On the Endogeneity of International Trade Flows and Free Trade Agreements

Scott L. Baier
The John E. Walker Department of Economics
Clemson University
Clemson, SC 29634
and the Federal Reserve Bank of Atlanta
Atlanta, GA 30309
sbaier@clemson.edu

Jeffrey H. Bergstrand*
Department of Finance and Business Economics
Mendoza College of Business and Kellogg Institute for International Studies
University of Notre Dame
Notre Dame, IN 46556
bergstrand.1@nd.edu

Abstract

For 40 years, the gravity equation has been a workhorse for cross-country empirical analyses of international trade flows and – in particular – the effects of free trade agreements (FTAs) on trade flows. However, the gravity equation is subject to the same econometric critique as early cross-industry studies of U.S. tariff and nontariff barriers and U.S. multilateral imports: trade policy is *not* exogenous. The potential endogeneity of FTA binary right-hand-side variables motivates developing a theory of FTA determination compatible with the theory of international trade flows. Allowing econometrically for the FTA variable's potential endogeneity yields striking empirical results: the effect of FTAs on trade flows is *quadrupled*. Moreover, in the context of our estimation framework which accounts for selection bias, the results suggest that countries that select into FTAs have "chosen well."

August 2002

*Corresponding author. An earlier draft was presented at the American Economic Association annual meeting, January 2002, in Atlanta.

On the Endogeneity of International Trade Flows and Free Trade Agreements

> Trade theorists continue to puzzle over their surprisingly small estimates of the impact of trade liberalization on imports. All explanations of the puzzle treat trade liberalization as a given. But the level of trade protection is not exogenous. (Trefler, 1993, p. 138)

> The gravity equation constitutes . . . one of the most important results about trade flows. (Evenett and Keller, 2002, p. 282)

> The issue of exogeneity may also be an important problem when dummy variables are used to estimate the effects of free trade areas (in a gravity equation). . . . If we find a large coefficient on a particular free trade area (dummy variable), is that an indication the agreement has strong effects or simply that the countries that have formed the agreement have *chosen well* ? (Lawrence, 1998, p. 59)

For 40 years, the "gravity equation" has been used in international trade to study the effects of free trade agreements (FTAs) and customs unions on bilateral merchandise trade flows. The gravity equation typically explains cross-sectional variation in country pairs' trade flows in terms of the countries' incomes, populations, bilateral distance, and dummy variables for common languages, for common land borders, and for the presence or absence of an FTA. Eichengreen and Irwin (1997, p. 142) note that the gravity equation has "long been the workhorse for empirical studies of the pattern of trade."

Nobel laureate Jan Tinbergen (1962) was the first to provide an econometric study using the gravity equation for international trade flows (including dummies for FTAs), and the study has spawned over one hundred similar empirical analyses. Yet, since Tinbergen trade economists typically have found a mix of significant and insignificant coefficient estimates for the binary variables representing FTAs in the gravity equation. For instance, Tinbergen found a significant effect for membership in the British Commonwealth, but an insignificant effect for membership in the Benelux FTA (after accounting for the influences of incomes and bilateral distance). Aitken (1973), Abrams (1980), and Brada and Mendez (1983) found the EC to have an economically and statistically significant effect on trade flows among members, whereas Bergstrand (1985) and Frankel, Stein and Wei (1995) found insignificant effects. Frankel (1997) found positive significant effects from Mersosur and insignificant effects from the Andean Pact. In a pooled regression, Frankel found a significant *negative* effect from membership in the EC.[1] Thus, empirical support for whether FTAs have a positive effect on trade flows is still quite mixed.

However, all these studies typically assume an *exogenous* RHS dummy variable to represent the FTA "treatment effect," that is, the effect of an FTA on the bilateral trade flow. In reality, FTA dummies are not exogenous random variables; rather, countries likely select endogenously into FTAs for reasons possibly related to the level of trade. In the related endogenous U.S. trade policy literature, for instance, Trefler (1993) addressed systematically the simultaneous determination of U.S. multilateral imports and U.S. multilateral nontariff barriers in

[1] Frankel (1997) and Oguledo and MacPhee (1994) provide summaries of FTA coefficient estimates across studies. In 1992, the European Union commissioned a major study to analyze the potential enlargement of the union into central and eastern Europe. That study, now Baldwin (1994), used the gravity equation as the primary empirical tool.

a cross-industry analysis. Trefler found that, after accounting for the endogeneity of trade policies, the effect of policies on U.S. imports increased tenfold. Clearly, the literature on bilateral international trade flows and bilateral FTAs is subject to the same critique that Trefler raised: the presence or absence of an FTA is *not exogenous*. The issue is important because – if FTAs are endogenous – previous cross-section and panel empirical analyses of the effects of FTAs on trade flows may be biased and inconsistent, and the effects of FTAs on trade may be seriously over- or under-estimated. Put quite simply, the issues of endogeneity and selection bias in FTA binary variable coefficients have *never* been addressed before in the international trade gravity equation literature.

In the end, we intend this paper to contribute to the literature in three ways. First, despite a voluminous number of empirical studies using the gravity equation in international trade, no study has *articulated* clearly the possible reasons for estimation bias in the effect of an FTA on trade flows due to the potential endogeneity of this RHS variable.[2] We delineate the potential sources of this endogeneity, with an emphasis on selection bias. Thus, our paper complements a recent articulation of the potential bias introduced in gravity models by endogenous multilateral price resistance terms, cf., Anderson and van Wincoop (2001). Second, we provide systematic empirical evidence that previous estimates of the effects of FTAs on international trade flows over the past forty years have been considerably *under*-estimated by about 75 percent. Thus, our paper complements such cross-industry tariff- and nontariff-policy studies as Trefler (1993) and Lee and Swagel (1997) that showed previous estimates of the impact of trade liberalization on imports had been considerably underestimated. Third, using the "treatment effects" methodology, we answer Lawrence's (1998) question stipulated above – but in a surprising way. The literature on treatment effects, developed in the context of numerous labor economics studies (cf., Heckman, 2001), provides rich tools that have not previously been used for analyzing the effects of bilateral trade policies on international bilateral trade flows. The treatment-effects methodology allows us to demonstrate that a large coefficient estimate for an FTA dummy variable indicates that the FTA has a large effect on bilateral trade *and* that the two countries that have formed an FTA have "chosen well."

Section I provides motivations for suspecting endogeneity of FTA dummies in the typical gravity equation. Section II discusses the econometric issues associated with estimating average treatment effects with potentially endogenous FTAs. If FTA binary variables are endogenous, one needs a theory of FTAs to understand their determinants (to form instruments). Section III summarizes a theoretical framework for understanding the "pure economic" determinants of FTAs in a manner consistent with economic factors determining the pattern and volume of international trade. In constructing this framework, we generate a new theoretical foundation for the gravity equation most commonly estimated and provide a numerical analysis of the effects of various economic characteristics on a social planner's decision to form an FTA. Section IV addresses the data requirements. Section V provides the main empirical results, a sensitivity analysis, and interpretation. Section VI concludes.

[2]We will show that the sole empirical attempt to allow for simultaneity bias, in particular, which ignored the econometric theory of dummy endogenous variables, produced estimates of a logically-inconsistent model.

I. On the Potential Endogeneity of the FTA Binary Variable

The gravity equation in international trade estimated most commonly using cross-country data is:

$$PX_{ij}^{g} = \beta_0 (GDP_i)^{\beta 1} (GDP_j)^{\beta 2} (POP_i)^{\beta 3} (POP)^{\beta 4} (DIST_{ij})^{\beta 5} e^{\beta 6(LANGij)} e^{\beta 7(ADJij)} e^{\beta 8(FTAij)} \varepsilon_{ij} \quad (1)$$

or

$$SPX_{ij}^{g} = \beta_0 (GDP_i\ GDP_j)^{\beta 1} (POP_i\ POP_j)^{\beta 2} (DIST_{ij})^{\beta 3} e^{\beta 4(LANGij)} e^{\beta 5(ADJij)} e^{\beta 6(FTAij)} \varepsilon_{ij} \quad (2)$$

where PX_{ij}^{g} is the value of the merchandise trade flow imported by country j from exporter i, GDP_i (GDP_j) is the level of nominal gross domestic product in country i (j), POP_i (POP_j) is population in country i (j), $DIST_{ij}$ is the distance between the economic centers of countries i and j, $LANG_{ij}$ is a binary variable assuming the value 1 if countries i and j share a common language and 0 otherwise, ADJ_{ij} is a binary variable assuming the value 1 if countries i and j share a common land border and 0 otherwise, FTA_{ij} is a binary variable assuming the value 1 if countries i and j have a free trade agreement and 0 otherwise, e is the natural logarithm base, and ε_{ij} is assumed to be a log-normally distributed error term.[3] Alternative specification (2) examines the total trade between the country pair, where SPX_{ij}^{g} is usually the value of the total bilateral trade between countries i and j (i.e., $PX_{ij}^{g} + PX_{ji}^{g}$).

Typically, the effects of FTAs (or customs unions) on trade flows between member countries are estimated by the coefficient on the FTA binary variable. Linnemann (1966), Aitken (1973), and Bergstrand (1985, 1989), for example, estimated the effects of the European Community (EC) and the European Free Trade Association (EFTA) on trade flows among members. Sapir (1981) estimated the effects of the Generalized System of Preferences (GSP) on trade flows from qualified exporters to developed economies. In the 1980s and 1990s, a plethora of empirical trade studies have used the gravity equation, summarized in Baldwin (1994), Oguledo and MacPhee (1994), and Frankel (1997). As discussed in the introduction, coefficient estimates on FTAs have often been positive but statistically insignificant, and even negative. For instance, using a representative sample of 53 countries,[4] an estimate of gravity equation (2) is:

$$\begin{aligned} \ln(SPX_{ij}^{g}) = &\underset{(-0.13)}{-0.22} + \underset{(28.17)}{1.15} \ln(GDP_i\ GDP_j) - \underset{(-5.61)}{0.28} \ln(POP_i\ POP_j) - \underset{(-16.64)}{0.95} \ln(DIST_{ij}) \\ &+ \underset{(6.91)}{0.66} (LANG_{ij}) + \underset{(4.31)}{0.66} (ADJ_{ij}) + \underset{(1.67)}{0.18} (FTA_{ij}) \end{aligned} \quad (3)$$

$R^2 = 0.8485$; RMSE = 1.0199; n = 1378

Our results are consistent with the literature summarized in the introduction. All coefficient estimates have signs and values consistent with earlier studies. Similarly, the coefficient estimate for the FTA dummy is statistically

[3] The standard gravity equation sometimes includes the exporter and importer per capita GDPs instead of populations, which naturally alters their coefficients' estimates and interpretations. We will address this issue later in the empirical results.

[4] Data sources can be found in section IV.

significant at the 10 percent level, but not at the 5 percent level.

A standard problem in cross-sectional empirical work is the potential endogeneity of the RHS variables. If any of the RHS variables in, say, equation (2) is correlated with the error term, ε_{ij}, that variable is considered endogenous and ordinary least squares (OLS) may yield biased and inconsistent estimates of the coefficients.

Casual inspection of equation (2) reveals that several RHS variables are plausibly exogenous and some are plausibly endogenous. For instance, geographic factors such as distance and countries' adjacency can be considered exogenous for representative economic agents. Moreover, economists typically assume countries' populations are exogenous, and it is reasonable to assume language is exogenous. However, GDP – a function of exports and imports – is potentially endogenous to bilateral trade flows. Moreover, as emphasized in Trefler's introductory quote, trade policies are not exogenous.

In this paper, we deal primarily with the endogeneity of trade policies, and only secondarily with that of national incomes. Three reasons surface for largely ignoring potential endogeneity of incomes in this paper. First, GDP is a function of *net* exports, not gross exports. Typically, net exports tend to be less than 5 percent of a country's GDP. Second, the gravity equation relates *bilateral* trade flows to the countries' incomes. Trade between any pair of countries tends to be a very small share of any country's exports, much less its GDP. Third, Frankel (1997) and others have previously accounted for the potential endogeneity of national incomes econometrically using instruments including labor forces and stocks of human and physical capital. Frankel (1997) reported that coefficient estimates in gravity equations change insignificantly using these IV techniques, and concluded "Evidently, the endogeneity of income makes little difference" (p. 135).[5]

However, no study has addressed systematically the potential endogeneity of FTA dummies. Potential sources of endogeneity of RHS variables generally fall under the categories of measurement error, omitted variables, and simultaneity bias (cf., Wooldridge, 2002). Consider, in turn, the potential bias caused by each source. Measurement error in an explanatory variable, such as an FTA dummy, is generally associated with negative bias (in absolute terms) in the variable's coefficient. For instance, with the classical "errors-in-variables" assumption, the 0-1 FTA dummy variable (*FTA*) would be correlated positively with the measurement error (ς) if the true trade-policy variable (say, the tariff rate, t) was assumed uncorrelated with ς ($\varsigma = FTA - t$). In equation (2)'s context, the correlation between *FTA* and ε would be negative, leading to the classical "attenuation bias" of *FTA*'s coefficient estimate toward zero.[6] This suggests that OLS estimates of *FTA*'s effect have been systematically underestimated.

Second, consider the potential endogeneity created by omitted variables bias (in particular, selection bias). Countries' governments generally choose to enter into an FTA based upon economic welfare of consumers, political considerations, or both. As in the labor economics literature on treatment effects, actual participation is voluntary.

[5] Nevertheless, in the robustness analysis in section V, we will account for any potential endogeneity of GDPs.

[6] Even without the classical errors-in-variable assumption, the correlation between *FTA* and ε is likely negative. Suppose the true trade policy variable is the bilateral tariff rate, t, where $t > 0$. If an FTA exists, $FTA = 1$, $t = 0$, and ς consequently equals 1. If no FTA exists, $FTA = 0$, $t > 0$, and consequently $\varsigma < 1$. Thus, *FTA* and ς are positively correlated.

Consequently, the FTA dummy variable may be endogenous by being correlated with unobservable (omitted) variables that are correlated also with the decision to trade. For instance, error term ε_{ij} in equation (2) may be representing unobservable policy-related bilateral barriers influencing trade between two countries that are not accounted for by standard gravity equation RHS variables. If two countries have high tariffs at the border and strong domestic nontariff policy regulations, there will be little trade. An agreement that mutually eliminates tariffs only will not cause trade, but an agreement also encompassing liberalization of domestic regulations will tend to enhance trade. The likelihood of the two countries' governments selecting into (or forming) an FTA is high since there is a large potential welfare gain from bilateral trade creation if the FTA also broadens liberalization beyond tariff barriers. This reason suggests that ε and *FTA* are negatively correlated, and the *FTA* coefficient will tend to be underestimated.

In support of this argument, numerous authors have noted that one of the major benefits of increased "regionalism" is the potential for "deeper integration." Lawrence (1996, p. xvii) distinguishes between "international policies" that deal with border barriers, such as tariffs, and "domestic policies" that are concerned with everything "behind the nation's borders, such as competition and antitrust rules, corporate governance, product standards, worker safety, regulation and supervision of financial institutions, environmental protection, tax codes ..." and other national issues. The GATT and WTO have been remarkably effective in the postwar era reducing border barriers. However, these institutions have been much less effective in liberalizing the domestic policies just itemized. As Lawrence states it, "Once tariffs are removed, complex problems remain because of differing regulatory policies among nations" (p. 7). He argues that in many cases, FTA "agreements are also meant to achieve deeper integration of international competition and investment" (p. 7). Gilpin (2000) echos this argument: "Yet, the inability to agree on international rules or to increase international cooperation in this area has contributed to the development of both managed trade *and regional arrangements*" (p. 108; italics added). Preeg (1998) concludes:

> [Free] trade agreements over time, however, have tended to include a broader and broader scope of other trade-related policies. This trend is a reflection, in part, of the fact that as border restrictions [tariffs] are reduced or eliminated, other policies become relatively more important in influencing trade flows and thus need to be assimilated in the trade relationship (p. 50).

Third, consider the potential endogeneity created by simultaneity. Trefler's (1993) cross-industry study of U.S. multilateral imports and trade barriers and Lee and Swagel's (1997) cross-industry and cross-country study of trade barriers and trade flows addressed this issue. In these two empirical studies, policy-induced (nontariff) trade barriers were expected to have a negative impact on imports. However, imports (specifically, more import penetration or a higher import/output ratio) were expected to have a positive effect on trade barriers, due to political pressures. In the context of a simultaneous-equations system, the consequence of ignoring this simultaneity is an underestimated impact of trade barriers on imports. However, the potential simultaneity bias need not be negative, nor must it appeal to a political economy framework. As Lawrence's (1998) quote above suggests, if FTAs are endogenous, a large coefficient on an FTA dummy variable may indicate that the FTA has a large effect *or* that countries that trade large volumes have "chosen well." The natural inclination is to estimate – analogous to Trefler (1993) and Lee and Swagel (1997) – a system of simultaneous equations treating bilateral trade and FTAs as endogenous. Unfortunately, as established in Heckman (1978) and Maddala (1983), we will show that the conventional approach yields a system that is logically inconsistent. However, using the treatment-effects methodology, we can provide evidence that – while FTA coefficient estimates have been *under*-estimated systematically – countries that have formed an agreement have, on average, "chosen well."

As is well known, the standard econometric approach to address the problem of an endogenous RHS variable is instrumental variables, cf., Angrist and Krueger (2001) and Wooldridge (2002). However, as discussed in Heckman (1997), Vella and Verbeek (1999), and Wooldridge (2002, Ch. 18), in the presence of the selection bias issues noted above instrumental variables may need to be augmented with control functions. We discuss these econometric issues in the next section.

II. Econometric Issues

Methodological issues regarding the cross-section estimation of the partial effects of a binary variable (such as *FTA*) on a continuous variable (such as trade flows) fall under the "treatment effect" literature in econometrics. The "average treatment effect" refers to the notion that the trade flow between two countries will differ depending upon whether the countries share an FTA or not. The fundamental econometric dilemma is that one can only observe one situation or the other. The treatment effect literature addresses several methodological issues related to how to estimate the effect of the "treatment" (the presence of an FTA) on the trade flow values. This issue has *never* been addressed in forty years of estimating the effects of FTAs on international trade flows. An excellent summary of recent developments in the treatment-effect literature is in Heckman (2001) and Wooldridge (2002).

The average treatment effect (ATE) of an FTA between a country pair is defined as:

$$ATE \ / \ E(spx_1 - spx_0 * FTA) \tag{4}$$

where spx_1 (spx_0) denotes the logarithm of the sum of the trade flows between countries i and j with (without) the FTA and E denotes the expectation operator. We assume that observations (spx) are independently and identically distributed across country pairs; this assumption ensures that the treatment of one pair does not affect another pair's trade flow.[7]

As previous discussion suggests, we assume that the level of trade between two countries depends upon an

[7] This is a strong assumption because it suggests that an FTA between countries i and j does not have a trade-diverting effect on other bilateral relations, such as countries i and m or j and n. However, we maintain this assumption in this paper in order to engage this methodology, concerning ourselves in another paper with the secondary estimation effects of violating this assumption.

array of exogenous "covariates" other than the treatment – the standard set of gravity equation variables in levels or log levels (GDPs, populations, bilateral distance, adjacency, common language, and possibly remoteness) – which we denote by $\boldsymbol{q}$:

$$ATE(\boldsymbol{q}) \equiv E(spx_1 - spx_0 \mid \boldsymbol{q}, FTA) \tag{5}$$

In section III, we rationalize theoretically factors composing $\boldsymbol{q}$. We can also estimate the average treatment effect on countries that have formed an FTA:

$$TTE(\boldsymbol{q}, FTA=1) \equiv E(spx_1 - spx_0 \mid \boldsymbol{q}, FTA=1) \tag{6}$$

This is known as the "treatment effect on the treated" (TTE), but TTE will only be discussed much later.

Consistent estimation of the average treatment effect depends upon the assumptions made about relationships among the variables. For instance, a common assumption in this literature is termed the "ignorability of treatment" assumption. This assumption is that FTA, spx_1, and spx_0 are independent, conditioned on $\boldsymbol{q}$. Under this assumption, the average treatment effect simplifies to the difference in the conditional means of outcomes spx_1 and spx_0 :

$$\begin{aligned} ATE(\boldsymbol{q}) &\equiv E(spx_1 - spx_0 \mid \boldsymbol{q}, FTA) \equiv E(spx_1 \mid \boldsymbol{q}, FTA) - E(spx_0 \mid \boldsymbol{q}, FTA) \\ &= E(spx_1 \mid \boldsymbol{q}) - E(spx_0 \mid \boldsymbol{q}) \end{aligned} \tag{7}$$

However, $E(spx_1 \mid \boldsymbol{q}, FTA)$ does not likely equal $E(spx_1 \mid \boldsymbol{q})$ because the decision to enter an FTA likely depends on more factors than typically controlled for in the gravity equation. Consequently, the ignorability-of-treatment assumption seems inappropriate.

Since one can only observe a trade flow in the presence *or* absence of an FTA, we define the observed outcome (spx) for a country pair (i, j) as:

$$spx \equiv (FTA)\, spx_1 + (1 - FTA)\, spx_0 \tag{8}$$

where FTA = 1 if an FTA exists between the pair, and 0 otherwise. As noted in Heckman (1997) and Wooldridge (2002), instrumental variables (IV) provides a method of consistent estimation of the average treatment effect of an FTA in the presence of endogeneity, but only under certain assumptions. To illustrate, assume trade flows spx_0 and spx_1 have the standard linear form as in a gravity equation:

$$spx_0 = \mu_0 + \boldsymbol{\beta_0}'\boldsymbol{q} + \varepsilon_0 \tag{9}$$
$$spx_1 = \mu_1 + \boldsymbol{\beta_1}'\boldsymbol{q} + \varepsilon_1 \tag{10}$$

Substituting equations (9) and (10) in equation (8) yields:

$$spx = \mu_0 + \boldsymbol{\beta_0}'\boldsymbol{q} + \alpha FTA + \varepsilon_0 + FTA(\boldsymbol{\beta_1}'-\boldsymbol{\beta_0}')(\boldsymbol{q} - \bar{\boldsymbol{q}}) + FTA(\varepsilon_1 - \varepsilon_0) \tag{11}$$

where $(\boldsymbol{q} - \bar{\boldsymbol{q}})$ denote the demeaned values of $\boldsymbol{q}$ so that α corresponds to the average treatment effect. Consistent estimation of $\boldsymbol{\beta_0}$ and α depends upon the correlation of: (1) *FTA* with error term ε_0; (2) *FTA* with differences of economic factors influencing trade flows for partners with FTAs versus those without FTAs $(\boldsymbol{q} - \bar{\boldsymbol{q}})$; and (3) *FTA* with differences in unobservables for partners with FTAs versus partners without FTAs $(\varepsilon_1 - \varepsilon_0)$. In the case where *FTA* is uncorrelated with all three factors, $\boldsymbol{\beta_0}$ and α are estimated consistently using OLS.

Consider now several alternative cases. First, consider where *FTA* is correlated with ε_0, but *FTA* is uncorrelated with $(\boldsymbol{q} - \bar{\boldsymbol{q}})$ and $(\varepsilon_1 - \varepsilon_0)$. This is the classic case of an endogenous RHS variable, arising potentially from measurement error, omitted variables, or simultaneity. Instrumental variables (IV) can be applied to generate consistent estimates of $\boldsymbol{\beta_0}$ and α, where the instrument for *FTA* is determined in a first-stage, based upon some economic model. If *FTA* is correlated with ε_0 and $(\boldsymbol{q} - \bar{\boldsymbol{q}})$, but not with $(\varepsilon_1 - \varepsilon_0)$, once again IV can be used to generate consistent estimates of $\boldsymbol{\beta_0}$, $\boldsymbol{\beta_1}$ and α.

Finally, suppose *FTA* is correlated with ε_0, $(\boldsymbol{q} - \bar{\boldsymbol{q}})$, and $(\varepsilon_1 - \varepsilon_0)$. As Heckman (1997) suggests, instrumental variables only identifies the parameters consistently when economic agents' decisions to select into a "program" are unrelated to (or ignore) unobservable factors influencing the outcome. However, as discussed earlier, many trade-policy observers have noted that unobservable policies tending to inhibit trade, such as nontariff barriers and domestic regulations, *may* be one of the main reasons governments have selected into FTAs; the intent is that FTAs will lead to reductions in domestic barriers that multilateral agreements have been unable to attain. In the context of this study, instrumental variables will not yield consistent estimates in the presence of selection bias, that is, if the unobservable component of economic factors influencing the decision to form an FTA are correlated with unobservable economic factors influencing trade flows. In this instance, we apply Heckman's procedure to control for selection.

To demonstrate the potential relevance of alternative sources of endogeneity, we provide three alternative ways of estimating equation (11). As a benchmark, we estimate the standard gravity equation by OLS. Second, we show how a standard instrumental variables technique suggests that the average treatment effect of an FTA has been considerably underestimated owing to the conventional sources of measurement error, omitted variables, and simultaneity. Third, we show how the Heckman approach suggests evidence of selection bias and illustrates how estimated treatment effects vary according to whom is "treated."

For our IV approach, we use multi-stage estimation of equation (11) to generate unbiased and consistent estimates of the gravity equation. First, assume $(\boldsymbol{\beta_1} - \boldsymbol{\beta_0}) = (\varepsilon_1 - \varepsilon_0) = 0$. Second, assume there exists a set of "instrumental variables" $\boldsymbol{z}$ that are uncorrelated with ε_0, i.e., $E(\varepsilon_0 * \boldsymbol{q}, \boldsymbol{z}) = E(\varepsilon_0 * \boldsymbol{q})$. Third, assume that this set of variables $\boldsymbol{z}$ is significantly related to *FTA*, i.e., $E(FTA * \boldsymbol{q}, \boldsymbol{z}) \neq E(FTA * \boldsymbol{q})$. Then 2SLS estimation of equation (11) yields consistent, asymptotically normal estimates of α and $\boldsymbol{\beta_0}$.

The addition of two further assumptions yields an even more efficient IV estimator. First, assume the probability of *FTA* (P*(FTA)*) can be estimated by a known parametric form (such as probit) such that $P(FTA = 1 * \boldsymbol{q},$

$\boldsymbol{z}) = \Phi(\pi_0 + \boldsymbol{\pi_1}'\boldsymbol{q} + \boldsymbol{\pi_2}'\boldsymbol{z})$ where $P(FTA = 1 * \boldsymbol{q}, \boldsymbol{z}) \ldots P(FTA = 1 * \boldsymbol{q})$.[8] Second, assume the variance of ε_0 is a constant. This suggests a multi-step IV method with the following steps: (i) estimate a binary response model, such as probit, $P(FTA = 1 * \boldsymbol{q}, \boldsymbol{z}) = \Phi(\pi_0 + \boldsymbol{\pi_1}'\boldsymbol{q} + \boldsymbol{\pi_2}'\boldsymbol{z})$ by maximum likelihood to generate predicted probabilities Φ^P; and (ii) estimate equation (11) by IV, using instruments 1, Φ^P, and $\boldsymbol{q}$. The estimator is consistent, asymptotically efficient, and the usual 2SLS standard errors and test statistics are asymptotically valid. We will refer to this instrumental variables approach as our IV technique.[9]

In our final approach, we relax the assumptions $\boldsymbol{\beta_1} = \boldsymbol{\beta_0}$ and $\varepsilon_1 = \varepsilon_0$. First, relaxing $\boldsymbol{\beta_1} = \boldsymbol{\beta_0}$ allows the coefficient estimates for the "treated" and "untreated" to differ. A consistent and efficient IV estimator results by "de-meaning" the set of variables in $\boldsymbol{q}$. The first step stays the same. The second step in the previous IV estimator is modified by first adding to the existing instruments (1, Φ^P, and $\boldsymbol{q}$) a set of interactive terms $[\Phi^P(\boldsymbol{q} - \bar{\boldsymbol{q}})]$ where $\boldsymbol{q} - \bar{\boldsymbol{q}}$ are the de-meaned values of $\boldsymbol{q}$. Given the similarity to the previous IV technique, for brevity we omit estimating this here. However, if ε_1 and ε_0 are allowed to differ also, we must account for selection bias by using instead Heckman's (1978) technique and introduce control functions.

Intuitively, think of the sample as decomposable into two subsets, spx_1 and spx_0. Define $E(spx_1)$ and $E(spx_0)$ as:

$$E(spx_1 * \boldsymbol{q}, \boldsymbol{z}, FTA{=}1) = \mu_1 + \boldsymbol{\beta_1}'\boldsymbol{q} + E(\varepsilon_1 * \boldsymbol{q}, \boldsymbol{z}, FTA{=}1) \tag{12}$$

and

$$E(spx_0 * \boldsymbol{q}, \boldsymbol{z}, FTA{=}0) = \mu_0 + \boldsymbol{\beta_0}'\boldsymbol{q} + E(\varepsilon_0 * \boldsymbol{q}, \boldsymbol{z}, FTA{=}0) \tag{13}$$

Then the expectation of equation (8) becomes:

$$E(spx * \boldsymbol{q}, \boldsymbol{z}) = FTA\ E(spx_1 * \boldsymbol{q}, \boldsymbol{z}, FTA{=}1) + (1\text{-}FTA)\ E(spx_0 * \boldsymbol{q}, \boldsymbol{z}, FTA{=}0) \tag{14}$$

or

$$\begin{aligned} E(spx * \boldsymbol{q}, \boldsymbol{z}) = {} & \xi + \alpha\, FTA + \boldsymbol{\beta_0}'\boldsymbol{q} + \boldsymbol{\delta}'[FTA(\boldsymbol{q} - \bar{\boldsymbol{q}})] \\ & + FTA\ E(\varepsilon_1 * \boldsymbol{q}, \boldsymbol{z}, FTA{=}1) + (1\text{-}FTA)\ E(\varepsilon_0 * \boldsymbol{q}, \boldsymbol{z}, FTA{=}0) \end{aligned} \tag{15}$$

where $\boldsymbol{\delta}' = (\boldsymbol{\beta_1}'\text{-}\boldsymbol{\beta_0}')$. Drawing on our theoretical model to be discussed shortly, FTA can be interpreted as an indicator function, $FTA[\ ]$, where $FTA[\ ]$ equals 0 or 1. Define fta as a latent variable representing the difference in utility levels for the two countries' consumers from the formation of an FTA. $FTA[fta > 0] = 1$ and 0 otherwise. As will be shown, the gain or loss in utility can be shown to be a function of numerous economic characteristics. Assume $FTA[fta = \pi_0 + \boldsymbol{\pi_1}'\boldsymbol{q} + \boldsymbol{\pi_2}'\boldsymbol{z} + v \ \$\ 0]$ where v represents unobservable factors influencing the governments' FTA decision, $E(v * \boldsymbol{q}, \boldsymbol{z}) = 0$, and v is distributed Normal (0,1). If we assume v, ε_1 and ε_0 are trivariate normally distributed, then:

---

[8] If a probit function, then $\Phi(\ )$ is the standard normal cumulative distribution function.

[9] Wooldridge (2002) refers to this as a two-step estimator. However, for clarity, we refer to three steps. The first stage is the estimation of the predicted probabilities, Φ^P. The second stage is a linear regression of FTA on a constant, Φ^P, and $\boldsymbol{q}$. The third stage is estimation of the gravity equation substituting the predicted values from the second-stage regression for FTA.

$$\mathrm{E}(spx * \boldsymbol{q}, \boldsymbol{z}) = \xi + \alpha\, FTA + \boldsymbol{\beta_0}'\boldsymbol{q} + \boldsymbol{\delta}'[FTA(\boldsymbol{q} - \bar{\boldsymbol{q}})] + \rho_1\, FTA[\varphi(\boldsymbol{q}, \boldsymbol{z}; \boldsymbol{\pi}) / \Phi(\boldsymbol{q}, \boldsymbol{z}; \boldsymbol{\pi})] - \rho_0(1\text{-}FTA)[-\varphi(\boldsymbol{q}, \boldsymbol{z}; \boldsymbol{\pi})/\{1 - \Phi(\boldsymbol{q}, \boldsymbol{z}; \boldsymbol{\pi})\}] \quad (16)$$

where $\varphi(\boldsymbol{q}, \boldsymbol{z}; \boldsymbol{\pi}) = \varphi(\pi_0 + \boldsymbol{\pi_1}'\boldsymbol{q} + \boldsymbol{\pi_2}'\boldsymbol{z})$ is the standard normal probability density function, $\Phi(\boldsymbol{q}, \boldsymbol{z}; \boldsymbol{\pi}) = \Phi(\pi_0 + \boldsymbol{\pi_1}'\boldsymbol{q} + \boldsymbol{\pi_2}'\boldsymbol{z})$ is the standard normal cumulative distribution function, and ρ_1 and ρ_0 are parameters (positively defined). $FTA[\varphi(\,)/\Phi(\,)]$ and $(1\text{-}FTA)[-\varphi(\,)/\{1-\Phi(\,)\}]$ serve as control functions (positively defined) for the unobservable heterogeneity and potential selection bias, and will occasionally be referred to as "hazard rates.".

This analysis suggests a two-step estimator. First, estimate π_0, $\boldsymbol{\pi_1}$, and $\boldsymbol{\pi_2}$ from a probit of *FTA* on $\boldsymbol{q}$, $\boldsymbol{z}$, and a constant to form predicted probabilities Φ^P along with predicted φ^P for all observations. Second, estimate by OLS:

$$spx = \xi + \alpha\, FTA + \boldsymbol{\beta_0}'\boldsymbol{q} + \boldsymbol{\delta}'[FTA(\boldsymbol{q} - \bar{\boldsymbol{q}})] + \rho_1\, FTA\, [\varphi^P / \Phi^P] - \rho_0\, (1\text{-}FTA)[-\varphi^P/(1 - \Phi^P)] + u \quad (17)$$

where u is a normally distributed error term. We will refer to this estimation technique as the "Heckman" procedure. The coefficients will be consistently estimated; see Heckman (1997), Vella and Verbeek (1999), and Wooldridge (2002, Ch. 18). In the next section, we provide a theoretical model to suggest variables to include in $\boldsymbol{q}$ and $\boldsymbol{z}$.

III. Theoretical Issues

Angrist and Krueger (2001) remind us that good instruments "often come from detailed knowledge of the economic mechanism and institutions determining the regressor of interest" (p. 73). Thus, an economic model of determinants of FTAs is needed. In modeling the determinants of FTAs, we note the reasonable standard articulated in Trefler (1993):

> The ideal trade model from which to derive the estimating equation for imports must satisfy two criteria. First, it must predict the pattern of trade. Second, it must be compatible with the theory of endogenous protection (p. 142).

In this section, we summarize a model of international trade that satisfies both of Trefler's standards. The theoretical model yields a gravity equation that "predicts the pattern of trade." Then, we show how the same model that predicts trade flows is "compatible" with a theory of endogenous FTA policy. Using some additional assumptions for tractability, we calibrate the model to determine economic factors in the model that tend to lead countries' policymakers to form an FTA.

As the "ideal" trade model should predict both the pattern of trade *and* the pattern of FTAs, we construct an economic model that synthesizes aspects of the trade volume and the FTA literatures. First, a theoretical gravity equation for trade flows has typically been derived in a competitive setting, either perfect or monopolistic competition. In fact, the results in Evenett and Keller (2002) recently "underscore the importance of both relative factor abundance and IRS (increasing returns to scale) as determinants of the extent of specialization and international trade." However, Evenett and Keller ignore intra- and inter-continental transport costs. As economic geography influences trade levels *and* most FTAs are intra-continental, the model should recognize explicitly these costs.

Second, FTAs are selected into by governments. Governments likely weigh both economic welfare and political special interests in determining whether or not to form (or enforce) an FTA. However, two recent econometric studies based upon welfare-maximizing theoretical frameworks indicate that consumer welfare tends to dominate political interests in U.S. trade-policy decisions. Goldberg and Maggi (1999) found "the weight of [consumer economic] welfare in the government's objective function is many times larger than the weight of [political] contributions" (p. 1135). Specifically, they estimated the weight of consumer welfare (political contributions) in U.S. government tariff decisions to be 98 (2) percent. Gawande and Bandyopadhyay (2000) estimated a comparable weight of 99 percent. Thus, for two reasons we assume each country's government chooses to form an FTA based solely upon maximizing its respective agent's welfare. This is consistent with recent empirical evidence and is consistent with our trade model, which assumes competitive markets; in a competitive setting, there is no role for lobbying activities of firms. (Later, in the sensitivity analysis of the empirical results, we do, however, consider political-economy factors.)

This section is comprised of three parts. In part A, we summarize a general equilibrium model of world production, consumption, and trade. In part B, we show how a typical gravity equation can be derived from the model, suggesting variables for inclusion in $\boldsymbol{q}$. In part C, we illustrate in the model's context economic factors influencing governments' decisions to form an FTA, suggesting variables for $\boldsymbol{z}$.

A. The Model

The model constructed is a generalization of the bilateral trade flow models of Anderson (1979), Helpman and Krugman (1985), Bergstrand (1985, 1989, 1990), Baier and Bergstrand (2001), Head and Ries (2001), and Evenett and Keller (2002), and of FTA behavior examined in Krugman (1991a, b), Frankel (1997), and Frankel, Stein, and Wei (1995, 1996, 1998). Because the bulk of international trade appears to be inter-industry trade driven by relative factor abundances and intra-industry trade in differentiated goods produced under IRS in imperfectly competitive markets, we provide a model of world trade with two industries, two factors, differentiated products produced under increasing returns to scale, and multiple countries with international transportation costs. International trade within each of the two monopolistically-competitive sectors is generated by the interaction of consumers having tastes for diversity and production being characterized by economies of scale. We assume two factors of production, capital and labor, each perfectly mobile between sectors within a country but each immobile internationally.

We label the two sectors goods and services. The main reason is that the gravity equation is typically estimated using *goods* (merchandise) trade flows. Hence, we must distinguish between "goods" trade and other trade; the natural distinction is goods versus services. Data on bilateral trade flows in services is not yet widely available for analysis, so the empirical work cannot be extended for services trade. Initially in the theoretical analysis, the two sectors are assumed symmetric. However, later in the analysis we will differentiate the two sectors along one conventional (Balassa-Samuelson) line: goods (services) will be assumed to be capital (labor) intensive in production.[10] The monopolistic-competition framework is standard in modeling international trade in goods (e.g., manufactures) in the context of the new trade theory.[11] Within each sector, a taste for diversity exists, captured formally by Dixit-Stiglitz preferences. Increasing returns to scale internal to the firm are captured with a fixed cost term (say, representing marketing and distribution costs) absorbing an exogenous amount of both factors.

1. Consumers

Each country has a representative consumer who derives utility from consuming goods and services (g and s, respectively) based upon Cobb-Douglas preferences. Within each sector, the consumer's taste for diversity is captured formally by Dixit-Stiglitz preferences. Thus, the representative consumer in each country has the nested utility function:

$$U_j = [(\Sigma_{i=1}^{N}\Sigma_{k=1}^{n_i^g} g_{ijk}^{\theta^g})^{1/\theta^g}]^{\gamma}\ [(\Sigma_{i=1}^{N}\Sigma_{k=1}^{n_i^s} s_{ijk}^{\theta^s})^{1/\theta^s}]^{1-\gamma} \qquad j = 1,...,N \tag{18}$$

where U_j denotes the utility of the representative household in country j. Let g_{ijk} be household consumption in country j of (differentiated) good k produced in country i. Similarly, s_{ijk} is household consumption in country j of (differentiated) service k produced in country i. Let θ^g (θ^s) denote the parameter determining the elasticity of substitution in consumption in goods (services) with $0 < \theta^g, \theta^s < 1$. Let γ ($1-\gamma$) be the Cobb-Douglas preference parameter for goods (services). Finally, let n_j^g (n_j^s) be the number of varieties of goods (services) produced in country j.

Within any country, households and firms are assumed symmetric; hence, we may replace the notation $\Sigma_{k=1}^{n_i^g}$ ($\Sigma_{k=1}^{n_i^s}$) by n_i^g (n_i^s), and suppress subscript k. Consequently, the budget constraint for the representative consumer in country j is:

$$w_j + r_j(K_j/L_j) + T_j = \Sigma_{i=1}^{N} n_i^g p_{ij}^g(1+t_{ij}^g)g_{ij} + \Sigma_{i=1}^{N} n_{ij}^s p_{ij}^s(1+t_{ij}^s)s_{ij} \tag{19}$$

where w_j is the wage rate of the representative consumer-worker (or household) in country j, r_j is the rental rate on capital per household, K_j/L_j is the amount of capital exogenously supplied (or endowed) per household, T_j is tariff revenue redistributed back to households in a lump sum, p_{ij}^g (p_{ij}^s) is the c.i.f. price in country j of the good (service)

[10] For empirical evidence supporting this assumption, see Bergstrand (1991).

[11] Sapir and Winter (1994) note that "Most service sectors operate under conditions of imperfect competition resulting from various degrees of market power on the part of producers" (p. 277).

produced in country i, and t_{ij}^{g} (t_{ij}^{s}) is the *ad valorem* tariff rate by country j on the good (service) produced in country i; assume $t_{jj} = 0$.

Following most recent models of trade, we assume that the international transportation costs of shipping goods and services can be modeled using Samuelson-type "iceberg" costs. Let δ_{ij} represent the fraction of output exported by country i that is "consumed" (or lost) due to international transport to country j; hence, only a portion (1- δ_{ij}) arrives. It will be useful to define c_{ij} / $\delta_{ij}/(1-\delta_{ij})$ so that $(1-\delta_{ij})$ / $1/\{1 + [\delta_{ij}/(1-\delta_{ij})]\}$ / $1/(1 + c_{ij})$; assume $\delta_{jj} = c_{jj} = 0$. With positive transport costs, the price level of the representative good of country i in country j, p_{ij}^{g} is:

$$p_{ij}^{g} = p_{i}^{g}/(1-\delta_{ij}^{g}) = p_{i}^{g}(1 + c_{ij}^{g}) \quad (20)$$

In the following, symmetric equations hold for services but are omitted for brevity. Tariff rates and transport costs are allowed to differ between sectors and country pairs.

For each country's consumer, maximizing (18) subject to equation (19) yields:

$$g_{ij} = \gamma \, (py_{j} / P_{j}^{g})[p_{ij}^{g} \, (1+t_{ij}^{g})/P_{j}^{g}]^{-\sigma g} \quad (21)$$

where py_{j} is nominal income of the representative consumer in j, $\sigma^{g} = 1/(1-\varphi^{g})$, and

$$P_{j}^{g} = \{G_{k=1}^{N} \, n_{k}^{g}[p_{k} \, (1+c_{kj}^{g}) \, (1+t_{kj}^{g})]^{1-\sigma g}\}^{1/(1-\sigma g)} \quad (22)$$

P_{j}^{g} has been referred to in the literature as a "multilateral resistance term," interpreted as an output-weighted measure of the remoteness (in terms of trade costs) of country j.

2. Firms

Each firm in the goods industry in each country is assumed to produce subject to the technology:

$$g_{i} = z_{i}^{g}(k_{i}^{g})^{\alpha g}(l_{i}^{g})^{1-\alpha g} - \varphi^{g} \quad (23)$$

where g_{i} denotes output of the representative firm in country i, z_{i}^{g} is an exogenous productivity term for goods producers, k_{i}^{g} is the amount of capital used by the representative firm, l_{i}^{g} is the amount of labor used by the representative firm, and φ^{g} represents a fixed cost facing each firm (absorbing both capital and labor), the latter assumed identical across countries for simplicity. Factor intensities α^{g} and α^{s} (the latter in the symmetric services production function) will be allowed to differ.

Firms in each industry in each country maximize profits subject to the technology just defined, given market demand schedules implied in the previous section. Equilibrium in these types of models is characterized by two conditions. First, profit maximization ensures that prices are a markup over marginal production costs. Second, under monopolistic competition firms earn zero profits. As common to these models, output of the firm in each industry is determined parametrically. However, we can show that the numbers of firms in the goods industry in i is endogenous:

$$n_i^g = (\sigma^g \varphi^g)^{-1}\, z_i^g (K_i^g)^{\alpha g} (L_i^g)^{1-\alpha g} \tag{24}$$

3. Factor Endowment Constraints

As is standard, we assume that endowments of capital (K_i) and labor (L_i) are exogenous, with both factors internationally immobile. Assuming full employment:

$$K_i = K_i^g + K_i^s = n_i^g\, k_i^g + n_i^s\, k_i^s \tag{25}$$
$$L_i = L_j^g + L_i^s = n_i^g\, l_i^g + n_i^s\, l_i^s \tag{26}$$

4. Equilibrium

The factor employments and prices in each industry and country, consumption of each good, and product prices can be determined uniquely given parameters of the model ($\gamma, \theta^g, \theta^s$, α^g, α^s, φ^g, φ^s) and initial tariffs (t_{ij}^g, t_{ij}^s), transport costs (c_{ij}^g, c_{ij}^s), and factor endowments (K_i, L_i) for i, j = 1, ..., N.

B. Trade Flow Determinants (***q***)

The model allows us to solve for a "gravity" equation for the trade flow from country i to country j in each industry. Empirically, international trade economists have studied merchandise (goods) trade flows as these are measured accurately through customs; bilateral trade flows in services have not been studied empirically due to measurement difficulties. Aggregating equation (21) across all L_j consumers in country j, noting that the representative profit-maximizing firms in country i in the goods industry will set the prices of their products delivered to market j according equation (20), yields the *equilibrium* trade flow for each goods-producing firm in country i to market j:

$$x_{ij}^g = \gamma\, [(GDP_j + T_j L_j)/P_j^g][p_i^g(1+c_{ij}^g)(1+t_{ij}^g)/P_j^g]^{-\sigma g} \tag{27}$$

where GDP is nominal gross domestic product and $\sigma^g = 1/(1-\theta^g)$. Since we assume $\theta^g < 1$, then $\sigma^g > 1$.

Since the gravity equation usually estimates the cross-sectional relationship between the *c.i.f. value* of aggregate merchandise trade between two countries and its determinants, we multiply both sides of equation (27) by p_{ij}^g and n_i^g. With some algebraic manipulation (details in Appendix A), we can find:

$$PX_{ij}^g = [\gamma / \varphi^g(\sigma^g\text{-}1)]\; (GDP_i / p_i^g)\; GDP_j\; [p_i^g(1+ c_{ij}^g)/P_j^g]^{1-\sigma g}\; \{s_i^g\, [(1+ t_j)(1+ t_{ij}^g)^{-\sigma g}]\} \tag{28}$$

where s_i^g is the real share of goods output in national product in country i ($s_i^g = n_i^g g_i/[n_i^g g_i + (p_i^s/p_i^g) n_i^s s_i]$), t_j is the share of tariff revenue relative to income ($t_j = T_j L_j / GDP_j$). Theoretical equation (28) has become a standard way of expressing the gravity equation in the presence of transportation costs and tariffs, cf., Feenstra (2003, Ch. 5). With the exception of the last RHS term in { }, the equation is identical to Feenstra's equation (26).[12]

Since our empirical analysis in this study focuses upon bilateral trade policy and trade flows, we are interested in the gravity equation for the sum of the two countries' trade flows, SPX_{ij}. First, as in most studies, we assume that bilateral transaction costs are symmetric, i.e., $c_{ij}^g = c_{ji}^g$ and $t_{ij}^g = t_{ji}^g$. Second, in most countries, tariff revenue is a trivial share of GDP; hence, we assume $t_i = t_j = 0$. Using equation (28) and its analogue for PX_{ji}^g, we can derive:

$$SPX_{ij}^g = [\gamma /\varphi^g(\sigma^g\text{-}1)]\ GDP_i\ GDP_j\ (1+ c_{ij}^g)^{1-\sigma g}\ (1+ t_{ij}^g)^{-\sigma g}\ [s_i^g(p_i^g)^{-\sigma g}\ (P_j^g)^{1-\sigma g} + s_j^g(p_j^g)^{-\sigma g}\ (P_i^g)^{1-\sigma g}] \tag{29}$$

The RHS last term in equation (29) is cumbersome. Anderson and Van Wincoop (2001) have shown using an implicit solution technique for such a system that $p_i^g{*} = (GDP_i^g/n_i^g\ GDP_W^g)^{1/(1-\sigma g)}/P_i^{g*}$, where GDP_W^g is world output of goods, and $p_i^g{*}$ and $P_i^g{*}$ denote the implicit solutions for p_i^g and P_i^g, respectively (see Appendix B). Substituting $p_i^g{*}$ into equation (28), the analogue for $p_j^g{*}$ into the analogue for equation (28), and summing PX_{ij}^g and PX_{ji}^g yields:

$$SPX_{ij}^g = [\gamma /GDP_W^g\ \varphi^g(\sigma^g\text{-}1)]\ GDP_i\ GDP_j\ (s_i^g + s_j^g)\ (1+ c_{ij}^g)^{-(\sigma g-1)}\ (1+ t_{ij})^{-\sigma g}\ (P_i^g{*}\ P_j^g{*})^{\sigma g-1} \tag{30}$$

where $P_i^g{*}$ and $P_j^g{*}$ are interpreted generally as "multilateral price resistance terms," estimated in Anderson and van Wincoop using a customized nonlinear least squares procedure.

Equation (30) resembles remarkably closely gravity models (2) and (3). It provides a rationale for including the logarithm of the product of nominal GDPs in $\boldsymbol{q}$. The term $(1+c_{ij}^g)$ provides a rationale for including bilateral distance, a dummy variable for border adjacency, and a dummy variable for common language in $\boldsymbol{q}$. The term $(s_i^g + s_j^g)$ provides a rationale for including the average of the per capita GDPs (or, as in equation (2), populations) of the two countries as proxies for the average capital-labor endowment ratios of the two countries (which determine the endogenous shares of goods in national output).[13] Hence, $\boldsymbol{q}$ will include nominal GDPs,

---

[12]A key difference of solving for gravity equation (28) versus earlier papers is that it is derived from a two-sector, two-factor model of world trade, whereas most gravity equations are derived in a single-industry setting (except for Bergstrand, 1989). Consequently, equation (28) includes the share of national output in goods, s_i^g, which is a function of the capital-labor endowment ratio; thus, equation (28) provides a rationale for including exporter per capita GDP in the standard gravity equation. Furthermore, our model allows for lump-sum redistributions of tariffs to consumers. This explains the presence of $(1+t_j)$ in the last RHS term; in a single-industry setting (and assuming MFN duties, or $t_{ij} = t_j$, for all i partners), tariff rates would enter in the same form as in Feenstra (2003). Thus, the model here generalizes that in Baier and Bergstrand (2001) and Feenstra (2003) to a two-sector, two-factor world.

[13]Populations (or alternatively per capita incomes) now have a straightforward interpretation in the gravity equation. Higher populations (for given incomes) reduce the capital-labor endowment ratios of the two countries, tending to reduce the capital-intensive industry's share of national output in both countries. If goods are capital intensive, then goods trade should fall (relative to national outputs).

populations, bilateral distance, an adjacency dummy, and a language dummy.[14]

The gravity equation's error term would be interpreted, in the model's context, as the product of the multilateral resistance term for the country pair. Recent applications of the gravity model introduce measures of multilateral resistance (cf., Helliwell (1998) and Anderson and van Wincoop (2001)) or fixed effects (cf., Rose and van Wincoop (2001) and Feenstra (2003)) to avoid an omitted variables bias. Anderson and van Wincoop (2001) and Feenstra (2003) show that both approaches yield consistent estimates in a gravity equation, and both studies illustrate empirically that the simpler fixed-effects approach yields similar estimates to the more complex nonlinear least squares procedure.

C. Economic Determinants of FTAs (***z***)

In this section, we use the theoretical model to help identify a plausible set of instruments $\boldsymbol{z}$ for the first-stage probit regression, $P(FTA = 1^* \boldsymbol{q}, \boldsymbol{z}) = \Phi(\pi_0 + \boldsymbol{\pi_1}'\boldsymbol{q} + \boldsymbol{\pi_2}'\boldsymbol{z})$, to generate predicted probabilities Φ^P. Econometrically, the factors composing $\boldsymbol{z}$ need to be correlated with FTA, but uncorrelated with the error term in the gravity equation.

First, we show how various economic factors likely influence the decision by a pair of governments to form an FTA. As discussed earlier, we assume that each country's government acts as a social planner, in light the competitive setting described above and empirical evidence. Thus, the social planner for each country should decide whether or not to form an FTA based upon the net welfare gains (or losses) of that decision.[15] For a bilateral FTA to be formed, it must be the case that the change in utility is positive for *both* countries' agents. If the change in utility is negative for either country, we assume an FTA is not formed. The greater the increase in utility for the agents in each country from an FTA, the more likely that each country's government would enter an FTA.

The economic rationale for estimating the first-stage probit regression is based on the qualitative choice model of McFadden (1975, 1976). A qualitative choice model can be derived from an underlying latent variable model. As in section II, let *fta* denote an unobserved (or latent) variable, where for simplicity we ignore the observation subscript. Let *fta* in the present context represent the difference in utility levels from an action (the

---

[14]Note that a specification even closer to (2) evolves if we use the sum of the logarithms of the two countries trade flows, rather than the log of the sum. In this case, all terms *including* s_i^g and s_j^g enter multiplicatively. Since the correlation in our sample between the sum of the logs and the log of the sum is over 90 percent, the results will be qualitatively identical and nearly identical quantitatively.

[15]We note that we do not formulate any explicit coalition-formation structure to determine tariff levels; this is beyond the scope of this primarily empirical paper. We assume that the formation of an FTA is consistent with the social planner's objectives. In reality, the gains (losses) from an FTA may be sensitive to the initial level of tariffs. In the simulations that will follow, we do not solve for the "optimal" tariffs. Instead, following Frankel (1997, pp. 167-168), we use initial tariff levels of 0.30; see Frankel (1997) for empirical justification. Moreover, in reality the Nash equilibrium pre-integration and post-integration tariff levels are likely to differ, suggesting that the familiar "inverse-elasticity" formula for designating tariffs would be inappropriate. Addressing this limitation is also beyond the scope of this paper, especially due to the emphasis here on asymmetric economies with intra- and inter-continental transport costs; most papers solving for Nash equilibrium tariffs benefit from an assumption of symmetric economies.

formation of an FTA), where:

$$fta = \pi_0 + \boldsymbol{\pi_1}'\boldsymbol{q} + \boldsymbol{\pi_2}'\boldsymbol{z} + v \qquad (31)$$

where $\boldsymbol{q}$ and $\boldsymbol{z}$ are vectors of explanatory variables, $\boldsymbol{\pi_1}$ and $\boldsymbol{\pi_2}$ are vectors of parameters, and error term v is assumed to be independent of $\boldsymbol{q}$ and $\boldsymbol{z}$ and to have a standard normal distribution. In the context of our model, formally fta = min(ΔU_i , ΔU_j). Hence, *both* countries' consumers need to benefit from an FTA for their representative countries to form one.

Since *fta* is unobservable, *FTA* serves as an indicator variable, which takes the value 1 if two countries have a free trade agreement (indicating $fta > 0$), and 0 otherwise (indicating fta # 0). We can derive the response probability, P, for *FTA* as:

$$P(FTA = 1 * \boldsymbol{q}, \boldsymbol{z}) = P(fta > 0 * \boldsymbol{q}, \boldsymbol{z}) = \Phi(\pi_0 + \boldsymbol{\pi_1}'\boldsymbol{q} + \boldsymbol{\pi_2}'\boldsymbol{z}) \qquad (32)$$

where Φ( ) is the standard normal cumulative distribution function, which ensures that P(*FTA* = 1) lies between zero and unity. The standard errors of the estimates of $\boldsymbol{\pi_1}$ and $\boldsymbol{\pi_2}$ are asymptotically normally distributed. Thus, standard z-statistics are reported in Table 1 in section V and indicate whether estimates of $\boldsymbol{\pi_1}$ and $\boldsymbol{\pi_2}$ are statistically significant.

1. Variables in $\boldsymbol{q}$

The first-stage probit regression includes variables in $\boldsymbol{q}$. In this section, we show the linkage theoretically between some variables in $\boldsymbol{q}$ and latent variable *fta*. This will establish expected qualitative statistical relationships between variables in $\boldsymbol{q}$ and P(*FTA*), i.e., signs for $\boldsymbol{\pi_1}$.

The complexity of the model described in section A precludes deriving comparative statics associated the net welfare gains of an FTA. In order to establish relationships between economic factors and the net welfare gains or losses from an FTA, we need to narrow the general model described in Section A – which has N countries and N(N-1)/2 potential transport costs – to a limited set of countries on a limited set of continents to conduct feasible simulations with a computable version of our general equilibrium model (CGE). For simulation purposes, following Baier and Bergstrand (2002) we assume only three continents (1, 2, 3) with only two countries (A, B) on each continent (i.e., 1A, 1B, 2A, 2B, 3A, 3B). We assume an identical intra-continental transport-cost factor between any two countries on the same continent (a^g, a^s) and an identical inter-continental transport-cost factor between any two continents (b^g, b^s); the differing transport costs between countries on the same and different continents suggest that countries on the same continent are "closer" in distance. In other respects, the model is identical to that in part A.[16] The remainder of the equations that describe the model are found in Appendix C.

[16]Note these transport costs are of the hub-and-spoke variety discussed in Frankel, Stein, and Wei (1995) where each continent represents a hub. For intercontinental shipments, costs are broken down into two components. The cost of transporting a good (service) from one hub to another is given by b^g (b^s) and the cost to distribute the good (service) to each

(a) Distance

Frankel, Stein, and Wei (1995, 1996, and 1998) showed theoretically in a world with symmetric economies that two countries that are "natural" trading partners (i.e., close in distance) will benefit more from an FTA than two countries that are "unnatural" partners (i.e., far apart). Assuming initially that the two countries are identical in all respects,[17] Figure 1 illustrates clearly for any values of intra- or inter-continental transport costs that the consumer welfare effects of a *natural* FTA exceed (or equal) those from an *unnatural* FTA. The top (bottom) surface is the net welfare gain from a natural (unnatural) FTA. This suggests – in a more general context – that the gains from an FTA are greater the smaller the distance between two countries due to more trade creation.[18] This suggests a negative relationship between distance and the probability of an FTA. Moreover, distance should be uncorrelated with the error term in the gravity equation (ε_{ij}) as distance is presumed exogenous in the typical gravity equation.

(b) Nominal GDPs

Intuitively, the welfare gains from FTAs should be higher for countries with larger absolute factor endowments (and consequently larger real GDPs). The formation of an FTA between two (economically) large partners creates trade in more varieties than an FTA between two small partners, and diverts trade from nonmembers in fewer varieties than two small partners, improving utility more in large countries relative to small ones. To illustrate, we examine the relationship between economic size and the net welfare gain from an FTA first for natural trading partners and then for unnatural partners. In the first exercise, we allow countries on continent 1 (1A, 1B) to have larger absolute endowments of capital and labor than countries on continent 2 (2A, 2B), and countries 2A and 2B to have larger absolute endowments than countries on continent 3 (3A, 3B). Countries on the same continent have identical factor endowments. Assume for now relative factor endowments in every country are identical.

Figure 2a illustrates that welfare gains from natural FTAs on all three continents are monotonically higher the larger the endowments and real GDPs of the countries. In Figure 2a, the plane for continent 1 is unambiguously above that for continent 2, and the plane for continent 2 is unambiguously above that for continent 3.

Analogous reasoning applies to an unnatural FTA, the second exercise. If the two unnatural partners have a higher average absolute factor endowment, they will enjoy relatively higher net gains from an FTA. The second

---

spoke is a^g (a^s). Transportation costs of shipping goods (services) intra-continentally only consist of the cost of shipping the good from spoke-to-spoke.

[17] Assume identical absolute and relative factor endowments, multi-factor productivity terms, and parameters γ, σ^g, σ^s, α^g, α^s, φ^g, and φ^s with $\sigma^g = \sigma^s$, $\alpha^g = \alpha^s$, $\varphi^g = \varphi^s$, $a^g = a^s$, and $b^g = b^s$.

[18] It is useful to note that Figure 1 is a generalization of the two-dimensional "Figure 2" in Frankel, Stein, and Wei (1995), where they assumed $a = 0$. In fact, FSW's "Figure 2" is a special case of Figure 1, evaluated at a equal to zero; this special case is shown in our Figure 1 by the plane relating "% Change in Welfare" to "Intercontinental T.C. Factor."

exercise reorients absolute factor endowments so that countries 1A and 2A (on different continents) have the largest endowments, countries 1B and 3A have the medium endowments, and countries 2B and 3B have the smallest endowments. In Figure 2b, the top (bottom) plane represents the welfare benefits to an unnatural FTA for the two largest (smallest) countries.

Thus, larger exporter and importer real GDPs should be associated with a higher likelihood of an FTA. Since the typical gravity equation uses *nominal* rather than real GDPs (likely due to estimation with bilateral trade flow *values*), we include in $\boldsymbol{q}$ the (sum of the logs of the) two countries' nominal GDPs (denoted *NGDP*). Just as exporter and importer GDPs are assumed exogenous in a typical gravity equation, it is reasonable to assume GDPs are uncorrelated with ε_{ij}.

2. Variables in $\boldsymbol{z}$

For econometric purposes, the variables composing $\boldsymbol{z}$ must be correlated with P(*FTA*), but uncorrelated with the error term in the gravity equation (ε_{ij}).

(a) Remoteness of Continental FTA Partners

A major theoretical contribution of the CGE in Frankel, Stein, and Wei (1995, 1996, 1998) is that the net welfare gain from a continental (or natural) FTA is positively related to the remoteness of the FTA partners from the ROW. Frankel, Stein, and Wei also found a positive relationship for unnatural trading partners; however, we will show that the relationship does not generalize as well.

As discussed earlier, a smaller bilateral distance between a country pair causes an FTA to have a larger potential welfare gain; this is due to greater potential trade creation between the pair. Analogously, two continental trading partners facing higher intercontinental transport costs tend to benefit more from an FTA as there will be less trade diversion with non-continental trading partners. This is confirmed in Figure 2a. For any intracontinental transport cost factor, the higher are intercontinental transport costs, the greater the net welfare gain. However, Figure 2b illustrates that the relationship between intercontinental transport costs and the net welfare gains from an FTA for *unnatural* partners may be positive, negative, or even quadratic.[19] Thus, our model suggests only a monotonic relationship between remoteness and the net gains from an FTA for *continental* FTA partners.

The theory suggests that the likelihood of an FTA is related to a variable not typically included in the gravity equation, *CONTREMOTE*. This variable is defined as the remoteness of a pair of countries if they have a continental FTA, and zero otherwise. Since this variable is constructed from the interaction of a dummy for sharing

[19]Consider, for instance, the case of large factor endowments and initially low intercontinental transport costs (*b*). With low *b*, the large trade creation from an FTA among large unnatural partners producing many varieties more than offsets the trade diversion with smaller closer countries on the continent. However, as *b* increases, the trade diversion effect grows, diminishing the unnatural FTA's gains until they eventually reach zero because of no FTA partner trade. By contrast, with small unnatural FTA partners, at low *b* trade diversion offsets trade creation, but as *b* increases the net loss is eliminated as FTA partner trade dwindles to zero.

a continent with a measure of physical remoteness (defined in section IV), it is reasonable to assume that $CONTREMOTE_{ij}$ is uncorrelated with ε_{ij}, especially in a gravity equation including multilateral resistance terms discussed in section A.

(b) Relative Factor-Endowment Differences

In traditional trade models (ignoring transport costs), the benefits of an FTA between a pair of countries should be enhanced the wider their relative factor endowments because traditional comparative advantages would be exploited more fully. Figure 3 illustrates the relationship between differences in capital-labor endowment ratios and the net welfare benefits from a natural FTA for the capital-abundant country (1A).[20] At high intercontinental transport costs, the relationship is positive and monotonic. At high values of *b*, there is little intercontinental trade. Consequently, most variety is exchanged intra-continentally. As relative factor endowments widen, both countries specialize more in the industry where they have comparative advantages and enjoy more net welfare gains from an FTA. There is little loss of variety (trade diversion) for 1A as relatively capital-abundant 1A produces more goods and less services, but relatively labor-abundant 1B produces more services and less goods. However, at low intercontinental transport costs, the net gains from an FTA increase at first with wider relative factor endowments, but eventually decline. At low values of *b*, there is considerable intercontinental trade in goods and services. As relative factor endowments widen, country 1A gains initially from specialization in goods. Yet, at high levels of specialization, 1A relies more on intra- and inter-continental trade to meet its demand for varieties of services. Hence, at low values of *b*, a natural FTA causes considerable trade diversion of services intercontinentally. Thus, with increasing specialization, the net welfare gains from inter-industry trade are eventually offset by the net trade diversion due to intra-industry trade.[21] A qualitatively similar relationship holds for an unnatural FTA (not shown, for brevity). Consequently, the relationship between the net benefits from an FTA and wider relative factor endowments is likely positive, but possibly quadratic (if intercontinental transport costs are very low).

We measure relative factor endowment differences as the absolute value of the difference in the logs of

[20] Up to this point, asymmetries have been introduced to the CGE between *economies*, but not between *industries*; a limitation of the model up to this point is the assumption of only one factor and one industry. We now introduce different parameters in the two industries' production functions. Following most work, we assume "goods" are capital-intensive and "services" are labor-intensive in production in the spirit of the Balassa-Samuelson model, cf., Kravis and Lipsey (1987, 1988) and Bergstrand (1991, 1992). Following Roland-Holst, Reinhart, and Sheills (1994), we set $\alpha_g = 0.36$ and $\alpha_s = 0.27$ (the capital shares in production). All other consumption and production parameters for the two sectors are assumed identical. Starting with all countries initially having identical capital and labor endowments, we increase the capital stock of country 1A to initiate a difference between the capital-labor endowment ratios of 1A and 1B, but reduce proportionately the productivity of 1A in both sectors (z^g, z^s) to hold 1A's real GDP constant (to suppress scale-economies effects).

[21] The second traditional distinction between goods and services is that goods (services) historically have been relatively more (less) tradable. Services historically have often been regarded as "nontradables." On the other hand, recent advances in communication technology raise questions about services' relative nontradability. An exhaustive examination of the effects of higher or lower relative services inter- and intra-continental transport costs is beyond the scope of this paper. However, for the present section we note that higher intercontinental transport costs for services relative to goods will tend to diminish the initial level of intercontinental trade in services. Consequently, the trade-diversion effects from an FTA in this example will tend to be muted with higher relative intercontinental transport costs for services.

the two countries' capital-labor endowment ratios, denoted *DKL*. We assume *DKL* and its square (*SQDKL*) are uncorrelated with ε_{ij}.

(c) Relative Factor-Endowment Differences with ROW

In the part above, the hypothesis concerns only trade creation between the country pair and the pair's capital-labor ratios. Trade diversion for the pair will be higher the more Heckscher-Ohlin trade is foregone (up to a point). Trade diversion from an FTA between a country pair is less the smaller the difference between the relative factor-endowment ratio of the country pair and that of the ROW. Figure 4 illustrates that, at various levels of *b*, the net welfare gains from an FTA decrease as the capital-labor ratio of the pair of FTA members deviates from that of the ROW. A variable measuring this difference, *DROWKL*, is expected to be negatively related to the likelihood of an FTA. As with *DKL*, *DROWKL* is expected to be uncorrelated with ε_{ij}.

Hence, *CONTREMOTE, DKL*, and *DROWKL* compose $\boldsymbol{z}$. In the empirical results in section V, we will show that these variables are correlated with P(*FTA*) and – using a test of overidentifying restrictions – likely uncorrelated with ε_{ij}.

IV. Data Issues

The first challenge was to create an index of FTAs. Going back to Linnemann (1966) and Aitken (1973), a plethora of international trade studies have measured the presence or absence of an FTA between a pair of countries using a binary variable; see Frankel (1997, Ch. 4) for a thorough survey. Following those studies, the variable FTA_{ij} will have the value 1 for a pair of countries (i,j) with a free trade agreement in 1996, and 0 otherwise. This variable was constructed for the pairings of 53 countries [hence, (53x52)/2 or 1378 pairings] using appendices in Lawrence (1996) and Frankel (1997), and FTAs notified to the GATT/WTO under GATT Article XXIV or the Enabling Clause for developing economies as of November 1999 (WTO, 1999); we included only full (no partial) FTAs and customs unions.

Consider now the variables composing $\boldsymbol{q}$. We calculated 1378 bilateral distances among 53 countries' economic centers. Distances were calculated in nautical miles using U.S. Department of the Navy (1965) for sea distances and Rand McNally (1988) for land distances (the latter multiplied by a standard factor of 2 for the transport-cost differential between land and sea transport; see Bergstrand, 1985). We used Linnemann (1966) for economic centers of countries, as in Bergstrand (1985, 1989). The variable $DIST_{\mathrm{ij}}$ is the natural logarithm of the distance between the economic centers of i and j.

Nominal GDPs (*NGDP*) used in a standard gravity model such as (3) come from World Bank (2001). Population data comes from the same source; the sum of the logs of populations is denoted *POP*.

As suggested in theoretical gravity equation (30), we need to account for multilateral resistance. Put simply, accounting for the roles of multilateral price terms such as p_i^g, p_j^g, P_i^g, and P_j^g in equation (29), or P_i^g* and P_j^g* in equation (30), has always been a difficult issue empirically, as no such data exist. As discussed in Anderson and van Wincoop (2001), proper estimation of P_i^g* and P_j^g* requires a custom nonlinear estimation technique, which is well beyond the scope of the analysis at hand. To focus on the issues at hand, we construct proxies for the two multilateral resistance terms using various elasticities of substitution. We demonstrate later that the results central to our paper are robust to various elasticities of substitution. Moreover, we show later that an alternative procedure to Anderson and van Wincoop (2001) suggested in Rose and van Wincoop (2001) and Feenstra (2003) – using fixed effects – suggests our results are robust to alternative means to account for multilateral resistance. In section V, we employ the multilateral resistance proxies:

$$P_i^g* = \{G^N_{k=1,\, k\ldots i}\ GDP_k\ (Distance_{ki})^{1-\sigma g}\}^{1/(1-\sigma g)} \qquad (33a)$$
$$P_j^g* = \{G^N_{k=1,\, k\ldots j}\ GDP_k\ (Distance_{kj})^{1-\sigma g}\}^{1/(1-\sigma g)} \qquad (33b)$$

Note that equations (33a) and (33b) are analogous to equation (22) normalizing prices and gross tariff rates to unity, substituting GDP_k for n_k^g, and using bilateral distance to proxy for the bilateral c.i.f.-f.o.b. transport factor. We estimate (33a) and (33b) using a conventional range of elasticities between 2 and 6. We will also use fixed effects to capture these terms; fixed effects can also generate consistent estimates.

Consider now the variables composing $\boldsymbol{z}$. We constructed *CONTREMOTE* as:

$$CONTREMOTE_{ij} = DCONT_{ij} \text{ x } \{[\log (\Sigma^N_{k=1,\, k\ldots j}\mathrm{Distance}_{ik}/N\text{-}1) + \log (\Sigma^N_{k=1,\, k\ldots i}\mathrm{Distance}_{jk}/N\text{-}1)]/2\}$$

The interpretation of *CONTREMOTE* is as follows. First, *DCONT* is a binary variable assuming the value 1 if both countries are on the same continent, and 0 otherwise. If two countries (i, j) are on the same continent, *CONTREMOTE* measures the simple average of (the natural logarithms of) the mean distance of country i from all of its trading partners except j and the mean distance of country j from all of its trading partners except i. If two countries (i,j) are on different continents, then *CONTREMOTE* has a value 0. This measure captures the spirit of *b* for natural FTAs because it measures how far two countries on the same continent are from other countries, but it has no value for unnatural trading partners. As discussed in section III, for any given value of intracontinental transport costs (*a*), only the welfare gains from a continental FTA increase monotonically with increases in intercontinental transport costs (*b*).

The two other variables in $\boldsymbol{z}$ are the (absolute value of the) difference in the (logs of the) two countries' capital-labor ratios, *DKL*, and the average difference between each member country's capital-labor ratio with respect to the ROW's average capital-labor ratio. The latter variable – $DROWKL_{ij}$ – is:

$$DROWKL_{ij} = \{|\log [(\Sigma^N_{k=1,\, k\ldots i} K_k)/(\Sigma^N_{k=1,\, k\ldots i} L_k)] - \log [K_i/L_i]| + |\log [(\Sigma^N_{k=1,\, k\ldots j} K_k)/(\Sigma^N_{k=1,\, k\ldots j} L_k)] - \log [K_j/L_j]|\}/2$$

Data on per worker physical capital stocks (all in international dollars) and labor stocks are from Baier, Dwyer, and Tamura (2001), assembled from primary data in Mitchell (1992, 1993, 1995); availability of capital stock data determined the sample of countries.[22]

The only other variable we needed for the empirical analysis was nominal trade flow data. For the trade flow equations, we employ aggregate bilateral trade flow data for 1996 from the IMF's *Direction of Trade Yearbook* statistics to construct *spx*.

V. Empirical Results

This section has six parts. In part A, we present the results of estimating the first-stage probit regression. In part B, we provide the main results for the standard gravity equation, using OLS, IV, and controls for selection bias. Since the results in part B may be sensitive to the absence of measures of multilateral resistance, we provide evidence in part C that the results are robust to inclusion of explicit measures of remoteness and to fixed effects. In part D, using Heckman's methodology, we draw some inferences regarding whether or not countries that have selected into FTAs have, in Lawrence's words, "chosen well." In part E, we estimate treatment effects on the "treated" for several particular FTAs to illustrate how endogeneity, selection bias, and differing gravity coefficient estimates for the treated and untreated may have influenced estimates of the effects of FTAs. In part F, we show that the results are robust to the additional inclusion of numerous "political-economy" variables that might influence the probability of an FTA.

A. First-Stage Probit Results

Table 1 presents the results of estimating the first-stage probit regression. First, several variables in $\boldsymbol{q}$ have coefficient estimates consistent with the theoretical model and are statistically significant. Greater bilateral distance between two countries lowers the probability of an FTA. Country pairs with larger GDPs have a positive relationship with P(*FTA*), suggesting greater trade creation between the pair and greater welfare benefits from an FTA. Dummy variables for adjacency and language are statistically insignificant.

Second, all three variables composing $\boldsymbol{z}$ have coefficient estimates consistent with the theoretical model and are statistically significant. The more remote two *continental* trading partners are from the ROW, the greater the probability of an FTA. A wider bilateral difference in capital-labor ratios tends to increase P(*FTA*), while a larger difference of this ratio with that of the ROW tends to decrease P(*FTA*), consistent with the model.[23]

Third, by two alternative measures there is a reasonably good fit of the probit regression. The pseudo-R^2 is 70 percent; this is the measure used in McFadden (1974), one minus the ratio of the log-likelihood value for the

---

[22] The 53 countries include Algeria, Egypt, Nigeria, South Africa, Hong Kong, Iran, Japan, Singapore, Austria, Belgium, Denmark, France, Germany, Greece, Ireland, Italy, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, United Kingdom, Canada, Costa Rica, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama, United States, Argentina, Bolivia, Brazil, Chile, Colombia, Ecuador, Paraguay, Peru, Uruguay, Venezuela, Bulgaria, Czech Republic, Hungary, Poland, Romania, South Korea, Philippines, Thailand, Indonesia, Australia.

[23] We also ran the model including *SQDKL* based upon the theoretical model, but the coefficient estimate was trivially small and statistically insignificant. See Baier and Bergstrand (2002) for details.

estimated model to that for the model with only an intercept. An alternative measure of fit for probit models is the "percent correctly predicted," discussed in Wooldridge (2002). If the estimated probability of the pair exceeds 0.5, we define a variable *PredFTA* to be one; if the probability is less than 0.5, *PredFTA* takes the value zero. The percentage of times that *PredFTA* matches *FTA* (which equals 1 if an FTA exists, and 0 otherwise) is an alternative measure of goodness of fit. However, Wooldridge notes that it is even more useful to report the percent correctly predicted for *each* of the two possible outcomes, for the following reason. With 1378 country pairs and 285 FTAs, a probit specification of *FTA* on a constant only would result in predicted values of *FTA* of 0.21 for every observation (i.e., 285/1378). In this naive specification, however, *PredFTA* would still match *FTA* for 1093 of the 1378 cases, or 79 percent of the time. Even if the model failed to predict *even one FTA correctly*, this goodness-of-fit measure would suggest a predictive power of 79 percent. Consequently, we report the percent correctly predicted for each of the two possible outcomes. In our sample of 1378 pairs, 285 pairs have an FTA and 1093 pairs do not have an FTA ("No-FTA"). Using the rule described, 222 of the 285 FTAs are predicted correctly, or *78 percent*. Also, 1055 of the 1093 pairs without an FTA are predicted correctly, or *97 percent.*

Thus, using alternative criteria the model appears to have a reasonably good fit, and suggests a useful instrument for the second stage. The results suggest that the instruments in $\boldsymbol{z}$ (chosen based upon the theory) are correlated with P(*FTA*). However, it would also be useful to know that $\boldsymbol{z}$ is uncorrelated with the gravity equation error term ε_o. While this can never be known, under certain circumstances a test of overidentifying restrictions can suggest whether some of the variables in $\boldsymbol{z}$ are uncorrelated with ε_o (see Wooldridge, 2000, p. 484). We address this issue later in section C.

B. Gravity Equation Estimates

Table 2 presents the results of estimating the basic gravity equation using OLS, instrumental variables (IV), and finally the Heckman procedure to control for selection bias. The first specification provides the same gravity equation OLS estimates from section I for equation (3). The second specification uses the IV technique described in section II. The contrast between these two specifications is the *first important result* of this paper: the coefficient estimate for the FTA dummy variable *more than quadruples* relative to the OLS estimate. This suggests that previous gravity equation estimates of the effects of an FTA on trade have been systematically *underestimated* due to endogeneity of the FTA variable. Moreover, note that none of the other gravity equation coefficient estimates changes materially.

The third specification controls for unobserved heterogeneity and allows coefficient estimates for the gravity equation variables to differ depending upon whether the pair has an FTA or not. The contrast between specifications (1) and (3) yields the *second important result* of this paper: the coefficient estimate for the FTA dummy using the Heckman procedure more than quadruples relative to the the OLS estimate, *and* a classical hypothesis test rejects the null of no selection bias. Moreover, the negative coefficient estimate (with statistical significance at 10 percent) for the *FTA*HAZARD* variable can reveal an answer to Lawrence's introductory quote;

we discuss this in section D. The third specification also suggests that coefficient estimates for the nominal GDP and population variables do differ depending upon whether the pair has an FTA or not. The adjusted R^2 for specification (3) compared with that for (1) suggests that specification (3) has the higher explanatory power.

We know from section A that the variables in $\boldsymbol{z}$ are correlated with P(*FTA*). It would also be useful to know that $\boldsymbol{z}$ is uncorrelated with the gravity equation error term. While this can never to known, tests using "overidentifying restrictions" have surfaced in recent years to try to estimate this correlation using residuals, cf., Wooldridge (2000). However, such tests require, as the name suggests, a sufficient number of variables in the instrumental variables regression (in our IV procedure, actually, the second-stage of three stages) to "overidentify" the third stage regression. However, the instruments in our second stage include only Φ^P and $\boldsymbol{q}$; consequently, the third-stage regression is just identified. While this precludes an overidentifying-restrictions test here, we will find in the next section with fixed effects that such a test is feasible.

C. Sensitivity Analysis for "Multilateral Resistance"

As discussed earlier, recent work by Helliwell (1998) and Anderson and van Wincoop (2001) suggests that an important omitted variable in gravity equations has been the multilateral resistance terms described in our theoretical model by P_i^g and P_j^g. As summarized in Feenstra (2003, Ch. 5), there are generally three approaches that have been acknowledged to address these variables. First, Bergstrand (1985, 1989) and Baier and Bergstrand (2001) used price indexes to try to account for variation in the multilateral (price) resistance terms. Second, Anderson and van Wincoop (2001) suggest a procedure to estimate a gravity equation using nonlinear least squares to take into account an implicit solution for these terms. Third, Redding and Venables (2000), Rose and van Wincoop (2001), and Feenstra (2003) have employed fixed effects to deal with these terms.

In this part, we use two alternative techniques to acknowledge these terms. First, we note that the complex nonlinear estimation technique in Anderson and van Wincoop (2001), which accounts for the endogeneity of σ^g in the terms P_i^g and P_i^g, is beyond the scope of this paper. However, we constructed proxies for these multilateral terms using plausible alternative values of σ^g suggested in the literature, ranging between 2 and 6. Second, as suggested in Redding and Venables (2000), Rose and van Wincoop (2001), and Feenstra (2003), we used fixed effects. Feenstra (2003) illustrates that the fixed effects yield consistent and similar parameter estimates.

Table 3 reports the results using OLS, IV, and the Heckman procedure adding a variable *REMOTE2*, which is the logarithm of the product of the variables in equations (33a) and (33b) for $\sigma^g = 2$ (the results for other values of σ^g are qualitatively identical and quantitatively similar but are omitted here for brevity). First, in Table 3 we find that the multilateral resistance term (*REMOTE2*) is significant with the positive sign suggested by theoretical equation (30). Second, comparing the respective columns in Tables 2 and 3, none of the coefficient estimates of the standard variables in $\boldsymbol{q}$ change materially with *REMOTE2*'s inclusion. Third, the OLS estimate for the FTA coefficient is slightly larger with *REMOTE2* than without it. The IV estimate of the FTA coefficient is *twice* the OLS estimate in Table 3. Using the Heckman procedure, the coefficient estimate of *FTA* in Table 3 is *triple* the

OLS estimate.

The reminder of this section addresses estimation of the model using fixed effects. As mentioned earlier, Feenstra (2003) suggests that a simpler alternative to the nonlinear estimation approach of Anderson and van Wincoop (2001) for handling the multilateral resistance terms is fixed effects; a gravity equation with fixed effects yields consistent coefficient estimates (ignoring the endogeneity of FTAs). However, in the presence of fixed effects, probit estimates of the probability of an FTA yield inconsistent estimates, cf., Maddala (1983); this ruled out the Heckman procedure. However, we could use a linear probability model (LPM) to generate the predicted probabilities for the IV technique. A linear probability model with fixed effects yields consistent estimates.

As in Rose and van Wincoop (2001) and Feenstra (2003), the fixed effects introduced are country-specific fixed effects. With such effects, variables such as *NGDP, POP,* and *REMOTE2* (which are the sums of the logs of country-specific variables) are perfectly correlated with the fixed-effect dummies; consequently, these variables are excluded. The remaining gravity variables – distance, adjacency, language, and *FTA* – are bilateral specific and not perfectly correlated with the fixed effects terms. Note that by using fixed effects, we also eliminate potentially endogenous *NGDP* as well as *REMOTE2*.

Table 4 reports the results of estimating the gravity equation with fixed effects using OLS and IV. First, OLS estimates of the gravity equation with fixed effects generally yield similar results for coefficient estimates of standard bilateral-specific RHS gravity variables; compare the first columns in Tables 2 and 4. Second, IV estimation with fixed effects was robust. The coefficient estimate for the FTA dummy (0.54) using IV was double the OLS estimate, and close in value to the estimate in Table 3 using the Heckman procedure (0.65).

Moreover, with the linear probability model, it is now feasible to test the overidentifying restrictions. Following the IV estimation with fixed effects, we obtained the residuals from the gravity equation and regressed these on all the exogenous variables to obtain the R^2 value. The test statistic for the null hypothesis that instruments $\boldsymbol{z}$ are uncorrelated with the residuals is the number of observations (1378) times the R^2 value (0.0025); this is distributed $\chi^2(h=2)$ where h is the number of instrumental variables outside the model (3) minus the number of endogenous explanatory variables (1). The test statistic computed was 3.45. The $\chi^2(2; 0.90)$ is 4.61, implying we could not reject the null of no correlation between $\boldsymbol{z}$ and the gravity equation residuals at even the 10 percent significance level. Thus, we have provided evidence that the variables in $\boldsymbol{z}$ are correlated with the P(FTA) and they are uncorrelated with the gravity equation error term.

D. Have Countries that have Selected into FTAs "Chosen Well"?

The previous two sections provide strong evidence that typical gravity equation estimates of FTA effects, treating the dummies as exogenous, have been underestimated. This is consistent with the evidence for U.S. nontariff barriers in Trefler (1993), as addressed in the introductory quote. But we have not yet addressed Lawrence's (1998) question: "If we find a large coefficient on a free trade area, is that an indication the agreement has strong effects or simply that the countries that have formed the agreement have *chosen well*?" The answer to the question is that – after accounting for endogeneity – a large coefficient estimate indicates that the agreement has strong effects. *However,* in the context of our estimation technique, we can still illustrate that countries that have formed FTAs have "chosen well" – in the sense implied by Lawrence that countries that select into FTAs would trade extensively even if they did not have an FTA. Moreover, we will provide evidence suggesting that the decision to form an FTA is related more to potential "trade creation" than "trade diversion."

First, when reading Lawrence's introductory quote, the natural inclination is to consider the simple simultaneous equations system:

$$spx = \mu_1 + \alpha\, FTA + \boldsymbol{\beta' q} + u_{spx} \tag{34}$$
$$fta = \mu_2 + \lambda\, spx + \boldsymbol{\delta' z} + u_{fta} \tag{35}$$

where *fta* (as defined in section II) is a latent variable representing the utility gain or loss for the representative consumer of an FTA, and the other variables are defined as before. Utility gains or losses from an FTA may be a function of variables in $\boldsymbol{z}$, but also may be a function of *spx*. However, we cannot observe *fta*, so instead we assume the indicator variable, *FTA*, is a function of *fta*, $FTA(fta>0)$ = 1 and zero otherwise, and

$$\mathrm{P}(FTA = 1) = \mathrm{P}(fta > 0) = \Phi(\ )$$

where Φ( ) is the standard normal cumulative distribution function. This is the approach in Magee (2002).

However, as summarized in Maddala (1983, p. 118), for logical consistency either λ or α must equal zero. If $\alpha = 0$, then studying the effect of an FTA on trade is irrelevant. So, logically, λ must equal 0. To see this, substitute equation (34) into (35) to yield:

$$fta = \mu_2 + \lambda\mu_1 + \lambda\alpha\, FTA + \lambda\, \boldsymbol{\beta' q} + \boldsymbol{\delta' z} + \lambda u_{spx} + u_{fta} \tag{36}$$

If *fta* > 0, then *FTA* = 1 implying $\mu_2 + \lambda\mu_2 + \lambda\alpha + \lambda\, \boldsymbol{\beta' q} + \boldsymbol{\delta' z} + \lambda u_{spx} > -u_{fta}$ and:

$$\mathrm{P}(FTA = 1) = \mathrm{P}(fta > 0) = \Phi[\mu_2 + \lambda\mu_1 + \lambda\alpha + \lambda\, \boldsymbol{\beta' q} + \boldsymbol{\delta' z} + \lambda u_{spx}]\ .$$

If *fta* # 0, then *FTA* = 0 implying $\mu_2 + \lambda\mu_1 + \lambda\, \boldsymbol{\beta' q} + \boldsymbol{\delta' z} + \lambda u_{spx}\ \#\ -u_{fta}$ and:

$$\mathrm{P}(FTA = 0) = \mathrm{P}(fta\ \#\ 0) = 1 - \Phi[\mu_2 + \lambda\mu_1 + \lambda\, \boldsymbol{\beta' q} + \boldsymbol{\delta' z} + \lambda u_{spx}]\ .$$

For logical consistency, P(*FTA* = 0) + P(*FTA* = 1) = 1. Hence, the model is consistent only if:

$$1 - \Phi[\mu_2 + \lambda\mu_1 + \lambda\, \boldsymbol{\beta' q} + \boldsymbol{\delta' z} + \lambda u_{spx}] + \Phi[\lambda\alpha + \mu_2 + \lambda\mu_1 + \lambda\, \boldsymbol{\beta' q} + \boldsymbol{\delta' z} + \lambda u_{spx}] = 1\ .$$

This condition holds if and only if either λ or α equals zero (so that $\lambda\alpha = 0$). Thus, estimation of the system in equations (34) and (35) is logically inconsistent unless λ or α equals zero.[24]

However, the treatment-effects methodology used here allows us to show that – even as evidence suggests traditional FTA coefficient estimates have been underestimated – the results are consistent with the notion that countries that select into FTAs would trade extensively *even without* an FTA. Using Heckman's technique, the coefficient estimates of the "hazard rates" (*FTA*HAZARD, NFTA*HAZARD*) reveal the likely "structure" of sorting in underlying the data. Following Vella and Verbeek (1999), under certain assumptions one can infer whether the sorting structure is a "hierarchical" or a "comparative advantage" structure. In a hierarchical structure in this paper's context, country pairs that trade extensively with the "treatment" (FTAs) are also pairs that would trade extensively *without* the treatment. In a hierarchical structure, countries that have formed an FTA have, in the context of Lawrence's quote, "chosen well" in the sense that they already trade extensively. By contrast, in a comparative-advantage structure, country pairs that trade extensively with FTAs are pairs that would not trade very much without FTAs.

In the context of this literature, the coefficient estimates on the hazard-rate variables indicate a comparative advantage or hierarchical structure. A comparative advantage structure requires a positive coefficient for *FTA*HAZARD* and a negative coefficient for *NFTA*HAZARD*. However, a hierarchical structure can have two negative coefficients, as found in our results. In particular, the statistically significant negative coefficient estimate for *FTA*HAZARD* in Table 2 (at the 10 percent level) implies that the structure here is *hierarchical*. Thus, on average, countries that have formed FTAs would have traded extensively even without an agreement. In this sense, countries in FTAs have, on average, "chosen well."

Returning to the empirical results in Table 1, one finds evidence, moreover, that FTAs have been formed where the effects of potential trade creation tend to dominate the effects of potential trade diversion. These empirical results, combined with the simulation results from our CGE model, provide evidence that FTAs have been formed (on average) when trade-creating economic factors are high (on average) and trade-diverting economic factors are low (on average). The CGE results implied that two countries consumers' welfare gain from an FTA is greater the smaller the distance between them, the larger their economic size, and the wider the difference in their capital-labor endowment ratios, consistent with greater bilateral trade "creation." The probit results in Table 1 confirmed this. The CGE results also implied that two countries consumers' welfare gain from an FTA is greater

[24]The logical inconsistency is especially problematic in Magee (2002) because estimation is primarily by maximum likelihood. Even if one ignores this logical inconsistency, there are two reasons to suggest that the results in Magee are biased and inconsistent (Magee found that coefficient estimates of *FTA* were overestimated). First, when using IV, Magee does not use a three-stage procedure as in our study; he calculates the predicted probability of an FTA using a probit, and then inserts this point estimate into the trade flow equation. As discussed in Angrist and Krueger (2001, p. 80), this will yield inconsistent estimates in the final-stage regression *unless* the probit function "happens to be exactly right." In our model, the predicted probabilities are used in a linear regression along with $\boldsymbol{q}$ to generate consistent final-stage estimates. Second, Magee includes in his $\boldsymbol{z}$ variables the average (of the logs) of per capita incomes. However, as is well known and evident in equation (3) earlier in this paper, the average of per capita incomes is traditionally an important explanatory variable in the gravity equation. Its absence in his gravity model implies that the error term in his gravity equation is likely to be highly correlated with $\boldsymbol{z}$, violating a necessary condition that $\boldsymbol{z}$ must satisfy for IV, and which he cannot test.

the more remote two (continental) partners are from the ROW and the smaller the difference between the partners' average capital-labor ratio and that of the ROW, consistent with less trade diversion. The probit results confirmed this. Thus, the relationships between each of these variables and the likelihood of an FTA are all consistent with Lawrence's suggestion that countries with FTAs have "chosen well."

The answer then to Lawrence's question is the following. After accounting for self-selection bias, large coefficient estimates on FTA dummy variables indicate a large effect of an FTA on trade. However, our estimation procedure suggests that countries that have selected into FTAs have, on average, "chosen well."

E. Estimates of the Treatment Effect on the "Treated"

In this part, we extend the model to estimate treatment effects for several specific FTAs – notably, the EU, NAFTA, MERCOSUR, the Central American Common Market (CACM), and the Andean Pact – and compare these effects to the average treatment effects (ATEs) from OLS ignoring endogeneity. Since many countries in 1996 had bilateral FTAs with EU members, we define a dummy variable EUBroad to assume the value 1 if either the country pair were both EU members or a non-EU country had an FTA with an EU member. For NAFTA, MERCOSUR, CACM, and the Andean Pact countries, we assume the value 1 if both countries were members of the FTA.

Note that the ATEs we have addressed up to now average effects over country pairs with FTAs as well as those pairs without FTAs. Policymakers may want to know what the estimated effect of a particular FTA is on trade among those members (on average). Following equation (6) in section II, we can estimate the treatment effect on the treated (TTE), which provides such information. For brevity, we only report the TTEs for one specification; the other estimates are available on request. A representative specification is the Heckman specification in Table 3, column 3; to account for multilateral resistance, we include *REMOTE2*. We then compare these TTEs against the coefficient estimates from a standard OLS estimation ignoring endogeneity.

First, for only those countries with FTAs, the TTE is estimated to be 0.57, slightly smaller than the ATE of 0.65. These two values are not significantly different statistically as they are only one-half of a standard error apart.[25] However, for the groups of countries with FTAs that are smaller economically – such as the Andean Pact, CACM, and MERCOSUR – the TTEs are much larger. The Andean Pact TTE is 1.45, suggesting that this pact increases trade (on average) among these countries by 326 percent ($e^{1.45}$). Our TTE estimates suggest that trade increases from membership in the CACM by 395 percent ($e^{1.60}$) for those countries, and by 222 percent from membership in MERCOSUR ($e^{1.17}$). The presence of NAFTA increases trade by an estimated 86 percent (on average) among Canada, Mexico, and the United States. Finally, membership in the EU – broadly defined, as above – is estimated to increase trade by only 38 percent. In the context of our framework, the smaller estimate can be explained by the differing estimated coefficients for *NGDP* and *POP* in our regressions for the treated and the untreated (see Table 3); richer countries tend to have smaller estimated effects from FTAs.

---

[25] The standard errors for the TTE estimates were obtained using a bootstrapping procedure drawing one hundred random samples from our population.

Second, Table 5 summarizes these estimates of TTEs. Two points are worth emphasizing. First, for three of the five FTAs, the estimated treatment effects on the treated exceed the estimated ATEs using OLS; for the remaining two (NAFTA and Andean Pact), there is little difference. Second, while the estimated standard errors for the TTEs are larger than for the ATEs using OLS, the increases are not very large. Consequently, the estimated TTEs for the EU and NAFTA are statistically significant (at the 10 percent level, one-tail test), whereas the OLS-estimated ATEs (ignoring endogeneity) for the EU and NAFTA were statistically insignificant. The most notable difference is the *EUBroad* effect, which was negative but insignificant under OLS, but is positive and statistically significant with a plausible value using the TTE estimate.

F. Sensitivity Analysis for "Political-Economy" Variables and Grossman-Helpman's "Politics of FTAs"

We have endeavored in this study to estimate the endogeneity bias of the *FTA* coefficient estimate in a typical gravity equation using a parsimonious model (only three variables in $\boldsymbol{z}$) based upon economic theory. In reality, of course, FTAs are formed by governments based upon political, institutional, and social considerations as well as consumer welfare. These omitted factors may influence the predicted FTA probabilities, and consequently influence the effect of an FTA on trade in the final-stage regression. While a veritable plethora of variables might influence the probability of an FTA, we considered six broad categories of factors that might influence governments' decisions: short-run displacement costs, distributional preferences toward income equality, common legal origins, national defense interests, national labor standards, and environmental policies.

One major factor that tends to prevent governments from liberalizing trade policies is the potential short-run "displacement," or adjustment, costs associated with workers changing jobs and industries as countries specialize in their comparative advantages. The theoretical model presented earlier assumes away such costs, but they likely exist. In the context of our model, if one holds constant the distance between two countries, their remoteness, and their absolute and relative factor endowments, the theoretical model predicts that the shares of each country's labor force in each of the two countries' goods (services) industries will be identical. This suggests that any *observed* differences between the two countries' shares of labor in an industry would require adjustment to equality; the greater the difference, the larger the adjustment costs, and the less likely each government would form an FTA. We use data on differences in labor shares of two countries in agriculture (*DAG*), mining and manufacturing (*DMM*), non-transport services (*DS*), and transport services (*DTS*), and add these to $\boldsymbol{z}$.

Second, as Rodrik (1995) noted, one of the factors that may weigh into a government's trade policy decision is "distributional preferences." In many countries with wide income inequality, there is considerable unskilled labor. A government may be adverse to liberalizing trade via an FTA because of actual or perceived effects on increasing income inequality. We include the average Gini index (*GINI*) of the pair of countries to reflect the influence of such distributional preferences.

Third, since the decision to form an FTA is made by countries' governments, it is possible that a pair of countries that share "legal origins" may have a higher likelihood of forming an FTA. A common legal origin may facilitate dispute resolutions. We include a dummy variable (*LEGAL*), which assumes the value 1 if a common legal origin (i.e., British law, etc.) and 0 otherwise.

Fourth, governments also pay attention to national security issues in developing FTAs; the original EEC was likely formed as much for defense purposes as economic ones. To capture this influence, we include the average share of GDP of the countries in defense expenditures and the (absolute value of the) difference in these shares (*DEF* and *DDEF*, respectively).

Fifth, some governments are more concerned than others with high labor standards for their countries, and differences in standards might prevent the formation of an FTA. To capture this difference, we included the (absolute value of the) difference in the fraction of children in each country aged 10 to 14 in that country's labor force (*DCHILD*). For completeness, we also included the average of these fractions (*CHILD*).

Sixth, environmental policies possibly influence FTA formations. Again, similarities in such policies might increase the likelihood of an FTA. To address this issue, we included in the first-stage probit regression the average per capita CO2 emissions in the pair of countries and the (absolute value of the) difference in these shares (*CO2* and *DCO2*, respectively).

There are, of course, numerous other factors that likely influence the probability of an FTA beyond those mentioned, such as the level of tariffs in the pair and non-member countries in the ROW and the market size of other countries in existing FTAs (i.e., bloc size). However, we are not concerned with an exhaustive list of factors determining FTAs; the key is simply to find a set of instruments $\boldsymbol{z}$ such that Φ^P is a good instrument in the second stage.

Specification 2 in Table 6 provides the results of estimating the first-stage probit regression using the Heckman procedure including these additional 12 regressors; for brevity, we omit the results for IV. For comparison, specification 1 is the main probit regression from Table 1, but extended to include *REMOTE2* (this probit was estimated for the first-stage of Column 3 in Table 3). Three points are worth noting. First, of these 12 additional regressors, six are statistically significant at the 5 percent level. Three of the four short-run displacement cost variables were statistically significant. Labor share differences in manufacturing (and mining) and in services have significant negative effects on the probability of an FTA; in contrast to conjecture, wider labor share differences in transport services increased the probability of an FTA. Of the remaining eight variables, wider differences in child labor standards have a significant negative relationship with P(*FTA*), while greater per capita CO2 emissions (wider differences in per capita emissions) have a significant negative (positive) relationship with P(*FTA*). Second, as a summary statistic, the pseudo-R^2 in specification 2 was only 0.02 higher than in specification 1 (with only three variables in $\boldsymbol{z}$). Third, we computed the percent of FTAs correctly predicted as well as the percent of No-FTAs correctly predicted. Of 285 FTAs, 239 are correctly predicted with the new specification vs. 222 with specification 1. Of 1093 No-FTAs, 1064 are correctly predicted with the new specification vs. 1055 with specification 1. There is a minor improvement in predictions using the enhanced specification.

Table 7 reveals that the estimated effect of an FTA in the gravity equation using the enhanced specification is not materially different from that with only three variables in ***z*** using the Heckman procedure. Comparing the *FTA* coefficient estimates between columns 1 and 2 shows a small deterioration in the *FTA* coefficient; however, once we introduce a channel discussed in Grossman and Helpman (1995), this small deterioration is eliminated.

In the remainder of this section, we discuss a potential simultaneity bias that might arise via a political economy channel such as motivated in Grossman and Helpman (1995). Their model suggested that – in a setting with potential economic rents due to a limited number of producers in each country with the ability to form lobbies – the potential success of an FTA arising politically depends upon "balance" in the trade between the two parties.

Operationally, a fairly ready measure of the Grossman-Helpman political-pressure concept exists. Ideally, an instrument to include in ***z*** is the (absolute value of the) difference in (the logs of) the two countries' gross trade flows in the particular year examined, i.e., $IMBALANCE = * \ln PX_{ij}^{g} - \ln PX_{ji}^{g} *$. Rivers and Vuong (1988) proposed a two-step estimator for estimating simultaneous equations models where one of the endogenous variables is qualitative, as here. This test is implemented by first estimating the reduced-form equation for *IMBALANCE* incorporating all the variables in ***q*** and ***z*** to generate its predicted values. The predicted values are used to construct the *IMBALANCE* equation residuals. The second stage of the estimator is to estimate the *FTA* probit equation including the actual values of *IMBALANCE* and the residuals. These results are shown in specification 3 of Table 6.

Inclusion of the residual from the first-stage of this estimator along with the variable *IMBALANCE* in the structural *FTA* equation yields a test of the endogeneity of *IMBALANCE*. The coefficient estimate of the residual is 8.58 with a z-statistic of 6.37, suggesting that we can reject that *IMBALANCE* is exogenous. Also, the coefficient estimate of *IMBALANCE* is -8.81 with a z-statistic of -6.58. Under the assumption that the variance of the error term in this structural *FTA* equation is unity, then -8.81 is a consistent estimate of the effect of the imbalance in trade flows on FTAs. This result is consistent with the presence of a simultaneous-equations bias in the model and the Grossman-Helpman (1995) conjecture that imbalance in potential trade between two trading partners will tend to have a negative impact on the likelihood of an FTA between them.

Specification 3 in Table 7 reports the results of the gravity equation using this procedure.[26] Earlier results are robust to accounting for the *IMBALANCE* variable. The coefficient estimate for the FTA dummy (including *REMOTE2*) is 0.68. This value is similar with that estimated earlier using only three variables in ***z***, as shown in the first column of Table 7. Thus, the result that the estimated coefficient of *FTA* has been considerably underestimated in earlier work on the gravity equation is robust to a variety of econometric estimates and variables determining FTAs. In the baseline gravity equation (say, with *REMOTE2*), the OLS estimate is 0.21 (Table 3, column 1), implying that the average (treatment) effect of having an FTA increases the value of trade by 23 percent ($e^{0.21}$). By

[26]For clarity, we note that there are effectively three stages in the estimation. The first stage is to regress *IMBALANCE* on ***q, z,*** and a constant to obtain the predicted values of *IMBALANCE* and the regression's residuals. The second stage is to estimate a probit of *FTA* on ***q, z,*** *IMBALANCE*, and these residuals (the estimated coefficient on the residuals is used to test for endogeneity of *IMBALANCE*); this generates predicted probabilities (Φ^P and φ^P). The predicted values of Φ^P and φ^P from this second stage are used in the third stage to estimate gravity equation (17). We show in Appendix D that estimation of this model is logically consistent.

contrast, using either the "narrow" set of instruments z or the "broader" set, the Heckman procedure yields an average FTA treatment effect of 92 percent ($e^{0.65}$), *quadruple* the OLS estimated effect. Moreover, the hazard rates' coefficient estimates are statistically significant and suggest the same interpretation of Lawrence as before.

We note two important conclusions. First, the inclusion of additional "political-economy" variables (beyond the theoretical model's scope) improves the predictability of FTAs. We can predict larger percentages of FTAs and No-FTAs correctly. However, we also find that these results had a minimal impact on the properly-estimated effect of an FTA on trade flows. Thus, the parsimonious econometric model of FTAs suggested by the theoretical model was sufficient.

VI. Conclusions

For forty years, the gravity equation has been used in international trade to estimate cross-sectionally the effects of preferential trade agreements, free trade agreements, or customs unions on the value of bilateral merchandise trade. As Eichengreen and Irwin (1997) have noted, the gravity equation has become the "workhorse for empirical studies of the pattern of trade." Yet over this entire period, with one recent exception, no one has attempted to determine the potential bias in coefficient estimates from the likely *endogeneity* of FTAs.

This paper has addressed the potential bias in estimating the effects of FTAs on trade flows arising from endogeneity in the form of measurement error, simultaneity, and omitted variables. While we found evidence of all three factors, we considered additionally the potential bias introduced by *self-selection*. As many trade-policy observers have argued, countries' governments have steered trade policy increasingly toward "deeper integration." This motivation raises the possibility that unobservable factors (to the econometrician), such as nontariff barriers and domestic regulations, that tend to inhibit bilateral trade may be a distinctive force in motivating policymakers to form FTAs. Such "unobservable heterogeneity" introduces potential selection bias into estimates of FTAs on trade flows. Using estimation techniques developed by James Heckman for the labor economics literature, we find the effect of FTAs on trade flows has been systematically *underestimated by 75 percent*. On average, when ignoring the endogeneity of an FTA, the agreement tends to increase the value of trade by 23 percent. By contrast, when acknowledging the endogeneity of an FTA, the agreement tends to increase the value of trade by *92 percent*. Moreover, these empirical estimates are robust to the additional inclusion of multilateral resistance terms, fixed effects, and political factors potentially influencing trade and FTA formation.

Finally, our estimation procedure based upon the "treatment effects" econometric methodology allows us to also draw inferences about "sorting" into FTAs. Using this methodology, our parameter estimates of hazard rates suggest that country pairs that tend to select into FTAs on average would trade extensively *even if* they did not have FTAs. In the context of vernacular familiar to most trade economists, our results suggest that, on average, FTAs have been selected into where there is more "trade creation" than "trade diversion." This result is plausibly interpreted – in the words of Lawrence (1998) – that countries that have selected into FTAs have "chosen well."

Table 1

FTA Probit Equation (First-Stage) Results[1]

Variable Category	Variable	Coeff. (z-statistic)
Gravity Equation Variables (***q***)	CONSTANT	25.77 (5.51)*
	DIST	-1.64 (-13.34)*
	NGDP	0.55 (4.62)*
	POP	-0.51 (-3.77)*
	ADJACENCY	-0.26 (-0.92)
	LANGUAGE	-0.05 (-0.23)
Variables in ***z***	CONTREMOTE	0.16 (9.45)*
	DKL	0.48 (3.60)*
	DROWKL	-0.76 (-2.82)*
Pseudo R^2		0.7016
Log likelihood		-209.57
Number of observations		1378

[1]z-statistics are in parentheses. * denotes statistically significant at the 5 percent level in a two-tailed test.

Table 2

Basic Gravity Equation[1]

Variables	OLS	IV	HECKMAN
CONSTANT	-0.22 (-0.13)	-3.35 (-1.55)	-0.13 (-0.07)
NGDP	1.15 (28.17)*	1.11 (23.07)*	1.22 (26.75)*
POP	-0.28 (-5.61)	-0.24 (-4.10)*	-0.35 (-6.51)*
DIST	-0.95 (-16.64)	-0.81 (-10.25)*	-0.71 (-7.39)*
ADJACENT	0.66 (4.31)	0.67 (4.70)*	0.45 (1.35)
LANGUAGE	0.66 (6.91)	0.70 (8.83)*	0.50 (4.55)*
FTA	0.18 (1.67)	0.81 (4.15)*	0.83 (3.40)*
FTA*NGDP			-0.46 (-4.10)*
FTA*POP			0.41 (3.06)*
FTA*DIST			-0.06 (-0.35)
FTA*ADJ			0.18 (0.47)
FTA*LANG			0.29 (1.21)
FTA*HAZARD			-0.36 (-1.86)
NFTA*HAZARD			-0.63 (-5.12)*
R^2	0.8485	0.8446	0.8562
Adjusted R^2	0.8478		0.8548
RMSE	1.0199	1.0327	0.9961
Number of observations	1378	1378	1378

[1]t-statistics are in parentheses. * denotes statistically significant at the 5 percent level in a two-tailed test.

Table 3

Gravity Equation with *REMOTE2*[1]

Variables	OLS	IV	HECKMAN
CONSTANT	4.23 (1.94)	3.55 (1.45)	5.31 (2.15)*
NGDP	1.19 (28.03)*	1.18 (24.14)*	1.27 (26.99)*
POP	-0.33 (-6.36)*	-0.32 (-5.34)*	-0.41 (-7.25)*
DIST	-0.99 (-17.03)*	-0.92 (-12.39)*	-0.90 (-9.97)*
ADJACENT	0.60 (3.89)*	0.60 (4.21)*	0.38 (1.10)
LANGUAGE	0.59 (6.14)*	0.61 (7.53)*	0.40 (3.60)*
REMOTE2	21.05 (3.36)*	22.12 (3.83)*	16.83 (2.56)*
FTA	0.21 (2.00)*	0.40 (2.77)*	0.65 (2.91)*
FTA*NGDP			-0.41 (-3.68)*
FTA*POP			0.37 (2.75)*
FTA*DIST			0.0001 (0.00)
FTA*ADJ			0.16 (0.41)
FTA*LANG			0.21 (0.84)
FTA*REMOTE2			19.93 (0.88)
FTA*HAZARD			-0.23 (-1.29)
NFTA*HAZARD			-0.56 (-3.19)*
R^2	0.8497	0.8494	0.8552
Adjusted R^2	0.8489		0.8536
RMSE	1.0161	1.0172	1.0004
Number of observations	1378	1378	1378

[1]t-statistics are in parentheses. * denotes statistically significant at the 5 percent level in a two-tailed test.

Table 4

Gravity Equation with Fixed Effects [1]

Variables	OLS	IV with LPM
DIST	-1.02 (-18.53)*	-0.92 (-10.21)*
ADJACENT	0.59 (4.44)*	0.59 (4.42)*
LANGUAGE	0.90 (9.02)*	0.90 (8.99)*
FTA	0.26 (2.65)*	0.54 (2.42)*
RMSE	0.8438	0.8476
Number of observations	1378	1378

[1]Fixed effects coefficient estimates are not reported. t-statistics are in parentheses. * denotes statistically significant at the 5 percent level in a two-tailed test.

Table 5

Estimated Treatment Effects on the Treated (TTE)

FTA	TTE Mean	TTE Stan. Dev.	OLS ATE	OLS Stan. Dev.
Broad EU	0.32	0.21	-0.09	0.12
NAFTA	0.62	0.47	0.63	0.60
MERCOSUR	1.17	0.54	0.74	0.43
CACM	1.60	0.42	0.78	0.34
Andean Pact	1.45	0.41	1.49	0.36

Table 6
FTA Probit Equation Results*

Variable Category	Variable	Specif. 1	Specif. 2	Specif. 3
	CONSTANT	-16.45 (-2.24)*	-12.96 (-1.56)	19.04 (2.89)*
Gravity Equation Var. (***q***)				
	DIST	-1.42 (-9.92)*	-1.61 (-9.31)*	-1.27 (-6.83)*
	NGDP	0.36 (2.58)*	0.31 (2.06)*	0.24 (1.57)
	POP	-0.18 (-1.08)	-0.13 (-0.74)	-0.26 (-1.48)
	ADJACENCY	0.03 (0.09)	-0.19 (-0.58)	-1.73 (-4.31)*
	LANGUAGE	0.83 (3.59)*	0.99 (3.39)*	1.52 (4.71)*
	REMOTE2	-222.28 (-7.93)*	-215.54 (-6.64)*	Collinear (deleted)
Baseline Variables in ***z***				
	CONTREMOTE	0.24 (10.61)*	0.26 (9.92)*	0.24 (9.67)*
	DKL	0.24 (1.76)	0.44 (2.25)*	2.30 (7.05)*
	DROWKL	0.23 (0.76)	-0.33 (-0.93)	-1.94 (-5.37)*
"Other" Variables in ***z***				
Short-Run Displacement Costs	DAG		-0.02 (-1.12)	-0.07 (-3.66)*
	DMM		-0.03 (-2.78)*	-0.07 (-5.13)*
	DSE		-0.02 (-2.64)*	-0.03 (-3.45)*
	DTS		0.07 (2.31)*	0.23 (6.02)*
Government Preferences				
	GINI		-0.01 (-0.91)	0.06 (4.03)*
Institutional Structure				
	LEGAL		-0.07 (-0.34)	0.28 (1.27)
National Defense				
	DEF		-0.12 (-1.53)	0.43 (3.78)*
	DDEF		0.02 (0.17)	-0.50 (-3.96)*
Labor Standards				
	CHILD		0.01 (0.77)	0.05 (2.52)*
	DCHILD		-0.07 (-2.87)*	-0.09 (-3.77)*
Environmental Policies				
	CO2		-0.06 (-3.01)*	-0.11 (-5.27)*
	DCO2		0.07 (2.39)*	0.03 (1.08)
Grossman-Helpman Variables	IMBALANCE			-8.81 (-6.58)*
	RESIDUAL			8.58 (6.37)*
Pseudo R^2		0.7573	0.7819	0.7865
Log likelihood		-170.46	-153.21	-149.96
Number of observations		1378	1378	1378

[1]z-statistics are in parentheses. * denotes statistically significant at the 5 percent level in a two-tailed test.

Table 7

Gravity Equation with "Political-Economy" and "Politics-of-FTAs" Variables[1]

Variables	Heckman from Table 3, Col. 3	Heckman with Political Economy Variables	Heckman with Political Economy and IMBALANCE Variables
CONSTANT	5.31 (2.15)*	5.77 (2.35)*	5.48 (2.24)*
NGDP	1.27 (26.99)*	1.28 (27.09)*	1.28 (27.13)*
POP	-0.41 (-7.25)*	-0.41 (-7.29)*	-0.41 (-7.32)*
DIST	-0.90 (-9.97)*	-0.95 (-10.84)*	-0.91 (-10.39)*
ADJACENT	0.38 (1.11)	0.46 (1.34)	0.42 (1.24)
LANGUAGE	0.40 (3.60)*	0.41 (3.64)*	0.41 (3.73)*
REMOTE2	16.83 (2.56)*	16.78 (2.55)*	17.05 (2.59)*
FTA	0.65 (2.91)*	0.56 (2.75)*	0.68 (3.44)*
FTA*NGDP	-0.41 (-3.68)*	-0.42 (-3.69)*	-0.44 (-3.87)*
FTA*POP	0.37 (2.75)*	0.37 (2.74)*	0.40 (2.93)*
FTA*DIST	0.0001 (0.00)	0.02 (0.15)	0.01 (0.10)
FTA*ADJ	0.16 (0.41)	0.08 (0.20)	0.12 (0.31)
FTA*LANG	0.22 (0.84)	0.20 (0.79)	0.18 (0.72)
FTA*REMOTE2	19.93 (0.88)	17.17 (0.77)	21.26 (0.96)
FTA*HAZARD	-0.23 (-1.29)	-0.20 (-1.13)	-0.32 (-1.86)
NFTA*HAZARD	-0.56 (-3.19)*	-0.44 (-2.61)*	-0.60 (-3.49)*
R^2	0.8552	0.8548	0.8556
Adjusted R^2	0.8536	0.8532	0.8540
RMSE	1.0004	1.0018	0.9990
Number of observations	1378	1378	1378

[1]t-statistics are in parentheses. * denotes statistically significant at the 5 percent level in a two-tailed test.

Figure 1: Net Welfare Gains from Natural and Unnatural FTAs

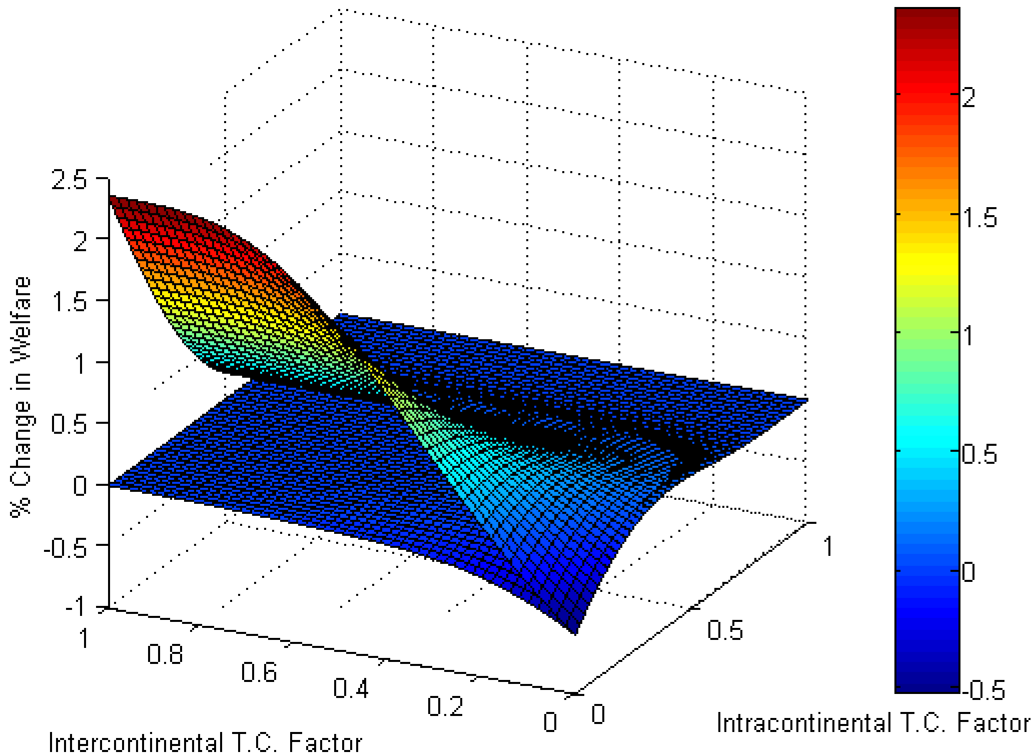


Figure 2a: Net Welfare Gains from Natural FTAs and Economic Size

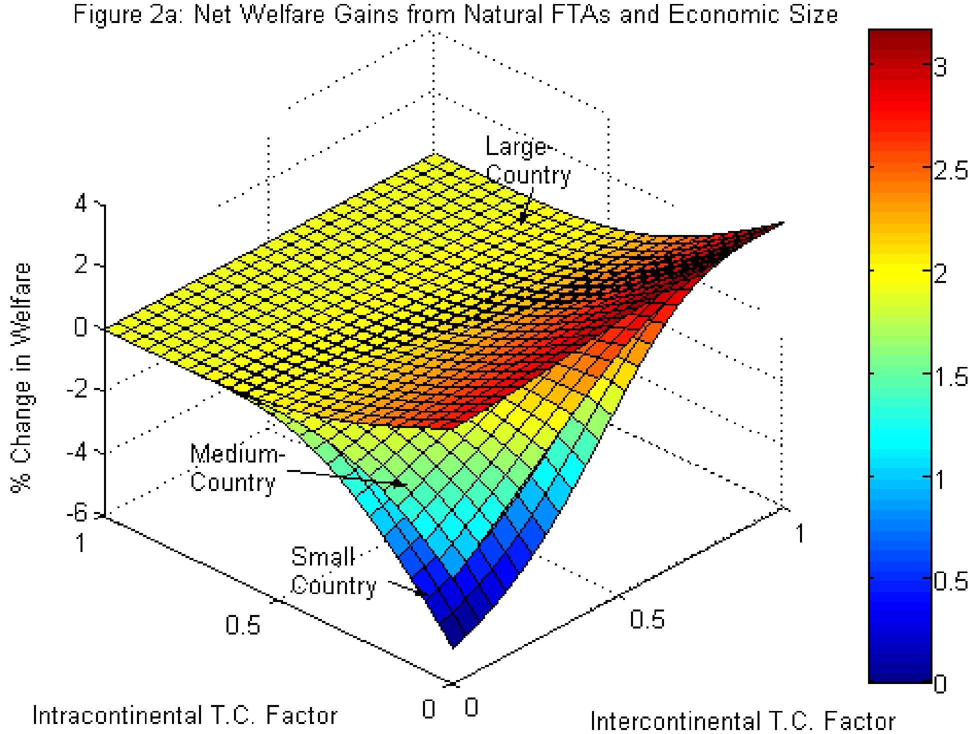


Figure 2b: Net Welfare Gains from Unnatural FTAs and Economic Size

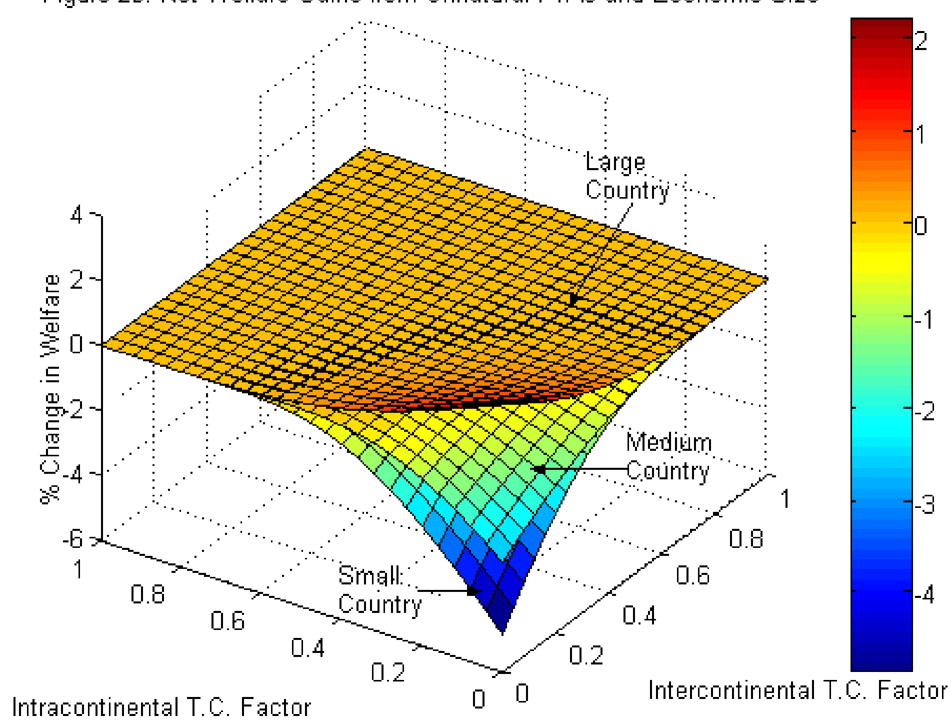


Figure 3: A Natural FTA and Relative Factor-Endowment Differences

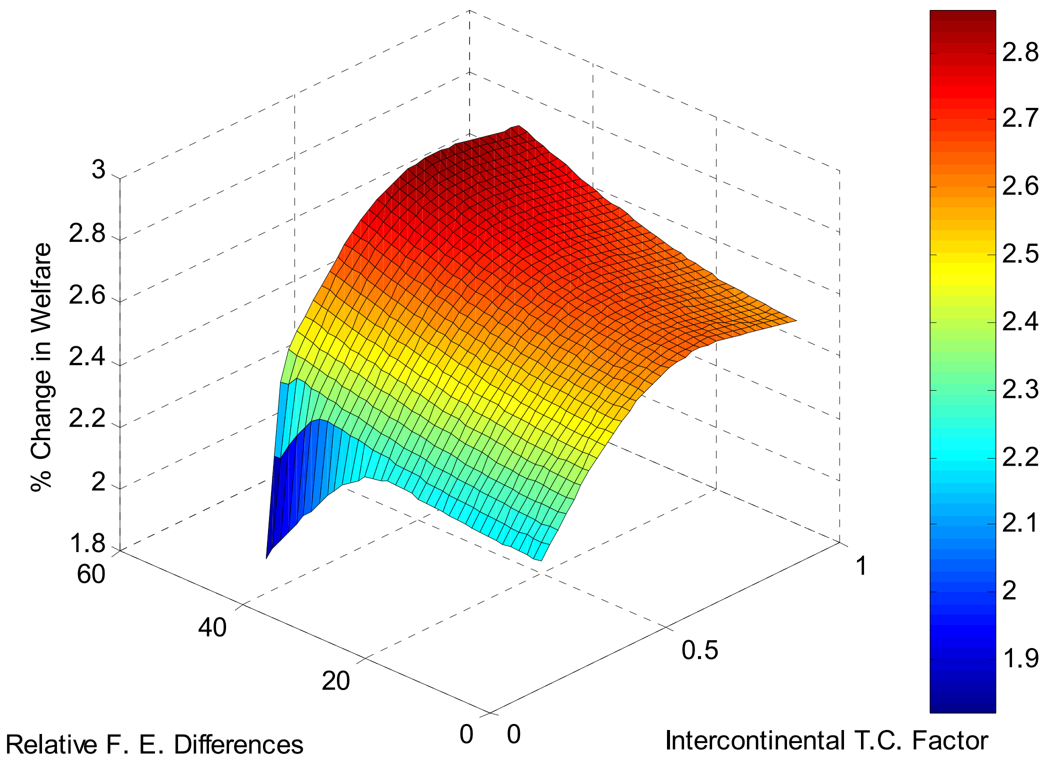
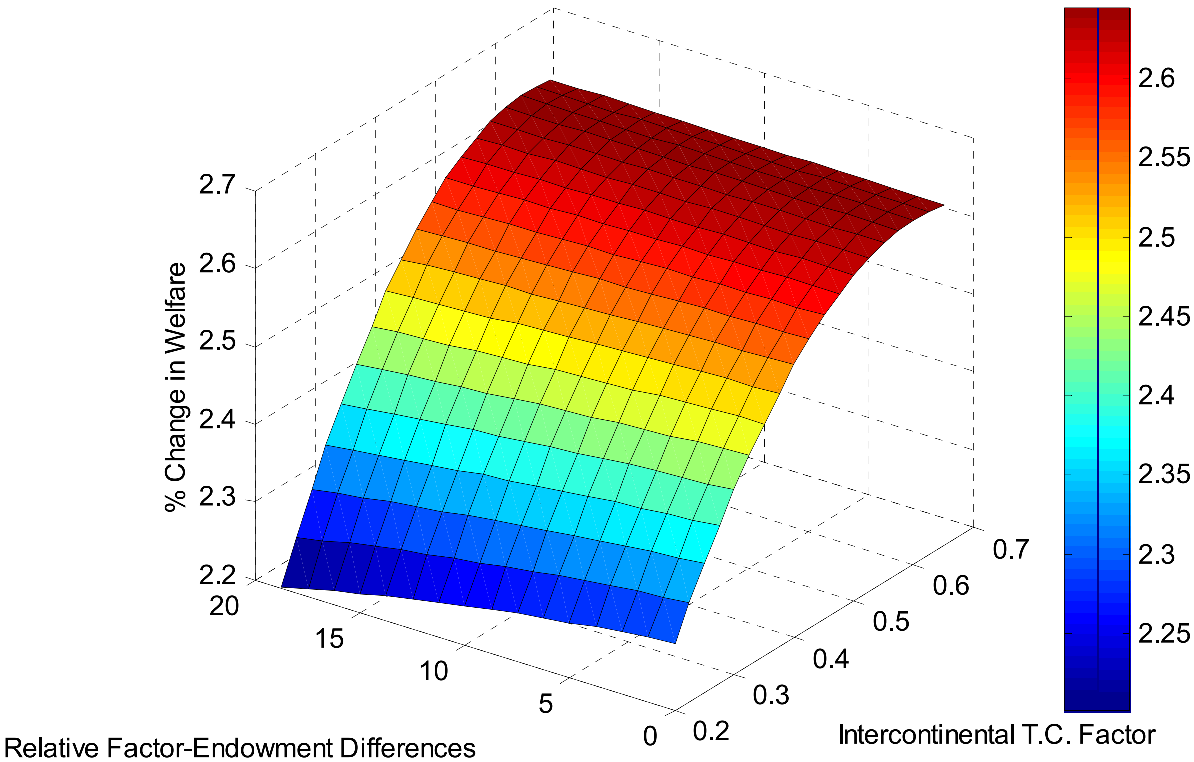


Figure 4: A Natural FTA and World Factor-Endowment Differences



References

Abrams, R.K. "International Trade Flows under Flexible Exchange Rates." Federal Reserve Bank of Kansas City *Economic Review* 65, no. 3 (1980): 3-10.

Aitken, Norman D. "The Effect of the EEC and EFTA on European Trade: A Temporal Cross-Section Analysis." *American Economic Review* 5 (December 1973): 881-892.

Anderson, James E. "A Theoretical Foundation for the Gravity Equation." *American Economic Review* 69, no. 1 (March 1979): 106-116.

Anderson, James E., and van Wincoop, Eric. "Gravity with Gravitas: A Solution to the Border Puzzle." Manuscript, October 2001.

Angrist, Joshua D., and Krueger, Alan B. "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments." *Journal of Economic Perspectives* 15, no. 4 (Fall 2001): 69-86.

Baier, Scott L., and Bergstrand, Jeffrey H. "Economic Determinants of Free Trade Agreements." Manuscript, University of Notre Dame (www.nd.edu/~jbergstr/working_papers.html) , June 2002.

Baier, Scott L., and Bergstrand, Jeffrey H. "The Growth of World Trade: Tariffs, Transport Costs, and Income Similarity." *Journal of International Economics* 53, no. 1 (February 2001): 1-27.

Baier, Scott L., Gerald Dwyer, and Robert Tamura. "The Growth and Development of Nations." Manuscript, 2000.

Baldwin, Richard E. *Towards and Integrated Europe*. London: Centre for Economic Policy Research, 1994.

Bergstrand, Jeffrey H. "The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence." *Review of Economics and Statistics* 67, no. 3 (August 1985): 474-481.

Bergstrand, Jeffrey H. "The Generalized Gravity Equation, Monopolistic Competition, and the Factor-Proportions Theory in International Trade." *Review of Economics and Statistics* 71, no. 1 (February 1989): 143-153.

Bergstrand, Jeffrey H. "The Heckscher-Ohlin-Samuelson Model, the Linder Hypothesis, and the Determinants of Bilateral Intra-Industry Trade." *Economic Journal* 100, no. 4 (December 1990), 1216-1229.

Bergstrand, Jeffrey H. "Structural Determinants of Real Exchange Rates and National Price Levels: Some Empirical Evidence." *American Economic Review* 81, no. 1 (March 1991): 325-334.

Bergstrand, Jeffrey H. "Real Exchange Rates, National Price Levels, and the Peace Dividend." *American Economic Review Papers and Proceedings* 82, no. 2 (May 1992): 55-61.

Brada, Josef C., and Mendez, Jose A. "Economic Integration among Developed, Developing and Centrally Planned Economies: A Comparative Analysis." *Review of Economics and Statistics* 67, no. 4 (November 1985): 549-556.

Eichengreen, Barry, and Irwin, Douglas. "The Role of History in Bilateral Trade Flows." In *The Regionalization of the World Economy*, edited by Jeffrey A. Frankel. Chicago: The University of Chicago Press, 1998.

Evenett, Simon, and Keller, Wolfgang. "On Theories Explaining the Success of the Gravity Equation." *Journal of Political Economy* (2001, forthcoming).

Feenstra, Robert. *Advanced International Trade: Theory and Evidence*. Princeton, NJ: Princeton University Press, forthcoming 2003.

Frankel, Jeffrey A. *Regional Trading Blocs*. Washington, DC: Institute for International Economics, 1997.

Frankel, Jeffrey A., Stein, Ernesto; and Wei, Shang-Jin. "Trading Blocs and the Americas: The Natural, the Unnatural, and the Super-Natural." *Journal of Development Economics* 47, no. 1 (June 1995): 61-95.

Frankel, Jeffrey A., Ernesto Stein, and Shang-Jin Wei. "Regional Trading Arrangements: Natural or Supernatural?" *American Economic Review* 86, no. 2 (May 1996): 52-56.

Frankel, Jeffrey A., Ernesto Stein, and Shang-Jin Wei. "Continental Trading Blocs: Are They Natural or Supernatural?" In *The Regionalization of the World Economy*, edited by Jeffrey A. Frankel. Chicago: The University of Chicago Press, 1998, 91-113.

Gawande, Kishore, and Bandyopadhyay, Usree. "Is Protection for Sale? Evidence on the Grossman-Helpman Theory of Endogenous Protection." *Review of Economics and Statistics* 82, no. 1 (February 2000): 139-152.

Gilpin, Robert. *The Challenge of Global Capitalism*. Princeton, NJ: Princeton University Press, 2000.

Goldberg, Pinelopi Koujianou, and Maggi, Giovanni. "Protection for Sale: An Empirical Investigation." *American Economic Review* 89, no. 5 (December 1999): 1135-1155.

Grossman, Gene M., and Helpman, Elhanan. "The Politics of Free-Trade Agreements." *American Economic Review* 85, no. 4 (September 1995): 667-690.

Head, Keith, and Ries, John. "Increasing Returns vs. National Product Differentiation as an Explanation for the Pattern of U.S.-Canada Trade." *American Economic Review* 91, no. 4 (September 2001): 858-876.

Heckman, James J. "Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture." *Journal of Political Economy* 109, no. 4 (August 2001): 673-748.

Heckman, James J. "Dummy Endogenous Variables in a Simultaneous Equation System." *Econometrica* 46, no. 1 (July 1978): 931-959.

Heckman, James J. "Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations." *Journal of Human Resources* 32, no. 3 (Summer 1997): 441-462.

Helliwell, John F. *How Much Do National Borders Matter?* Washington, DC: The Brookings Institution, 1998.

Helpman, Elhanan, and Krugman, Paul. *Market Structure and Foreign Trade*. Cambridge, MA: MIT Press, 1985.

Kravis, Irving B., and Lipsey, Robert E.. "The Assessment of National Price Levels." In *Real-Financial Linkages among Open Economies*, edited by Sven W. Arndt and J. David Richardson. Cambridge, MA: MIT Press,1987, 97-134.

Kravis, Irving B., and Lipsey, Robert E. "National Price Levels and the Prices of Tradeables and Nontradeables." *American Economic Review Papers and Proceedings*, 78, no. 2 (May 1988): 474-8.

Krugman, Paul. "Is Bilateralism Bad?" In *International Trade and Trade Policy*, edited by Elhanan Helpman and Assaf Razin. Cambridge, MA: MIT Press, 1991a, 9-23.

Krugman, Paul. "The Move Toward Free Trade Zones." In *Policy Implications of Trade and Currency Zones*, proceedings of a symposium sponsored by the Federal Reserve Bank of Kansas City, 1991b, 7-41.

Lawrence, Robert Z. *Regionalism, Multilateralism, and Deeper Integration*. Washington, DC: The Brookings Institution, 1996.

Lawrence, Robert Z. "Comment on 'The Role of History in Bilateral Trade Flows'." In *The Regionalization of the World Economy*, edited by Jeffrey A. Frankel. Chicago: The University of Chicago Press, 1998, 57-59.

Lee, Jong-Wha, and Swagel, Phillip. "Trade Barriers and Trade Flows Across Countries and Industries." *Review of Economics and Statistics* 79 (1997): 372- 382.

Linnemann, Hans. *An Econometric Study of International Trade Flows*. Amsterdam: North-Holland, 1966.

Maddala, G.S. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge, MA: Cambridge University Press, 1983.

Magee, Chris. "Endogenous Preferential Trade Agreements: An Empirical Analysis." Manuscript, Bucknell University, March 2002.

McCallum, John. "National Borders Matter: Canada-U.S. Regional Trade Patterns." *American Economic Review* 85, no. 3 (June 1995): 615-623.

McFadden, Daniel. "Quantal Choice Analysis: A Survey." *Annals of Economic and Social Measurement* 5, no. 4 (1976): 363-390.

McFadden, Daniel. "The Revealed Preferences of a Government Bureaucracy: Theory." *Bell Journal of Economics* 6, no. 2 (1975): 401-416.

McFadden, Daniel. "Conditional Logit Analysis of Qualitative Choice Analysis." In *Frontiers of Econometrics,* edited by Paul Zarembka. New York: Academic Press, 1974.

Mitchell, Brian R. *International Historical Statistics: Europe 1750-1988*. New York: Stockton Press, 1992.

Mitchell, Brian R. *International Historical Statistics: The Americas 1750-1988*. New York: Stockton Press, 1993.

Mitchell, Brian R. *International Historical Statistics: Africa, Asia and Oceania 1750-1988*. New York: Stockton Press, 1995.

Oguledo, Victor I., and MacPhee, Craig R. "Gravity Models: A Reformulation and an Application to Discriminatory Arrangements." *Applied Economics* 26 (1994): 107-120.

Preeg, Ernest H. *From Here to Free Trade*. Chicago: The University of Chicago Press, 1998.

Redding, Stephen, and Venables, Anthony J. "Economic Geography and International Inequality." Centre for Economic Policy Research Discussion Paper No. 2568, 2000.

Rivers, Douglas, and Vuong, Quang H. "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models." *Journal of Econometrics* 39 (1988): 347-366.

Road Atlas of Europe. Chicago: Rand McNally and Co., 1988.

Rodrik, Dani. "Political Economy of Trade Policy" In *Handbook of International Economics, Volume 3*, edited by Gene M. Grossman and Kenneth Rogoff. Amsterdam: Elsevier Science, 1995, 1457-1494.

Roland-Holst, David W.; Reinhart, Kenneth A.; and Sheills, Clinto R. "A General Equilibrium Analysis of North American Economics Integration." In *Modeling Trade Policy: Applied General Equilibrium Assessments of North American Free Trade*, edited by Joseph F. Francois and Clinton R. Shiells. Cambridge, UK: Cambridge University Press, 1994.

Rose, Andrew, and van Wincoop, Eric. "National Money as a Barrier to International Trade: The Real Case for Currency Union." *American Economic Review Papers and Proceedings* 91, no. 2 (May 2001): 386-390.

Sapir, Andre. "Trade Benefits under the EEC Generalized System of Preferences." *European Economic Review* 15 (1981): 339-355.

Sapir, Andre, and Winter, Chantal. "Services Trade." In *Surveys in International Trade*, edited by David Greenaway and L. Alan Winters. Oxford, UK: Blackwell, 1994.

Tinbergen, Jan. *Shaping the World Economy*. New York: The Twentieth Century Fund, 1962.

Trefler, Daniel. "Trade Liberalization and the Theory of Endogenous Protection: An Econometric Study of U.S. Import Policy." *Journal of Political Economy* 101, no. 1 (February 1993): 138-160.

U.S. Department of the Navy Oceanographic Office. *Distance Between Ports*. H.O. Publication No. 15, U.S. Government Printing Office, 1965.

Vella, Francis, and Verbeek, Marno. "Estimating and Interpreting Models with Endogenous Treatment Effects." *Journal of Business and Economic Statistics* 17, no. 4 (October 1999): 473-478.

Wooldridge, Jeffrey M. *Introductory Econometrics*. Cincinnati: South-Western Publishing, 2000.

Wooldridge, Jeffrey M. *Econometric Analysis of Cross-Section and Panel Data*. Cambridge, MA: MIT Press, 2002.

World Bank. *World Development Indicators*, CD-ROM. Washington, DC: World Bank, 2001.

World Trade Organization. http://www.wto.org , November 1999.

Appendix A
(Not intended for publication)

The following shows the derivation of eq. (28) from eq. (27):

$$x_{ij}^{g} = \gamma\left[\left(GDP_j + T_j L_j\right)/P_j^g\right]\left[p_i^g\left(1+c_{ij}^g\right)\left(1+t_{ij}^g\right)/P_j^g\right]^{-\sigma g} \tag{27}$$

A gravity equation for the gross bilateral trade flow from *i* to *j* is found by first converting the real trade flow for each of *i*'s firms to a value using p_i^g, where p_i^g is expressed in terms of the numeraire (country 1's good). To obtain p_i^g, we note from section III.A.2. that profit maximization ensures that prices are a markup over marginal production costs:

$$p_i^g = \left[\sigma^g/\left(\sigma^g - 1\right)\right]\left[\left(C/z_i^g\right)r_i^{\alpha g} w_i^{1-\alpha g}\right] \tag{A1}$$

where $C = (\alpha^g)^{-\alpha g}(1-\alpha^g)^{-(1-\alpha g)}$. Then, defining $px_{ij}^{g,fob} = p_i^g\, x_{ij}^g$, the f.o.b. value of the trade flow of a representative firm in i is:

$$\begin{aligned} px_{ij}^{g,fob} &= \left[\gamma\sigma^g/\left(\sigma^g - 1\right)\right]\left[\left(C/z_i^g\right)r_i^{\alpha g} w_i^{1-\alpha g}\right]\left[GDP_j\left(1+t_j\right)/P_j^g\right] \\ &\bullet\left[p_i^g\left(1+c_{ij}^g\right)\left(1+t_{ij}^g\right)/P_j^g\right]^{-\alpha g} \end{aligned} \tag{A2}$$

where t_j (= T_jL_j/GDP_j) is tariff revenue in j as a fraction of factor income. To derive the *aggregate* f.o.b. trade flow, multiply both sides of (A2) by eq. (24), letting $PX_{ij}^{g,fob} = n_i^g\, px_{ij}^{g,fob}$:

$$\begin{aligned} PX_{ij}^{g,fob} &= \left[\gamma/\Phi^g\left(\sigma^g - 1\right)\right]\left[C\left(r_i K_i^g\right)^{\alpha g}\left(w_i L_i^g\right)^{1-\alpha g}\right]GDP_j \\ &\bullet\left[\left(1+t_j\right)/P_j^g\right]\left[p_j^g\left(1+c_{ij}^g\right)\left(1+t_{ij}^g\right)/P_j^g\right]^{-\sigma g} \end{aligned} \tag{A3}$$

However, $\left[C\left(r_i K_i^g\right)^{\alpha g}\left(w_i L_i^g\right)^{1-\alpha g}\right]$ is country i's national output of goods, or $s_i^g GDP_i$, where s_i^g is the share of national output in goods production (and is likely related to i's capital-labor endowment ratio; see Bergstrand (1991, 1992)). Substituting $s_i^g GDP_i$ in, and noting equation (20), the aggregate c.i.f. bilateral trade flow (PX_{ij}^g) is:

$$\begin{aligned} PX_{ij}^{g} &= \left[\gamma/\Phi^g\left(\sigma^g - 1\right)\right]\left(s_i^g GDP_i\, GDP_j\right)\left[\left(1+c_{ij}^g\right)\left(1+t_j\right)/P_j^g\right] \\ &\bullet\left[p_j^g\left(1+c_{ij}^g\right)\left(1+t_{ij}^g\right)/P_j^g\right]^{-\sigma g} \end{aligned} \tag{A4}$$

Consolidating the s_i^g, $(1+t_j)$, and $(1+t_{ij}^g)^{-\sigma g}$ terms and multiplying the RHS by unity in the form p_i^g/p_i^g yields equation (28) in the text:

$$\begin{aligned} PX_{ij}^{g} &= \left[\gamma/\Phi^g\left(\sigma^g - 1\right)\right]\left(GDP_i/p_i^g\right)GDP_j\left[p_i^g\left(1+c_{ij}^g\right)/P_j^g\right]^{1-\sigma g} \\ &\bullet\left\{s_i^g\left[\left(1+t_j\right)\left(1+t_{ij}^g\right)^{-\sigma g}\right]\right\} \end{aligned} \tag{28}$$

Appendix B

(Not intended for publication)

The following shows the derivation of eq. (30) from eq. (29). First, rewrite (29) as:

$$SPX_{ij}^{g} = \left[\gamma / \Phi^{g}\left(\sigma^{g} - 1\right)\right] GDP_{i}\ GDP_{j}\left(1 + c_{ij}^{g}\right)$$
$$\left\{ s_{i}^{g}\left[p_{i}^{g}\left(1 + c_{ij}^{g}\right)\left(1 + t_{ij}^{g}\right)\right]^{-\sigma g} + s_{j}^{g}\left[p_{j}^{g}\left(1 + c_{ji}^{g}\right)\left(1 + t_{ji}^{g}\right)\right]^{-\sigma g} \right\} \tag{29}$$

A condition of market clearing in the goods sector for each firm requires:

$$p_{i}^{g} g_{i} = \sum_{k=1}^{N} p_{ik}^{g} x_{ik}^{g} \qquad i = 1, \ldots, N \tag{B1}$$

where g_i (see eq. (23)) is output of the representative firm in i and p_{ik}^{g} and x_{ik}^{g} are determined by eqs. (20) and (27), respectively. Assuming, as earlier, that $c_{ij}^{g} = c_{ji}^{g}$ and $t_{ij}^{g} = t_{ji}^{g}$, then an implicit solution to (B1) and (B2):

$$P_{i}^{g*} = \left\{ \sum_{k=1}^{N} n_{k}^{g}\left[p_{k}^{g}\left(1 + c_{ki}^{g}\right)\left(1 + t_{ki}^{g}\right)\right]^{1-\sigma g} \right\}^{1/(1-\sigma g)} \tag{B2}$$

is:

$$p_{i}^{g*} = \left(GDP_{i}^{g} / n_{i}^{g} GDP_{W}^{g}\right)^{1/(1-\sigma g)} / P_{i}^{g*} \tag{B3}$$

where GDP_{i}^{g} is the value of goods output in country i and GDP_{W}^{g} is the value of goods output in the world (a constant).

Proof:

We will show that substitution of eq. (B3) into country j's analogue for eq. (B2) yields equation (B6). Then we show that substitution of eq. (B3) into market clearing eq. (B1) also yields eq. (B6). First, multiply both sides of (B3) by $(1+c_{ij}^{g})(1+t_{ij}^{g})$:

$$p_{i}^{g*}\left(1 + c_{ij}^{g}\right)\left(1 + t_{ij}^{g}\right) = \left(GDP_{i}^{g} / n_{i}^{g} GDP_{W}^{g}\right)^{1/(1-\sigma g)}\left(1 + c_{ij}^{g}\right)\left(1 + t_{ij}^{g}\right) / P_{i}^{g*} \tag{B4}$$

Rewrite eq. (B2) for country j:

$$P_{j}^{g*} = \left\{ \sum_{k=1}^{N} n_{k}^{g}\left[p_{k}^{g}\left(1 + c_{kj}^{g}\right)\left(1 + t_{kj}^{g}\right)\right]^{1-\sigma g} \right\}^{1/(1-\sigma g)} \tag{B5}$$

Substituting (B4) into (B5) yields:

$$P_{j}^{g*} = \left\{ \sum_{k=1}^{N} \left(\frac{GDP_{k}^{g}}{GDP_{W}^{g}}\right)\left[\frac{\left(1 + c_{kj}^{g}\right)\left(1 + t_{kj}^{g}\right)}{P_{k}^{g*}}\right]^{1-\sigma g} \right\}^{1/(1-\sigma g)} \tag{B6}$$

Second, rewrite market clearing condition (B1) as:

$$GDP_i^g = n_i^g p_i^g g_i = \sum_{k=1}^{N} n_i^g p_{ik}^g x_{ik}^g$$

Using eq. (27):

$$GDP_i^g = n_i^g \sum_{k=1}^{N} p_i^g\left(1+c_{ik}^g\right)\left(1+t_{ik}^g\right)\left[\gamma GDP_k\left(1+t_k\right)/P_k^g\right]\left[p_i^g\left(1+c_{ik}^g\right)\left(1+t_{ik}^g\right)/P_k^g\right]^{-\sigma g} \qquad \text{(B7)}$$

Substitute eq. (B3) into eq. (B7) to obtain:

$$GDP_i^g = n_i^g \sum_{k=1}^{N}\left[\left(GDP_i^g / n_i^g GDP_W^g\right)^{1/(1-\sigma g)}\left(1+c_{ik}^g\right)\left(1+t_{ik}^g\right)/P_i^{g*}P_k^{g*}\right]^{1-\sigma}\gamma\ GDP_k\left(1+t_k\right)$$

or

$$GDP_i^g = \frac{GDP_i^g}{GDP_W^g}\sum_{k=1}^{N}\left(\frac{\left(1+c_{ik}^g\right)\left(1+t_{ik}^g\right)}{P_i^{g*}P_k^{g*}}\right)^{1-\sigma g}\left[\gamma\ GDP_k\left(1+t_k\right)\right]$$

If $t_k = 0$, then since $\gamma\ GDP_k = GDP_k^g$:

$$GDP_W^g = \sum_{k=1}^{N}\left[\frac{\left(1+c_{ik}^g\right)\left(1+t_{ik}^g\right)}{P_i^{g*}P_k^{g*}}\right]^{1-\sigma g} GDP_k^g$$

or

$$P_i^{g*} = \left\{\sum_{k=1}^{N}\left(\frac{GDP_k^g}{GDP_w^g}\right)\left[\frac{\left(1+c_{ik}^g\right)\left(1+t_{ik}^g\right)}{P_k^{g*}}\right]^{1-\sigma g}\right\}^{1/(1-\sigma g)}$$

or, for country j:

$$P_j^{g*} = \left\{\sum_{k=1}^{N}\left(\frac{GDP_k^g}{GDP_W^g}\right)\left[\frac{\left(1+c_{jk}^g\right)\left(1+t_{jk}^g\right)}{P_k^{g*}}\right]^{1-\sigma g}\right\}^{1/(1-\sigma g)}$$

which is identical to eq. (B6) if $c_{jk}^g = c_{kj}^g$ and $t_{jk}^g = t_{kj}^g$. QED.

Since eq. (B3) is a solution, then substitute eq. (B3) for p_i^g and its analogue for p_j^g into gravity equations. For instance, substituting eq. (B3) into equation (28) rewritten as:

$$PX_{ij}^g = \left[\gamma / \Phi^g\left(\sigma^g - 1\right)\right]\left(s_i^g GDP_i / p_i^g\right) GDP_j\left(1+c_{ij}^g\right)^{1-\sigma g}\left(1+t_i^g\right)^{-\sigma g}\left(1+t_j\right)\left(p_i^g / p_j^g\right)^{1-\sigma g}$$

assuming $t_j = 0$, and recalling $s_i^g\, GDP_{\mathrm{i}} / p_i^g = n_i^g$, yields:

$$PX_{ij}^{g} = \left[\gamma / \Phi^{g}\left(\sigma^{g} - 1\right)\right] n_i^g GDP_j \left(1 + c_{ij}^{g}\right)^{1-\sigma g}\left(1 + t_{ij}^{g}\right)^{-\sigma g}\left[\frac{\left(GDP_i^g / n_i^g GDP_W^g\right)^{1/(1-\sigma g)}}{P_i^{g^*} P_j^{g^*}}\right]^{1-\sigma g}$$

or

$$PX_{ij}^{g} = \left[\gamma / GDP_W^g \Phi^g\left(\sigma^{g} - 1\right)\right] s_i^g GDP_i\ GDP_j \left(1 + c_{ij}^{g}\right)^{-(\sigma g-1)}\left(1 + t_{ij}^{g}\right)^{-\sigma g}\left(P_i^{g^*} P_j^{g^*}\right)^{\sigma g-1} \tag{B8}$$

The sum of eq. (B8) and its analogue for PX_{ji}^{g} (assuming $c_{ij}^{g} = c_{ji}^{g}$ and $t_{ij}^{g} = t_{ji}^{g}$) yields eq. (30) in the text.

Appendix C
Endogenous Equations
(Not intended for publication)

Consumptions of the Home Good or Service by the Home Country

(T1-T6)
$$g_{jj} = \frac{\gamma[w_j + r_j(K_j/L_j)](\Psi_j^g)^{-1}}{n_j^g p_j^g[\gamma\Omega_j^g(\Psi_j^g)^{-1} + (1-\gamma)\Omega_j^s(\Psi_j^s)^{-1}]} \quad j = 1A,1B,2A,2B,3A,3B$$

(T7-T12)
$$s_{jj} = \frac{(1-\gamma)[w_j + r_j(K_j/L_j)](\Psi_j^s)^{-1}}{n_j^s p_j^s[\gamma\Omega_j^g(\Psi_j^g)^{-1} + (1-\gamma)\Omega_j^s(\Psi_j^s)^{-1}]} \quad j = 1A,\ldots,3B$$

Numbers of Varieties

(T13-T18)
$$n_i^g = \frac{z^g(K_i^g)^{\alpha g}(L_i^g)^{1-\alpha g}}{g + \varphi^g} \quad i = 1A,\ldots,3B$$

(T19-T24)
$$n_i^s = \frac{z^s(K_i^s)^{\alpha s}(L_i^s)^{1-\alpha s}}{s + \varphi^s} \quad i = 1A,\ldots,3B$$

Equality of Factor Prices and Marginal Products between Sectors within a Country

(T25-T36)
$$r_i = \frac{\alpha^g\theta^g p_i^g z_i^g (K_i^g)^{\alpha g}(L_i^g)^{1-\alpha g}}{K_i^g} = \frac{\alpha^s\theta^s p_i^s z_i^s (K_i^g)^{\alpha s}(L_i^s)^{1-\alpha s}}{K_i^s} \quad i = 1A, \ldots, 3B$$

(T37-T48)
$$w_i = \frac{(1-\alpha^s)\theta^s z_i^s (K_i^s)^{\alpha s}(L_i^s)^{1-\alpha s} p_i^s}{L_i^s} = \frac{(1-\alpha^g)\theta^g z_i^g (K_i^g)(L_i^g)^{1-\alpha g} p_i^g}{L_i^g} \quad i = 1A, \ldots, 3B$$

Market-Clearing Conditions in Each Sector

(T49-T54)
$$g_i = \frac{\varphi^g\theta^g}{1-\theta^g} = L_i\, g_{ii} + L_{i'}\left[\frac{(1-a^g)^{\theta^g}}{1+(1-a^g)t_{ii'}^g}\;\frac{p_{i'}^g}{p_i^g}\right]^{\frac{1}{1-\theta^g}} g_{i'i'} +$$

$$\sum_{j \neq i,i'} L_j\left[\frac{((1-a^g)(1-b^g))^{\theta^g}}{1 + (1-a^g)(1-b^g)t_{ij}^g}\;\frac{p_j^g}{p_i^g}\right]^{\frac{1}{1-\theta^g}} g_{jj} \qquad i = 1A, \ldots, 3B$$

(T55-T60)

$$s_i = \frac{\varphi^s\theta^s}{1-\theta^s} = L_i\, s_{ii} + L_{i'}\left[\frac{(1-a^s)^{\theta^s}}{1+(1-a^s)t^s_{ii'}}\;\frac{p^s_{i'}}{p^s_i}\right]^{\frac{1}{1-\theta^s}} s_{i'i'} +$$

$$\sum_{j\neq i,i'} L_j\left[\frac{((1-a^s)(1-b^s))^{\theta s}}{1+(1-a^s)(1-b^s)t^s_{ij}}\;\frac{p^s_j}{p^s_i}\right]^{\frac{1}{1-\theta^s}} s_{jj} \qquad i = 1A,\ .\ .\ .\ ,3B$$

In above equation (T49-T60), for each country i (i=1A, ..., 3B), i' denotes the other country on the same continent and j ($j\neq i, i'$) denotes countries on different continents.

In equations (T61-T84) below, for each country j (j=1A, ..., 3B), j' denotes the other country on the same continent and i ($i\neq j, j'$) denotes countries on different continents.

Identities
(T61-T84)
i=1A,. . . .3B

$$\Psi^g_j = 1 + \left(\frac{(1-a^g)}{1+(1-a^g)t^g_{j'j}}\frac{p^g_j}{p^g_{j'}}\right)^{\frac{\theta^g}{1-\theta^g}}\left(\frac{z^g_{j'}(K^g_{j'})^{\alpha g}(L^g_{j'})^{1-\alpha g}}{z^g_j(K^g_j)^{\alpha g}(L^g_j)^{1-\alpha g}}\right)$$
$$+\sum_{i\neq j,j'}\left[\left(\frac{(1-a^g)(1-b^g)}{1+(1-a^g)(1-b^g)t^g_{ij}}\right)\left(\frac{p^g_j}{p^g_i}\right)\right]^{\frac{\theta^g}{1-\theta^g}}\left(\frac{z^g_i(K^g_i)^{\alpha g}(L_i)^{1-\alpha g}}{z^g_j(K^g_j)^{\alpha g}(L^g_j)^{1-\alpha g}}\right)$$

$$\Psi^s_j = 1 + \left(\frac{(1-a^s)}{1+(1-a^g)t^s_{j'j}}\frac{p^s_j}{p^s_{j'}}\right)^{\frac{\theta^s}{1-\theta^s}}\left(\frac{z^s_{j'}(K^s_{j'})^{\alpha s}(L^s_{j'})^{1-\alpha s}}{z^s_j(K^s_j)^{\alpha s}(L^s_j)^{1-\alpha s}}\right)$$
$$+\sum_{i\neq j,j'}\left[\left(\frac{(1-a^s)(1\text{-}b^s)}{1+(1-a^s)(1\text{-}b^s)t^s_{ij}}\right)\left(\frac{p^s_j}{p^s_i}\right)\right]^{\frac{\theta^s}{1-\theta^s}}\left(\frac{z^s_i(K^s_i)^{\alpha s}(L^s_i)^{1-\alpha s}}{z^s_j(K^s_j)^{\alpha s}(L^s_j)^{1-\alpha s}}\right)$$

$$\Omega^g_j = 1 + \left(\frac{(1-a^g)^{\theta g}}{1+(1-a^g)t^g_{j'j}}\left(\frac{p^g_j}{p^g_{j'}}\right)^{\theta g}\right)^{\frac{1}{1-\theta^g}}\left(\frac{z^g_{j'}(K^g_{j'})^{\alpha g}(L^g_{j'})^{1-\alpha g}}{z^g_j(K^g_j)^{\alpha g}(L^g_j)^{1-\alpha g}}\right)$$
$$+\sum_{i\neq j,j'}\left[\left(\frac{(1-a^g)(1-b^g)^{\theta g}}{1+(1-a^g)(1-b^g)t^g_{ij}}\right)\left(\frac{p^g_j}{p^g_i}\right)^{\theta g}\right]^{\frac{1}{1-\theta^g}}\left(\frac{z^g_i(K^g_i)^{\alpha g}(L^g_i)^{1-\alpha g}}{z^g_j(K^g_j)^{\alpha g}(L^g_j)^{1-\alpha g}}\right)$$

$$\Omega^s_j = 1 + \left(\frac{(1-a^s)^{\theta s}}{1+(1-a^s)t^s_{j'j}}\left(\frac{p^s_j}{p^s_{j'}}\right)\right)^{\frac{1}{1-\theta^s}}\left(\frac{z^s_{j'}(K^s_{j'})^{\alpha s}(L^s_{j'})^{1-\alpha s}}{z^s_j(K^s_j)^{\alpha s}(L^s_j)^{1-\alpha s}}\right)$$
$$+\sum_{i\neq j,j'}\left[\left(\frac{(1-a^s)(1-b^s)^{\theta s}}{1+(1-a^s)(1-b^s)t^s_{ij}}\right)\left(\frac{p^s_j}{p^s_i}\right)\right]^{\frac{1}{1-\theta^s}}\left(\frac{z^s_i(K^s_i)^{\alpha s}(L^s_i)^{1-\alpha s}}{z^s_j(K^s_j)^{\alpha s}(L^s_j)^{1-\alpha s}}\right)$$

Appendix D

(Not intended for publication)

We show here that the inclusion of *IMBALANCE* (= * $\ln PX_{ij}^{g}$ - $\ln PX_{ji}^{g}$ *) does *not* generate a logical inconsistency in the estimation. To see this, let:

$$\ln PX_{ij}^{g} = \mu_1 + \alpha\, FTA_{ij} + \boldsymbol{\beta' q_{ij}} + u_{ij} \tag{D1}$$

$$\ln PX_{ji}^{g} = \mu_1 + \alpha\, FTA_{ji} + \boldsymbol{\beta' q_{ji}} + u_{ji} \tag{D2}$$

and assume $FTA_{ij} = FTA_{ji}$ and:

$$fta = \mu_2 + \boldsymbol{\delta' z_{ij}} + \theta * \ln PX_{ij}^{g} - \ln PX_{ji}^{g} * + w_{ij} \tag{D3}$$

Substitute equations (D1) and (D2) into (D3) to yield:

$$fta = \mu_2 + \boldsymbol{\delta' z_{ij}} + \theta * \boldsymbol{\beta'(q_{ij} - q_{ji})} + (u_{ij} - u_{ji}) * + w_{ij} \tag{37}$$

If $fta > 0$, $FTA = 1$ and:

$$P(FTA = 1) = P(fta > 0) = \Phi[\mu_2 + \boldsymbol{\delta' z_{ij}} + \theta * \boldsymbol{\beta'(q_{ij} - q_{ji})} + (u_{ij} - u_{ji}) *] .$$

If fta # 0, $FTA = 0$ and:

$$P(FTA = 0) = P(fta \# 0) = 1 - \Phi[\mu_2 + \boldsymbol{\delta' z_{ij}} + \theta * \boldsymbol{\beta'(q_{ij} - q_{ji})} + (u_{ij} - u_{ji}) *] .$$

For internal consistency, $P(FTA = 0) + P(FTA = 1) = 1$. Hence, the model is only logically consistent if:

$$1 - \Phi[\mu_2 + \boldsymbol{\delta' z_{ij}} + \theta * \boldsymbol{\beta'(q_{ij} - q_{ji})} + (u_{ij} - u_{ji}) *] + \Phi[\mu_2 + \boldsymbol{\delta' z_{ij}} + \theta * \boldsymbol{\beta'(q_{ij} - q_{ji})} + (u_{ij} - u_{ji}) *] = 1 .$$

which holds for all parameter values (since no term $\alpha\theta$ shows up).