Persistent Patterns of International Commerce

Michael D. Ward and Peter D. Hoff

Journal of Peace Research 2007 44: 157

DOI: 10.1177/0022343307075119

journal of peace RESEARCH

© 2007 Journal of Peace Research,
vol. 44, no. 2, 2007, pp. 157–175
Sage Publications (Los Angeles, London, New Delhi and Singapore) http://jpr.sagepub.com
DOI 10.1177/0022343307075119

Persistent Patterns of International Commerce*

MICHAEL D. WARD

Department of Political Science, University of Washington

PETER D. HOFF

Departments of Statistics and Biostatistics, University of Washington

The authors examine a standard gravity model of international commerce augmented to include political as well as institutional influences on bilateral trade. Using annual data from 1980-2001, they estimate regression coefficients and residual dependencies using a hierarchy of models in each year. Rather than gauge the generalizability of these patterns via traditional measures of statistical significance such as *p*-values, this article develops and employs a strategy to evaluate the out-of-sample predictive strength of various models. The analysis of recent international commerce shows that in addition to a typical gravity-model specification, political and institutional variables are important. The article also demonstrates that the often-reported link between international conflict and bilateral trade is elusive, and that inclusion of conflict in a trade model can sometimes lead to reduced out-of-sample predictive performance. Further, this article illustrates that there are substantial, persistent residual exporter- and importer-specific effects, and that ignoring such patterns in relational trade data results in an incomplete picture of international commerce, even in the context of a well-established framework such as the gravity model.

Introduction

There is longstanding interest in understanding international commerce and its relationship with international conflict. The *Journal of Peace Research* has published a series of important articles analyzing the effect of conflict on international commerce, including Barbieri (1996), Barbieri & Schneider (1999), Polachek, Robst & Chang (1999), Oneal & Russett (1999), Dorussen (1999, 2002), Morrow (1999), Hegre (2000, 2002), Barbieri & Levy (1999, 2001), Anderton & Carter (2001a,b), Long (2003), Gartzke & Li (2003), and Goenner (2004), among others.

It is well recognized and generally accepted that conditional on a variety of local and global contexts, countries will have levels of bilateral commerce roughly proportional to their combined share of the global market, but inversely proportional to their distance from one another. This idea is known as the gravity model and stands at the core of modern understanding of the patterns of international commerce (Feenstra, Rose & Markusen, 2001): countries

* We appreciate the comments of Janet Box-Steffensmeier, Nils Petter Gleditsch, Kristian Skrede Gleditsch, Brian Pollins, and Indra de Soysa on an earlier version of this manuscript, as well as the helpful comments on behalf of JPR by reviewers and Håvard Hegre. Peter Hoff's research was partially supported by Office of Naval Research grant N00014-02-1-1011. Hoff and Ward were also supported by a grant from the Methods, Measurement, and Statistics Program at the National Science Foundation, grant number SES-0417559. The data used in this article and a web appendix can be found at http://www.prio.no/jpr/datasets. Correspondence to mdw@u.washington.edu.

with large economies have a lot of trade with each other, especially if they are neighbors. A gravity model can typically explain about one-half the variation in bilateral international commerce and is widely considered a benchmark if not the gold standard. However, few political economists believe that international commerce occurs absent the friction and lubrication created by individuals, firms, and institutions which are not represented completely by size and distance alone. Thus, discerning the additional forces that affect international commerce has been an important item in this longstanding research agenda.

As a result, many scholars have introduced into the gravity model a variety of factors thought to affect the bilateral flow of goods and services. Following the pioneering works of Williamson (1985) and North (1990), scholars have focused on a wide variety of international as well as domestic institutions that affect the bilateral level of commerce. Such institutions are those that are thought to provide accountable, stable, corruption-free, and effective governance, especially in so far as they provide an enforceable legal framework. Recent work by De Groot et al. summarizes these findings in the sphere of the domestic political economy:

> We find that [domestic] institutional quality has a significant, positive and substantial impact on bilateral trade flows. The same goes for similar quality of governance. These results support the hypothesis that institutional variation is an important determinant of informal barriers to trade. The positive correlation between income per capita and quality of institutions gives rise to an explanation of why high-income countries trade disproportionately amongst each other, while the same does not hold for low-income countries. Generally good governance lowers the transaction costs for trade between high-income countries, while trade between low-income countries suffers from high insecurity and transaction costs. (De Groot et al., 2004: 119)

Rose (2004a,b, 2005a,b, 2006) has carefully investigated whether and which international organizations may provide similar positive benefits for international commerce, finding that many well-worn adages about the international institutions are, if not simply wrong, certainly more nuanced than previously thought. Rose's work is hard to summarize, but he finds that not all international organizations have substantial nor similar impacts on international commerce. Indeed, the impacts of the GATT/WTO and IMF are generally quite small and often negative, while the impact of membership in the OECD is both substantial and positive (Rose, 2005b: 692).

But incorporating domestic and international institutions does not tell the entire story of international commerce. Does international collaboration spill over into the economic realm beyond what an augmented standard gravity model predicts? One perplexing claim is whether political cooperation and conflict have substantial impacts on international commerce. There is a considerable lineage of studies that show that international political conflict attenuates bilateral commerce and that international cooperation promotes it. For example, Barbieri (1996) and Barbieri & Levy (2001) find that trade is not necessarily diminished by war among trading partners, but Anderton & Carter (2001a,b) find evidence that war often interrupts international commerce at the bilateral level. Many scholars believe that since democracies are thought to have less conflict with one another, this will have positive externalities upon their bilateral commerce. Bliss & Russett (1998) argue, for example, that private actors in democratic states will prefer commerce with partners also in democratic states because democracies rarely go to war with one another and the attendant risk ratio of having business interrupted by international conflict is very low. Since democratic societies are viewed as less threatening in the foreign policy domain, international commerce with partners in democratic societies will rarely be viewed as threatening national security. Mansfield, Milner & Rosendorff (2000) have suggested that because governments are subject to domestic pressures in developing trade policies, democratic countries are more likely to agree to liberalizing trade agreements than non-democratic countries. Consequently, democratic countries will also trade more with one another. Morrow, Siverson & Tabares (1999) also show that pairs of countries that are democratic trade more with each other. However, the role of mass-based domestic politics in poor, authoritarian countries is unclear. Keshk, Pollins & Reuveny (2004) recently suggested that the widespread notion that conflict inhibits trade does not appear to be a very robust finding. Stated differently, there is no consensus on the linkage between conflict and commerce, either in terms of theoretical expectations or in terms of empirical findings: prominent published studies have concluded that bilateral international conflict either stimulates, attenuates, or has no direct effect on bilateral international commerce.

Such discrepancies can arise for a variety of reasons, two of which we address in this article. First of all, despite their prominence and utility over the past four decades, statistical implementations of gravity models have misestimated the flow of goods among countries, often by a lot.[1] One main reason is that the typical linear model specification assumes that the residuals are 'pattern-less', or statistically independent and identically distributed. This is implausible, since it assumes that there is no correlation of trade flows having the same source, nor any correlation of flows having the same destination. Given the asymmetry of international commerce, in which a few states account for a large share of international commerce, such an assumption is untenable. Such correlation calls into question the validity of any *p*-values or standard errors associated with regression coefficients, and therefore the validity of any theory of trade based on such a statistical analysis.

A second explanation for inconsistent results across studies is that the relationships between trade and explanatory variables may vary over time, and a strong relationship between variables in one year does not necessitate the same relationship in subsequent years. Strictly speaking, there is no statistical theory to justify the claim that a significant result in one year should generalize to other years: Even when the above-mentioned correlations are properly accounted for, the strict sampling-theory interpretation of confidence intervals and *p*-values is that the data from a given year gives information about the super-population from which that year's data were sampled. Since the data used in the estimation of a gravity model typically include all data from a given year and are therefore not randomly sampled, the existence of such a super-population and the connection between it and data from another year are hypothetical. Because of this, whether or not an explanatory variable should be part of a general, broadly applicable theory of trade is not something that can be addressed with one year's worth of data. Typically, this is addressed by panel methods, which do not work well in the case of trade, as we demonstrate below.

To address these issues, we suggest an alternative evaluation of gravity models in which

(a) residual statistical dependencies in trade involving a common exporter or importer are estimated via latent additive and multiplicative effects, and
(b) the generalizability of regression and latent effects are evaluated by a type of out-of-sample predictive performance, which we refer to as temporal persistence.

[1] McCallum (1995) found that borders between the United States and Canada were associated with a huge distortion in US commerce with Canada: trade among provinces was 22 times larger than trade between provinces and states. This was called the paradox of the missing trade, especially since there are few formal barriers to commerce across this border.

The former builds on work on the bilinear random effects model developed earlier (Hoff & Ward, 2004; Hoff, 2005), while the latter provides a re-evaluation of the potential impact of institutional context and conflict patterns on trade, and gives a description of international commerce that exposes many heretofore hidden dependencies in trade data. This should assist us in developing a clearer, more complete picture of the evolution of bilateral trade in the contemporary world and enable a more detailed evaluation of the major claims which are addressed in the context of the gravity model.

The Gravity Model of International Commerce

Sir Isaac Newton developed the first gravity model when he noted in *Philosophiae Naturalis Principia Mathematica* (1687) that 'every object in the Universe attracts every other object with a force directed along the line of centres for the two objects that is proportional to the product of their masses and inversely proportional to the square of the separation between the two objects'. That is, the gravitational force between two objects i and j, given their masses M_i,M_j and the distance $D_{i,j}$ between them, can be expressed as

$$F_{i,j} \propto \frac{M_i \times M_j}{D_{i,j}^2}.$$

The gravity model for international trade can be conceptualized analogously: given some definition of the 'economic mass' of each country, some additional 'gravitational' influences $Z_{i,j}$ and a measure of multiplicative error (e^{ϵ}), a multiplicative model for trade flow between countries i and j is

$$F_{i,j} \propto \frac{M_i^{\beta_1} \times M_j^{\beta_2} \times Z_{i,j}^{\beta_4} \times e^{\epsilon_{i,j}}}{D_{i,j}^{-\beta_3}}.$$

Taking logarithms reveals the standard linear form of the gravity model typically employed in statistical studies of bilateral trade:

$$\underbrace{\ln F_{i,j}}_{\text{bilateral trade}} = \beta_0 + \underbrace{\beta_1 \ln M_i + \beta_2 \ln M_j}_{\text{mass}} + \underbrace{\beta_3 \ln D_{i,j}}_{\text{distance}} + \underbrace{\beta_4 \ln Z_{i,j}}_{\text{other forces}} + \epsilon_{i,j}$$

where, making the analogy with Newton's law of gravitation, we expect $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_3 < 0$.[2]

Such models for trade are widely employed and have typically been evaluated using confidence intervals and p-values for the regression parameters, obtained using ordinary least-squares regression and assuming the error terms are independent and identically distributed. In particular, they assume as follows:

$$\begin{aligned} E(\epsilon_{i,j} \times \epsilon_{j,i}) &= E(\epsilon_{i,j} \times \epsilon_{i,k}) \\ &= E(\epsilon_{i,j} \times \epsilon_{k,j}) = 0 \end{aligned} \tag{1}$$

meaning that the correlation of residual trade flow between two countries is zero and that the correlation of exports and the correlation of imports from a given country are also both zero. In the context of international trade, it is dubious that these assumptions are valid. There is good reason to expect that a wide range of countries will import certain goods from specific, large exporters, such as the United States or China. It is also to be expected that many countries will export to specific, large importers, such as Japan and the United States. Finally, we may expect that if one country has certain trade agreements with another, then it will have both large residual exports to and imports from that country. There should be a large amount of residual dependence that is attributable to the sender, the receiver, and the dyad that the standard formulation ignores. We explore this in the

[2] $\beta_3 < 0$ following the ratio specification of the gravity model. This reflects the idea that countries far apart will have less trade with one another, other things being equal.

context of a single year's data in the next section, and over a 20-year period below.

Analysis of 1981 Data

We begin with a descriptive analysis of the 1981 trade data using the following linear model:

$$\begin{aligned} \ln F_{i,j} = \beta_0 + \beta_1 \ln G_i + \beta_2 \ln G_j \\ + \beta_3 \ln D_{i,j} + \beta_4 P_i + \beta_5 P_j \\ + \beta_6 P_i \times P_j + \beta_7 C_{i,j} + \epsilon_{i,j} \end{aligned} \quad (2)$$

where $F_{i,j}$ portray the annual exports of country i to country j, G_i is the gross domestic product of country i, $D_{i,j}$ is the distance from the capital of i to the capital of j, P_i is the polity score of country i, and $C_{i,j}$ measures the amount of cooperation between i and j in conflictual international disputes. Details on the data employed are provided in Appendix A and in the web appendix for this article. The basic substantive motivation of this model is that the flow of trade between two countries is log proportional to the size of their individual economic outputs (G_i, G_j) and to their geographic distance from one another ($D_{i,j}$). Additionally, it is generally hypothesized that countries with developed democratic institutions (large values of P_i) will not only export more goods but will also import more. Further, countries with similar institutional environments which are 'jointly' democratic or autocratic (large values of $P_i \times P_j$) are expected to have a tendency to trade more with one another. Finally, following the literature, it is generally hypothesized that countries that are cooperative in militarized interstate disputes (large values of $C_{i,j}$) will trade more with one another, while those in conflicts will tend to trade less with each other.[3] This representation of log trade in terms of linear combinations of explanatory variables falls into the class of standard gravity models of international commerce, though individual variable choices may differ slightly

Table I presents estimated parameters, nominal standard errors, and bivariate correlations $\hat{\rho}_{yx}$ between trade and the explanatory variables in Equation (2). The signs on the coefficients for GDP and distance are as expected, indicating positive relationships between trade and GDP as well as proximity. There is also a positive relationship between trade and the polity scores of both the exporter and importer. The variable measuring the degree of cooperation in conflict episodes does not have strong independent association with the volume of trade. But, by and large, the standard gravity model is shown to conform to prior expectations, with plausible estimated parameters which have small variances around them.

As described in Hoff & Ward (2004), the residuals obtained from analyses of dyadic data such as these often display a large amount of structure. Each country in the 1981 database is represented many times, both as an exporter and an importer. This suggests examining the residuals from the above model with the following 'row and column effects' model:

$$\hat{\epsilon}_{i,j} = a_i + b_j + \gamma_{i,j}, \quad (3)$$

where i and j range over the indices of the countries. This fits an additive exporter effect a_i and an importer effect b_j for each country. Fitted values for these coefficients are shown in Figure 1.

The results indicate that, in 1981, countries such as Singapore, Panama, and South Korea exported more than other countries having similar explanatory variables, whereas Nigeria, Algeria, Niger, and Oman exported less. In terms of imports, Singapore, the Netherlands, and Chile imported more and Turkey, Nigeria, and Peru less than would be expected under the

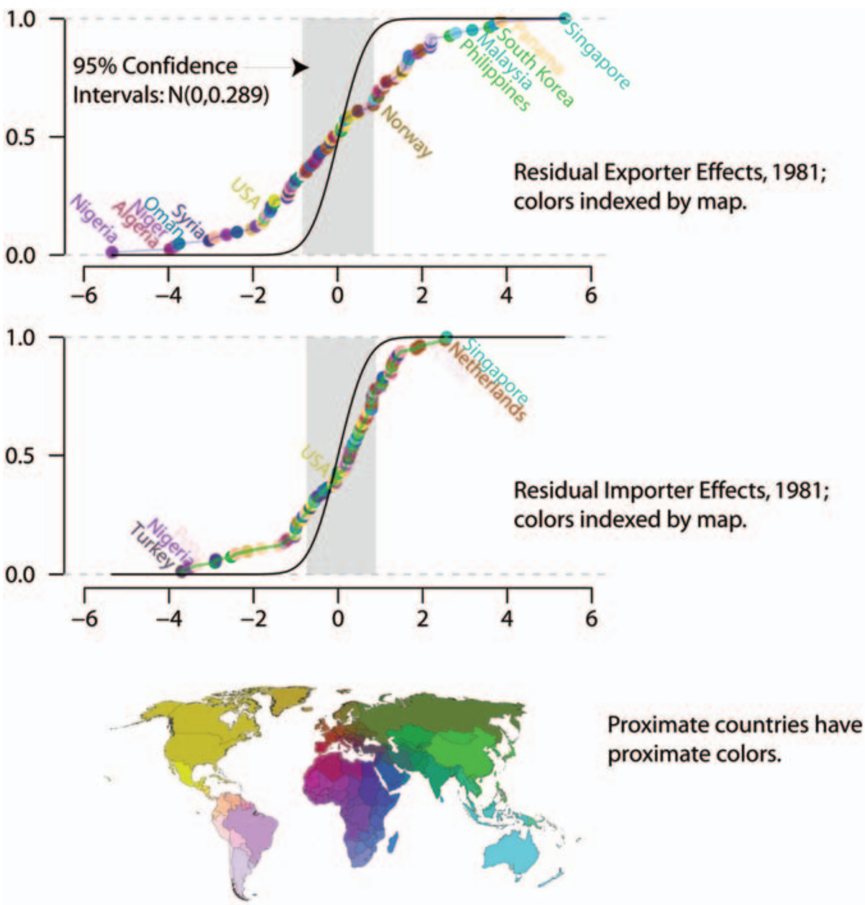
[3] Our measure is a nuanced version of the Militarized Interstate Dispute data. It does not explicitly measure war, though wars are a subset of the militarized disputes.

Table I. Least Squares Estimates of a Gravity Model as Specified in Equation (1) Estimated with Data from 1981.

Variable	p_{yx}	β	$se(\beta)$
(Intercept)		–64.237	1.103
GDP of exporter, logged	0.554	1.897	0.032
GDP of importer, logged	0.431	1.476	0.032
Distance, logged	–0.285	–2.134	0.075
Regime type of exporter	0.295	0.0998	0.0080
Regime type of importer	0.231	0.079	0.0080
Similarity of regime types	0.088	–0.0022	0.00090
Cooperation in conflict	–0.110	1.181	0.604

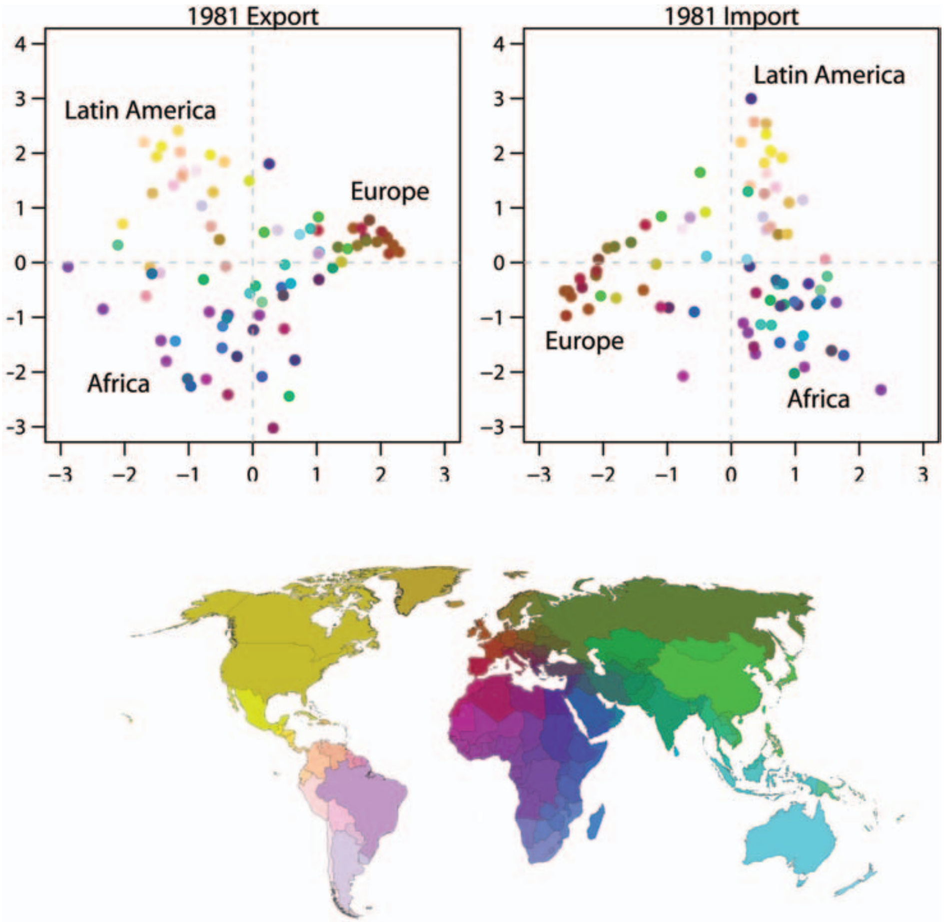
Residual standard error is 4.811 on 6634 degrees of freedom; $R^2 = 0.57$.

Figure 1 Residual Exporter and Importer Effects Not Negligible, as Assumed by Standard Gravity Models.



In the top two panels, the exporter and importer effects are scaled and arrayed from low to high. A normal distribution would fall along the black line in each panel. The third panel shows a color key for each of the observations in these plots, based on geography. These graphics illustrate the magnitude of the exporter and importer and importer residual effects.

Figure 2 Residual Multiplicative (Inner Product) Effects.



Effects are portrayed in two dimensions, with the left panel presenting exports and the right imports. These data show the remaining latent structure of international trade in 1981.

model specified by Equation (2) and given their covariates.[4]

Are these effects large compared with the error in the model? If the residuals $\epsilon_{i,j}$ were really uncorrelated, as standard analyses of gravity models typically presume, then the distribution of these effects should be roughly normally distributed with mean zero and variance approximately equal to $\hat{\sigma}^2/(n-1) = 4.811^2/80 = 0.29$. If this were the case, we would expect to see only about 6% of the estimated effects outside of –1 and 1. However, we observe more than 50% of the effects outside of this bound, indicating there is a substantial amount of structure in the data that is unexplained by the standard gravity formulation.

Finally, dyadic data such as these often display higher-order patterns (Hoff, Raftery & Handcock, 2002; Hoff & Ward, 2004). Essentially, if country A trades a lot with country B, and country B trades a lot with country C, then it is quite likely that countries

[4] Pinheiro & Bates (2000) and Snijders & Bosker (1999) describe multilevel, mixed-effects models.

A and C will also be frequent trading partners. Some aspects of these kinds of patterns can be represented with latent, multiplicative effects (Hoff, 2005) that are similar to bi-plots (Gabriel, 1971, 1978; Bradu & Gabriel, 1978) and other matrix representation methods. We represent the residuals from Equation (3) as equal to the inner product of exporter- and importer-specific vectors of latent attributes, plus a disturbance term:

$$\hat{\gamma}_{i,j} = u_i' v_j + \delta_{i,j} \quad (4)$$

The inner product $u_i'v_j$ can be thought of as a reduced-rank interaction term, decomposed into exporter (u_i) and importer (v_j) components. This inner product is simply a multiplicative term that captures these dependencies. Countries that trade heavily with each other will have similar inner product terms. For now, we fit u_i and v_j as each being of length two. A method for selecting the dimension of these vectors is discussed in Hoff (2005), and details for obtaining least-squares estimates of u_i and v_j are discussed in Appendix B. These multiplicative effects are presented graphically in Figure 2, where the plotting colors for the countries are based on geographic locations, with nearby countries being of similar color.

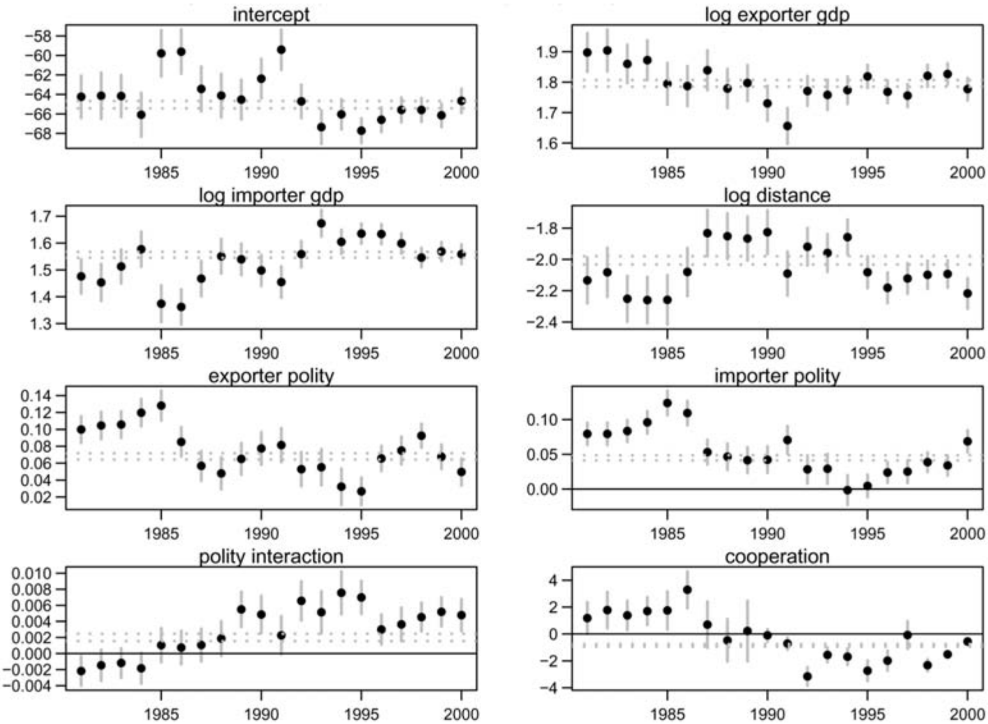
The interpretation of the fitted values of u_i and v_j is that if country i exports more to country j than predicted by the gravity model and after controlling for the additive importer and exporter effects in Equation (3), then the vectors u_i and v_j will have a similar direction. For example, note that many European countries are in the middle right-hand side of the plot of u (exporter) vectors, whereas they are on the opposite side of the v (importer) plot. This means that the residuals are negative, which, in turn, suggests that the gravity model drastically over-predicts the amount of actual trade among this set of countries, since the residuals are predominantly negative. This phenomenon can be confirmed by observing that the residuals from Equation (3) between pairs of OECD countries tend to be negative. Overall, these inner-product representations also show considerable clustering of countries by geography. In particular, European countries strongly tend to group together. To a lesser extent, so do African countries and South American countries. This is remarkable, especially given that geographic distance is already included explicitly in the gravity model.

Persistence of Effects and Change over Time

One goal of statistical inference is to make general conclusions based on specific data. With this in mind, it would be desirable to know which patterns in the 1981 trade data generalize to other years and which patterns change substantially over time. We initially examine the change over time by fitting Equation (2) to trade data from 1981 to 2000 for each year separately.[5] Year-specific least-squares regression parameters are shown in Figure 3. A dot illustrates the point estimate of the parameter, while the gray vertical line denotes the nominal 95% intervals based on ordinary least-squares regression. The gray band of dotted lines in each panel indicates the nominal 95% confidence interval of the pooled estimate, based on ordinary least-squares regression. This graphic illustrates that parameter estimates and confidence intervals obtained from any given

[5] All of our estimates and forecasts are undertaken in terms of dollars current in the year in which estimates/forecasts are made, not in terms of constant dollars for the trade volumes in the year to which the forecast applies. As a result, our forecast trade is likely on average to underestimate actual future trade by the average amount of inflation from one year to the next. This imparts some conservative bias into our forecasts. We choose to stay with this approach, since we do not wish to incorporate or develop a global model of inflation; nor do we believe that aggregate trade will be undertaken with a precise model of future inflationary pressures. Over the period studied, US inflation, for example, ranges from about 14% in 1980 to approximately 1.5% in 1998. Ignoring inflation should make our model perform less well than if forecasts were made with complete knowledge of inflationary pressures, as would be implied by converting forecasts into constant dollars. We also estimated a pooled model that is presented visually below.

Figure 3 Yearly Least-Squares Fits of the Gravity Model to Trade Data.



yearly slice are inconsistent with the pooled data. Conversely, the graphic shows that confidence intervals obtained from the pooled data will fail to represent the year-to-year variation in the parameter estimates.

Note that the variables in the standard gravity model (distance and GDP) all fluctuate around values that are far from zero, compared with their variance across years. In contrast, the coefficients for cooperation fluctuate near zero, even changing sign over the course of the 20-year period. The figure indicates substantial correlation of regression coefficients over time, also indicated in Table II. This correlation is not too surprising, as we would expect the mechanisms underlying trade to not change too drastically from one year to the next.

Interpretation of such estimates and confidence intervals has typically not been descriptive, but rather has provided the basis for statistical inference and tests of hypotheses. For example, the nominal confidence intervals from the 1981 analysis suggests that all explanatory variables are significantly different from zero at level $\alpha = 0.05$. Traditionally, such a result would be interpreted as meaning that the non-zero relationships between trade and explanatory variables in 1981 are generalizable to some larger super-population of trade data. However, difference between the 1981 estimates and the others calls into question to what putative super-population the significant results from any given year generalize.

As we see it, there are at least two additional problems with the analysis of trade data using a hypothesis-testing approach applied to ordinary least-squares regression models. One is that the analysis assumes that all the observations are independent, an assumption that cannot in principle be correct, since individual countries are included many times in the data, both as exporters and as importers. The second is that these analyses generally

Table II. First-Order Serial Correlations of Regression Coefficients and Exporter and Importer Effects Illustrating Persistence of Modeled Components in Gravity Model of International Trade, 1981–2000.

Parameter	*Lag correlation*
Intercept	0.47
Exporter GDP	0.59
Importer GDP	0.59
Distance	0.61
Exporter polity	0.73
Importer polity	0.77
Polity interaction	0.78
Cooperation	0.72
Additive exporter effect	0.95
Additive importer effect	0.79
First multiplicative exporter effect	0.75
Second multiplicative exporter effect	0.71
First multiplicative importer effect	0.71
Second multiplicative importer effect	0.69

use all available data in a given year or set of years, calling into question any measure of parameter salience based on statistical sampling theory: the theory behind a confidence interval is that, in repeated experiments or random samples from a population, a 95% confidence interval will cover the true parameter value 95% of the time. In contrast, analyses such as the ones described above use the complete data from each year under study, and so, in a sense, the parameter estimates provide the 'true' measure of the linear relationship between trade and the explanatory variables within a year. As such, the value of a regression coefficient is never actually zero, and regression coefficients from two different years are never exactly the same. With this in mind, hypothesis tests of the form $H_0 : \beta = 0$ or $H_0 : \beta_{1980} = \beta_{1981}$ are useful neither as measures of the salience of certain variables nor as measures of the amount of change in them over time. To quote Savage (1957: 332), 'null hypotheses of no difference are usually known to be false before the data are collected; when they are, the rejection or acceptance simply reflects the size of the sample and the power of the test, and is not a contribution to science'. In the analysis of a single year's data with 100 countries, there will be about 10,000 observations, and so it is not surprising that using a hypothesis-testing approach leads to the declaration that almost all of the estimated parameters are statistically significant.

Out-of-Sample Predictive Persistence

We suggest an alternative measure of the explanatory power and temporal persistence of a variable or set of variables, based on out-of-sample predictive performance. Measuring such performance is a standard method of model validation, in which one subset of a database (the training set) is used to estimate parameters in a model, and a separate subset (the test set) is used to evaluate the model. Roughly speaking, a model is 'validated' if the model estimated using the training set provides reasonable predicted values for the test set. Such an approach is also often used for variable selection in regression models. Suppose we want to decide whether or not to choose between two regression models: Model 1: $E[y_i] = \beta_1 x_{1,i}$ and Model 2: $E[y_i] = \beta_1 x_{1,i} + \beta_2 x_{2,i}$. In effect, we are trying to decide whether or not to include x_2 as an explanatory variable. Using only observations from the training set, we can obtain regression coefficients $\hat{\beta}_{(1),1}$ via Model 1 and $\hat{\beta}_{(2),1}$, $\hat{\beta}_{(2),2}$ via Model 2. The two models can then be evaluated in terms of how well they predict the test data:

$$\mathrm{R}(M1, M2) = \frac{\mathrm{Error}(M1)}{\mathrm{Error}(M2)} = \frac{\Sigma_{i \in test}(y_i - \hat{\beta}_{(1),1} x_{i,1})^2}{\Sigma_{i \in test}(y_i - [\hat{\beta}_{(2),1} x_{i,1} + \hat{\beta}_{(2),2} x_{i,2}])^2}.$$

If the sum were taken over the training set, then the ratio would always be greater than 1, as including an additional term in the model always improves the residual sum-of-squares. In contrast, by summing over the test data, if x_2 offers little explanatory power, or the relationship between x_2 and y is substantially different in the test and training data, then the above statistic could be less than 1. This method of comparison is valid regardless of the number of parameters in the models being compared.

In many applications, the training set is a random sample from an available database and the test set is its complement. In this case, the goal (besides prediction) is, in some sense, to identify a model that generalizes across random subsets of the data. In contrast, our goal is to identify patterns in trade that are generalizable across time, so we modify the cross-validation approach to this end. Given two models of trade, $M1$ and $M2$, let $\hat{\beta}^t_{(1)}$ and $\hat{\beta}^t_{(2)}$ be vectors of the corresponding regression coefficients estimated using data from year t. Let $R_{t,t+1}(M1,M2)$ be the ratio of the sum-of-squared error in using $\hat{\beta}^t_{(1)}$ to predict data from year $t + 1$ relative to the error in using $\hat{\beta}^t_{(2)}$. Table III shows the use of this statistic in assessing the explanatory power of the variables in the expanded gravity model (1).

In each year t, we fit the full gravity model of trade (1) as well as four reduced models, each lacking one of the log GDP variables, log distance, the polity variables, or cooperation. For each reduced model, regression coefficients based on data from year t are then used to make predictions for the trade in year $t + 1$ (using the x values from year $t + 1$), and the sum-of-squared prediction error is compared to that of the full model.

Not surprisingly, GDP and distance are both important, based on this measure, with the GDP variables typically halving the one-

Table III. Predictive Performance of Last Year's Gravity Model, and Relative Performance Attributable to Measures of GDP, Geographic Distance, Regime Characteristics of Exporters and Importers (Polity Main Effects and Interaction) and Level of Cooperative and Conflictual Interactions Among Trading Partners.

Year	*Full model predictive* R^2	*GDP*	*Distance*	*Polity*	*Cooperation*
1982	0.525	1.738	1.108	1.040	1.001
1983	0.567	1.837	1.136	1.044	1.001
1984	0.563	1.845	1.143	1.054	1.001
1985	0.515	1.656	1.129	1.069	1.001
1986	0.512	1.677	1.110	1.038	1.003
1987	0.567	1.906	1.105	1.009	0.998
1988	0.580	1.914	1.112	1.013	1.000
1989	0.594	1.978	1.107	1.024	1.000
1990	0.557	1.863	1.128	1.028	0.999
1991	0.538	1.696	1.111	1.023	1.000
1992	0.591	2.115	1.119	1.018	1.005
1993	0.593	2.091	1.119	1.016	1.000
1994	0.597	2.168	1.117	1.013	1.003
1995	0.609	2.268	1.121	1.013	1.004
1996	0.606	2.182	1.141	1.012	1.002
1997	0.596	2.124	1.135	1.017	0.999
1998	0.620	2.179	1.149	1.025	1.000
1999	0.645	2.284	1.159	1.022	1.007
2000	0.600	2.069	1.138	1.020	0.987

year-ahead predictive sum-of-squared error. The results for polity and cooperation are more subtle: the results indicate that a model including regime-type variables always gives better one-year-ahead predictions than the standard gravity model. This suggests that the regime-type variables have predictive power and that the relationship between regime type and trade is not changing too quickly from year to year. In contrast, adding cooperation as an explanatory variable adds no predictive benefit in nine of the 19 comparisons made (out to three decimal places) and even *decreases* the one-year-ahead predictive performance of the gravity model in four of the comparisons. The fact that the ratios for these latter four years are only slightly lower than 1.0 is of little solace: in these four cases, inclusion of cooperation in the model gives worse predictions. This suggests that, for these pairs of consecutive years, the relationship between cooperation and trade, in the presence of the other variables, is not generalizable from one year to the next.[6]

To get an idea of how much a variable adds to one-year-out predictive performance 'on average', we can simply take the mean of each column in Table III. To assess how fast such a pattern changes over time, we can calculate $R(M1,M2)$ as above to see how well a model fit with data from year t predicts data from year $t + k$. We compute this ratio for each variable, each year t, and each possible lag in years k, and average the results over t. The result is a measure of the relative predictive performance of a model with and without a particular variable, in terms of predicting k years out. The results are shown in Figure 4.

On average, gravity models with the effects of regime type are better than gravity models without, in terms of predicting up to about eight years in the future. The effect of cooperation is small and not consistent enough over time to have a positive predictive effect. Indeed, even if the coefficient for cooperation were nominally significant in a given year, it is likely that removing it from the regression would give a more generalizable model, even just a few years out. Polity effects persist in improving out-of-sample predictions for almost a decade.

Persistence of Latent Structure

As shown above, dyadic trade data often exhibit substantial residual structure that can be represented by additive exporter- and importer-specific effects, as well as higher-order latent structure that can be analyzed via multiplicative effects. We evaluate the magnitude of these effects and their persistence over time using the approach outlined above. For each year t, we decompose the residuals into the additive exporter and importer effects described in Equation (3), as well as the multiplicative effects of Equation (4), setting the dimension of the latent multiplicative vectors to 3. These effects are then used to make predictions for trade in years $s : s \neq t$. Specifically, when assessing the persistence of the exporter and importer effects, we compute a set of predicted values for year s, based on Equation (1) as estimated from data in year t:

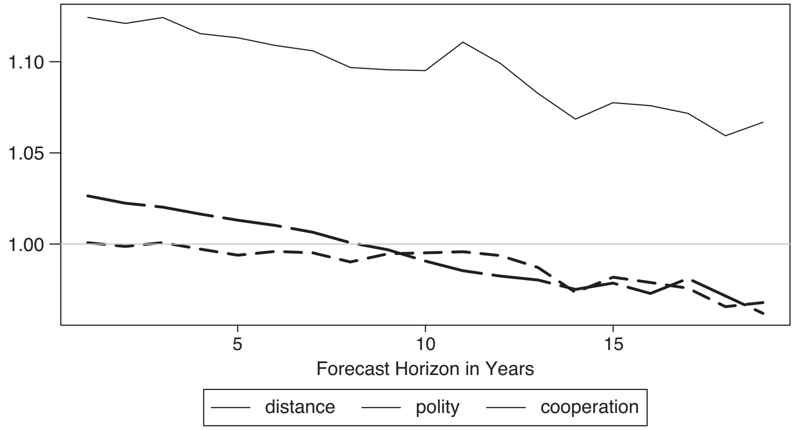
$$\begin{aligned}\ln \hat{F}_{i,j,s} = \hat{\beta}_0 &+ \hat{\beta}_1 \ln G_i + \hat{\beta}_2 \ln G_j \\ &+ \hat{\beta}_3 \ln D_{i,j} + \hat{\beta}_4 P_i + \hat{\beta}_5 P_j \\ &+ \hat{\beta}_6 P_i \times P_j + \hat{\beta}_7 C_{i,j} + \hat{a}_i \\ &+ \hat{b}_j + \hat{u}_i' \hat{v}_j,\end{aligned}$$

with $\hat{\beta}, \hat{a}_i, \hat{b}_j, \hat{u}_i, \hat{v}_j$, all being estimated from year t and the covariates being from years. The fit of these predicted values is then compared with that of the predicted values made from fitting the baseline gravity model of trade:

$$\begin{aligned}\ln F_{i,j} = \hat{\beta}_0 &+ \hat{\beta}_1 \ln G_i + \hat{\beta}_2 \ln G_j \\ &+ \hat{\beta}_3 \ln D_{i,j} + \epsilon_{i,j}\end{aligned} \tag{5}$$

[6] A similar forward selection type of approach could also be used to assess the temporal stability/predictive performance of each variable; the results give a similar ordering of the importance of the above variables.

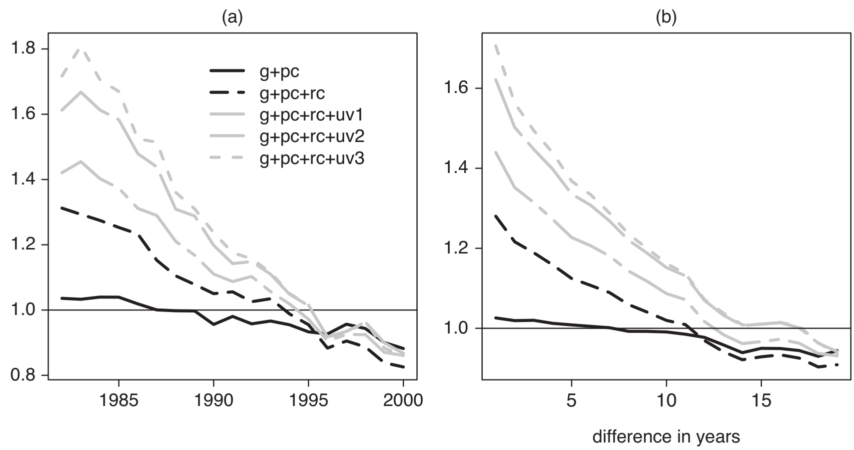
Figure 4 Average *k*-Year-Out Relative Predictive Performances of Gravity Models with and without Distance, Polity Effects, and Cooperation.



GDP and Distance effects persist in making out-of-sample predictions for more than two decades into the future. Polity effects persists in helping out-of-sample predictions for about eight years. On the other hand, the effects of cooperation in conflict are never appreciable and often actually hurt the out-of-sample prediction.

The first panel in Figure 5 shows how well the various models fit, in terms of their ability to do better than a baseline gravity model. Models fit with data from year 1981 are used to predict data in each subsequent year, 1982–2000. These fits are compared with the baseline gravity model in Equation (5). The second panel averages these curves over years to provide a measure of the average *k*-year-out predictive performance of the

Figure 5 Predictive Persistence of Different Model Components over 20 Years.

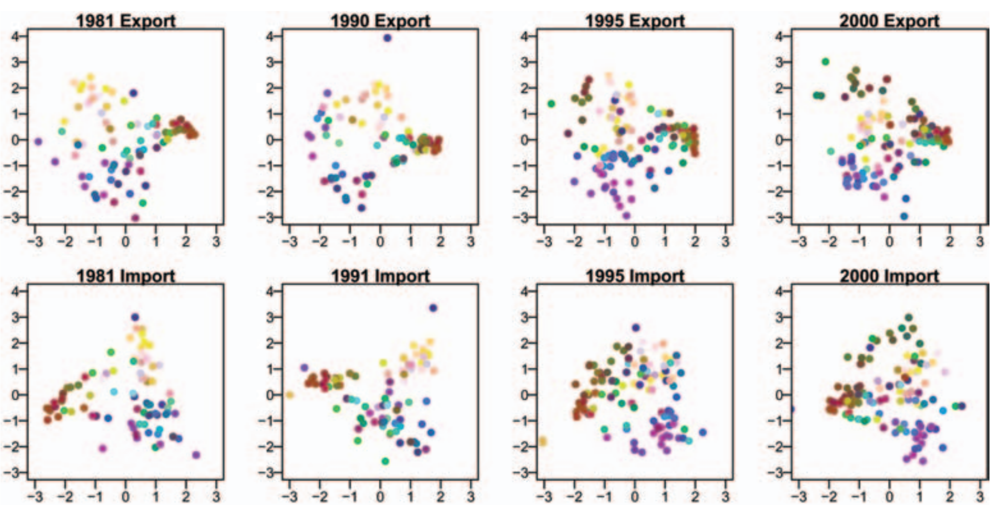


The solid black line, labeled g + pc, portays the standard gravity model plus the polity and cooperation variables. The solid dashed line (g + pc + rc) adds the additive exporter and importer effects. The three dashed gray lines illustrate the performance gained by adding the multiplicative latent effects to the model (g + pc + rc + uv1,2,3). As is clear from the graphic, the multiplicative effects provide a substantial benefit for over 10 years. Values above 1.0 illustrate better predictions, and values below 1.0 illustrate worse predictions.

various models. The black line gives the relative performance of the expanded gravity model that includes both the polity and cooperation variables. This model (denoted g + pc) improves predictions several years out, but the improvement is quite small compared with the improvement offered by the latent structure (shown in the gray lines). Additionally, note that the addition of the latent structure offers substantial predictive improvement many years out, indicating that the latent structure is real and persistent. What is most striking is that the gray lines showing the effects of adding the latent dimensions to the prediction generally fall above the predictive performance of the standard gravity specifications. This suggests that the latent structures of international trade persist and continue to aid out-of-sample prediction for well over a decade. If the residuals within a year were indeed pattern-less and not persistent from year to year, then the gray lines in the plot should all be near or below 1.

Within a given year, as the dimension of the multiplicative effects increases, the estimated latent structure $\hat{a}_i + \hat{b}_j + \hat{u}_i'\hat{v}_j$, becomes a closer approximation to the sociomatrix of residuals $\hat{\epsilon}_{i,j}$ for that year. One might argue that, if trade patterns are not changing that quickly, the best predictive model obtainable from a given year will be based on the regression parameters estimated that year and the residuals from that year, that is, representing the latent structure with the residuals. Indeed, such a model typically gives the 'best performance' one or two years out, but, after about two years, the simpler models having two- or three-dimensional multiplicative effects offer better predictions. This indicates that the sociomatrix of residuals from a given year contains some structure that is persistent over time (represented with the additive and multiplicative country-specific effects) in addition to 'noise', or non-persistent structure. For example, the country-specific effects estimated in 1981, 1990, 1995, and 2000 (as displayed in Figure 6) show some similar structure, even though they are based on data observed once a decade. Among other important effects, these latent dimensions capture the dependencies among European trade

Figure 6 Estimated Additive and Multiplicative Exporter- and Importer-Specific Effects for 1981, 1990, 1995, and 2000 Using the Same Color Scheme as Figure 2.



that are underestimated by standard gravity approaches.

Conclusion

The standard representation of bilateral international trade is the gravity model. In empirical tests, about half the variance in within-sample trade can be explained by this simple model that focuses on the sizes of the importing and exporting countries, as well as the distance between them. Many scholars have added to this standard model variables capturing domestic as well as international forces. The power of these additions has been typically gauged using traditional statistical sampling theory. Almost universally, the additions have been found to be statistically significant, suggesting the results are generalizable beyond the data from which they are estimated.

We have discussed why such an approach can be misleading, and we have developed and applied a framework to combat this tendency to affirm all new hypotheses about international commerce. Standard hypothesis-testing techniques based on ordinary linear-regression models are not appropriate for assessing the generalizability of the parameter estimates. We have proposed evaluating generalizability based on a simple measure of out-of-sample predictive performance, and we have used this metric to examine some empirical claims found in the extant literature. We found that the domestic political framework (polity) of the exporter and importer are important, persistent factors relating to trade. We also found that the relationship between trade and international conflict (specifically, cooperation in international disputes) is small and not reliably generalizable from one year to the next.

In addition, we have shown that a substantial amount of trade variation is attributable to country-specific exporter and importer effects, effects typically assumed to be nonexistent. These effects indicate strong geographic patterns unexplained by the gravity model, most recognizably, the over-prediction of trade among OECD countries. Finally, the variation in trade attributable to exporter- and importer-specific effects is large and persistent over time. The estimation and reporting of such effects thus gives a more complete picture of patterns in international trade.

Apart from the development of a new methodology for (re)examining empirical results in the arena of dyadic models of international relations, what are the implications of these findings? First and foremost, bilateral trade is determined mostly by the strength of the large, vibrant economies, as suggested by macroeconomic theory. Democratization is certainly associated with increased trade, even in just one of the trading nations. Increased democratization will serve to expand bilateral trade at a global level, as well as within pairs of trading nations. These effects are important but are overwhelmed by the impact on economic productivity as a result of international commerce. For example, recent work suggests that even small increases in national productivity owing to increased trade will far outstrip the potential economic impact of any politically plausible level of foreign aid. Cline (2004: 1) estimates that 'global free trade would confer income gains of at least about $90 billion annually in developing countries ... and total long-term gains including dynamic effects would be about $200 billion annually'. The 9% growth in Chinese GDP in 2004, coupled with its 25% growth in exports during the same period, comes to mind. Thus, one strong benefit of increasing democratization is its direct impact on international commerce and the subsequent impetus that brings to national productivity.

Beyond this, there has been considerable discussion of the importance of international politics on commerce. Much of this discussion is a consequence of variable selection and over-fitting. By adopting an approach to look at the persistence of latent dependencies in international commerce and employing out-of-sample heuristics, we find that, first and foremost, the standard gravity model that is widely employed is exceptionally useful as a theoretical and practical tool. At the same time, because it ignores important dependencies that characterize international commerce, it produces results that are consistently biased. In particular, it overestimates the trade among rich, industrialized Western countries. Second, we find that regime characteristics of trading states are quite important in predicting trade flows among them. The more democratic a country, the more it exports; similarly, democratic orientation tends to enhance imports as well. Moreover, there is a substantial dyadic effect, such that pairs of highly democratic countries have considerably higher trade flows with one another, even after accounting for their economic size and their geographic distance from one another. International conflict does not bear much of an independent relationship to international commerce, however. Indeed, we found compelling evidence that trying to use the conflictual and cooperative behavior of nations in militarized interstate disputes to explain trade can lead to worse, not better, predictions. This does not mean that conflict is good for commerce; neither does it suggest that cooperation is bad for commerce. Our results are entirely in line with the suggestion that, under some circumstances, the impact of conflict on trade is indeterminate (Morrow, 1999) or highly conditional (Reuveny & Kang, 2003). There are many reasons to promote a peaceful world, but there is little evidence that increasing trade is one of them.

A better understanding of the enabling and constraining conditions of international commerce has important policy consequences. We have shown that global trade is highly structured and interdependent, even after taking standard macroeconomic and macro-political forces into account. As such, it is important to analyze trade flows in a way that recognizes the persistence of deep interdependencies in the global system of trade. It seems evident that the deep, latent structure of the international trade network is likely to remain an important force in the ebb and flow of daily international commerce.

Appendix A

Software and Data Documentation

Software All analyses were performed with *R* (2004). Codes to replicate the analyses herein are available on the replication website for the journal, along with a copy of the data.

Trade Data Trade data were taken from the UN Commodity Trade Statistics Database (UN Comtrade), available online via http://unstats.un.org/unsd/comtrade/. Data are from the Harmonized Commodity Description and Coding System (HS2002) classification and are annual bilateral exports and imports given in current US dollars.

Distance Distance was calculated using the Haversine formula with data on latitude and longitude of capital cities taken from the world.cities database maintained as part of the maps package in the *R* statistical programming package. These are available from cran.r-project.org. Distance was calculated in thousands of kilometers.

GDP Data on the annual Gross Domestic Product in current US dollars was taken from the World Bank's World Development Indicators, available online at http://devdata.worldbank.org/dataonline/.

Polity We use the annualized polity score, which ranges from −10 for highly authoritarian states to +10 for highly democratic societies, to gauge the domestic institutions in each country. These data are available from http://www.cidcm.umd.edu/inscr/polity/, with registration. Some of these data were updated with information on so-called 'micro-states' from the separate database maintained by Gleditsch (2003).

Militarized Interstate Disputes We employed the Militarized Interstate Disputes database (version 3.02), which is maintained by Ghosn & Bennett (2003). Our measure of international interactions is coded −1 for each militarized interstate dispute in which country i is on the opposite side of the dispute from country j; similarly, if i and j are on the same side in an interstate dispute, this is scored +1. We sum dispute scores for each pair of countries if there is more than one ongoing dispute in a given year.

Estimating Multiplicative Effects

The rank k least-squares approximation to an $n \times n$ matrix Γ can be obtained by the singular value decomposition of Γ (Householder & Young, 1938). This provides $n \times k$ matrices U and V such that $\hat{A} = UV'$ minimizes $||\Gamma - A||^2$ among all rank k matrices A. In the context of dyadic data, where the diagonal of Γ is undefined, the computational problem can be written in terms of weighted least-squares:

$$\text{minimize} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j}(\gamma_{i,j} - u_i' v_j)^2$$

subject to $u_i, v_j \in \mathrm{R}^k$, $w_{i,i} = 0$ and $w_{i,j} = 1$ if $i \neq j$. This problem has been considered by Gabriel & Zamir (1979), who provide an iterative 'criss-cross' regression technique for finding an approximate solution. We have found that combining the regression technique with a singular value decomposition gives an algorithm that converges much more quickly. The algorithm we employ is as follows:

Given current values of $U = \{u_1,\ldots,u_n\}$ and $V = \{v_1,\ldots,v_n\}$, iterate the following until convergence:

(1) for $i = 1,\ldots,n$:
  (a) given the v-vectors, update u_i as the vector of regression coefficients from $\gamma_{i,j} = u_i' v_j + \delta_{i,j}$, $j \neq i$.
  (b) given the u-vectors, update v_i by the vector of regression coefficients from $\gamma_{j,i} = v_i' u_j + \delta_{j,i}$, $j \neq i$.

(2) Obtain the $n \times n$ matrix $\hat{\Gamma}$ from Γ and having diagonal given by $\gamma_{i,i} = u_i' v_i$. Obtain new values of U and V from the singular value decomposition of $\hat{\Gamma}$.

R-code for obtaining U and V can be obtained at http://www.stat.washington.edu/hoff/research.html; current code is also available in the publication archive for the journal.

References

Anderton, Charles & John Carter, 2001a. 'The Impact of War on Trade: An Interrupted Times-Series Study', *Journal of Peace Research* 38(4): 445–457.

Anderton, Charles & John R. Carter, 2001b. 'On Disruption of Trade by War: A Reply to Barbieri and Levy', *Journal of Peace Research* 38(5): 625–628.

Barbieri, Katherine, 1996. 'Economic Interdependence: A Path to Peace or Source of Interstate Conflict?', *Journal of Peace Research* 33(1): 29–49.

Barbieri, Katherine & Jack S. Levy, 1999. 'Sleeping with the Enemy: The Impact of War on Trade', *Journal of Peace Research* 36(4): 463–479.

Barbieri, Katherine & Jack S. Levy, 2001. 'Does War Impede Trade? A Response to Anderton and Carter', *Journal of Peace Research* 38(5): 619–624.

Barbieri, Katherine & Gerald Schneider, 1999. 'Globalization and Peace: Assessing New Directions in the Study of Trade and Conflict', *Journal of Peace Research* 36(4): 387–404.

Bliss, Harry & Bruce M. Russett, 1998. 'Democratic Trading Partners: The Liberal Connection, 1962–1989', *Journal of Politics* 60(4): 1126–1147.

Bradu, Daniel & Kuno Ruben Gabriel, 1978. 'The Biplot as a Diagnosis Tool for Models of Two-Way Tables', *Technometrics* 20(1): 47–68.

Cline, William, 2004. *Trade Policy and Global Poverty*. Washington, DC: Center for Global Development & Institute for International Economics.

de Groot, Henri L. F.; Gert-Jan Linders, Piet Rietveld & Uma Subramanian, 2004. 'The Institutional Determinants of Bilateral Trade Patterns', *Kyklos* 57(2): 103–124.

Dorussen, Han, 1999. 'Balance of Power Revisited: A Multi-Country Model of Trade and Conflict', *Journal of Peace Research* 36(4): 443–462.

Dorussen, Han, 2002. 'Trade and Conflict in Multi-Country Models: A Rejoinder', *Journal of Peace Research* 39(1): 115–118.

Feenstra, Robert C.; Andrew K. Rose & James R. Markusen, 2001. 'Using the Gravity Model to Differentiate Among Alternative Theories of Trade', *Canadian Journal of Economics* 34(2): 430–447.

Gabriel, Kuno Ruben, 1971. 'The Biplot-Graphic Display of Matrices with Application to Principal Component Analysis', *Biometrika* 58(3): 453–467.

Gabriel, Kuno Ruben, 1978. 'Least Squares Approximation of Matrices by Additive and Multiplicative Models', *Journal of the Royal Statistical Society. Series, B. Methodological* 40(2): 186–196.

Gabriel, Kuno Ruben & Shmuel Zamir, 1979. 'Lower Rank Approximations of Matrices by Least Squares with Any Choice of Weights', *Technometrics* 21(4): 489–467.

Gartzke, Erik & Quan Li, 2003. 'Measure for Measure: Concept Operationalization and the Trade Interdependence–Conflict Debate', *Journal of Peace Research* 40(5): 553–571.

Ghosn, Faten & Scott Bennett, 2003. 'Codebook for the Dyadic Militarized Interstate Incident Data' (http://cow2.la.psu.edu).

Gleditsch, Kristian S., 2003. 'Modified Polity P4 and P4D Data, Version 1.0' (http://privatewww.essex.ac.uk/ksg/polity.html).

Goenner, Cullen F., 2004. 'Uncertainty of the Liberal Peace', *Journal of Peace Research* 41(5): 589–605.

Hegre, Håvard, 2000. 'Development and the Liberal Peace: What Does It Take To Be a Trading State?', *Journal of Peace Research* 37(1): 5–30.

Hegre, Håvard, 2002. 'Trade Decreases Conflict More in Multi-Actor Systems: A Comment on Dorussen', *Journal of Peace Research* 39(1): 109–114.

Hoff, Peter D., 2005. 'Bilinear Mixed Effects Models for Dyadic Data', *Journal of the American Statistical Association* 100(469): 286–295.

Hoff, Peter D. & Michael D. Ward, 2004. 'Modeling Dependencies in International Relations Networks', *Political Analysis* 12(2): 160–175.

Hoff, Peter D.; Adrian E. Raftery & Mark S. Handcock, 2002. 'Latent Space Approaches to Social Network Analysis', *Journal of the American Statistical Association* 97(460): 1090–1098.

Householder, Alston Scott & Gale Young, 1938. 'Matrix Approximations and Latent Roots', *American Mathematical Monthly* 45(3): 165–171.

Keshk, Omar M. G.; Brian M. Pollins & Rafael Reuveny, 2004. 'Trade Still Follows the Flag: The Primacy of Politics in a Simultaneous Model of Interdependence and Armed Conflict', *Journal of Politics* 66(4): 1153–1179.

Long, Andrew G., 2003. 'Defense Pacts and International Trade', *Journal of Peace Research* 40(5): 537–552.

McCallum, John, 1995. 'National Borders Matter: Canada–U.S. Regional Trade Patterns', *American Economic Review* 85(3): 615–623.

Mansfield, Edward; Helen V. Milner & B. Peter Rosendorff, 2000. 'Free To Trade? Democracies, Autocracies, and International Trade Negotiations', *American Political Science Review* 94(2): 305–321.

Morrow, James D., 1999a. 'How Could Trade Affect Conflict?', *Journal of Peace Research* 36(4): 481–489.

Morrow, James D.; Randolph M. Siverson & Tressa E. Tabares, 1999. 'Correction to: "The Political Determinants of International Trade"', *American Political Science Review* 94(4): 931–933.

North, Douglass C., 1990. *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press.

Oneal, John R. & Bruce M. Russett, 1999. 'Assessing the Liberal Peace with Alternative Specifications: Trade Still Reduces Conflict', *Journal of Peace Research* 36(4): 423–442.

Pinheiro, José C. & Douglas M. Bates, 2000. *Mixed-Effects Models in S and S-Plus*. Berlin: Springer.

Polachek, Solomon W.; John Robst & Yuan-Ching Chang, 1999. 'Liberalism and Interdependence: Extending the Trade–Conflict Model', *Journal of Peace Research* 36(4): 405–422.

R Development Core Team, 2004. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing (http://www.R-project.org).

Reuveny, Rafael & Heejoon Kang, 2003. 'A Simultaneous-Equations Model of Trade, Conflict, and Cooperation', *Review of International Economics* 11(2): 279–295.

Rose, Andrew K., 2004a. 'Do WTO Members Have More Liberal Trade Policy?', *Journal of International Economics* 63(2): 209–235.

Rose, Andrew K., 2004b. 'Does the WTO Really Increase Trade?', *American Economic Review* 94(1): 98–114.

Rose, Andrew K., 2005a. 'Does the WTO Make Trade More Stable?', *Open Economies Review* 17(1): 7–22.

Rose, Andrew K., 2005b. 'Which International Institutions Promote International Trade?', *Review of International Economics* 13(4): 682–698.

Rose, Andrew K., 2006. 'The Foreign Service and Foreign Trade: Embassies as Export Promotion', manuscript, University of California, Berkeley (http://www.nber.org/papers/w11111).

Savage, I. Richard, 1957. 'Nonparametric Statistics', *Journal of the American Statistical Association* 52(279): 331–344.

Snijders, Tom A. B. & Roel J. Bosker, 1999. *Multilevel Analysis: An Introduction to Basic and Advanced Modeling*. London: Sage.

Williamson, Oliver E., 1985. *The Economic Institutions of Capitalism*. New York: Free Press.

MICHAEL D. WARD, b. 1948, PhD in Political Science (Northwestern University, 1977); Professor, Department of Political Science, University of Washington (1997–). Current main interest: dynamic analysis of international networks.

PETER D. HOFF, b. 1971, PhD in Statistics (University of Wisconsin, 2000); Assistant Professor, Departments of Statistics, Biostatistics & Center for Statistics and the Social Sciences, University of Washington (2006–). Current main interest: latent variable methods for models of relational data.