No. 2004-05-A

OFFICE OF ECONOMICS WORKING PAPER

U.S. INTERNATIONAL TRADE COMMISSION

Monte Carlos Appraisals of Gravity Model Specifications

Michael A. Anderson
Washington and Lee University

Michael J. Ferrantino*
U.S. International Trade Commission

Kurt C. Schaefer
Calvin College

May 2004

*The author is with the Office of Economics of the U.S. International Trade Commission. Office of Economics working papers are the result of the ongoing professional research of USITC Staff and are solely meant to represent the opinions and professional research of individual authors. These papers are not meant to represent in any way the views of the U.S. International Trade Commission or any of its individual Commissioners. Working papers are circulated to promote the active exchange of ideas between USITC Staff and recognized experts outside the USITC, and to promote professional development of Office staff by encouraging outside professional critique of staff research.

Address correspondence to:
Office of Economics
U.S. International Trade Commission
Washington, DC 20436 USA

Monte Carlo Appraisals of Gravity Model Specifications

Michael Anderson
Washington and Lee University

Michael J. Ferrantino
U.S. International Trade Commission

Kurt C. Schaefer
Calvin College schk@calvin.edu

May 2004

COMMENTS WELCOME

ABSTRACT: Many improvements have been proposed for the basic gravity model specification, most of which are confirmed by standard statistical tests due to the large number of observations often used to estimate such models. We use Monte Carlo experiments to examine situations in which features of models may be found statistically significant (or insignificant) when it is known ex ante that they are absent (or present) in the underlying data process. Erroneous assumptions about the presence or absence of lagged dependent variables, fixed effects, free-trade associations and customs unions are shown to introduce economically important bias in estimates of the coefficients of interest, and in some cases to be confirmed spuriously. Policy effects, such as for free trade associations and currency unions, can also be confirmed spuriously when they do not exist in the data-generating process.

Monte Carlo Appraisals of Gravity Model Specifications

Introduction

We have argued that both theory and evidence suggest that history plays a role in shaping the direction of international trade. The standard gravity-model formation, which neglects the role of historical factors, suffers from omitted-variable bias...The implication is that we will never run another gravity equation that excludes lagged trade flows. If our paper is successful (and widely read), neither will other investigators.

Barry Eichengreen and Douglas A. Irwin (1998), pp. 55-56.

...the empirical application of the gravity model is still rather basic. As demonstrated by Cheng and Wall (1999), though providing a high R^2, the standard estimation method tends to underestimate trade between high-volume traders, and overestimate it between low-volume traders. They attribute this to heterogeneity bias...Their fixed-effects method, which I will use in this study...allows(s) for trading-pair heterogeneity and (is) statistically superior to the standard model.

Howard Wall (1999), pp. 35, 40

...I find a large positive effect of a currency union on international trade, and a small negative effect of exchange rate volatility, even after controlling for a host of features...These effects are statistically significant and imply that two countries that share the same currency trade three times as much as they would with different currencies. Currency unions like EMU may thus lead to a large increase in international trade, with all that entails.

Andrew K. Rose (2002), p. 7.

It has become increasingly common to analyze data on international trade flows with the help of gravity models. These models are simple in structure, fit the data well, and are in principle consistent with a wide range of theoretical underpinnings (Deardorff (1998)). The gravity model has been characterized as "a very simple model that explains the size of international trade between countries with a remarkably consistent (and thus, for economics, unusual) history of success as an empirical tool." (Rose, 2000, 14) Gravity models have provided "some of the clearest and most robust empirical findings in economics." (Leamer and Levinsohn, 1995)

Each of this paper's introductory quotations is a strong statement from a widely cited paper. While the first two are about issues of econometric hygiene (specification) and the third addresses an issue of public policy, all three claim that some omission of variables has important implications for the interpretation of results. Eichengreen and Irwin (1998) raise the specter that any regressions which ignore the effect of history by omitting a lagged dependent variable may be misspecified, with the result that policy conclusions drawn from such regressions may be misleading. Cheng and Wall (1999) state that the omission of fixed effects may lead to a pattern of over- and under-estimates of trade flows. Rose (2002) includes currency-union membership as an explanatory variable and finds a coefficient large enough to imply as much as a tripling of trade flows under common currencies—a result with important practical implications, as currency unions represent a policy option that is in some places being pursued with a great deal of public discussion.

This paper argues that many previous findings about gravity models in the literature suffer from a common problem which has not previously been addressed: The data may be plentiful and rich to a fault. Because there are annual (or higher-frequency) observations on many pairs of bilateral trading partners, sometimes for many product categories, it is not uncommon for gravity modelers to operate with thousands or tens of thousands of degrees of freedom in identifying a very few parameters.[1] This makes it relatively easy to obtain statistical verification (i.e. high t-statistics for individual variables or F-statistics for sets of variables) whenever the model specification is elaborated in any way. In turn, this easy verification creates incentives to elaborate the model further.

---

[1]For a panel of bilateral trade between n countries with observations for t time periods, there are $t(n^2\text{-}n)$ available observations. By the standards of gravity modeling, (n=35, t=10) is a relatively modest panel, and contains 11,900 observations. Feasible panels on the order of (n=100, t=30), which yield 297,000 observations, have been available in electronic form from a variety of sources (IMF Direction-of-Trade Statistics, UN COMTRADE, Statistics Canada) for about a decade now.

The irony in this is that the original appeal of the gravity model lay in the fact that it explained a high degree of variation in the data with a relatively small number of parameters. The original double-log gravity models of Tinbergen (1962) and Pöyhönen (1963) in fact required only three variables to explain trade flows, which were further simplified into only two regressors by multiplying together exporter's GDP and importer's GDP for a simple activity variable. Economic distance (usually measured as shipping distance) was the second regressor. This structure could only have been simplified further by dispensing with regression altogether, constraining the coefficients to equal those implied by the physical-science model of Newtonian gravity (1 for activity, -2 for distance). Actually letting the regression *pick* the coefficients was sufficient to cause the amazing goodness-of-fit, with two regressors often explaining 50-70 percent of the variation in large datasets. This is what attracted researchers to the methodology in the first place.

Given their early general success, gravity models have been required to answer ever more specific policy questions. For example, to assess the effects of removing an economic sanction, a fair degree of precision in the estimate of a residual or an out-of-sample projection may be required. To estimate the effects of international borders, non-tariff barriers, or new trade agreements, precision in the estimate of a dummy variable may be wanted. Identification issues also became important for policy purposes. For example, because countries that join trade agreements are usually near each other, it is difficult to separate the effect of distance from the effect of the trade agreement.

Thus, a variety of refined specifications of the gravity equation have been proposed. Moving from cross-sectional data to time-series panels has allowed the use of a lagged dependent variable, country fixed effects for exporters and/or importers, log-first-differences of variables, and estimations of time-varying regression parameters. The proliferation of specifications had unfortunately not been met by a rapid sifting of the possibilities; there is a great deal of controversy over the "correct" specification of the gravity model. There is also debate about just which things the model is able to discern and where its limitations lie. Given the very large volume of data available to gravity modelers, many different specifications of the gravity equation can be made to "work" in the sense that they generate statistical tests that tend to corroborate the specification's validity. So while some researchers have confidence in gravity models, others suggest these strong statistical results may be the result of poor measurement or incorrect specification of the model. In an application of Leamer's extreme bounds analysis, Ghosh and Yamarik (2003) raise doubts that much can be known with confidence from the application of gravity models, given the proliferation of proposed specifications. (The authors mention as many as eighteen candidate variables that have been used in the literature, without exhausting them all.) This literature has by no means reached a consensus, and many versions of the model currently coexist.

Method and Data

Monte Carlo simulations were created for just such agnostic moments. They allow the researcher to create hypothetical but reasonable data that emerge from a *known* data-generating process, and then to evaluate which models most nearly and frequently draw correct conclusions about the underlying structure of the data.

We use Monte Carlo simulations to evaluate a number of common gravity-model specifications:

A. We begin with panel data on trade flows and trade-related, country-pair-specific attributes provided by Andrew Rose (Rose, 2002; data available at http://faculty.haas.berkeley.edu/arose/RecRes.htm). The data comprise 33,903 bilateral trade observations for the years 1970, 1975, 1980, 1985, and 1990, for the

186 countries and other entities for which the United Nations Statistical Office reports trade data. Rose estimates that these data cover 98% of all trade (Rose, 11).[2]

B. We estimate several common specifications of the gravity model, which we number throughout the paper according to the following scheme:

1. A simple gravity model (which we call "the Newtonian model"): logged two-way trade as a function of logged product of GDPs, logged product of per-capita GDPs, logged distance, the standard deviation of changes in the relative exchange rate, and dummy variables for common languages, contiguous borders, having a common colonizer, being part of the same country (as with overseas departments of France), and for cases in which one trade partner colonized the other. This is perhaps the most common extant gravity model.
2. A trade-pair fixed-effects process, removing the time-invariant variables. This is similar to the model preferred by Howard Wall, whom we quote in the introduction.
3. A lagged dependent variable model, which simply adds a lagged dependent variable to the original Newtonian specification. With this model we aim to assess Eichengreen and Irwin's conclusion that they (and we) should never again run a gravity model that does not include a lagged dependent variable.
4. A model supplementing the Newtonian model with a dummy variable indicating if the trade partners are members of the same free-trade association. (No lagged dependent variable is included.) By comparing these results to a simulation that

[2] Trade data come from the World Trade Database. Population and real-GDP data are from the Penn World Table 5.6, filled in with World Bank World Development Indicator data where there were gaps. Data on location/distance, official language, colonial background, and related indicators came from the CIA's web site. The FTA variable was constructed from the WTO's web site.

includes a lagged dependent variable, we can test Eichengreen and Irwin's proposition that FTA variables only appear to be significant because they proxy an omitted lagged dependent variable.

5. A model with the FTA dummy and a common-currency-union dummy. (No lagged dependent variable is included.) This model is used by Rose to reach his provocative conclusion about the astonishingly-large trade-creating effects of currency unions.

C. We then use each of these model estimates to create sets of Monte Carlo data. Each dataset is comprised of 250 randomly-generated panels of trade data. In each case, these panels would be reasonable observations if the underlying particular specification of the gravity model represented the true world-trade data-generating process.

D. We then use each batch of Monte Carlo datasets to estimate parameters for some of the competing gravity model specifications.

This approach allows us to evaluate some of the questions raised in the literature. (For example, do we detect a lagged-dependent-variable coefficient when it's not there in the data-generating process, or do some specifications fail to detect it when it is present? Can the effects of FTAs be distinguished from those of distance?) Our aim is to make judgments about the relative strengths and weaknesses of the various specifications of gravity models, and if possible to suggest which specifications most frequently lead to correct judgments about the actual data-generating process.

Results

The Stata programs gmc0x0y.do (an example is included in Appendix One) perform the Monte Carlo simulations using Model 0x as the true data-generating process and Model 0y as the Monte Carlo estimation model:

A. Monte Carlo data sets are constructed by adding normally-distributed randomly-generated error terms to the trade flows predicted by Model 0x. The standard error of this normal distribution is set equal to the estimated standard error from the estimation of Model 0x.

B. Model 0y is estimated from these 250 Monte Carlo data sets. We report the resulting mean parameter estimates, the standard error of the 250 estimates of each parameter, and the means of the F-statistic, R-squared, and standard error of regression. We also report probability values for standard tests of significance of the parameter estimates, along with probability values for tests for differences between true data-generating parameters and the Monte Carlo estimates of these parameters.3

We discuss these results as a series of scenarios in which a particular data-generating process is investigated using a competing (mistaken) specification of the gravity model. Stata results are reported in Appendix Two in five tables—one table for each data-generating process. Each table reports all of the Monte Carlo simulations that relied upon that data-generating process.

[3] In reading regression results, one normally desires P-values near zero, since the maintained hypothesis is normally that the variables do not matter. But in the Monte Carlo simulation result columns of our tables, one hopes for P-values near one, indicating that a mis-specified model has nonetheless correctly estimated the data-generating-process parameter, since the hull hypothesis

Scenario One: Trade is generated by a Newtonian process, but we instead estimate a country fixed-effects model. We estimate this in two ways: Model 0102 compares the full Newtonian model with a model that estimates country fixed-effects. Model 0102*Lite* compares the fixed-effects model to a Newtonian model that does not include time-invariant variables (distance, contiguous border, common language, common country, common colonizer, colonizer-colony relationship). These time-invariant variables drop out of fixed-effects estimations, and by comparing the fixed-effects results to both versions of the Newtonian model we can obtain a fair comparison of the two models while also exploring the possibility of excluded-variables misspecification bias in the smaller Newtonian model. The results are found on page 24 in Table One Column Three, and page 25 Table One(Lite) Column Three.

All variables are highly significant in both the full Newtonian data-generating model and the FE simulations. The FE simulations' slope parameters all nail the true parameters, though the constant is of course significantly different than in the full Newtonian model. Note the low R-squared in the simulations, despite their respectable F-statistic. Note also that there is very little difference between the "regular" and "Lite" estimates for Model 1, so missing-variable misspecification does not seem to be a problem for the smaller Newtonian model. The fixed-effects specification cannot directly measure the effects of the categorical variables, though it appears to perform reasonably well otherwise--whether the effects of categorical variables are important or unimportant.

Scenario Two: Trade generated by a country fixed-effects process, but we instead estimate the Newtonian model. (Models 0201 and 0201W, reported in Table Two, Columns Two, Three and Four.)

---

underlying these P-values holds that the Monte Carlo simulation is telling the truth about the data-generating process.

Again, all slopes in the data-generating regression and all the simulations are highly significant. Regarding the data-generating regression (Column Two), note again the very low R-squared in the FE models, despite good F-stats.[4]

In simulation 0201 (Column Three), the Newtonian model is not coming anywhere close to the underlying FE model's coefficients. Apparently, if you have a FE data-generating process but estimate a standard Newtonian model, you get odd results.

One implication that might be drawn in comparing 0102 and 0201 (Tables One, One(Lite), and Two) is that if a FE model is capable of picking up something close to the true data-generating process (except for categorical variables) when fixed effects are not present, but the Newtonian OLS model is incapable of picking up the true data-generating process when fixed effects are present, one should always run fixed effects. Of course the categorical variables may be of independent interest, and cannot be obtained in the FE specification.

Another striking aspect of Models 0201, 0201(Lite) and 0203 (Table Two, Columns Three, Four and Five) is that each of them contains one or more variables which we "know" have

[4] Note also that in the data-generating regression results (Column Two) the coefficient on GDP in the FE regressions has a negative (i.e., the wrong) sign. We suspected that this was a complicated artifact of including both Real GDP and Per-Capita Real GDP in the same fixed-effects estimation, which we do throughout the paper in order to be consistent with Rose's original use of the data. It appears our suspicions are correct. When per-capita real GDP is dropped from the data-generating process, real GDP receives a positive coefficient. We removed per-capita real GDP from all of the data-generating processes and Monte Carlo simulations in this paper, and found that none of the results important for our conclusions were affected; there were no changes in the interpretations of the MC simulations. For completeness, we note the following effects when real per-capita GDP is dropped from the regressions and simulations:

Table 1: In the data-generating process, *contiguous border* and *common colonizer* shrink; *contiguous border* also becomes insignificant.

Table 3: The Newtonian and Newtonian-with-FTA MCs capture the *contiguous border* and *common colonizer* effects correctly now. The false-positive finding of the coefficient on FTA (column 5) is twice as large now, going from 0.70 to 1.4.

Table 4: The LDV model (column 4) does a worse job of estimating the *distance* coefficient than before; there's now a significant difference from the data-generating process coefficient. In the data generating process, *contiguous border* and *common colonizer* shrink; *contiguous border* also becomes insignificant.

Table 5: In the data generating process, *contiguous border* and *common colonizer* shrink, and both become insignificant; the coefficient on FTA doubles from estimations that included GDP per capita.

no influence whatsoever on the dependent variable—such things as distance and contiguity. Yet these included variables' coefficients are often identified with a high degree of false statistical significance. This is not the last simulation that will suggest the real possibility that adding variables to otherwise robust specifications can give rise to false positives.

Finally, one additional big-picture comment that could be drawn from these first scenarios concerns the value of traditional t-tests of significance. There is little relationship between the traditional t-statistic in these simulations and the ability of a model to precisely measure the actual population value of the parameter.

Scenario Three: Trade generated by the full Newtonian model, but instead estimated with a lagged-dependent variable model. (Model 0103, in Table 1 Column 4) This and the next scenario allows us to explore whether hysteresis might appear significant when it is not present in the data-generating process, and to see if parameters are mismeasured when real hysteresis is ignored. (Note that our tables report both the "raw" or "impact" coefficients, and the long-run coefficients, in the LDV regressions.)

All coefficients are highly significant, except for the LDV coefficient in the simulation, which is good news because the data-generating process does not include an LDV. All coefficients are also not statistically-significantly different from the true data-generating coefficients. Essentially, if you have a Newtonian world, and instead apply a LDV model, the LDV model measures the coefficients correctly. Even the R-squared is approximately the same.

Scenario Four: Trade generated by a lagged-dependent variable process, but estimated with the full Newtonian model. (Model 0301, reported in Table Three, Column Three)

First compare Column Two of Table Three (the LDV estimates from world trade data) with Column Two of Table One (the Newtonian estimates from world trade data). The LDV model is essentially trying to pick up the effect of all the time-invariant variables; that is why all of them lose their statistical significance. In the process, the LDV model achieves a much higher R^2 than the Newtonian estimations using the same world trade data.

Now consider Column Three of Table Three. The long-run effect of an independent variable in the LDV model is calculated by dividing the coefficient by (one minus the coefficient on the LDV variable). These results, reported in Appendix Table Three Column Two, are the coefficients that should be compared to those of the other models, and the t-statistics and probability values in Columns Three through Five report tests of Monte Carlo estimates relative to the long-run data-generating-process coefficients. In all cases the simulation coefficients are statistically significant, but also significantly different from the long-run effects in the data-generating process. The Newtonian model is lying.

Scenario Five: Trade generated by a process involving country fixed-effects, estimated with a model that ignores fixed effects but includes a lagged dependent variable. (Model 0203, Table Two, Column Five) This and the next scenario allow us to see if hysteresis can be distinguished from or substituted for trade-pair fixed effects.

GDP, per-capita GDP and standard deviation of exchange rate are all significant in all regressions, but the simulation gives estimates significantly different—strikingly different—from the actual data-generating process. The simulation goes on to report four false-positives: distance, common colonizer, colony of the trade partner, and LDV. The coefficients seem fairly large—potentially significant for policy, especially in the latter three cases. It seems that, if country fixed effects matter but are modeled by a LDV, the influence of GDP, per-capita GDP, common colonizer, mutual-colony status, and to some extent distance are all exaggerated. This gives reason for some pause before routinely including lagged dependent variables in gravity models.

Comparing scenario 0203 to 0201, the false positives in 0203 are (with the exception of distance) much less severe than in 0201, which one might expect. If there are country fixed effects in the real world, but they are ignored by the researcher and replaced by a LDV, that LDV is able to pick up some of the fixed effects so that they are not falsely pinned to other variables' coefficients.

Scenario Six: Trade generated by a process involving a lagged dependent variable, estimated with a model that ignores the lagged dependent variable but includes country fixed effects. (Model 0302, table Three, Column Four)

In the initial data-generating regression using world trade data, the LDV is acquiring all of the influences of the non-time-varying variables, which we might expect; thus six variables turn up with statistically insignificant coefficients. The Monte Carlo simulation reports a poor R-squared, along with coefficients that are all statistically significant but also significantly different from the true parameter values. The errors in the GDP and per-capita GDP coefficients are striking. These results give some caution that counterbalances our earlier positive results for fixed-effects specifications. Taken with the last scenario, these results also warrant some agnosticism; it appears that fixed effects are not easily distinguished from hysteresis.

Scenario Seven: Trade generated by a lagged-dependent variable process, but estimated with an FTA model. (Model 0304, Table Three, Column Five) In this and the next scenario, we incorporate FTAs in Model 04 to evaluate the Eichengreen/Irwin-inspired discussion about whether measured FTA effects may actually be caused by hysteresis.

All of the coefficients in the FTA model are statistically significant, but significantly different from the coefficients in the true data generating process. This is not a good sign; the differences are strikingly large. If the world is LDV but is modeled as a FTA world, it appears that the results are not reliable, and apparent FTA effects may be the result of hysteresis.

Scenario Eight: Trade generated by a process influenced by FTAs, but estimated with a lagged-dependent variable without an FTA variable. (Model 0403, Table Four, Column Four)

All coefficients are significant in all regressions, except the LDV variable in the simulation—which is good, because this variable isn't in the data-generating process. None of the coefficients in the simulation are statistically different from those in the data generating process. If the world is FTA and you estimate it with a LDV model, everything seems to be fine. Even the R-squares are about the same between the first estimation and the simulation. In effect, 0403 confirms 0103. Inclusion or exclusion of the additional LVD variable is unlikely to cause spurious findings of significant lags, or to move the other coefficients around much. Similarly, 0401 and 0104 tend to confirm each other – the coefficients on other variables are robust to the inclusion or exclusion of the FTA variable (with the exception of the Distance coefficient in Simulation 0401).

Scenario Nine: Trade generated by the full Newtonian model, but instead estimated with an FTA model. (Model 0104, Table One, Column Five) This scenario and the next allow us to consider whether FTAs can be distinguished from time-invariant factors like distance.

All coefficients are significant in all regressions, except the FTA variable in the simulation—which is good, because this variable is not in the data-generating process. None of the coefficients in the simulation are statistically different from those in the data generating process. It appears that no harm is done in this case by including an irrelevant FTA variable.

Scenario Ten: Trade generated by a process influenced by FTAs, but estimated with the full Newtonian model. (Model 0401, Table Four, Column Three)

All coefficients are significant in all regressions, and the simulation results in estimates that are not significantly different from the true data-generating process, except that the coefficient on distance is overestimated in the simulation that deletes FTAs. The last several paragraphs give some confidence that Rose's FTA variable measures what it purports to measure, and that the influence of FTAs can be distinguished from the effects of distance.

Scenario Eleven: Trade generated by a process influenced by FTAs, estimated with a model that adds currency unions with FTAs. (Model 0405, Table Four, Column Five) This and the next scenario allow us to evaluate claims for large independent currency-union effects. (e.g., Rose, 2000)5

All coefficients are significant in all regressions—not good news, because currency unions are registering a false positive here in the simulation regression. CUs are not in the data-generating process, but show up as statistically significant, with a fairly large coefficient. None of the other simulation coefficients are significantly different from the data-generating process. If an irrelevant CU variable is mistakenly included, it appears that none of the other coefficients are biased, but there is risk of a false-positive on the CU variable.

Scenario Twelve: Trade generated by a process influenced by FTAs and currency unions, estimated with a model that ignores currency unions. (Model 0504, Table Five, Column Three)

[5] We believe we used the same data, model and programming language as Rose (2000), but were unable to exactly replicate his results. Fortunately exact replication is not necessary for the purposes of our paper. Here we report the names of the model parameters, followed by our parameter estimates (from our Table Five, Column Two, starred when our estimate is 5%-statistically-significantly different from the Rose estimate), followed by Rose's parameter estimates (Rose (2000), Table One, "Pooled" column):

Log-Real GDP: 0.790, 0.80.	Log-Real GDP per capita: 0.584*, 0.66.
Log-Distance: -1.089, -1.09.	Contiguous Border dummy: 0.594, 0.53.
Common Language dummy: 0.447, 0.40.	Common Country dummy: 1.398, 1.29.
Common Colonizer dummy: 0.511, 0.63.	Colonial Relationship dummy: 2.151, 2.20.
Std. Dev. of Exchange Rate: -0.043*, -0.017.	FTA dummy: 0.837, 0.99.
CU dummy: 0.903*, 1.21.	N: 22,948, 22,948.
R^2: 0.605, 0.63.	RMSE: 2.093, 2.02.

All coefficients are significant in all regressions, and none in the simulation are significantly different from those in the data-generating process. Of course, the simulation does not generate an estimate for CUs, which do matter to the underlying data-generating process. The influence of CUs seems to be picked up mainly by the Common Country and FTA coefficients—only these estimates are particularly different from those in the data-generating process, but they are still not significantly different (in the statistical sense) from the true parameters. Taken with the last scenario, this is a bit troubling: If CUs are omitted when they should not be, the remaining coefficients are not biased, but if an irrelevant CU variable is included there is risk getting a false positive. Perhaps we are forced to remain a bit agnostic about the importance of CUs.

Conclusions

We have presented evidence based on Monte Carlo analysis that the inclusion of a lagged dependent variable in a gravity equation may indeed prove healthy, though this is complicated by negative results if trade-pair fixed effects affect the data generation process. We have noted several cases in which the inclusion of additional variables can lead to false positives. The most striking finding is a false positive for the Rose currency-union effect when the data-generating process does not contain it.

The finding on fixed-effects coefficients in the presence of hysteresis may be the least popular. It has become increasingly popular to include fixed effects for both exporting and importing countries, or for country pairs in panels with a time-series dimension, while omitting the usual GDP and distance variables. This is usually done so that some new variable of interest can be investigated – the argument being that the fixed effects "soak up" the effects of the traditional gravity variables. However, as fewer of the original variables are included, it becomes more likely that any new variable will prove (falsely) significant, as with the currency-union variable.

There is no guarantee that the results found here will generalize to other datasets, or indeed to other variables added to the present data set. However, we hope that the present results will provide researchers with encouragement in some directions, and caution in others.

References

Cheng, I.H. and Howard Wall (1999). “Controlling for Heterogeneity in Gravity Models of Trade.” *Federal Reserve Bank of St. Louis Working Paper* 1999-010D.

Deardorff, Alan V. (1998). "Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World?" In J.A. Frankel, ed., *The Regionalization of the World Economy*, University of Chicago Press, 7-32.

Eichengreen, Barry and Douglas A. Irwin (1996). "The Role of History in Bilateral Trade Flows." *NBER Working Papers*, 5565, National Bureau of Economic Research, Inc.

Ghosh, S. and S. Yamarik, forthcoming. "Are Regional Trade Arrangements Trade-Creating? An Application of Extreme Bounds Analysis." *Journal of International Economics*.

Leamer, Edward and James Levinsohn (1995). "International Trade Theory: The Evidence." In G. Grossman and K. Rogoff, eds., *The Handbook of International Economics: Vol. III*, Elsevier Science B.V., 1339-1394.

Poyhonen, Pentti (1963). “A Tentative Model for the Volume of Trade Between Countries.” *Weltwirtschaftliches Archiv* 90(1), 93-99.

Rose, Andrew K. (2002). “One Money, One Market: The Effect of Common Currencies on Trade.” Economic Policy, 15:30, 7-46. Data posted at http://faculty.haas.berkeley.edu/arose/RecRes.htm .

Tinbergen, Jan (1962). *Shaping The World Economy: Suggestions for an International Economic Policy*. New York: 20th Century Fund.

Wall, Howard I. (1999). “Using the Gravity Model to Estimate the Costs of Protection.” *Federal Reserve Bank of St. Louis Review* 81:1 (January/February), 33-40.

Appendix One: Typical Stata program file

```
*Stata SE for UNIX 7.0    program to
*    generate gravity model Monte Carlo estimates
*    using gravdata.dta (Rose dataset from 2000)
*    for a full-Newtonian world
*    with country-pair dummies but no FTA CU LDV
*        being modeled instead as a country-pair FE
*        world, no country-pair dummies, CU or LDV
*Simulation results are saved in gmc0102.dta
*Written by Kurt Schaefer, Nov 2003

*Local filenames
*    may need to be replaced by other users

*--->get data
version 7.0
set mem 40m
use gravdata
set more off

*estimate Model 03: Full Newtonian Model:

log using gmc0102.log, replace
display "Program 0102"
display "Regression for the data-generating process: "
display "Full Newtonian Model "

reg lvalue lrgdp lrgdppc ldist border comlang /*
   */ comctry comcol colonial sdd
log off

*save predicted LHS values, std error
*    eit:
predict y1hat, xb
predict ser, stdr

tempname myb
matrix `myb'=e(b)
gen tbrgdp=`myb'[1,1]
gen tbrgdppc=`myb'[1,2]
gen tbdist=`myb'[1,3]
gen tbborder=`myb'[1,4]
gen tbclang=`myb'[1,5]
gen tbcctry=`myb'[1,6]
gen tbccol=`myb'[1,7]
gen tbcolon=`myb'[1,8]
gen tbsdd=`myb'[1,9]
gen tbconst=`myb'[1,10]
preserve
keep tbrgdp tbrgdppc tbdist tbborder tbclang /*
   */tbcctry tbccol tbcolon tbsdd  /*
   */ tbconst
save gmc0102t, replace
restore
```

```
*gen variable for replacement in the following program
gen error=1
gen y1hatnew=1

*1st construct id nums for
*one-way and total-trade flows:
egen c1=group(cty1)
egen c2=group(cty2)
gen onewayti=(1000*c1)+ c2
gen tfid=(1000*c1)+ c2
replace tfid=(1000*c2)+c1 if c2<c1
sort year
egen time=group(year)
tsset onewayti time

* define MC program:
program define pilot
    version 7.0
    if "`1'" == "?" {
        #delimit ;
        global S_1 "rsq rmse fst
                    bgdp segdp bgdppc segdppc
                    bdist sedist
                    bsdd sesdd
                    bconst seconst " ;

        #delimit cr
        exit
    }

*generate random errors with se equal to se of FE
*    regression (See 16.3.2, p. 117):
replace error=ser*invnorm(uniform())

replace y1hatnew=y1hat + error

xtreg y1hatnew lrgdp lrgdppc ldist sdd, fe

scalar rmse=e(rmse)
scalar Fstat=e(F)
scalar rsq=e(r2)

#delimit ;
post `1' (rsq) (rmse) (Fstat)
     (_b[lrgdp])    (_se[lrgdp])
     (_b[lrgdppc])  (_se[lrgdppc])
     (_b[ldist])    (_se[ldist])
     (_b[sdd])      (_se[sdd])
     (_b[_cons])    (_se[_cons]) ;
#delimit cr
scalar drop rmse Fstat rsq
end

*do MC simulations:
simul pilot, reps(250) /*
    */ saving(gmc0102) /*
    */ replace noisily
```

```
clear
*combine MC ests with ests from "true" reg:
use gmc0102
merge using gmc0102t, nokeep

*make scalars for means+se-s; names from "global"
* and "keep" command after 1st regression:

*First from KEEP command, the true betas:
tempname 1b1 1b2 1b3 1b4 1b5 1b6 1b7 1b8 1b9 1b10
tempname 1b11 1b12 1b13

sum tbrgdp
scalar `1b1'=r(mean)

sum tbrgdppc
scalar `1b2'=r(mean)

sum tbdist
scalar `1b3'=r(mean)

sum tbborder
scalar `1b4'=r(mean)

sum  tbclang
scalar `1b5'=r(mean)

sum  tbcctry
scalar `1b6'=r(mean)

sum  tbccol
scalar `1b7'=r(mean)

sum  tbcolon
scalar `1b8'=r(mean)

sum tbsdd
scalar `1b9'=r(mean)


sum tbconst
scalar `1b11'=r(mean)

*Now from the MC simulations:

tempname 2b1 2b2 2b3 2b4 2b5 2b6 2b7 2b8 2b9
tempname 2b10 2b11 2b12 2b13
tempname 2se1 2se2 2se3 2se4 2se5 2se6 2se7
tempname 2se8 2se9 2se10 2se11 2se12 2se13

sum bgdp
scalar `2b1'=r(mean)

sum segdp
scalar `2se1'=r(mean)
```

```
sum bgdppc
scalar `2b2'=r(mean)

sum segdppc
scalar `2se2'=r(mean)

sum bdist
scalar `2b3'=r(mean)

sum sedist
scalar `2se3'=r(mean)

sum bsdd
scalar `2b9'=r(mean)

sum sesdd
scalar `2se9'=r(mean)

sum bconst
scalar `2b11'=r(mean)

sum seconst
scalar `2se11'=r(mean)

tempname r2 ser2
sum rsq
scalar `r2'=r(mean)
scalar `ser2'=r(sd)

tempname rse serse nobs
sum rmse
scalar `rse'=r(mean)
scalar `serse'=r(sd)
scalar `nobs'=r(N)

tempname f sef
sum fst
scalar `f'=r(mean)
scalar `sef'=r(sd)

*Add MC results to gmc0x0y.log file:
log on
display "Monte-Carlo Simulation Results: "
display "FE Model, no country-pair binary vars:"

display "All P(t) below are one-tail tests; "
display "P-values for differences are two-tail."

display "Number of MC iterations: " `nobs'
display "MC mean rmse: " `rse' "  Standard error: " `serse'
display "MC mean F-stat: " `f' "  Standard error: " `sef'
display "MC mean R-sq: " `r2' "  Standard error: " `ser2'

display "All Coefficients do exist in the true"
display "   data-generating process:  "

tempname 2b1tprb
```

```
scalar `2b1tprb'=norm( -(abs(`2b1')/`2se1'))
display "MC real gdp coefficient: " `2b1'
display "Mean standard error: " `2se1' "  P(t)= " `2b1tprb'
tempname 1b1t 1b1tprb
scalar `1b1t'=(`2b1'-`1b1')/`2se1'
scalar `1b1tprb'=2*norm( -(abs(`2b1'-`1b1')/`2se1'))
display "True value: " `1b1' "t for difference: " `1b1t'
display "P-value for difference: " `1b1tprb'


tempname 2b2tprb
scalar `2b2tprb'=norm( -(abs(`2b2')/`2se2'))
display "MC real per-cap gdp coefficient: " `2b2'
display "Mean standard error: " `2se2' "  P(t)= " `2b2tprb'
tempname 1b2t 1b2tprb
scalar `1b2t'=(`2b2'-`1b2')/`2se2'
scalar `1b2tprb'=2*norm( -(abs(`2b2'-`1b2')/`2se2'))
display "True value: " `1b2' "t for difference: " `1b2t'
display "P-value for difference: " `1b2tprb'

tempname 2b3tprb
scalar `2b3tprb'=norm( -(abs(`2b3')/`2se3'))
display "MC distance coefficient: " `2b3'
display "Mean standard error: " `2se3' "  P(t)= " `2b3tprb'
tempname 1b3t 1b3tprb
scalar `1b3t'=(`2b3'-`1b3')/`2se3'
scalar `1b3tprb'=2*norm( -(abs(`2b3'-`1b3')/`2se3'))
display "True value: " `1b3' "t for difference: " `1b3t'
display "P-value for difference: " `1b3tprb'

tempname 2b9tprb
scalar `2b9tprb'=norm( -(abs(`2b9')/`2se9'))
display "MC s.d. of Xchange coefficient: " `2b9'
display "Mean standard error: " `2se9' "  P(t)= " `2b9tprb'
tempname 1b9t 1b9tprb
scalar `1b9t'=(`2b9'-`1b9')/`2se9'
scalar `1b9tprb'=2*norm( -(abs(`2b9'-`1b9')/`2se9'))
display "True value: " `1b9' "t for difference: " `1b9t'
display "P-value for difference: " `1b9tprb'

tempname 2b11tprb
scalar `2b11tprb'=norm( -(abs(`2b11')/`2se11'))
display "MC Constant coefficient: " `2b11'
display "Mean standard error: " `2se11' "  P(t)= " `2b11tprb'
tempname 1b11t 1b11tprb
scalar `1b11t'=(`2b11'-`1b11')/`2se11'
scalar `1b11tprb'=2*norm( -(abs(`2b11'-`1b11')/`2se11'))
display "True value: " `1b11' "t for difference: " `1b11t'
display "P-value for difference: " `1b11tprb'

log off
```

Appendix Two: Simulation Results

Variable definitions:

lrgdp:	log of product of country real GDPs
lrgdppc:	log of ((product of country real GDPs)/(product of country populations))
ldist:	log of distance between the trade pair
contig:	binary variable, =1 if pair shares a contiguous border
comlang:	binary variable, =1 if pair shares a common language
comctry:	binary variable, =1 if pair is part of the same nation (e.g., France and overseas departments)
comcol:	binary variable, =1 if pair were colonies after 1945 with same colonizer
colonial:	binary variable, =1 if one in pair was colonized by the other
sdd:	standard deviation of change in bilateral exchange rate
FTA:	binary variable, =1 if pair shares a FTA
CU:	binary variable, =1 if pair shares a currency union agreement

Table One: Newtonian Data-Generating Process (DGP)

Column two shows the parameters that generate the data.
Columns three through five show the simulation estimates generated from these data. Numbers in **bold** in columns three through five show each estimated coefficient, and the probability value for a test that this coefficient is identical to the true data-generating parameter. (For estimated coefficients that do not appear in the data-generating process, the emboldened probability value presents a test that the estimated coefficient is significantly non-zero.)

(1)	*(2)* **Newtonian DGP**	*(3)* 0102: Fixed effects estimation		*(4)* 0103: LDV estimation*		*(5)* 0104: FTA estimation	
Variable	*Coefficient* *(t-stat)* *Prob{t}*	*Coefficient* *(t-stat)* *Prob{t}*	*t(error)* *Prob{t}*	*Coefficient* *(t-stat)* *Prob{t}*	*t(error)* *Prob{t}*	*Coefficient* *(t-stat)* *(Prob{t})*	*t(error)* *Prob{t}*
lrgdp	**0.785** (135.9) (0.00)	**0.805** (10.5) (0.00)	0.3 **(0.79)**	**0.785** (56.1) (0.00)	0.4 **(0.97)**	**0.785** (135.3) (0.00)	-0.02 **(0.99)**
lrgdppc	**0.596** (55.2) (0.00)	**0.558** (4.8) (0.00)	-0.3 **(0.74)**	**0.594** (4.8) (0.00)	-0.1 **(0.94)**	**0.595** (54.1) (0.00)	-0.07 **(0.94)**
ldist	**-1.148** (-60.4) (0.00)	---	---	**-1.150** (35.9) (0.00)	-0.05 **(0.96)**	**-1.149** (58.3) (0.00)	-0.03 **(0.98)**
contig	**0.587** (6.08) (0.00)	---	---	**0.587** (4.1) (0.00)	-0.00 **(0.99)**	**0.584** (6.1) (0.00)	-0.04 **(0.97)**
comlang	**0.490** (11.7) (0.00)	---	---	**0.490** (7.9) (0.00)	0.01 **(0.98)**	**0.488** (11.6) (0.00)	-0.07 **(0.95)**
comctry	**2.009** (7.6) (0.00)	---	---	**1.948** (3.3) (0.00)	-0.10 **(0.92)**	**1.97** (7.4) (0.00)	-0.15 **(0.88)**
comcol	**0.584** (11.0) (0.00)	---	---	**0.584** (6.7) (0.00)	0.00 **(0.99)**	**0.588** (11.0) (0.00)	0.07 **(0.94)**
colonial	**2.099** (17.9) (0.00)	---	---	**2.095** (12.8) (0.00)	-0.02 **(0.98)**	**2.090** (17.8) (0.00)	-0.08 **(0.94)**
sdd	**-0.043** (-21.2) (0.00)	**-0.043** (14.3) (0.00)	-0.1 **(0.91)**	**-0.043** (15.4) (0.00)	-0.04 **(0.97)**	**-0.043** (21.5) (0.00)	-0.03 **(0.97)**
LDV	---	---	---	**-0.001** (0.08) **(0.47)**	---	---	---
FTA	---	---	---	---	---	**-0.007** (0.07) **(0.47)**	---
constant	**-17.763** (-69.9) (0.00)		-8.0 **(0.00)**	**-17.731** (36.9) (0.00)	-0.1 **(0.95)**	**-17.742** (69.4) (0.00)	0.08 **(0.93)**
$\overline{R^2}$	*0.603*	*0.089*		*0.574*		*0.603*	
F stat	*3870.49*	*526.63*		*1379.85*		*3481.27*	
rmse	*2.098*	*2.099*		*2.098*		*2.098*	
N	*22948*	*250 x 22948*		*250 x 22948*		*250 x 22948*	

*Only the impact coefficients are reported, since the LDV coefficient is statistically zero.

Table One (Lite): Lite-Newtonian Data-Generating Process (DGP)

Column two shows the parameters that generate the data.
Column three shows the simulation estimates generated from these data. Numbers in **bold** in column three show each estimated coefficient, and the probability value for a test that this coefficient is identical to the true data-generating parameter. (For estimated coefficients that do not appear in the data-generating process, the emboldened probability value presents a test that the estimated coefficient is significantly non-zero.)

(1)	*(2)* **Lite-Newtonian DGP**	*(3)* 0102Lite: Fixed effects estimation	
Variable	*Coefficient (t-stat) Prob{t)*	*Coefficient (t-stat) Prob{t}*	*t(error) Prob{t}*
lrgdp	**0.688** (113.0) (0.00)	**0.694** (8.0) (0.00)	0.07 **(0.94)**
lrgdppc	**0.594** (49.7) (0.00)	**0.589** (4.5) (0.00)	-0.04 **(0.97)**
ldist	**---**	---	---
contig	---	---	---
comlang	---	---	---
comctry	---	---	---
comcol	---	---	---
colonial	---	---	---
sdd	**-0.055** (-24.0) (0.00)	**-0.055** (16.2) (0.00)	-0.05 **(0.96)**
LDV	---	---	---
FTA	---	---	---
constant	**-15.739** (-67.2) (0.00)	**-23.711** (18.2) (0.00)	---
$\overline{R^2}$	*0.497*	*0.066*	
F stat	*7557.33*	*384.77*	
rmse	*2.3614*	*2.362*	
N	*22948*	*250 x 22948*	

Table Two: Fixed Effects Data-Generating Process (DGP)

Column two shows the parameters that generate the data.
Columns three through five show the simulation estimates generated from these data. Numbers in **bold** in columns three through five show each estimated coefficient, and the probability value for a test that this coefficient is identical to the true data-generating parameter. (For estimated coefficients that do not appear in the data-generating process, the emboldened probability value presents a test that the estimated coefficient is significantly non-zero.)

(1)	*(2)* **Fixed Effects DGP**	*(3)* 0201: Newtonian estimation		*(4)* 0201Lite: Newtonian estimation, Lite version		*(5)* 0203: LDV estimation		
Variable	*Coefficient* *(t-stat) Prob{t)*	*Coefficient* *(t-stat)* *Prob{t}*	*t(error)* *Prob{t}*	*Coefficient* *(t-stat)* *Prob{t}*	*t(error)* *Prob{t}*	*Impact Coefficient* *(t-stat)* *(Prob{t})*	*Long-run Coefficient* *=impact/* *(1-.935)*	*t(error,LR)* *Prob{t}*
lrgdp	**-0.438** (-9.4) (0.00)	**0.785** (135.9) (0.00)	211.8 **(0.00)**	**0.688** (114.7) (0.00)	185.0 **(0.00)**	**0.038** (5.3) (0.00)	**0.585**	9.3 **(0.00)**
lrgdppc	**1.224** (17.4) (0.00)	**0.596** (55.2) (0.00)	-58.2 **(0.00)**	**0.593** (49.6) (0.00)	-52.7 **(0.00)**	**0.120** (12.0) (0.00)	**1.846**	4.0 **(0.00)**
ldist	---	**-1.148** (60.4) **(0.00)**	---	**---**	---	**-0.074** (4.5) (0.00)	**-1.138**	---
contig	---	**0.588** (6.1) **(0.00)**	---	---	---	**-0.072** (0.95) (0.17)	**-1.108**	---
comlang	---	**0.493** (24.3) **(0.00)**	---	---	---	**0.007** (0.21) (0.41)	**0.108**	---
comctry	---	**1.857** (7.0) **(0.00)**	---	---	---	**0.271** (0.88) (0.19)	**4.169**	---
comcol	---	0.582 (10.9) (0.00)	---	---	---	**0.226** (4.96) (0.00)	**3.477**	---
colonial	---	**2.109** (17.9) **(0.00)**	---	---	---	**0.312** (3.6) (0.00)	**4.800**	---
sdd	**-0.012** (-6.7) (0.00)	**-0.043** (21.5) (0.00)	-15.2 **(0.00)**	**-0.054** (23.8) (0.00)	-18.6 **(0.00)**	**-0.008** (5.3) (0.00)	**-0.123**	4.8 **(0.00)**
LDV	---	---	---	---	---	**0.935** (132.0) (0.00)	**---**	---
FTA	---	---	---	---	---	---	---	---
constant	**-4.662** (6.64) (0.00)	**-17.757** (69.8)**(0.00)**	-8.0 **(0.00)**	**-23.578** (103.9)(0.0)	---	**-2.077** (8.3) (0.00)	**---**	---
$\overline{R^2}$	*0.041(w/in)*	*0.603*		*0.497*		*0.867*	*0.867*	
F stat	*233.51*	*3869.39*		*7559.76*		*6650.51*	*6650.51*	
rmse	*1.273*	*2.099*		*2.361*		*1.101*	*1.101*	
N/#groups	*22948/6707*	*250 x 22948*		*250 x 22948*		*250 x 22948*	*250 x 22948*	

Table Three: Lagged Dependent Variable Data-Generating Process (DGP)

Column two shows the parameters that generate the data—both impact and long-run coefficients. Columns three through five show the simulation estimates generated from these data. Numbers in **bold** in columns three through five show each estimated coefficient, and the probability value for a test that this coefficient is identical to the true data-generating parameter—the *long run* parameter, wherever a long-run parameter is available. (For estimated coefficients that do not appear in the data-generating process, the emboldened probability value presents a test that the estimated coefficient is significantly non-zero.)

(1)	*(2)* **LDV DGP**		*(3)* 0301: Newtonian estimation		*(4)* 0302: FE estimation		*(5)* 0304: FTA estimation	
Variable	*Impact Coefficient (t-stat) Prob{t}*	*Long-run coefficient =impact/ (1-.958)*	*Coefficient (t-stat) Prob{t}*	*t(error,LR) Prob{t}*	*Coefficient (t-stat) Prob{t}*	*t(error,LR) Prob{t}*	*Coefficient (t-stat) (Prob{t})*	*t(error,LR) Prob{t}*
lrgdp	**0.036** (5.89) (0.00)	**0.857**	**0.776** (101.81) (0.00)	-10.6 **(0.00)**	**-0.711** (8.3) (0.00)	18.3 **(0.00)**	**0.779** (102.2) (0.00)	-10.2 **(0.00)**
lrgdppc	**0.064** (7.59) (0.00)	**1.524**	**0.647** (45.22) (0.00)	-61.3 **(0.00)**	**1.000** (8.2) (0.00)	-4.3 **(0.00)**	**0.633** (43.9) (0.00)	-61.8 **(0.00)**
ldist	**-0.020** (-1.42) (0.16)	**-0.476**	**-1.066** (45.38) (0.00)	25.1 **(0.00)**	---	---	**-1.024** (50.9) (0.00)	27.2 **(0.00)**
contig	**0.082** (1.29) (0.20)	**1.952**	**0.503** (4.21) (0.00)	-12.1 **(0.00)**	---	---	**0.469** (3.9) (0.00)	-12.3 **(0.00)**
comlang	**-0.040** (-1.43) (0.15)	**-0.952**	**0.461** (8.98) (0.00)	27.5 **(0.00)**	---	--	**0.441** (8.6) (0.00)	27.2 **(0.00)**
comctry	**0.327** (1.3) (0.21)	**7.786**	**1.737** (3.55) (0.00)	-12.4 **(0.00)**	---	---	**1.554** (3.2) (0.00)	-12.8 **(0.00)**
comcol	**0.030** (0.79) (0.43)	**0.714**	**0.534** (7.43) (0.00)	-2.5 **(0.01)**	---	---	**0.496** (6.88) (0.00)	-3.0 **(0.00)**
colonial	**0.024** (.3) (0.75)	**0.571**	**1.770** (13.15) (0.00)	8.9 **(0.00)**	---	---	**1.782** (13.3) (0.00)	9.0 **(0.00)**
sdd	**-0.008** (-6.64) (0.00)	**-0.190**	**-0.026** (11.06) (0.00)	-69.8 **(0.00)**	**-0.015** (7.3) (0.00)	-85.2 **(0.00)**	**-0.026** (11.2) (0.00)	-70.6 **(0.00)**
LDV	**0.958** (159.8)(0.00)		---	---	---	---	---	---
FTA	---		---	---	---	---	**0.705** (6.49)(**0.00**)	---
constant	**-1.805** (-8.5) (0.00)		**-19.109** (55.75) (0.00)	-50.5 **(0.00)**	**18.816** (11.8)(0.00)	12.9 **(0.00)**	**-19.364** (56.2)(0.00)	-51.00 **(0.00)**
$\overline{R^2}$	*0.902*		*0.658*		*0.036*		*0.659*	
F stat	*9419.20*		*2184.43*		*74.61*		*1978.20*	
rmse	*0.932*		*1.7420*		*0.963*		*1.738*	
N	*10246*		*250 x 10246*		*250 x 10246*		*250 x 10246*	

Table Four: FTA Data-Generating Process (DGP)

Column two shows the parameters that generate the data.
Columns three through five show the simulation estimates generated from these data. Numbers in **bold** in columns three through five show each estimated coefficient, and the probability value for a test that this coefficient is identical to the true data-generating parameter. (For estimated coefficients that do not appear in the data-generating process, the emboldened probability value presents a test that the estimated coefficient is significantly non-zero.)

(1)	*(2)* **FTA DGP**	*(3)* 0401: Newtonian estimation		*(4)* 0403: LDV estimation*		*(5)* 0405: FTA + CU estimation	
Variable	*Coefficient (t-stat) Prob{t}*	*Coefficient (t-stat) Prob{t}*	*t(error) Prob{t}*	*Coefficient (t-stat) Prob{t}*	*t(error) Prob{t}*	*Coefficient (t-stat) (Prob{t})*	*t(error) Prob{t}*
lrgdp	**0.788** (136.4) (0.00)	**0.785** (135.9) (0.00)	-0.5 **(0.63)**	**0.778** (56.0) (0.00)	-0.7 **(0.497)**	**0.790** (136.6) (0.00)	0.46 **(0.64)**
lrgdppc	**0.583** (53.6) (0.00)	**0.595** (55.1) (0.00)	1.1 **(0.26)**	**0.594** (31.1) (0.00)	0.6 **(0.54)**	**0.584** (53.7) (0.00)	0.12 **(0.91)**
ldist	**-1.104** (-56.1) (0.00)	**-1.149** (60.5) (0.00)	-2.4 **(0.02)**	**-1.150** (36.1) (0.00)	-1.5 **(0.14)**	**-1.089** (55.0) (0.00)	0.76 **(0.44)**
contig	**0.584** (6.06) (0.00)	**0.584** (6.05) (0.00)	0.0 **(0.999)**	**0.624** (4.3) (0.00)	0.28 **(0.78)**	**0.594** (6.2) (0.00)	0.11 **(0.91)**
comlang	**0.470** (11.2) (0.00)	**0.487** (11.6) (0.00)	0.4 **(0.67)**	**0.491** (7.9) (0.00)	0.3 **(0.73)**	**0.447** (10.6) (0.00)	-0.54 **(0.59)**
comctry	**1.825** (6.9) (0.00)	**1.967** (7.4) (0.00)	0.5 **(0.59)**	**1.988** (3.4) (0.00)	0.28 **(0.78)**	**1.40** (5.1) (0.00)	-1.56 **(0.12)**
comcol	**0.548** (10.3) (0.00)	**0.588** (11.0) (0.00)	0.7 **(0.46)**	**0.593** (6.8) (0.00)	0.52 **(0.60)**	**0.511** (9.5) (0.00)	-0.70 **(0.48)**
colonial	**2.123** (18.1) (0.00)	**2.090** (17.8) (0.00)	-0.3 **(0.78)**	**2.091** (12.8) (0.00)	-0.2 **(0.84)**	**2.151** (18.3) (0.00)	0.24 **(0.81)**
sdd	**-0.043** (-21.5) (0.00)	**-0.043** (21.6) (0.00)	0.2 **(0.87)**	**-0.043** (15.3) (0.00)	0.1 **(0.92)**	**-0.043** (21.5) (0.00)	0.25 **(0.80)**
LDV	---	---	---	**0.006** (0.44) **(0.33)**	---	---	---
FTA	**0.903** (8.8) (0.00)	---	---	---	---	**0.837** (8.1) (0.00)	-0.65 **(0.52)**
CU	---	---	---	---	---	**0.903** (6.2) (**0.00**)	
constant	**-18.023** (-70.6) (0.00)	**-17.744** (69.9) (0.00)	1.1 **(0.27)**	**-17.535** (36.5) (0.00)	1.0 **(0.31)**	**-18.262** (70.8) (0.00)	-0.93 **(0.35)**
$\overline{R^2}$	*0.604*	*0.603*		*0.574*		*0.605*	
F stat	*3502.88*	*3868.17*		*1380.99*		*3193.05*	
rmse	*2.095*	*2.098*		*2.099*		*2.093*	
N	*22948*	*250 x 22948*		*250 x 22948*		*250 x 22948*	

*Only the impact coefficients are reported, since the LDV coefficient is statistically zero.

Table Five: FTA + CU Data-Generating Process (DGP)

Column two shows the parameters that generate the data.
Column three shows the simulation estimates generated from these data. Numbers in **bold** in column three show each estimated coefficient, and the probability value for a test that this coefficient is identical to the true data-generating parameter. (For estimated coefficients that do not appear in the data-generating process, the emboldened probability value presents a test that the estimated coefficient is significantly non-zero.)

(1)	*(2)* **FTA + CU DGP**	*(3)* 0504: FTA estimation	
Variable	*Coefficient* *(t-stat)* *Prob{t)*	*Coefficient* *(t-stat)* *(Prob{t})*	*t(error)* *Prob{t}*
lrgdp	**0.790** (136.6) (0.00)	**0.788** (136.5) (0.00)	-0.46 **(0.64)**
lrgdppc	**0.584** (53.7) (0.00)	**0.583** (53.6) (0.00)	-0.12 **(0.91)**
ldist	**-1.089** (-55.0) (0.00)	**-1.104** (56.1) (0.00)	-0.77 **(0.44)**
contig	**0.594** (6.17) (0.00)	**0.584** (6.1) (0.00)	-0.11 **(0.91)**
comlang	**0.447** (10.6) (0.00)	**0.470** (11.2) (0.00)	0.55 **(0.59)**
comctry	**1.398** (5.1) (0.00)	**1.825** (6.9) (0.00)	1.61 **(0.11)**
comcol	**0.511** (9.5) (0.00)	**0.548** (10.3) (0.00)	0.70 **(0.48)**
colonial	**2.151** (18.3) (0.00)	**2.123** (18.1) (0.00)	-0.24 **(0.81)**
sdd	**-0.043** (-21.2) (0.00)	**-0.043** (21.5) (0.00)	-0.25 **(0.80)**
LDV	**---**	**---**	**---**
FTA	**0.837** (8.1) (0.00)	**0.903** (8.8) (0.00)	0.65 **(0.52)**
CU	**0.903** (6.2) (0.00)	**---**	**---**
constant	**-18.262** (-70.8) (0.00)	**-18.023** (70.6) (0.00)	
$\overline{R^2}$	*0.605*	*0.604*	
F stat	*3193.05*	*3502.88*	
rmse	*2.093*	*2.095*	
N	*22948*	*250 x 22948*	