Bayesian Inference for the Gravity Model

Author(s): Priya Ranjan and Justin L. Tobias

Source: *Journal of Applied Econometrics*, Vol. 22, No. 4 (Jun. - Jul., 2007), pp. 817-838
Published by: Wiley
Stable URL: http://www.jstor.org/stable/25146547

JOURNAL OF APPLIED ECONOMETRICS
J. Appl. Econ. **22**: 817–838 (2007)
Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/jae.926

BAYESIAN INFERENCE FOR THE GRAVITY MODEL

PRIYA RANJAN[a] AND JUSTIN L. TOBIAS[b]*

[a] *Department of Economics, University of California–Irvine, USA*
[b] *Department of Economics, Iowa State University, Ames, Iowa, USA*

SUMMARY

This paper seeks to empirically extend the *gravity model*, which has been widely used to analyze volumes of trade between pairs of countries. We generalize the basic threshold tobit model by allowing for the inclusion of country-specific effects into the analysis and also show how one can explore the relationship between trade volumes and a given covariate via a non-parametric approach. We use our derived methodology to investigate the impact of a particular aspect of institutions—the enforcement of contracts—on bilateral trade. We find that contract enforcement matters in predicting trade volumes for all types of goods, that it matters most for the trade of differentiated goods, and that the relationship between contract enforcement and trade in our threshold tobit exhibits some nonlinearities. Copyright © 2007 John Wiley & Sons, Ltd.

Received 4 October 2005; Revised 5 January 2006

1. INTRODUCTION

Empirical work modeling volumes of trade between pairs of countries commonly makes use of the *gravity equation*. In its basic form, the gravity equation writes the volume of exchange as a function of the sizes of the trading countries as well as the distance separating them. Recent empirical work on international trade, however, suggests that trade barriers not only include directly measured transport costs (which a distance measure would intend to capture), and government-imposed non-tariff barriers (NTBs), but also includes other factors, such as contracting costs.[1] These contracting costs are a form of transactions costs, and can be affected, among other factors, by the trading countries sharing a common language or possessing colonial ties.[2]

This paper continues in the tradition of examining the effect of trade costs on trade volumes by exploring the relationship between a potential type of contracting cost—country-specific measures of contract enforcement—on the volumes of trade in different types of goods. North (1990), for example, describes the importance of reliable institutions and contract enforcement as determinants of international trade as he writes: 'The greater the specialization and the number and variability of valuable attributes, the more weight must be put on reliable institutions that allow individuals to engage in complex contracting with a minimum of uncertainty about whether the terms of the contract can be realized.'

The foregoing quote not only suggests that contract enforcement should matter in international trade relationships, but that it should matter most for the trade of heterogeneous or differentiated

* Correspondence to: Justin L. Tobias, Department of Economics, Iowa State University, 260 Heady Hall, Ames, IA 50011, USA. E-mail: tobiasj@iastate.edu

[1] See an excellent survey by Anderson and van Wincoop (2004) on trade costs.
[2] See, for example, Rauch (1999) for a study involving the impact of common language and colonial ties on trade volumes in different types of goods.

goods. Using this insight, Ranjan and Lee (2003) construct a simple theoretical model to show why the degree of contract enforcement may be more important for trade in differentiated goods than homogeneous goods.[3] Their empirical strategy is to extend the gravity equation to incorporate proxies for the enforcement of contracts and to see if, as claimed by North (1990), they have differential effects on the volume of trade in differentiated goods compared to the volume of trade in homogeneous goods. This empirical strategy is similar to that used by Rauch (1999) and Rauch and Trindade (2002), who test for the differential implications of networks and ethnic Chinese networks, respectively, for the volume of trade in differentiated goods compared to homogeneous goods.

These papers, however, as well as others in the literature which make use of the gravity equation, face some econometric challenges when applied to international trade data, as we describe below. It is our intent in this paper to introduce a methodology which can overcome some of these econometric issues, to argue that the methods derived here are useful for other studies in international trade which make use of the gravity equation, and to apply these techniques to explore the relationship between contract enforcement and the volumes of trade in different types of goods.

When faced with modeling bilateral trade across countries, it is important to recognize that many observations contain identically zero trade values. As a result, one is naturally led to adopt a tobit-type specification. In particular (though its use remains rather rare in applied work), one can adopt the *threshold tobit model*, originally suggested for use in trade applications by Eaton and Tamura (1994), where the volume of trade between a pair of countries shows up as a positive quantity only if desired trade exceeds some threshold. The adoption of the threshold tobit enables us to assign meaningful probabilities to the event of no trade, and also helps us to avoid the problem of taking the log of zero, which is commonly encountered in these types of models.[4] To remain true to the mixed discrete–continuous nature of our trade data, we use the threshold tobit as the starting point of our analysis.

Second, like other recent studies in this literature, we seek to investigate the impact of one aspect of contracting costs (in our case contract enforcement) on bilateral trade volumes. While economic theory has been reasonably clear regarding the specification of variables like GDP and distance in the gravity model, the theory remains ambiguous on the exact way in which contract enforcement or other measures of contracting costs would enter a gravity equation determining bilateral trade. In light of this functional form uncertainty, we allow a non-parametric specification of contract enforcement, and do so within the basic threshold tobit framework.

Finally, in a recent influential paper on the gravity model, Anderson and van Wincoop (2003) have shown that many estimated gravity equations are misspecified to the extent that they ignore a 'multilateral resistance' term for each country. These terms are country-specific implicit price indexes which depend on the trade barriers between a country and all of its trading partners. The intuition behind these terms is that observed bilateral trade between a pair of countries depends not only on the trade barriers between those countries, but also the barriers for that pair relative to the barriers with other trading partners.

[3] Anderson and Marcouiller (2002) also test for the implications of institutions or contract enforcement for the volume of trade, though they make no distinction between its effect on different types of goods. The theoretical motivation of Anderson and Marcouiller (2002), unlike Ranjan and Lee (2003), comes from the insecurity of exchange, which should not affect the trade in differentiated and homogeneous goods differentially.

[4] In most empirical trade papers which employ a tobit-type specification, latent trade is modeled in logarithmic form. Thus, when observed trade is zero, the typical implication is that latent trade is non-positive, leading to a conflict with the logarithmic model.

DOI: 10.1002/jae

One way to control for these multilateral resistance terms is to include an exhaustive set of country indicator variables into the gravity model. This desire to control for country-specific effects has typically forced researchers to abandon the tobit or threshold tobit model, and instead run a linear regression analysis, commonly using log$(1 + \text{trade})$ as the dependent variable (see the Anderson and van Wincoop, 2004, survey). This approach substitutes computational simplicity for remaining true to the observed data; such a regression ignores the mixed discrete–continuous nature of trade volumes and arbitrarily adds unity to the dependent variable to avoid taking the log of zero. An alternative and more desirable approach, of course, is to include the country effects within the threshold tobit framework. With a reasonably large number of countries, adding these terms to the threshold tobit model can significantly slow down a direct maximum likelihood procedure, which probably accounts for its abandonment in favor of simpler, though inappropriate, linear regression alternatives.

To overcome these issues, we propose a new Bayesian procedure for estimating a generalized threshold tobit model. Specifically, we consider the gravity equation as a hierarchical regression model on suitably defined latent data, with the country effects accounting for the 'clustering' patterns in this data. Characteristics that are specific to each country, such as log GDP, and in our application, measures of contract enforcement, are assumed to enter the regression equation describing these country-specific effects. The model we propose is fit using MCMC methods, and throughout we also discuss testing, prediction and diagnostic checking. The model and associated MCMC algorithm adequately handles the incidence of zeros in our trade data, permits a non-parametric specification of the contract enforcement variable, and allows for the inclusion of country-specific effects within the threshold tobit framework. We stress that this model and estimation algorithm should offer a contribution beyond this empirical exploration, as it will appeal to other researchers for future work involving the gravity equation.

In terms of our empirical results, we find that the degree of contract enforcement significantly affects the volumes of bilateral trade in all types of goods. Tests of the non-parametric specification against parametric alternatives reveal that contract enforcement is 'significant' in our analysis, and also reveal modest evidence in favor of nonlinearities in this relationship. These nonlinearities would not have been revealed if simpler parametric models were employed. Specifically, we find that the functions relating contract enforcement to bilateral trade resemble those implied by a linear regression model with a single changepoint. The shapes of these relationships suggest that increasing the effectiveness of governance matters comparatively little (in terms of trade volumes) when the overall quality of governance is poor relative to increases in the effectiveness of governance when the overall quality of governance is reasonably high. Given the results of our non-parametric exploration, we then re-estimate a restricted parametric model which imposes piecewise linearity of contract enforcement with a single, unknown, changepoint. These functions were found to mimic those of the non-parametric model quite well. Perhaps most importantly, we find that the impact of our measures of contract enforcement is greatest for the volume of trade in differentiated goods compared to other types of goods. This also supports the claim of North (1990) mentioned earlier, and the predictions of previous theoretical work, that the trade of complex goods demands more reliable institutions.

The plan of rest of the paper is as follows. The following section describes our empirical model and the associated posterior simulator. Section 3 describes the data employed in our analysis, while Section 4 presents our empirical results. The paper concludes with a summary in Section 5.

2. EMPIRICAL MODEL

The standard gravity equation (e.g., McCallum, 1995) expresses the volume of trade between two countries as a function of the sizes of those countries and the distance between them. Country size is typically proxied by variables such as GDP or population, and in what follows we adopt the convention of using GDP as our size measure. In the following section we first review the standard threshold tobit variant of the gravity equation, which can be used to accommodate the incidence of numerous zeros in our trade data. We then show how this specification can be generalized to incorporate country-specific effects and a non-parametric specification of a variable of interest. Sections 2.3 and 2.4 then describe a posterior simulator for fitting the primary model, while Section 2.5 provides a procedure for testing the non-parametric specification against various parametric alternatives.

2.1. The Basic Threshold Tobit Model

Following the framework of Eaton and Tamura (1994), we model T_{ijk}, the quantity of bilateral trade between countries i and j in goods of type k, as follows:[5]

$$\log(T^*_{ijk} + \tau_k) = \alpha_k + \beta_k \ln(\mathrm{GDP}_i) + \beta_k \ln(\mathrm{GDP}_j) + z_{ij}\theta_k + \varepsilon_{ijk}, \quad \varepsilon_{ijk} \overset{\text{iid}}{\sim} N(0, \sigma^2_{\varepsilon k}) \tag{1}$$

where

$$T_{ijk} = \begin{cases} T^*_{ijk} & \text{if } T^*_{ijk} > 0 \\ 0 & \text{if } -\tau_k < T^*_{ijk} \le 0 \end{cases} \tag{2}$$

In the above, GDP_i denotes the gross domestic product of country i, and z_{ij} denotes a vector of other characteristics which vary across the (i, j) pairs, including the distance between countries i and j and an indicator which denotes if i, j share a common border. All parameters of the model are allowed to vary across type of good k and, ultimately, we will conduct separate analyses for each type of good k.

The threshold tobit model differs in an important way from the standard tobit through the *threshold parameter* τ_k. When $\tau_k = 0$, we are back into the framework of a standard tobit model. However, in this case, we run into problems associated with taking the log of zero or negative values in (1), and the latent data T^* become a redundant modeling of observed trade values T. More generally, when $\tau_k > 0$, the model enables us to place a discrete mass over zero trade, as found in the raw data, while retaining the basic and well-established log-linear specification of the gravity equation in (1).

The threshold parameter τ_k can also be given an economic interpretation. In our formulation of the model, $W^*_{ijk} \equiv T^*_{ijk} + \tau_k$ can be interpreted as the *desired amount of bilateral trade*[6] between

[5] The normality assumption is also imposed in Eaton and Tamura (1994). It would be possible through simple extensions of the posterior simulator to allow for heteroscedasticity of unknown form by including gamma mixing variables to the disturbance variance, thereby generalizing to the class of Student-t errors. We do not, however, pursue this in the present paper; posterior predictive checks under normality, upon including county-specific effects, we found to fit the data reasonably well (see Section 4.2).

[6] Strictly speaking, this interpretation is not perfect (though perhaps still reasonable) since the model does not permit desired trade to equal zero; all countries are assumed to have some positive amount of desired trade, and if this amount is sufficiently small no trade will occur.

DOI: 10.1002/jae

countries i and j in good type k. To see this more clearly, note that we can rewrite (1) and (2) in the following form:

$$\log(W^*_{ijk}) = \alpha_k + \beta_k \ln(\mathrm{GDP}_i) + \beta_k \ln(\mathrm{GDP}_j) + z_{ij}\theta_k + \varepsilon_{ijk} \tag{3}$$

where

$$T_{ijk} = \begin{cases} W^*_{ijk} - \tau_k & \text{if } W^*_{ijk} > \tau_k \\ 0 & \text{if } 0 < W^*_{ijk} \leq \tau_k \end{cases} \tag{4}$$

In this equivalent formulation of the model, the actual trade volume observed, T_{ijk}, equals zero if desired trade falls below the threshold τ_k and equals $W^*_{ijk} - \tau_k$ if desired trade exceeds the threshold. The threshold parameter τ_k can be interpreted as the amount of trade that is lost in transit or the amount that 'melts away': trade will occur (and equal $W^*_{ijk} - \tau_k$) if desired trade exceeds the amount that will be lost, while trade will not occur if the desired amount of trade is less than the amount that will be lost.[7]

The above model appropriately accounts for the discrete–continuous nature of the dependent variable; for a non-trivial number of countries in our data, we observe identically zero trade. In many cases, empirical trade researchers often seek to generalize the model in (3) by adding country-specific effects to the set of controls. In such cases, researchers often ignore the tobit nature of the model and simply use OLS as an estimation method, presumably because of the increased computational difficulty associated with including a large number of parameters in the nonlinear threshold tobit. In the following section, we describe a Bayesian estimation algorithm which properly accounts for the discrete–continuous nature of trade volumes while allowing for the inclusion of country-specific effects. Additionally, our model permits a new explanatory variable, contract enforcement, to enter the equation non-parametrically, given *a priori* functional form uncertainty regarding the specification of this variable.

2.2. Adding a Non-parametric Component and Country Effects

Before describing our Bayesian approach to estimation, we first note that equation (1) can be generalized in the following way:

$$\ln(T^*_{ijk} + \tau_k) = \gamma^k_i + \gamma^k_j + z_{ij}\theta_k + \varepsilon_{ijk} \tag{5}$$

with γ^k_i denoting a specific effect for country i (in type of good k), and γ^k_j interpreted similarly. Equation (5) reduces to equation (1) under the restriction $\gamma^k_c = (\alpha_k/2) + \beta_k \ln(\mathrm{GDP}_c)$, $c = 1, 2, \cdots, C$, with C denoting the total number of countries in the sample.[8] Although the ij subscript in our bilateral trade model can, of course, be reordered to ji without altering the values of our variables, individual country terms are still identifiable and thus country effects can be

[7] Note the similarity between this version of the model and a standard tobit with an unknown censoring point. In the standard tobit, the condition in (4) $T_{ijk} = W^*_{ijk} - \tau_k$ if $W^*_{ijk} > \tau_k$ would be replaced by $T_{ijk} = W^*_{ijk}$ if $W^*_{ijk} > \tau_k$. Though either version of the model could be employed, the empirical trade literature has adopted the version in (4), perhaps because of the 'melting away' interpretation of τ_k. The version in (4) is, however, likely to imply more mass being placed over low values of trade for some countries. This is not inconsistent with our data—for many smaller countries, we observe bilateral trade volumes measuring as little as US $5000.

[8] We assume that the number of countries is the same across different types of goods.

DOI: 10.1002/jae

added to the regression model. In recent work, for example, Anderson and van Wincoop (2003) provide a theoretical justification for the gravity model, and argue that country-specific 'multilateral resistance' terms should be added to the empirical specification. The incorporation of these and other country-specific factors motivate the inclusion of the country-level parameters γ_i and γ_j.

To tie (5) and (1) together, we introduce the following equation for the country effects:

$$\gamma_c^k = w_c\pi^k + u_c^k, \quad c = 1, 2, \cdots C, \quad u_c^k \overset{\text{iid}}{\sim} N(0, \sigma_{\gamma^k}^2) \tag{6}$$

where w_c denotes a set of country characteristics including log GDP and other potential covariates like contract enforcement. Also note that (6) is not deterministic; we have added a random error in the equation generating the country effects to permit correlation patterns in trade volumes within countries.

The Role of Contract Enforcement: A Non-parametric Approach

In our empirical application, we are primarily interested in determining the role of contract enforcement on trade. Although the functional form of the basic gravity equation is well established, and the specification of covariates like log GDP are reasonably agreed upon, relatively little work has been done to investigate the impact of contract enforcement on trade.[9] In light of this functional form uncertainty, we might choose to permit our measure of contract enforcement to enter the gravity equation flexibly. To this end, we specify a generalized version of (6) of the form[10]

$$\gamma_c^k = \pi_1^k \log \text{GDP}_c + f^k(\text{Contract}_c) + u_c^k \tag{7}$$

Equations (5) and (7) constitute the empirical specification of interest—a *semi-parametric hierarchical threshold tobit model.*

The primary computational issue that arises in the estimation of this model is how to recover the function f. Our method for estimating the non-parametric component f follows that described in Koop and Poirier (2004). To this end, we first *sort the data by ascending values of contract enforcement* and number the countries consecutively once the sorting has taken place.[11] Once the data are sorted in this way, neighboring countries in our enumeration scheme have similar values of contract enforcement. We let $\psi_c^k = f^k(\text{Contract}_c)$ and write (7) as

$$\gamma_c^k = \pi_1^k \log \text{GDP}_c + \psi_c^k + u_c^k$$

Stacking over c we obtain

$$\gamma^k = \pi_1^k \log \text{GDP} + I_C\psi^k + u^k \tag{8}$$

$$= W\pi^k + u^k \tag{9}$$

where

$$W = [\log \text{GDP}\ I_C], \quad \pi^k = [\pi_1^k \psi^{k'}]', \quad \psi^k = [\psi_1^k\psi_2^k \cdots \psi_C^k]', \quad \gamma^k = [\gamma_1^k\gamma_2^k \cdots \gamma_C^k]'$$

[9] Exceptions include Ranjan and Lee (2003) and Anderson and Marcouiller (2002).

[10] Note that the intercept has been excluded from this equation, as it is absorbed in f.

[11] That is, country 1 in our data, which corresponds to parameter γ_1, has the lowest value of the contract enforcement variable, country 2 has the second lowest, and country C has the highest value.

DOI: 10.1002/jae

I_C denotes the $C \times C$ identity matrix and the variables log GDP and u have been stacked appropriately.

As discussed in Koop and Poirier (2004) in the context of a non-parametric regression problem, the model in this form can raise identification concerns as it contains more parameters than 'observations' at that stage of the hierarchy. To combat this issue, and more importantly for the purposes of this paper, to introduce the potential for *smoothing* the regression curve, one can place an informative prior over the elements of ψ^k. In what follows, we describe a prior which accomplishes such smoothing by centering differences of the pointwise slopes of the regression function over a mean of zero. A hyperparameter denoted η will then govern the strength of this prior information, and thereby the smoothness of the resulting regression function.

Before describing this prior in more detail, first let $\mathcal{C}_j$ denote the value of the contract enforcement variable for the jth country, (i.e., $\mathcal{C}_j = \text{Contract}_j$), once the data have been sorted, and let $\Delta_j = \mathcal{C}_j - \mathcal{C}_{j-1}$. Define the $C \times C$ differencing matrix H as follows:

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \Delta_2^{-1} & [-\Delta_2^{-1} - \Delta_3^{-1}] & \Delta_3^{-1} & 0 & \cdots & 0 & 0 & 0 \\ 0 & \Delta_3^{-1} & [-\Delta_4^{-1} - \Delta_3^{-1}] & \Delta_4^{-1} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \Delta_{C-1}^{-1} & [-\Delta_{C-1}^{-1} - \Delta_C^{-1}] & \Delta_C^{-1} \end{bmatrix}$$

We then specify a prior of the following form:[12]

$$H\psi^k | \eta^k \sim N(0_C, \eta^k V^*_{\psi^k}) \tag{10}$$

or equivalently,

$$\psi^k | \eta^k \sim N(0_C, \eta^k H^{-1} V^*_{\psi^k} [H^{-1}]') \tag{11}$$

where

$$V^*_{\psi^k} = \begin{bmatrix} cI_2 & 0 \\ 0 & I_{C-2} \end{bmatrix}$$

From (10), we see that the construction of H places a prior over the 'initial conditions' ψ_1^k and ψ_2^k, as well as first differences of the pointwise slopes, $[\psi_j^k]' - [\psi_{j-1}^k]'$, $j = 2, 3, \cdots, C-1$, where

$$[\psi_{j-1}^k]' \equiv \frac{\psi_j^k - \psi_{j-1}^k}{\mathcal{C}_j - \mathcal{C}_{j-1}}$$

In practice, we can specify a reasonably vague prior over the initial conditions ψ_1^k and ψ_2^k by choosing a value for the constant c that is large relative to the prior mean of η. The first

[12] Note that we are implicitly working with second rather than first differences. In this way, we can choose a prior to center differences between adjacent slopes over zero rather than center differences between adjacent function values over zero. The latter approach would not ideally accommodate situations where spacing between adjacent data points was not uniform. Koop and Poirier (2004) note that priors based on second differences can also match a natural cubic spline approach. The second differencing 'smoothness' prior has also been developed and applied in other related work, including Schiller (1973).

DOI: 10.1002/jae

differences of the pointwise slopes are centered over a prior mean of zero, and the strength of this prior information (and thereby degree of smoothness) is controlled by the parameter η. In the limiting case where $\eta \to 0$, the pointwise slopes are restricted to be equal, thereby smoothing the regression curve to a linear specification, with the initial conditions determining the slope and intercept of the line. In the other extreme where η is too large, essentially no prior information is imposed, the regression function will be excessively jumpy, and we will overfit the model.

To let the appropriate amount of smoothing be revised by the data, we include η^k as a parameter of the model and add a hierarchical prior of the form[13]

$$\eta^k \sim IG(c_1, c_2) \tag{12}$$

Finally, to complete the specification of our prior for all regression parameters π appearing in (9), define $V_{\psi_k}(\eta^k) = \eta^k H^{-1} V^*_{\psi^k}[H^{-1}]'$. Assuming independent normal priors for π_1 and ψ^k, we obtain

$$\pi^k|\eta^k \sim N[0, V_\pi(\eta)] \tag{13}$$

where

$$V_\pi(\eta) = \begin{bmatrix} v_{\pi_1} & 0 \\ 0 & V_{\psi_k}(\eta^k) \end{bmatrix}$$

2.3. The Joint Posterior Distribution and Posterior Simulator

To facilitate computation, we pursue a data augmentation approach and augment the parameter space with the latent data T^*_{ijk}. To simplify the model and associated notation, we assume that the parameters are independent across type of goods k *a priori*.[14] As such, given the assumptions of our model, the joint posteriors for our parameters will be independent across good types, and thus we can employ separate estimation procedures for each type of good. We thus simplify our notation by dropping the k subscript from our parameters and latent variables, noting that what follows applies for separate analyses of each type of good.

Let $\lambda = [\gamma\ \theta\ \sigma_\varepsilon^2\ \sigma_\gamma^2\ \pi\ \eta, \tau]$ denote the parameters of our model. We seek to characterize the augmented joint posterior distribution

$$p(\{T^*\}, \lambda|T) \tag{14}$$

where T^* denotes the vector of latent trade values T^*_{ij} stacked over countries and T is defined similarly.

The joint posterior in (14) can be written as

$$\begin{aligned} p(T^*, \lambda|T) &\propto p(T^*, T|\lambda)p(\lambda) \\ &= p(T|T^*, \lambda)p(T^*|\lambda)p(\lambda) \end{aligned}$$

[13] Alternatively, as discussed in Koop and Poirier (2004), one could choose η via an empirical Bayes approach, via an extreme bounds analysis or, potentially, using traditional frequentist criteria such as cross-validation. Here we pursue what seems to be a reasonable and computationally attractive alternative by specifying a prior for η and including it as a parameter of our model.

[14] One could, for example, relax this assumption by employing a prior which permits some type of correlation among the country parameters γ_i^k across goods k. We do not, however, explore this issue in the current paper.

$$= p(\lambda) \prod_{i,j=1}^{n} ([I(T_{ij} = T_{ij}^*)I(T_{ij}^* > 0) + I(T_{ij} = 0)I(-\tau < T_{ij}^* \leq 0)]p(T_{ij}^*|\lambda))$$

with $I(\cdot)$ denoting the standard indicator function. In the above, the conditional density $p(T_{ij}^*|\lambda)$ follows from a change of variable from ε_{ij} to $T_{ij}{}^*$ in (5), producing

$$p(T_{ij}^*|\lambda) \propto \frac{1}{T_{ij}^* + \tau} \exp\left(-\frac{1}{2\sigma_\varepsilon^2}[\ln(T_{ij}^* + \tau) - \gamma_i - \gamma_j - z_{ij}\theta]^2\right), \quad T_{ij}^* > -\tau \tag{15}$$

Working with this posterior distribution presents a small challenge, given the form of the conditional density in (15). We can, instead, reparameterize this joint posterior to a more 'traditional' form, and thereby gain some computational simplicity, by working in terms of $V_{ij}^* \equiv \ln(T_{ij}^* + \tau)$ rather than T_{ij}^*. Since the Jacobian of this transformation is just $\exp(V_{ij}^*)$, it follows that

$$p(V^*, \lambda|T) \propto \prod_{i,j=1}^{n} ([I[T_{ij} = \exp(V_{ij}^*) - \tau]I(V_{ij}^* > \ln(\tau)) + I(T_{ij} = 0)I[V_{ij}^* \leq \ln(\tau)]] \tag{16}$$

$$* \phi(V_{ij}^*; \gamma_i + \gamma_j + z_{ij}\theta, \sigma_\varepsilon^2))p(\lambda). \tag{17}$$

For our prior, $p(\lambda)$, we specify

$$\begin{aligned} \theta &\sim N(\mu_\theta, V_\theta) \\ \sigma_\varepsilon^2 &\sim IG(a_1, a_2) \\ \sigma_\gamma^2 &\sim IG(b_1, b_2) \\ \tau &\sim p(\tau), \end{aligned}$$

in addition to the priors already described in (7), (12) and (13). Equations (16) and (17), together with the priors above, define our complete model specification.

2.4. The Posterior Simulator

We fit the model using the Gibbs sampler, and to this end we derive and report the posterior conditionals of the model. To mitigate autocorrelations in our simulations, we employ several blocking steps. Before describing these, first let $\Gamma = [\lambda V^*]$ denote all the elements of the posterior in (16), and define Γ_{-x} as all the elements of Γ other than x.

Step 1: $\tau, V^*|\Gamma_{-\tau,V^*}, T$

Early work on this problem revealed high degrees of autocorrelation between the latent data V^* and threshold parameter τ, thus motivating our decision to group these objects into a single block. We draw from this joint conditional using the *method of composition*. Specifically, we write

$$p(\tau, V^*|\Gamma_{-\tau,V^*}, T) = p(\tau|\Gamma_{-\tau,V^*}, T)p(V^*|\Gamma_{-V^*}, T)$$

DOI: 10.1002/jae

The first density on the right-hand side of the equation above is proportional to the standard tobit likelihood $p(T|\Gamma_{-V^*})$ (which is marginalized over the latent data) times the prior $p(\tau)$. Thus, we obtain

$$p(\tau|\Gamma_{-\tau,V^*}, T) \propto p(\tau) \prod_{i,j:T_{ij}=0} \Phi\left(\frac{\ln(\tau) - \gamma_i - \gamma_j - z_{ij}\theta}{\sigma_\varepsilon}\right) \tag{18}$$
$$* \prod_{i,j:T_{ij}>0} \frac{1}{T_{ij}+\tau} \exp\left(-\frac{1}{2\sigma_\varepsilon^2}[\ln(T_{ij}+\tau) - \gamma_i - \gamma_j - z_{ij}\theta]^2\right)$$

The distribution above is not of a standard form. However, τ is a scalar, and thus a variety of methods can be used to generate variates from (18). In generated data experiments, we found that simulating draws from a fine discrete approximation of this density produced good results, and thus make use of this approach in our empirical work. For our prior, we choose $p(\tau)$ to be uniform over the discrete set of support points. If desired, one could instead employ a Metropolis–Hastings substep to generate draws from (18).

The complete posterior conditional for V^* is obtained as follows:

$$p(V_{ij}^*|\Gamma_{-V_{ij}^*}, T) \overset{\text{ind}}{\sim} \begin{cases} TN_{(-\infty,\ln(\tau)]}(\gamma_i + \gamma_j + z_{ij}\theta, \sigma_\varepsilon^2) & \text{if } T_{ij} = 0 \\ \ln(T_{ij}+\tau) & \text{if } T_{ij} > 0 \end{cases} \tag{19}$$

In the above, $TN_{(a,b)}(\mu, \sigma^2)$ denotes a normal density with mean μ and variance σ^2 which has been truncated to the interval (a, b). The conditional density in (19) resembles that of the standard tobit model (e.g., Chib, 1992) in that the latent data are drawn from and 'filled in' when no trade occurs ($T_{ij} = 0$), while V_{ij}^* is effectively known when positive trade occurs. Drawing from (18) and then independently from (19) gives a draw from the posterior conditional $p(\tau, V^*|\Gamma_{-\tau,V^*}, T)$.

Step 2: $\gamma, \theta|\Gamma_{-\gamma,\theta}, T$

For the second blocking step, note

$$\begin{aligned} V_{ij}^* &= \gamma_i + \gamma_i + z_{ij}\theta + \varepsilon_{ij} \\ &= d_{ij}\gamma + z_{ij}\theta + \varepsilon_{ij} \\ &= \overline{z}_{ij}\overline{\theta} + \varepsilon_{ij} \end{aligned}$$

where

$$\overline{z}_{ij} = [d_{ij}z_{ij}] \quad \text{and} \quad \overline{\theta} = [\gamma'\theta']'$$

In the above, d_{ij} is a $1 \times C$ vector containing two ones in the positions denoting the ith and jth countries and zeros elsewhere. Given this notation, it follows that

$$\overline{\theta}|\Gamma_{-\theta,\gamma}, T \sim N(Dd, D) \tag{20}$$

where

$$D = (\overline{Z}'\overline{Z}/\sigma_\varepsilon^2 + V_{\overline{\theta}}^{-1})^{-1}, \quad d = \overline{Z}'V^*/\sigma_\varepsilon^2 + V_{\overline{\theta}}^{-1}\mu_{\overline{\theta}}$$

$\overline{Z}$ denotes the matrix derived from stacking $\overline{z}_{ij}$ over (ij), V^* is similarly derived from staking V^*_{ij}, and

$$V_{\overline{\theta}} \equiv \begin{bmatrix} \sigma_\gamma^2 I_C & 0 \\ 0 & V_\theta \end{bmatrix}, \quad \mu_{\overline{\theta}} = \begin{bmatrix} W\pi \\ 0 \end{bmatrix}$$

with $W\pi$ defined in (9).

Step 3: $\sigma_\varepsilon^2 | \Gamma_{-\sigma_\varepsilon^2}, T$

$$\sigma_\varepsilon^2 | \Gamma_{-\sigma_\varepsilon^2}, T \sim IG\left[\frac{n}{2} + a_1, \left(a_2^{-1} + .5\sum_{i,j}(V^*_{ij} - \gamma_i - \gamma_j - z_{ij}\theta)^2\right)^{-1}\right] \tag{21}$$

with n denoting the total sample size.

Step 4: $\pi | \Gamma_{-\pi}, T$

$$\pi | \Gamma_{-\pi}, T \sim N(D_\pi d_\pi, D_\pi) \tag{22}$$

where

$$D_\pi = (W'W/\sigma_\gamma^2 + V_\pi(\eta)^{-1})^{-1}, \quad d_\pi = W'\gamma/\sigma_\gamma^2$$

and $V_\pi(\eta)$ is defined below (12).

Step 5: $\sigma_\gamma^2 | \Gamma_{-\sigma_\gamma^2}, T$

$$\sigma_\gamma^2 | \Gamma_{-\sigma_\gamma^2}, T \sim IG\left[\frac{C}{2} + b_1, (b_2^{-1} + .5(\gamma - W\pi)'(\gamma - W\pi))^{-1}\right] \tag{23}$$

Step 6: $\eta | \Gamma_{-\eta}, T$

$$\eta | \Gamma_{-\eta}, T \sim IG\left[\frac{C}{2} + c_1, (c_2^{-1} + (1/2)\psi' H' [V^*_\psi]^{-1} H\psi)^{-1}\right] \tag{24}$$

The posterior simulator involves iteratively drawing from (18)–(24).

2.5. Testing

In our particular application, it is of primary interest to conduct tests regarding the role of the contract enforcement variable. Specifically, we would like to determine if contract enforcement is a 'significant' predictor of bilateral trade, and if the function relating contact enforcement to trade is linear.

DOI: 10.1002/jae

To this end, we appeal to the Savage–Dickey density ratio. As shown in Verdinelli and Wasserman (1995), if one desires to test a restriction of the form $\delta = 0$, for some subvector δ of the parameter space, the Bayes factor of the restricted model $\mathcal{M}_1$, which imposes $\delta = 0$, in favor of the unrestricted $\mathcal{M}_2$ can be expressed as[15]

$$K_{12} = \frac{p(\delta = 0|y, \mathcal{M}_2)}{p(\delta = 0|\mathcal{M}_2)} \tag{25}$$

We now describe how the imposition of such zero subvector hypotheses allows us to test for the significance and linearity of the contract enforcement variable.

Test for Linearity

The model described in the previous section will be *linear* if the first differences of the pointwise slopes are equal to zero. In other words, letting $\delta^k = H\psi^k$, as in equation (10), the model is linear in contract enforcement if the last $C - 2$ elements of δ^k are equal to zero. The denominator of equation (25) is readily available since our prior is placed over the elements of δ^k. This ordinate can be calculated by Rao–Blackwellization, given the priors in (10) and (12):

$$p(\delta^k_{3:C} = 0|\mathcal{M}_2) \approx \frac{1}{J}\sum_{j=1}^{J}\phi(\delta^k_{3:C} = 0; 0, \eta^{k,(j)} I_{C-2})$$

where the 3: C notation is used to denote the last $C - 2$ elements of δ^k, $\phi(x^0; \mu, \Sigma)$ denotes a normal density with mean μ and covariance matrix Σ evaluated at x^0, and finally, $\eta^{k,(j)} \stackrel{\text{iid}}{\sim} IG(c_1, c_2)$.

The numerator of (25) is the marginal posterior ordinate at zero. To evaluate this quantity directly, we simply re-parameterize the model in terms of $\delta^k = H\psi^k$ and calculate the posterior ordinate in a similar manner via Rao–Blackwellization. Like (22), the conditional posterior for δ (which is an element of π) will be normal.[16] Thus, following similar arguments to the calculation of the prior ordinate above, the marginal posterior ordinate at zero can be calculated by averaging the conditional normal posterior ordinates, which are calculated at each iteration using the simulated posterior output.

Test for Significance

We would call the contract variable 'insignificant' if the estimated regression function were constant across the contract enforcement support. Given that we are working with derivatives of the function, this will be the case if the derivative at the first point is equal to zero and the differences between consecutive derivatives are also equal to zero. To this end, we now create a $C \times C$ matrix R, which equals the matrix H as defined above (10), except the second row of H is replaced with

$$[-\Delta_2 \quad \Delta_2 \quad 0 \quad 0 \quad \cdots \quad 0]$$

The test of significance then reduces to testing if all elements of $\delta^k = R\psi^k$, except for the first, is equal to zero. This test can be carried out in a similar manner, after reparameterizing in terms of δ^k and calculating the numerator and denominator of (25) via Rao–Blackwellization.

[15] This assumes that the same priors are employed for parameters common to $\mathcal{M}_1$ and $\mathcal{M}_2$.

[16] With the reparameterization, W in (9) becomes $W = [\log \text{GDP } H^{-1}]$, the prior for δ is given (10), and the posterior simulator is adjusted to reflect these definitions.

DOI: 10.1002/jae

3. THE DATA

Data on bilateral trade flows are taken from the NBER Trade Database, Disk 2: World Trade Flows, 1970–1992, which in turn is based on the World Trade Database from Statistics Canada. We employ data in this paper for the last year of the survey, 1992.

As described in the introduction of this paper, one question of interest is whether the trade of more complex goods demands more reliable and effective governance between the trading countries. By 'complex' goods, we simply mean goods which vary considerably in their attributes, and for whom a single price cannot adequately reflect all information relevant for international trade.[17] To divide goods into various categories along these lines, we follow the classification scheme of Rauch (1999), who categorizes goods into three types: homogeneous, reference and differentiated. Homogeneous goods are defined as products with little or no variation in attributes whose prices are quoted on organized international exchanges. Reference goods are goods for which no organized exchange exists, but 'reference prices' can be found in trade publications. Finally, differentiated goods do not possess such a reference price, are not traded on organized exchanges, and are taken to have sufficient variation in attributes that a single price cannot convey all information relevant for international trade. Further details of this classification process can be found in Rauch (1999). Our conjecture is that contract enforcement will matter most for the trade of differentiated products, and least for the trade of homogeneous goods.

We use the World Bank data on governance as our proxy for contract enforcement. The data are drawn from two types of sources: polls of experts, which reflect country ratings produced by commercial risk-rating agencies and other organizations, and cross-country surveys of residents carried out by international organizations and other non-governmental organizations. The disaggregated data are organized into six clusters and then for each cluster a process of aggregation is used to construct aggregate governance indicators (see Kaufman *et al.*, 1999a, 1999b, for details on data and aggregation). Of six aggregate indicators of governance available in the World Bank data, we focus on the following four: 'Government Effectiveness', 'Regulatory Burden', 'Rule of Law' and 'Graft'. The first two capture the capacity of the state to implement sound policies, while the last two capture the respect of citizens and the state for the rules which govern their interactions. We use an average of these four aggregated indicators as our measure of contract enforcement. The final variable ranges from 1.44 to 4.30 in our sample, with higher values indicating more effective governance.

Finally, our border indicator and measures of the distances between countries are taken from John Haveman,[18] who provides a variety of data useful for international trade studies. Real GDP (measured in thousands of 1992 US dollars) is taken from the Penn World Table version 5.6 for the year 1992. Our final sample consists of complete bilateral trade observations between 79 different countries, for a total of 3081 observations.

4. EMPIRICAL RESULTS

We employ the methodology discussed in Section 2 to address the following primary questions: (1) Does contract enforcement matter in predicting bilateral trade flows? (2) Is the relationship

[17] Rauch (1999), for example, uses footwear as an example of a differentiated product, and lead as an example of a homogeneous product.

[18] See http://www.macalester.edu/research/economics/PAGE/HAVEMAN/Trade.Resources/TradeData.html.

DOI: 10.1002/jae

between contract enforcement and bilateral trade volumes linear? and (3) Does contract enforcement matter most for describing the trade of differentiated goods?

4.1. Estimation and Testing

For each of our three types of goods (homogeneous, reference and differentiated), we implement the algorithm described in Section 2. We run our posterior simulator for 10 000 iterations and discard the first 200 of these as the burn-in. Results from these runs and other generated data experiments suggested that the chain mixed reasonably well and appeared to converge within a few hundred iterations. For our prior hyperparameters, we set $\mu_\theta = 0$, $V_\theta = 10I_2$, $a_1 = b_1 = 3$, $a_2 = 0.5$ and $b_2 = 1$. For the hyperparameters of the inverted gamma prior for the smoothing parameter η, we set $c_1 = 3$ and $c_2 = 5$. This sets the prior mean and standard deviation of η equal to 0.1, which we found in experimental work produced a reasonable amount of smoothing of the regression function. Of course, results will be most sensitive to the choices of c_1 and c_2, particularly when testing our hypotheses of interest, and thus, where appropriate, we describe how results change in accord with changes in these hyperparameters.

Presented in Table I are posterior means and standard deviations associated with the parameters of our model for each type of good. The results presented in the table are generally consistent with previous work on this topic, and thus will not be discussed at length: log GDP enters as a significant explanatory variable with a coefficient near one, while distance between countries negatively affects trade volumes. Somewhat surprisingly, we do not find strong and significant evidence of a border effect, which does depart somewhat from the existing literature on this topic. We have no convincing explanation why this may be the case, but simply note that the specification we employ—which accounts for both the discrete–continuous nature of bilateral trade and the inclusion of country-specific effects—has not been estimated in this literature, which may account for the difference in results.

Table I. Parameter posterior means (standard deviations in parentheses)

Variable/parameter	Type of good		
	Diff.	Hom.	Ref.
τ	277.62	259.78	376.07
	(22.44)	(23.05)	(30.10)
Log distance	−1.14	−1.13	−1.19
	(0.034)	(0.054)	(0.038)
Border	0.140	0.121	−0.123
	(0.162)	(0.247)	(0.169)
Log GDP	0.894	0.894	0.813
	(0.045)	(0.047)	(0.039)
σ_ε^2	1.25	2.75	1.35
	(0.049)	(0.121)	(0.056)
σ_γ^2	0.377	0.381	0.264
	(0.070)	(0.075)	(0.051)
η	0.0590	0.0586	0.0536
	(0.028)	(0.049)	(0.028)
N	3081	3081	3081

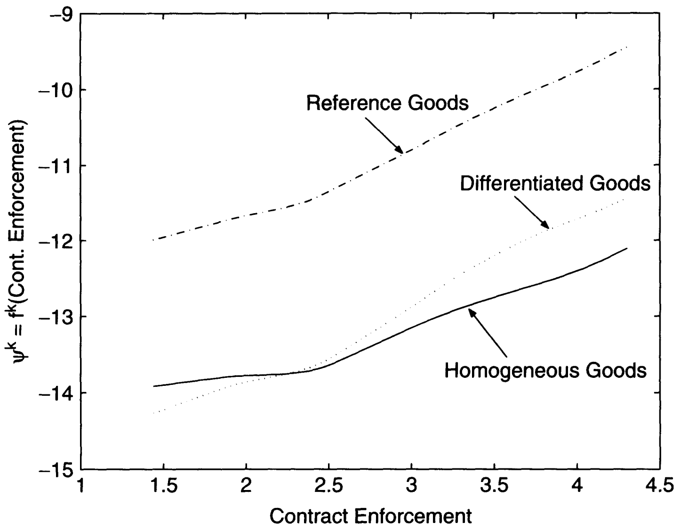

Figure 1. Estimation of regression functions (contract) across different types of goods

In Figure 1, we plot point estimates of our contract enforcement functions for each type of good. For clarity of presentation, we do not include standard error bands within the figure, as they obscure the point estimates provided in the graph. The figure suggests several important results. First, for each type of good, the estimated functions are increasing, suggesting that countries with more effective governance and greater ability to successfully enforce contracts engage in greater volumes of bilateral trade. Second, there is some evidence that the slope is greatest for differentiated goods, though it appears rather similar for both differentiated and reference goods. This provides some suggestive evidence that contract enforcement clearly matters more for the trade of differentiated than homogeneous goods, consistent with the speculation of North (1990), and previous theoretical work. We will revisit this issue and provide a more formal comparison of slopes later in this empirical section.

As described in Section 2, we can also use the Savage–Dickey density ratio to formally test for the significance and linearity of contract enforcement. To conduct a test of 'significance', we compare the unrestricted model with the given hyperparameters to a restricted model imposing $f(\text{Contract}) = c$, for some constant c. Calculating the Bayes factor for this case, we found strong evidence that contract enforcement is a significant predictor of bilateral trade for all reasonable hyperparameters c_1 and c_2. As for our linearity test, with hyperparameters $c_1 = 3$ and $c_2 = 5$, we calculate Bayes factors in favor of linearity equal to 2381, 8.31 and 139.1 for differentiated, reference and homogeneous goods, respectively. Thus, under this prior, linearity of contract enforcement is supported by the data. This is consistent with the results reported in Table I: the posterior means for η are approximately 0.05 or 0.06 for each type of good, with posterior standard deviations ranging from 0.028 to 0.049. The prior means and prior standard deviations, on the other hand, are all equal to 0.1. Thus, the data have revised our beliefs regarding the appropriate amount of smoothing, and have pulled us toward smaller values of η which, by construction, force the regression curve to be more linear. As a result, the Bayes factor calculation reveals a preference for linearity under these hyperparameter values.

DOI: 10.1002/jae

When the prior mean of η is chosen to be sufficiently small, however (keeping $c_1 = 3$ and changing c_2 accordingly), so as to virtually impose linearity through the prior, the posterior means of η are found to be *larger* than the prior means. In this case, the data pull us away from such an extreme amount of smoothing. Under these hyperparameter values, the Bayes factor rejects the linearity of contract enforcement, despite the fact that the estimated regression functions nonetheless appear linear as a consequence of our strongly informative prior.

The results of this testing analysis and the shapes of the curves presented in Figure 1 suggest that imposing a prior which forces the contract enforcement functions to be linear may not be supported by the data. Indeed, a quick inspection of Figure 1 suggests that the contract enforcement functions could be reasonably modeled as linear specifications, once an allowance has been made for a changepoint to occur around values of contract enforcement equal to 2.5. In economic terms, this suggests that improvements in the effectiveness of governance matters for all countries, but such improvements have the most impact on trade volumes for countries with already high values of contract enforcement. The shapes of the curves in Figure 1 suggest the potential existence of a *threshold*—once countries have established a minimum quality of governance, further improvements in governance quality result in larger impacts on trade volumes. This nonlinearity seems evident for each type of good, and reveals an empirical regularity that would not have been uncovered through a simple linear specification.

To examine this feature of our data in more detail, we decided to re-estimate our hierarchical threshold tobit model, and restrict the non-parametric specification to a piecewise linear one which allows for a changepoint in contract enforcement. That is, motivated by our non-parametric estimates in Figure 1, we estimate the model in (5) together with a restricted version of (7) of the form

$$\gamma_c = \pi_0 + \pi_1 \log \mathrm{GDP}_c + \pi_2\ \mathrm{Contract}_c + \pi_3(\mathrm{Contract}_c - \kappa)_+ + u_c, \quad c = 1, 2, \cdots, C$$

where κ denotes a *changepoint* of the model, and x_+ is defined as x if its value is positive, and 0 otherwise. Rather than impose the location of the changepoint κ *a priori*, we instead allow the location of κ to be revised by the data.

To fit this model, we employ a uniform prior over the discrete set of values $\kappa \in \{\mathcal{C}_1, \mathcal{C}_2, \cdots, \mathcal{C}_{C-1}\}$. Thus, we assume that a changepoint exists, but do not impose any informative prior beliefs regarding the location of that changepoint. It is also worth noting that if $\pi_3 = 0$, the model reduces to a specification that is linear in contract enforcement. This, again, enables us to employ the Savage–Dickey density ratio to test for linearity in our restricted model. The hierarchical tobit model with an unknown changepoint is then fit using MCMC methods.[19] In this case, the posterior conditional for the changepoint κ is not of standard form, but the support of κ is discrete, thus enabling straightforward sampling from its posterior conditional.

In Figure 2 we plot the estimated functions associated with our changepoint model for differentiated and reference goods.[20] For the sake of comparison, we also include the non-parametric estimates of Figure 1 for these types of goods in the graph. As you can see, the estimated regression functions are quite similar across approaches, suggesting that the single changepoint specification can indeed pick up the key features of this data. The restricted

[19] For this model specification, we employ a normal prior for $[\pi_0\ \pi_1\ \pi_2\ \pi_3]'$ with prior mean $\mu_\pi = [-15\ .9\ .7\ 0]'$ and diagonal prior covariance matrix with elements [100 0.5 4 0.5] distributed along the diagonal.

[20] Results for homogeneous goods showed a similar pattern, and are not included in Figure 2 for clarity.

DOI: 10.1002/jae

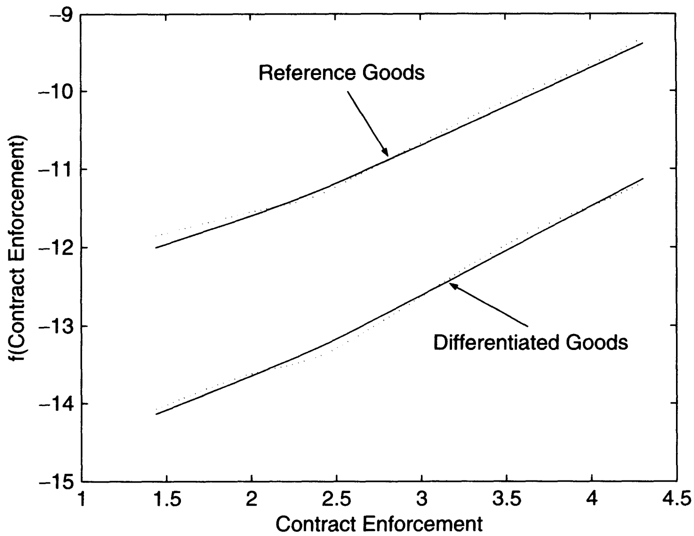

Figure 2. Regression functions (contract) for reference and differentiated goods. Non parametric model (dashed), changepoint model (solid)

changepoint specifications again provide evidence of a change in curvature throughout the contract enforcement support. It is also worth noting that the point estimates reported in Figure 2 average over uncertainty associated with the location of the changepoint, and thus the resulting posterior point estimate need not appear to be piecewise linear.

In terms of formal model comparison, the Bayes factors in favor of linearity under the changepoint model were 0.920, 1.04 and 0.899 for differentiated, reference and homogeneous goods, respectively. These results were obtained assuming independent priors for the elements of π, and specifically, assuming $\pi_3 \sim N(0, 0.5)$. With these prior values, our results suggest near indifference between the changepoint model and a linear specification. Since our Bayes factors will be sensitive to the prior employed (and this sensitivity is more pronounced in testing than in estimation), we conducted a sensitivity analysis using priors of the form $\pi_3 \sim N(0, \varphi 0.5)$. For $\varphi < 5$, our results consistently revealed approximate indifference between the linear and changepoint models. We begin to favor linearity only when our prior is quite non-informative. For $\varphi = 500$, for example, the Bayes factors for each type of good ranged from 10 to 11, indicating modest support for the linear specification.[21]

The main reason for the indifference under moderately informative priors arises from the fact that the parameter π_3 is not estimated precisely in this model, primarily because of uncertainty regarding the location of the changepoint. Figure 3 provides evidence of this, and plots a bar graph showing the realized values of the changepoint κ (for the homogeneous goods analysis) among a larger run of 80 000 posterior simulations. Recall that the prior for this parameter is uniform, and thus if no revision takes place from the data the heights of the vertical bars in Figure 3 should be around 1000 for each of the 78 points in the contract enforcement support. This is clearly not

[21] This support for the restricted model as the prior becomes non-informative is related to Bartlett's paradox (see, for example, Poirier, 1995, p. 390).

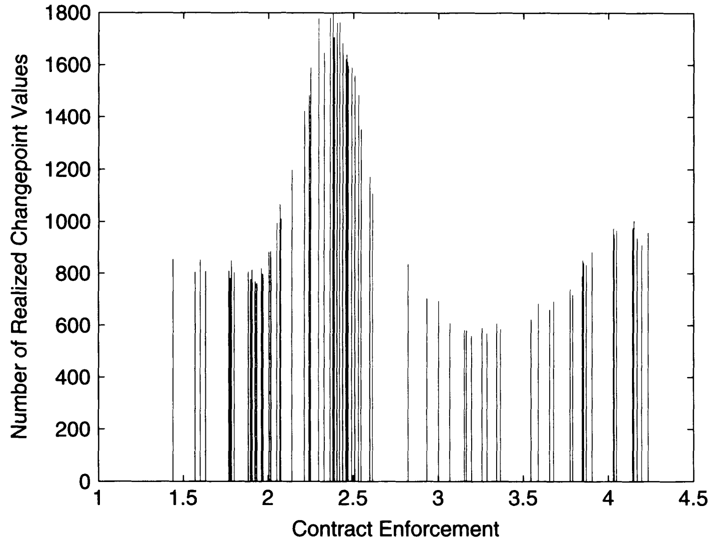

Figure 3. Posterior frequencies for changepoint κ : homogeneous goods

the case and, consistent with Figure 1, the data provide some evidence regarding a changepoint near contract enforcement values equal to 2.5. The posterior is not overwhelmingly informative, however, as the mass remains reasonably substantial throughout the support. This uncertainty associated with κ tends to increase the uncertainty associated with π_3. When fixing κ near 2.5, however, we begin to reject linearity with the given prior.

Taken together, the results of this section provide compelling evidence that contract enforcement matters in describing bilateral trade volumes for all types of goods. Non-parametric and restricted parametric analyses suggest some nonlinearities in this relationship which can be well approximated by a single changepoint model. Formal model comparison, however, provides inconclusive evidence regarding the nonlinearities given the posterior uncertainty associated with the parameters.

4.2. Prediction

A seemingly beneficial aspect to the methodology outlined in Section 2, which we feel should appeal to others conducting research on this topic, is the ease in which it can be used to generate predictions, both within and out of sample. That is, given some levels of GDP, distance, etc., how would one make use of the model and simulated parameter values to predict the volume of bilateral trade between two countries? These predictions should, of course, reflect the possibility that identically zero trade will occur, since it does happen for approximately 20% of our sample. These posterior predictive exercises can also be quite useful for diagnostic checking purposes.[22]

To this end, consider the case of out-of sample prediction, and suppose for simplicity that the covariates' values are given. Let T_{ij}^f denote a future, as yet unobserved value of bilateral trade

[22] See, for example, Lancaster (2004) and Geweke (2005).

DOI: 10.1002/jae

between countries i and j. Further, suppose that the model in (5) and (7) applies to this future observation. The posterior predictive density T_{ij}^f can be obtained as follows:

$$p(T_{ij}^f|T) = \int p(T_{ij}^f|V_{ij}^{f^*}, \lambda, T)p(V_{ij}^{f^*}|\lambda, T)p(\lambda|T)dV_{ij}^{f^*}\, d\lambda \tag{26}$$

This integration is easily approximated by the method of composition, since the posterior simulator provides draws from $p(\lambda|T)$, $p(V_{ij}^{f^*}|\lambda, T)$ is normal, and

$$T_{ij}^f|V_{ij}^{f^*}, \lambda, T = \begin{cases} \exp(V_{ij}^{f^*}) - \tau & \text{if} \exp(V_{ij}^{f^*}) > \tau \\ 0 & \text{if} \exp(V_{ij}^{f^*}) \leq \tau \end{cases} \tag{27}$$

To illustrate how this methodology can be used, we take up the following within-sample prediction exercise which is useful for diagnostic checking purposes. Consider, without loss of generality, obtaining the posterior predictive bilateral trade density between Algeria and Australia in differentiated goods.[23] Algeria, as it turns out, has the lowest value of contract enforcement in our sample (1.44), while Australia has a reasonably large value (3.90). The observed amount of 1992 bilateral trade in differentiated goods between these countries was US \$21 860 000. Using the output of our posterior simulator together with the steps in (26) and (27), we sample from the posterior predictive for trade between these two countries. The results are reported in Figure 4. Since the discrete mass at zero and the continuous density are difficult to depict in a single graph, Figure 4 simply smoothes all simulated values from the posterior predictive into a single density estimate.

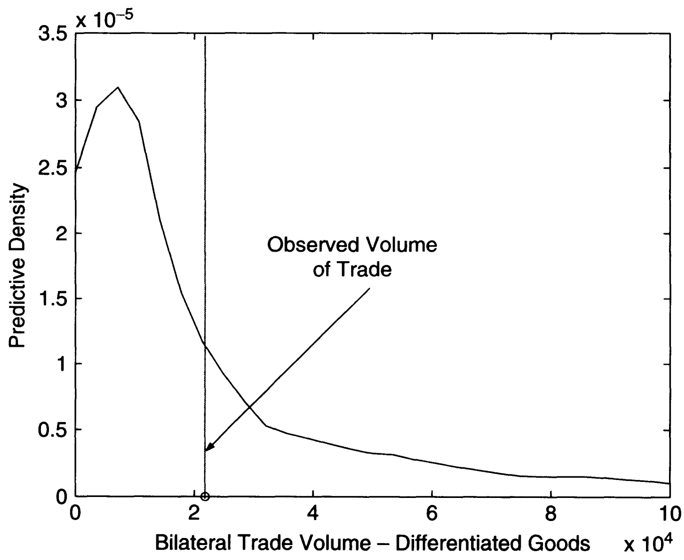

Figure 4. Posterior predictive density of bilateral trade (in 1,000's of 1992 U.S. dollars) in differentiated goods: Algeria and Australia

[23] This just happens to be the first observation in our sorted data.

DOI: 10.1002/jae

As you can see from Figure 4, the posterior predictive places appreciable mass over the actual amount of trade. Using our simulations from the posterior predictive, we find little probability associated with no trade: only 1.3% of the draws were equal to zero. This methodology can be used in an identical way to make predictions regarding the trade patterns of other countries, to conduct out-of-sample policy experiments (e.g., to examine the change in trade volumes in accord with improvements in governance quality), to forecast future, out-of-sample trade volumes (given the covariates' values) and to perform diagnostic checking regarding the model specification. For example, we repeated this process for a small set of randomly selected countries, and in these cases found that the posterior predictives placed considerable mass around the observed trade volumes. These exercises do not prove that the model is in fact 'true' (nor are these checks exhaustive), but at the same time, they neither point to any specific deficiencies associated with the model nor do they provide support that its predictions are inconsistent with the observed data.

As a final example, the largest amount of bilateral trade in our data occurred, not surprisingly, between the USA and Canada. The posterior predictive mean of trade in differentiated goods between these countries was approximately 160 billion, which is reasonably close to the observed amount of 145 billion. The fact that the generalized gravity model appears to perform well in the prediction of both small and large trade volumes adds support to the credibility of the model and its assumptions.

4.3. Does Contract Enforcement Matter Most for Differentiated Goods?

In the first section of our empirical investigation, we tested for the significance and linearity of contract enforcement. Another important question in our analysis is: Does contract enforcement matter more for the trade of differentiated than homogeneous goods? Said differently, we might ask ourselves if the slope associated with the contract enforcement variable is steepest in the trade of differentiated goods, suggesting that bilateral trade volumes are most responsive to changes in contract enforcement for differentiated products. While such a test is easily carried out when the model employs a linear specification of contract enforcement, its implementation requires a little more thought given our desire to treat the contract enforcement variable non-parametrically.

One possible solution in this regard is to focus on the *average derivative*, $E_{\text{Contract}}[f'(\text{Contract})]$, and compare its value across different types of goods. Given that our model is constructed to directly incorporate prior information on one-sided pointwise slopes of the form

$$f'(\mathcal{C}_j) \approx \frac{\psi_{j+1} - \psi_j}{\mathcal{C}_{j+1} - \mathcal{C}_j}$$

calculating the posterior distribution of this average derivative is straightforward in our analysis. That is, the posterior distribution associated with

$$g(\psi) = \frac{1}{C-1} \sum_{j=1}^{C-1} \frac{\psi_{j+1} - \psi_j}{\mathcal{C}_{j+1} - \mathcal{C}_j}$$

is readily calculable for each type of good.

In Figure 5 we plot the posterior distributions associated with these average derivatives. Consistent with our conjecture and results established in previous theoretical work, contract

DOI: 10.1002/jae

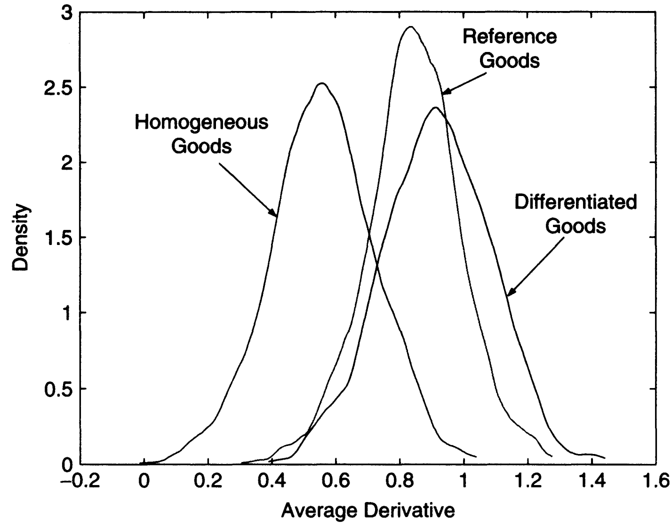

Figure 5

enforcement matters more (using our definition of the average derivative as the metric) for the trade of differentiated than homogeneous goods. In fact, $\Pr[g(\psi^{\mathrm{Diff}}) > g(\psi^{\mathrm{Hom}})|T] = 0.94$. The evidence that contract enforcement matters more for differentiated than reference goods, however, is less clear, as $\Pr[g(\psi^{\mathrm{Diff}}) > g(\psi^{\mathrm{Ref}})|T] = 0.64$. The ordering of these relationships, however, is exactly as we would expect, since reference goods, in the extent of variation of their attributes, fall somewhere in between homogeneous and differentiated goods.[24]

5. CONCLUSION

In this paper we introduced an empirical methodology that can be employed to extend the traditional gravity model, which has been widely used to analyze the volumes of trade between pairs of countries. In particular, we took up the case of a semi-parametric hierarchical threshold tobit model which allowed for country-specific trade effects, and also permitted a covariate to enter the gravity equation non-parametrically.

We used this methodology to investigate the impact of contract enforcement, or the effectiveness of governance, on bilateral trade volumes. Our empirical results, though not 'causal', suggested that measures of contract enforcement were important for describing bilateral trade in all types of goods, and that effective governance mattered most for the trade of differentiated products. The latter result is consistent with previous conjecture and established theoretical results in the literature.

[24] In a related analysis, Rauch (1999) investigates the impact of common language/colonial ties on trade in these different types of goods. He states (p. 9): 'The network/search model should apply most strongly to differentiated products and most weakly to products traded on organized exchanges, with its applicability to other homogeneous products unclear.' This is consistent with the ordering in our analysis, where contract enforcement clearly matters more for differentiated than homogeneous goods, and seems to matter more for differentiated than reference goods. (i.e., 'other homogeneous products').

Our empirical analysis also enabled us to test for the linearity of the contract enforcement variable and to make predictions which remain true to the mixed discrete–continuous nature of actual trade data. Interestingly, estimated functions relating contract enforcement and bilateral trade were often found to exhibit some nonlinearities, and these nonlinearities resembled the mean function of a linear regression model with a single changepoint. Formal comparison of the linear model against the non-parametric and changepoint alternatives, however, were not conclusive, given the posterior uncertainty surrounding the parameters of interest.

Our hope is that the methodology described here can be further refined and applied in other applications of the gravity equation. The model outlined here allows for the inclusion of country-specific effects within a threshold tobit model, which heretofore has not been done in the literature. Further, the ability to conduct a non-parametric analysis can also be used in future research which seeks to explore the relationship between different measures of trade costs and observed trade volumes.

ACKNOWLEDGEMENTS

We would like to thank two anonymous referees, the Associate Editor Herman van Dijk and seminar participants at Iowa State University for helpful comments and suggestions which significantly improved this paper. All errors, of course, remain our own.

REFERENCES

Anderson JE, Marcouiller DS. 2002. Insecurity and the pattern of trade: an empirical investigation. *Review of Economics and Statistics* **84**(2): 342–352.
Anderson JE, van Wincoop E. 2003. Gravity with gravitas: a solution to the border puzzle. *American Economic Review* **93**(1): 170–192.
Anderson JE, van Wincoop E. 2004. Trade costs. *Journal of Economic Literature* **43**(3): 691–751.
Chib S. 1992. Bayesian inference in the tobit censored regression model. *Journal of Econometrics* **51**: 79–99.
Eaton J, Tamura A. 1994. Bilateralism and regionalism in Japanese and US trade and direct foreign investment patterns. *Journal of the Japanese and International Economies* **8**: 478–510.
Geweke J. 2005. *Contemporary Bayesian Econometrics and Statistics.* Wiley: Hoboken, NJ.
Koop G, Poirier DJ. 2004. Bayesian variants of some classical semiparametric regression techniques. *Journal of Econometrics* **123**: 259–282.
Lancaster T. 2004. *An Introduction to Modern Bayesian Econometrics.* Blackwell: Oxford.
McCallum J. 1995. National borders matter: Canada–US regional trade patterns. *American Economic Review* **85**(3): 615–623.
North D. 1990. *Institutions, Institutional Change and Economic Performance.* Cambridge University Press: Cambridge, UK.
Poirier DJ. 1995. *Intermediate Statistics and Econometrics: A Comparative Approach.* MIT Press: Cambridge.
Ranjan P, Lee J. 2003. Contract enforcement and the volume of international trade in different types of goods. Mimeo, University of California–Irvine.
Rauch J. 1999. Networks versus markets in international trade. *Journal of International Economics* **48**: 7–35.
Rauch J, Trindade V. 2002. Ethnic Chinese networks in international trade. *Review of Economics and Statistics* **84**(1): 116–130.
Schiller R. 1973. A distributed lag estimator derived from smoothness priors. *Econometrica* **41**: 775–788.
Verdinelli I, Wasserman L. 1995. Computing Bayes factors using a generalization of the Savage–Dickey density ratio. *Journal of the American Statistical Association* **90**: 614–618.