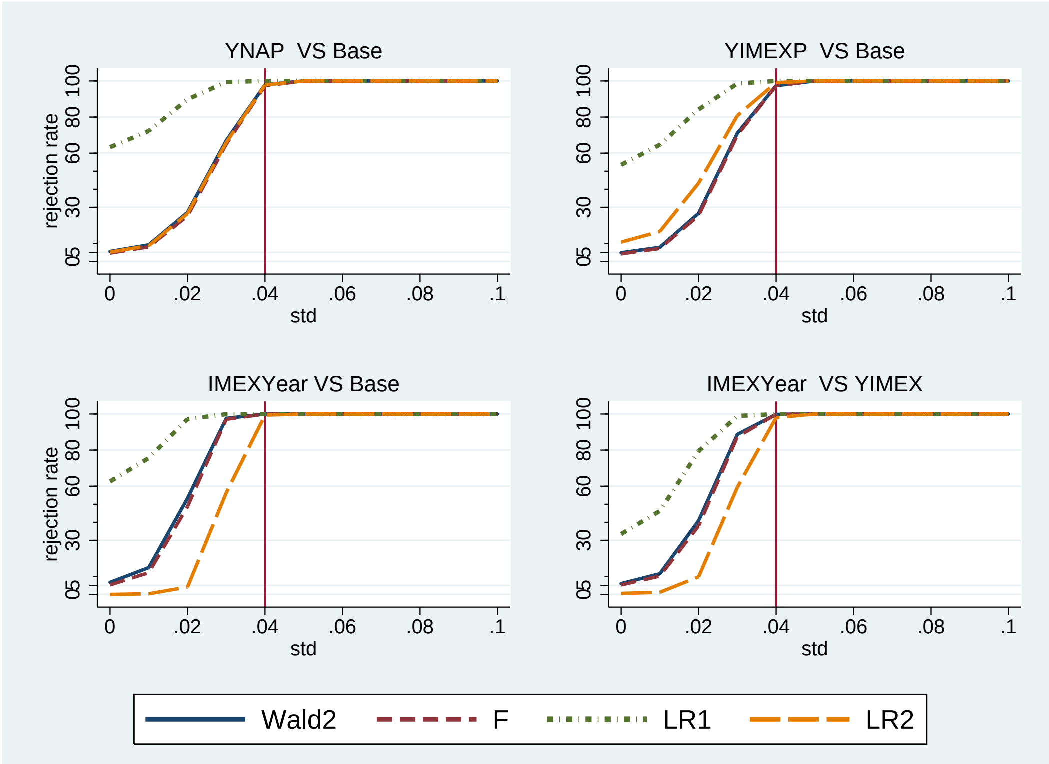

Figure 1: Power curves of three hypothesis tests

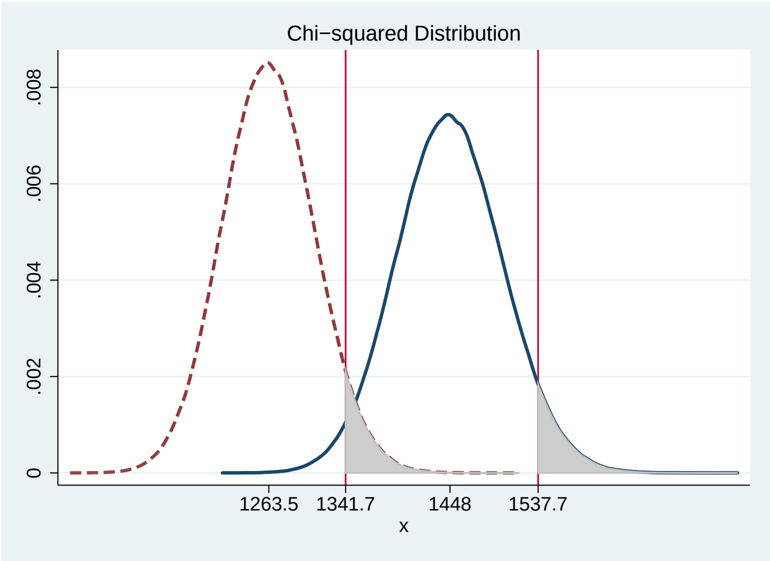

Figure 2: Chi-squared distributions with DF 1448

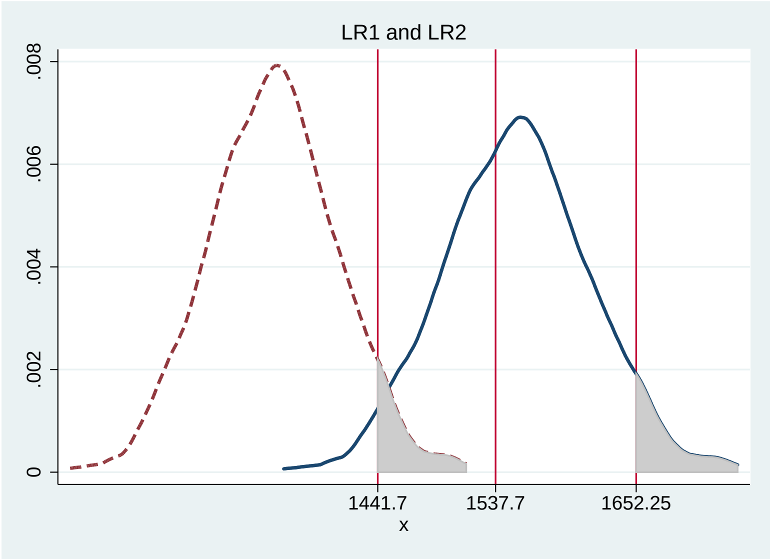

Figure 3: Kernel densities for LR1 and LR2: IMEX vs Base

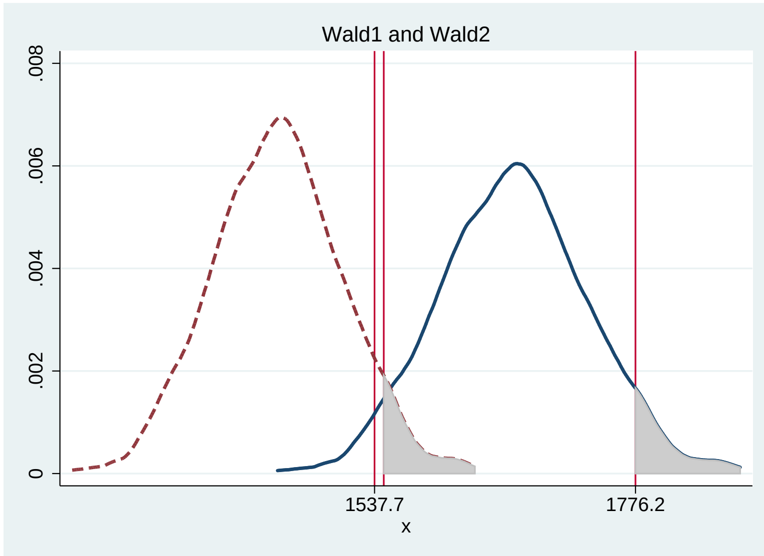

Figure 4: Kernel densities for Wald1 and Wald2: IMEX vs Base

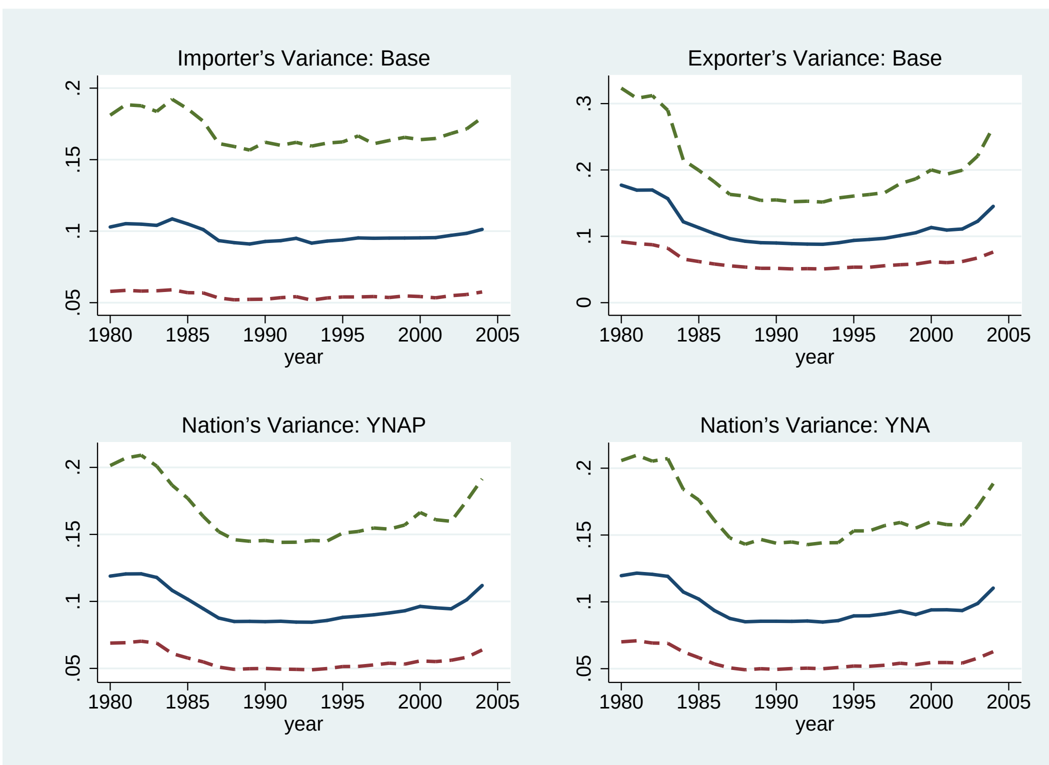

Figure 5: Posterior means and the 95% credible intervals for $\sigma^2_{\theta t}$ and $\sigma^2_{\phi t}$