

The Political Economy of Free Trade Agreements: an Empirical Investigation

Xuepeng Liu

Kennesaw State University

Abstract

This paper studies how income inequality affects the probability of signing free trade agreements (FTAs) in a political economy framework. The median voter approach to FTA formation within a Heckscher-Ohlin framework predicts that an increase in inequality lowers the desirability of an FTA if the partner country is relatively labor abundant, ceteris paribus, and raises the desirability of such an agreement if the partner country is relatively capital abundant. The lobbying model based on the argument of the free rider problem, however, offers the opposite predictions. This paper proposes a duration analysis for bilateral FTA panel data and finds strong support for the predictions from the median voter model.

- **JEL classification:** F15, F12
- **Keywords:** Free trade agreements, Political economy, Inequality, Median voter model, Lobbying model, Duration analysis

I. Introduction

Since the European Economic Community (EEC) came into force in 1958, there has been a steady rise in the number of regional trade agreements (RTAs) reported to the GATT/WTO. These RTAs violate the principle of nondiscrimination (the MFN Clause) and are fundamentally altering the world trade landscape. Same as

*Corresponding address: Department of Economics and Finance, Kennesaw State University, #0403, Burruss Bldg #4, Room 322, 1000 Chastain Road, Kennesaw, GA 30144, USA; Email: xliu6@kennesaw.edu; Tel: (770)423-6605.

tariffs and non-tariff barriers, signing an RTA is also endogenous. So what are the driving forces behind these RTAs? Besides some economic (e.g. the market sizes of partner countries) and geographic (e.g. bilateral distance) factors, political economy factors are also important in RTA formation. In this paper, we investigate RTA formation in a political economy framework. Only free trade areas and customs unions are considered and they are together called free trade agreements (FTAs) in this paper. The partial scope preferential arrangements signed under the Enabling Clause of the GATT are not included because these arrangements are not full-fledged and were usually signed among developing countries with very limited product coverage.¹ In addition, service agreements, a type of RTA notification to the WTO, are not treated as separate agreements in this paper because they were usually signed together or after the corresponding FTAs in goods.

According to the Stolper-Samuelson theorem, there are both winners and losers from FTAs. This redistributive consequence will present as long as a country differs from its trading partners in factor endowments. FTA decisions will be eventually determined by how the preferences of politicians and their parties, interest groups and voters are aggregated in political processes. To fully understand FTAs, it is important to open the black box of the political economy of FTA formation. Where theories deliver empirically testable predictions, there are two main branches in the literature of the political economy of trade policy (Gawande and Krishna, 2003). The first branch represents direct democracy or median voter approach with the implicit assumption that voters vote directly upon trade policy or government chooses trade policy according to majority opinions (general interest). The second is the lobbying or interest group approach (special interest).

The testable hypothesis from the median voter model in this paper is derived from Levy's political economy model of FTAs (Levy, 1997), which adopts the median voter approach proposed by Mayer (1984). The main theoretical proposition is driven by the Stolper-Samuelson theorem. According to this theorem, in capital-abundant countries, capitalists are winners and workers are losers from free trade; and vice versa in labor-abundant countries. In an unequal country, the distribution of capital stock is right-skewed because a small group of rich people almost always own disproportionately large share of capital. So the person with

¹The Enabling Clause of the GATT permits developing countries to exchange partial tariff preference within arrangement such as ASEAN (Association of Southeast Asian Nations). In a famous example, Indonesia, a member of ASEAN and a country on the equator, excluded major sectors but offered to liberalize imports of snow ploughs from other ASEAN countries (Panagariya, 1994).

median capital-labor ratio (K/L) is always capital-poor and labor-rich relative to the average K/L in the country. If trade policy is determined by the median voter, then it will be pro-labor. This in turn implies more trade protection in capital-abundant countries and freer trade in labor-abundant countries according to the Stolper-Samuelson theorem. Because the inequality level determines the relative position of the median voter in a country in terms of K/L , the changes in inequality will affect the median voter's trade preference. Higher inequality makes the median voter more capital-poor, which will make trade policies even more pro-labor. Hence an increase in inequality will cause more protection in capital-abundant countries, but more liberalization in labor-abundant countries.

A lobbying model within the same two-sector, two-factor Heckscher-Ohlin framework, however, offers the opposite predictions from the median voter model. The lobbying model, as in Rodrik (1986), assumes that it is easier for capitalists than workers to overcome the free-rider problem in lobbying because capitalists are more concentrated and more powerful individually than workers.² If capitalists lobby more effectively than workers, then trade policy will be pro-capital. An increase in inequality implies that an even larger share of capital is accumulated in the hands of fewer and bigger capitalists. Hence they can carry out collective lobbying more effectively and make trade policy even more pro-capital, which will lead to freer trade in capital-abundant countries, but more protection in labor-abundant countries.

It is largely an empirical question which model is correct. This paper contributes to the literature by empirically testing these conflicting theories in the context of FTAs. Our empirical strategy is similar to Dutt and Mitra (2002). They use aggregate trade policy data, such as tariffs, quotas and openness, and find empirical support for the median voter model proposed by Mayer (1984). The aggregate trade policy data, however, conceal the discriminatory nature of trade policy. This calls for a bilateral analysis. We use a large bilateral FTA panel dataset to test the two conflicting models. With a bilateral dataset, we can tell which country is relatively labor-abundant or capital-abundant and provide a more direct test of the models. In addition, with bilateral data, we can use the difference between two countries' capital labor ratios to capture more precisely the magnitude of the Heckscher-Ohlin distributional effects.

²In Rodrik's model, if we hold constant the aggregate capital stock and the total population, an increase (decrease) in inequality can be represented by a decrease (increase) in the number of capitalists.

Researchers usually resort to standard discrete choice model (e.g. probit or logit) to analyze binary FTA data. With panel data, however, this method is inappropriate because it requires the conditional independence assumption for FTAs. Due to the binary and irreversible nature of FTAs, this assumption is obviously violated. For example, the NAFTA after 1994 is, by no means, independent of the NAFTA in 1994. In this paper, we propose a duration analysis to address this problem in a panel data setting. Our results provide strong support for the predictions from the median voter model.

It is important to note that the support of the median voter model should not be taken too literally in the context of direct democracy and single-peaked preference. We can have exactly the same pro-labor policies as predicted by the median voter model if a government is concerned about inequality, as in Limao and Panagariya (2006). The intuition is clear: if a government is concerned about inequality, trade policies will be pro-labor as workers are poorer than capitalists. Therefore, the median voter model should be taken as a short-hand way to capture the effects of electoral competition, general concerns for popular support or partisan ideological preference for equality. As Alesina and Rodrik (1994) point out, “We appeal to this theorem to capture the basic idea that any government is likely to be responsive to the wishes of the majority when key distributional issues are at stake. Even a dictator cannot completely ignore social demands for fear of being overthrown. Thus, even in a dictatorship, distributional issues affecting the majority of the population will influence policy outcomes.”

The results in this paper deliver some policy implications for both the “static impact effect” of FTAs and the “dynamic time path” issues related to the move from regionalism to multilateralism. Coined by Bhagwati (1993), the “static impact effect” of regionalism is the immediate impact of FTAs on world welfare; and the “dynamic time path” discusses the role of FTAs as “stepping stones” or “stumbling blocks” towards multilateralism. “Static impact effect” arguments focus on trade creation and diversion. According to the classic analysis of Viner (1950), an FTA is welfare improving if the trade creation effect dominates the trade diversion effect. Manipulations by special interest groups representing import-competing sectors are often considered a source of trade diversion and stumbling block effect. Grossman and Helpman (1995) show in a lobbying model that the exclusion of some highly-protected sectors from an FTA can dilute the opposition of pressure groups and increase the probability that the FTA is adopted. Frankel, Wei and Stein (1997, chapter 10) and Schiff and Winters (2003, chapter 3) also argue that lobbies bias

FTAs toward trade diversion and may prevent an FTA from being a “stepping stone” toward multilateralism. These dangers are elegantly demonstrated by Krishna (1998) in a three-country model. The concern about adverse effects of lobbying is probably the primary reason why the GATT Article XXIV stipulates that free trade areas and custom unions should cover “substantially all the trade” within the blocs, not just a limited number of sectors. The support for the median voter model rather than the lobbying model sends an encouraging message to the current movement of regionalism.

This paper makes several contributions to the literature. Firstly, we derive a modified political economy model of FTAs proposed by Levy (1997) and provide the empirical support for the median voter model. Secondly, we propose a more appropriate empirical strategy for bilateral panel data analysis than has been seen so far in the empirical literature on FTAs. Thirdly, the problem of partial observability of FTA formation is also considered for the first time in the FTA literature. Finally, this paper is one of the first empirical studies on the political economy of FTAs and adds new evidences to the political economy determinants of FTAs.

The rest of this paper is organized as follows. In Section 2, we present a modified version of Levy’s median voter model of FTAs and perform some comparative statics; Section 3 lays out the empirical strategies; Section 4 discusses the econometric methodology. Section 5 describes the data. Regression results are presented in Section 6. Section 7 provides some robustness checks and discusses some related complexities. And Section VIII concludes.

II. Theoretical Framework

Levy (1997) develops a median voter theory of FTAs with differentiated products in a Heckscher-Ohlin framework. In the paper, Levy studies how FTAs may serve as stumbling blocks or stepping stones for the multilateral trading system. We instead focus on the model’s implications on how inequality might affect FTA formation. We modify Levy’s model by changing the production function of differentiated products to make sure that the equilibrium solution always exists.³

³Levy assumes a homothetic production function for differentiated product X as $x = \gamma_X K_X^{\xi\eta} L_X^{\xi(1-\eta)}$, where $\xi > 1$ is the IRS parameter. This production function leads to indeterminacy of the model as the profit maximization condition (MC=MR) and the free entry condition (p=AC) are parallel to each other and will never cross to give a solution in x . We solve this problem by adding a fixed cost.

The model assumes that countries differ only in factor endowments – capital (K) and labor (L) – and the distribution of factor ownership. There are two sectors of production: homogeneous products (Y) and differentiated products (X). The homogeneous products (Y , numeraire goods) are produced under constant returns to scale with a production function defined as $y = \gamma_Y K_Y^\mu L_Y^{1-\mu}$, where γ is a productivity parameter and $\mu \in (0, 1)$ is the elasticity of Y 's output with respect to capital. The differentiated products (X) are produced under increasing returns to scale (IRS). For each variety, the production function is $x = \gamma_X K_X^\eta L_X^{1-\eta} - a$, where $\eta \in (0, 1)$ is the elasticity of X 's output with respect to capital and a is the fixed cost measured in the unit of X . Having a fixed cost is the tradition of monopolistic competition trade literature, and this fixes the indeterminacy problem of solution in the original model by Levy. The cost functions can be derived from the production functions of X and Y as $C_Y(w, r) = C_Y(w, r)y$ and $C_X(w, r) = c_X(w, r)(x + a)$, where $c_Y(w, r)$ is the unit cost of the goods Y and $c_X(w, r)$ is the marginal cost function of product X . Both cost functions are separable.

Each agent i owns one unit of labor and k_i unit of capital. Hence the income of agent i is $I_i = rk_i + w$, ($i = 1, 2, \dots, L$), where r is interest rate and w is wage rate. Agents are assumed to have identical utility functions as $U = U_X^\alpha y^{1-\alpha}$, where U_X is the sub-utility function for consumptions in X , with a Dixit-Spence-Stiglitz type CES functional form.

$$U_X = \left(\sum_{j=1}^n D_j^\beta \right)^{\frac{1}{\beta}}, \beta = \left(1 - \frac{1}{\sigma} \right), \sigma > 1$$

where D_j is the consumption of variety x_j by an agent; n is the number of varieties; and σ is the elasticity of substitution between varieties. Following a two-stage budgeting process, agent i 's optimal consumption of Y and variety x_j are $y = I_i(1 - \alpha)$ and $D_j = \alpha I_i / (np)$ respectively, where p is the relative price of X in terms of Y . Substituting these optimal consumptions into the utility function for agent i yields $U_i = I_i(1 - \alpha)^{1-\alpha} \alpha^\alpha n^{\alpha/(\sigma-1)} p^{-\alpha}$. When agents vote for or against an FTA, they need to compare their utilities before the FTA to their utilities after the FTA. Agent i 's utility under FTA relative to autarky can be written as:

$$\frac{U_i^{FTA}}{U_i^{AUT}} = \left(\frac{I_i^{FTA}}{I_i^{AUT}}\right) \left(\frac{p^{FTA}}{p^{AUT}}\right)^{-\alpha} \left(\frac{n^{FTA}}{n^{AUT}}\right)^{\alpha/(\sigma-1)} \tag{1}$$

Levy (1997) calls the first two terms on the right hand side of equation (1) the “comparative advantage effect” (CAE) and calls the last term on the right hand side of equation (1) the “variety effect” (VE). The magnitude of the relative utility determines the desirability of an FTA to the agent. From now on, let us assume this agent is the median voter; then the relative utility of this “marginal” voter determines the desirability of an FTA to the country under majority voting. The reduced forms of CAE and VE for the median voter can be solved from a general equilibrium system.

$$CAE = \left(\frac{I_{mv}^{FTA}}{I_{mv}^{AUT}}\right) \left(\frac{p^{FTA}}{p^{AUT}}\right)^{-\alpha} = \left(\frac{b + \rho_{mv}/\varphi}{b + \rho_{mv}}\right) \varphi^{[\alpha\eta + (1-\alpha)\mu]} \tag{2}^4$$

$$VE = \left(\frac{n^{FTA}}{n^{AUT}}\right)^{\frac{\alpha}{\sigma-1}} = [(1 + \lambda_K)^\eta (1 + \lambda_L)^{1-\eta}]^{\frac{\alpha}{\sigma-1}} > 1 \tag{3}$$

where $b = [1 - \mu - \alpha\beta(\eta - \mu)] / [\mu + \alpha\beta(\eta - \mu)]$ is a function of parameters; $\lambda_K = (K^{FTA} / K^{AUT}) - 1$ and $\lambda_L = (L^{FTA} / L^{AUT}) - 1$ are the percentage increases in K and L when a country moves from autarky to an integrated economy resulting from an FTA; $\varphi = k^{FTA} / k^{AUT}$ is the capital labor ratio of the integrated economy under an FTA relative to the capital labor ratio of a country under autarky ($\varphi > 1$ if the country’s partner is relatively capital-abundant and $\varphi < 1$ if the partner is relatively labor-abundant); and $\rho_{mv} = k_{mv} / k^{AUT}$ is the median voter’s capital labor ratio relative to the overall capital labor ratio under autarky, which is always less than one for any unequal society and can be used to measure the degree of “equality”.⁵

Equation (2) offers a testable prediction on the relationship between equality and

⁴Our equation (2) is similar to equation (15) on page 516 in Levy (1997). His formula of parameter b , which is called θ in Levy (1997), is incorrect. Our equation (3) is same as equation (12) on page 515 in Levy (1997). The homothetic production function in Levy (1997) does not affect the main results of the model largely due to another mistake he makes on the free-entry condition. Specifically Levy’s formula for the optimal production of x^* on page 515 in Levy (1997) is incorrect based on his homothetic production function. Without the second mistake, he would not have been able to solve for x^* or have a “well-defined” equilibrium. Please refer to Liu (2007) for detailed discussion on these issues.

the desirability of an FTA conditional on a country's K/L . To see this more clearly, let us take natural logarithm of the CAE and obtain the derivative with respect to equality (ρ_{mv}).

$$\ln CAE = f(.) = [\alpha\eta + (1 - \alpha)\mu]\ln\varphi + \ln\left(b + \frac{\rho_{mv}}{\varphi}\right) - \ln(b + \rho_{mv})$$

$$\frac{\partial f}{\partial \rho_{mv}} = \frac{b(1 - \varphi)}{(\rho_{mv} + \varphi b)(\rho_{mv} + b)} < 0 \text{ if } \varphi > 1;$$

$$\frac{\partial f}{\partial \rho_{mv}} = \frac{b(1 - \varphi)}{(\rho_{mv} + \varphi b)(\rho_{mv} + b)} > 0 \text{ if } \varphi < 1.$$

These derivatives imply that an increase in equality ρ_{mv} reduces the CAE, hence reduces the desirability of an FTA, ceteris paribus, if the partner country is more capital-abundant ($\varphi > 1$), and vice versa if the partner country is more labor-abundant ($\varphi < 1$). Translating this in terms of inequality, we have the following proposition.

Proposition 1: An increase in inequality raises the desirability of an FTA, ceteris paribus, if the partner country is more capital abundant ($\varphi > 1$). An increase in inequality reduces the desirability of an FTA, ceteris paribus, if the partner country is more labor abundant ($\varphi < 1$).

The intuition of this proposition is as follows. With a higher degree of income inequality, the median voter is likely to be relatively better endowed with labor and would stand to benefit (lose) more from an FTA with country that is relatively more capital rich (poor). In the empirical analysis, we regress an FTA dummy on countries' inequality measure (*Gini* coefficient) or equality measure (the income share of third quintile, $Q3$). To take the magnitude of the Heckscher-Ohlin effect into account, *Gini* ($Q3$) is multiplied by the difference in capital labor ratios of the two countries in a pair.

The median voter theory shown above also offers another testable prediction. The variety effect in equation (3) is always greater than one because λ_K and λ_L

⁵For example, if we classify all citizens in a country into five quintiles according to their capital assets per capita and the median group (third quintile) has totally L_3 persons and K_3 capital assets, then we can show that ρ_3 (or ρ_{mv}) is equivalent to five times $Q3$ (the capital or income share of the third quintile – a often used measure of equality).

$$\rho_{mv} = \rho_3 = \frac{k_3}{k^{AUT}} = \frac{K_3/L_3}{K^{AUT}/L^{AUT}} = \frac{K_3/L_3}{K^{AUT}/(5L_3)} = 5 \frac{K_3}{K^{AUT}} = 5Q_3$$

are both positive. The larger the partner country in terms of K and L , the larger λ_K and λ_L are, hence the larger the variety effect will be. Therefore every country wants to sign FTAs with big partners. Because signing an FTA is a joint decision, another testable prediction can be stated as follows:

Proposition 2: An FTA is more likely signed between two big countries.

Equation (3) shows that an increase in n^{FTA} will raise VE , which implies that a country would prefer a big rather small country as its FTA partner. As a result of joint decision, an FTA is more likely signed between two big countries and less likely signed between countries with very different economic sizes. Empirically, we use the logarithm of the product of the two countries' GDPs in a pair (or the sum of $\log(\text{GDP})$) and the logarithm of the ratio of the two countries' GDPs (or the difference in $\log(\text{GDP})$) to test for the variety effect hypothesis. The sum of $\log(\text{GDP})$ is expected to be positively correlated with the probability of signing an FTA; and the difference in $\log(\text{GDP})$ is expected to be negatively correlated with the probability.

It is important to control for the variety effects and other factors in the analysis. For example, the median voter model alone suggests that no government will be willing to form an FTA with labor-abundant partners even if its income inequality is low because the median voter will lose from such an FTA. When we also consider the variety effects and other control factors, however, the median voter in a capital-abundant country can actually gain overall from an FTA with labor-abundant countries despite the loss from comparative advantage effect. Therefore both propositions 1 and 2 should be considered in the analysis of FTA formation.

III. Estimation Strategies

We are interested in testing how income inequality might affect the likelihood of FTAs based on the capital labor ratios of the two countries in a pair. In other words, how is the probability of an FTA affected by the interactions of the inequalities and capital labor ratios of the two countries? *Gini* coefficient is used as the measure of inequality. Alternatively, the income share of the third (median) quintile of the income distribution ($Q3$) is used as a measure of "equality". The two measures are usually highly negatively correlated. The expected signs of the coefficients of *Gini* and $Q3$ should be opposite to each other. The model specification in our empirical analysis is given by

$$\begin{aligned}
prob(FTA_{ij}) = & \beta_0 + \beta_1[\log(GDP_i) + \log(GDP_j)] + \beta_2|\log(GDP_i) - \log(GDP_j)| \\
& + \beta_3\log(Distance_{ij}) + \beta_4|\log(k_i) - \log(k_j)| + \beta_5GINI_i^*[\log(k_i) - \log(k_j)] \\
& + \beta_6GINI_j^*[\log(k_i) - \log(k_j)] + X_{ij}\gamma + \varepsilon_{ij}
\end{aligned} \tag{4}$$

The dependent variable is a dummy, indicating whether the two countries i and j had FTA relationship or not in a given year.⁶ The first six variables are of particular interests. X_{ij} is a vector of some geographic, institutional and political variables, and ε_{ij} is the error term. Except the geographic covariates, most variables are time-varying. The year subscripts, however, are omitted to simplify the formula.

The first two variables are the sum and the absolute difference of $\log(GDP)$ of the two countries for each observation, which are used to test for the variety effect as described in previous section. We expect that β_1 and β_2 are positive and negative respectively. The third variable is the geographic distance between the two countries. The fourth variable measures the absolute difference in the logarithms of the capital per worker (k) of the two countries in a pair. Based on the standard Heckscher-Ohlin model, β_4 should be positive; that is, the larger the difference in endowments the bigger the gain from an FTA.

The key variables of our interest are the fifth and sixth variables: the interactions between *Gini* ($Q3$) and the difference (*not the absolute difference*) in the logarithms of the two countries' capital per worker. Some discussions are warranted to determine the expected signs of β_5 and β_6 . For now, let us focus on the median voter model. The dataset used in our analysis is bilateral and each observation is associated with two countries, i and j . For each observation, the capital per worker of country i (k_i) can be either bigger or smaller than the capital per worker of country j (k_j). Fortunately, no matter country i is a rich country ($k_i > k_j$) or a poor country ($k_i < k_j$) compared to country j , the expected sign of β_5 is always negative according to the median voter model. If $k_i > k_j$, then β_5 should be negative so that the likelihood of an FTA is decreasing in $GINI_i$ (rich country's inequality). In other words, if $GINI_i$ (rich country) negatively affects the probability of an FTA and the term $[\log(k_i) - \log(k_j)]$ is positive, then β_5 must be negative. This implies that the rich partner is less willing to enter into an FTA if they are more unequal because the median voter holding relatively less capital will lose more from free

⁶Please note that i and j here are country subscripts and have nothing to do with the subscripts i and j in section II.

trade. This effect is augmented by $[\log(k_i) - \log(k_j)]$. The bigger the term $[\log(k_i) - \log(k_j)]$, the stronger the Heckscher-Ohlin distributional effects. On the contrary, if $k_i < k_j$, then $[\log(k_i) - \log(k_j)] < 0$ and β_5 should *still* be negative so that the likelihood of an FTA is *increasing* in $GINI_i$ (poor country's inequality). In other words, if $GINI_i$ (poor country) positively affects the probability of an FTA and the term $[\log(k_i) - \log(k_j)]$ is negative, then β_5 must be negative. This implies that the poor partner is more inclined to enter into an FTA if they are more unequal because the median voter holding relatively less capital will gain more from free trade. Similarly, no matter country i is a rich country ($k_i > k_j$) or poor country ($k_i < k_j$) compared to country j , the expected sign of β_6 is always positive according to the median voter model. In sum, when we use *Gini* coefficients in regressions, we should expect negative and positive signs on β_5 and β_6 respectively.⁷

The above discussions are about the median voter model. If the lobbying model is instead supported, all the sign patterns discussed above will be reversed. Apparently, the sign pattern should be exactly opposite if we substitute $Q3$ for *Gini* in the regressions.

IV. Econometric Methodology

Researchers usually resort to standard discrete choice model (e.g. probit or logit) to analyze binary FTA data. Baier and Bergstrand (2004) use a cross-sectional dataset and focus on the economic determinants of FTAs. Magee (2004) extends the analysis to a panel data setting and also considers some political variables. Both papers use bilateral data and run probit regressions on an FTA dummy variable. A cross section analysis as in Baier and Bergstrand (2004) seems acceptable if much of the variations come from cross sectional dimension. But panel data allow us to use more information on the time-varying variables and easily deal with unobserved heterogeneity. Therefore, the next natural step is to explore the richer information in panel data. The standard probit and logit, however, are inappropriate for panel data because they assume that the dependent variable (FTA dummy) is conditionally independent over time. To see this, let us take the NAFTA as an example. Mexico signed the NAFTA with the US and Canada in 1994, and they will have the NAFTA for all the following years. It is problematic to assume that

⁷To be exhaustive, we could also add the difference between countries' inequality. This variable is always highly insignificant and has little impact on the results, but makes the interpretation of the coefficients more complicated. Therefore the difference in inequality is not included in the regressions.

the NAFTA in 1995 and after is independent to the NAFTA in 1994, no matter how many variables are used as controls. Using standard probit or logit with FTA panel data, we would blindly predict something that is already predetermined and the results would be misleading. Traditional techniques dealing with temporal correlation will not help as the problem here is much more serious than serial correlation. This is a problem of exact correlation due to the irreversible nature of FTAs and will lead to serious bias in the estimation.⁸

In this paper, we propose a duration analysis to solve this problem. It is useful to note that we are only interested in the spells before the formation of FTAs, not the spells after FTAs. If we do not use the information after FTAs were formed, we can effectively avoid the conditional dependence problem as in the standard logit analysis. This insight leads us to consider the duration analysis (or survival analysis). Depending on the nature of time intervals, there are two types of duration analysis: continuous-time and discrete-time analysis. The continuous-time duration analysis, such as the Weibull model and the Cox proportional hazard model, has been widely used in economics literature, such as the studies on the durations of unemployment and Medicaid programs. The discrete-time duration analysis is relatively more recent, but becomes increasingly popular due to several advantages over the continuous-time analysis. For example, the discrete-time duration analysis combines both time-varying covariates and flexible specifications of duration dependence, while continuous-time duration analysis usually uses data at a point in time. The FTA data are discrete in time (or grouped by years) and many variables are time-varying. The time dimension is especially important as a fairly long time period (40 years) is covered in our FTA dataset. Strictly speaking, continuous-time duration analysis does not allow “ties” because it assumes that the time is continuous and two incidences such as signing FTAs should not happen exactly at the same moment or within the same time interval (i.e., year). A discrete-duration analysis, however, allows “ties” because it assumes that time is discrete or grouped and many incidences might happen simultaneously in a given time interval. Therefore, discrete-time duration analysis turns out to be a natural choice. Applications with similar data structure as FTAs can be found in international relations literature, such as Beck, Katz and Tucker (1998). They applied the duration analysis to international disputes (a binary dependent variable), using a

⁸Very few trade agreements were suspended so far. For example, Peru once suspended its participation in the Andean Community (CAN) in 1992, but the CAN is not covered in this paper as it is a partial scope preferential arrangement signed under the Enabling Clause of the GATT/WTO.

panel dataset over years 1951-1985.

A. Discrete Time Duration Analysis

The discrete time duration analysis is nothing but the standard complementary log-log (cloglog) regression with duration dependence. All we need is to amend the dataset by dropping all but the first positive outcomes of the dependent variable for each country pair. Dropping the repeated “1”s does not lose any useful information as we are only interested in the spells before FTAs were formed. Once the repeated “1”s are dropped, the problem of conditional dependence in the standard logit analysis disappears. Another advantage of the duration analysis is that it reduces the endogeneity problem of covariates. After dropping the repeated “1”s from the FTA variable, we can hardly think of any feedback from FTAs to inequality and other covariates. Of course, possible simultaneity problem may still exist, but can be alleviated by considering unobserved heterogeneity, as will be discussed later.

Following Jenkins (1995, 2004), we can show that the maximum likelihood function for a discrete time duration analysis has exactly the same form as the standard likelihood function for a binary choice model. For any country pair ij , the hazard rate to sign an FTA in year t is $h_{ijt} = Pr[T_{ij} = t | T_{ij} \geq t - 1]$, which is the probability the two countries signed an FTA in year t conditional on that they had not signed the agreement until year $t-1$. Let t_{ij} denotes the last year of a country pair ij 's spell in the dataset, which is different for different country pairs. In the last year (t_{ij}), country pair ij 's spell is either right censored ($c_{ij} = 0$) or complete ($c_{ij} = 1$). The likelihood contributions of a censored spell and a complete spell are given respectively by a discrete time survivor function and a discrete time density function as follows:

$$L_{ij}^{c_{ij}=0} = \prod_{t=1}^{t_{ij}} (1 - h_{ijt}) \text{ and } L_{ij}^{c_{ij}=1} = \frac{h_{ijt_{ij}}}{1 - h_{ijt_{ij}}} \prod_{t=1}^{t_{ij}} (1 - h_{ijt}) \tag{5}$$

Combining both contributions and then taking logarithm, we have the log likelihood as

$$\log L = \sum_{ij=1}^N c_{ij} \log\left(\frac{h_{ijt_{ij}}}{1 - h_{ijt_{ij}}}\right) + \sum_{ij=1}^N \sum_{t=1}^{t_{ij}} \log(1 - h_{ijt}) \tag{6}$$

Let y_{ijt} denotes the FTA dummy variable. The first term of the right hand side of

equation (6) will disappear if their spells are right censored ($c_{ij}=0$). If the spells are complete ($c_{ij}=1$), then we will have $y_{ijt} = 1$ when $t = t_{ij}$ and $y_{ijt} = 0$ otherwise. Hence equation (6) can be rewritten as

$$\begin{aligned}\log L &= \sum_{ij=1}^N \sum_{t=1}^{t_{ij}} y_{ijt} \log\left(\frac{h_{ijt}}{1-h_{ijt}}\right) + \sum_{ij=1}^N \sum_{t=1}^{t_{ij}} \log(1-h_{ijt}) \\ &= \sum_{ij=1}^N \sum_{t=1}^{t_{ij}} \left[y_{ijt} \log(h_{ijt}) + (1-y_{ijt}) \log(1-h_{ijt}) \right]\end{aligned}$$

which is exactly the traditional log likelihood function for binary response models. If we assume a cloglog hazard $h_{ijt} = 1 - \exp[-\exp(\beta' X_{ijt} + \theta_t)]$, where \mathbf{X} are country characteristics and θ_t is a transformation of the baseline hazard, we can run a cloglog regression on the amended dataset. This is an exact grouped duration analogue of the continuous time Cox proportional hazards model (Prentice and Gloeckler, 1978; Jenkins, 2004). Alternatively, we could assume a logistic hazard $h_{ijt} = 1/[1 + \exp(-\beta' X_{ijt} - \theta_t)]$ and run a logit regression. Logit link function also corresponds to an underlying continuous time duration model in which the within-interval durations follow a log-logistic distribution, but it imposes more complicated time-interactions than the proportional hazard model (Sueyoshi, 1995).

Most researchers are more familiar with logistic than cloglog function. Actually it is easy to show that, when the probability of positive outcomes is small as in the case of FTAs⁹, a cloglog link function is very close to a logistic link function. Hence the coefficients got from cloglog regressions can be exponentiated and understood in terms of odds ratio.

B. Duration Dependence

The duration dependence is analogous to the baseline hazard in the Cox model, which can be captured by time dummies or a polynomial of a time counter since the beginning of spells. The increasing trend in the number of FTAs over time suggests that the time trend must be taken into account.¹⁰ If year dummies are used to account for duration dependence, all the years without new FTAs will be

⁹After dropping all but the first “1”s, positive outcomes account for only 0.3% in the amended dataset.

¹⁰It also captures the “domino effect” of FTAs as suggested by Baldwin (1996).

dropped because the year dummies perfectly predict these zero outcomes. To avoid this problem, we create a time counter to count the number of years already past since 1960 (the beginning of the sample), as well as its square, cubic terms and so on. Then we use the polynomial to account for the time dependence. The order of the polynomial is determined by the best fit of regressions.

C. Unobserved Heterogeneity and Endogeneity

We may think that some country pairs are more likely to form an FTA due to omitted variables. The presence of unobserved heterogeneity attenuates the estimated parameters.¹¹ In duration analysis literature, the unobserved heterogeneity is called “frailty”, analogous to the fixed effects or random effects in panel data analyses. With frailty, the cloglog and logistic hazard functions become $h_{ijt} = 1 - \exp[-\exp(\beta' X_{ijt} + \theta_t + u)]$ and $h_{ijt} = 1/[1 + \exp(-\beta' X_{ijt} - \theta_t - u)]$ respectively. If we assume a normal distribution for the unobserved heterogeneity u , we can use random effects cloglog or random effects logit.¹² Alternatively, Meyer (1990) assumes that unobserved heterogeneity follows a gamma mixture distribution and proposes a model for discrete time proportional hazards regression on grouped duration data. Heckman and Singer (1984) instead propose a nonparametric approach to frailty. We will try both methods as robustness checks.

V. Data

We create a large bilateral panel dataset, covering the period 1960-2000 and 114 countries. Each observation is associated with two countries (i.e. a country pair). A country pair appears only once in the data (either ij or ji , single-dyad). The dependent variable is a dummy variable, which is equal to one if two countries had an FTA in a specific year and zero otherwise. Appendix 1 lists the FTAs covered in this paper and their data sources. We include the FTAs notified to the GATT/WTO between 1960 and 2000, as well as some FTAs which were implemented but have not yet been notified to the WTO.

¹¹Lancaster (1990, chapter 4) provides a proof for the case when unobserved heterogeneity follows a Gamma distribution.

¹²No consistent fixed effect cloglog procedure is available. The fixed effect logit procedure (i.e. conditional logit) is available but inappropriate because most of the country pairs have never signed FTAs and these observations will be dropped in fixed effect logit regressions and we will lose much of the cross sectional variations.

The covariates include economic, geographic, political, historical and military factors.

Capital-labor ratio (capital stock per worker or K/L) data for all years from 1960 to 2000 are from Baier, Dwyer and Tamura (2006). Income inequality data (*Gini* coefficient and $Q3$) for the period 1960-1998 are from Dollar and Kraay (2002).¹³ The inequality data after 1998 are from World Development Indicators (WDI 2003). The gaps of inequality data for some years are filled in by the weighted averages of the observed data in earlier and later years.¹⁴ Inequality data for some countries are extrapolated backward and forward to fill the missing data on both ends.¹⁵

GDP (PPP) data are from the Penn World Table (PWT 6.1). Bilateral trade flows are from the IMF Direction of Trade Statistics (DOT). Both GDP and trade data are measured in constant 1995 U.S. dollars. US GDP deflator data are from the IMF International Financial Statistics (IFS).

Child labor and carbon dioxide (CO₂) data are from the WDI 2003. Child labor is measured as the percentage of child labor in 10-14 age group; and CO₂ emission is measured in metric tons per capita. Polity score data, measuring the degree of democracy, are from the Polity IV Project (Marshall and Jaggers, 2002). The absolute differences in Polity scores, CO₂ per capita (in logarithms) and the shares of child labor are used to account for the concerns about democracy, environment and labor standards in FTA formation.

Geographic data, including latitudes, longitudes, land contiguity, landlocked status, and island status, are from the 2003 CIA Fact Book. The great circle distances are constructed from the latitudes and longitudes of country pairs. The “landlock” variable counts how many countries are landlocked in a pair (0/1/2). Similarly, the “island” variable counts how many countries are islands in a pair (0/1/2). Following Dalgin, Mitra and Trindade (2004), the “remoteness” of a country is defined as the distance of the country to the rest of the world weighted by all the other countries’ GDPs in a given year. The remoteness variable for a country pair

¹³Ideally, asset inequality would fit better into the theoretical model. However, due to data availability, only income inequality measures are used. Disregarding initial income, income inequality (flow measure) and asset inequality (stock measure) are closely related.

¹⁴This method is judged to be reasonable, given that changes in inequality generally follow smooth trajectories.

¹⁵The extrapolated data at least capture much of the between variations of Gini ($Q3$) across countries, which account for nearly 90% of total variations.

(*ij*) in year *t* is simply the product of the two countries' remotenesses:

$$Remoteness_{ijt} = \left(\frac{\sum_{m \neq i} Distance_{mi} GDP_{mt}}{\sum_{m \neq i} GDP_{mt}} \right) \left(\frac{\sum_{m \neq j} Distance_{mj} GDP_{mt}}{\sum_{m \neq j} GDP_{mt}} \right)$$

Theories in gravity models predict that remoteness, as an “index of multilateral resistance” (Anderson and Wincoop, 2003), increases bilateral trade. Applying the same logic to FTAs, two remote countries (e.g. Australia and New Zealand) are more likely to form an FTA because they are unlikely to form FTAs with the rest of the world.¹⁶

Other variables on colonial relationship, military conflict and alliance are also included. Common colony variable equals to unity if both have ever been colonized by the same third country and zero otherwise. Militarized Interstate Dispute Dataset (MID, Ghosn and Palmer, 2003) provides information on conflicts in which one or more states threatened, displayed, or used force against one or more other states between 1816 and 2001. They assign an ordinal number to each conflict, with higher number representing higher level of hostility.¹⁷ The Formal Alliance dataset (Gibler and Sarkees, 2004) seeks to identify each formal alliance between at least two states that fall into the classes of defense pact, neutrality or non-aggression treaty, or entente agreement.

Table 1 shows some descriptive statistics for these variables for year 2000.

VI. Empirical Results

This section shows the regression results from the standard logit and the duration analyses.¹⁸

¹⁶Baier and Bergstrand (2006) show using simulations that ad-hoc remoteness variables leads to biased coefficient estimates and therefore is of little use in the gravity equation. This criticism applies with less force here, however, because this is not a structural model and the remoteness variable is only trying to capture external trade diversion.

¹⁷Because the effect of military conflict may persist for many years after a conflict, the data should not be used on a yearly base. Instead, the average hostility level for country pairs over 1946-2001 is used to measure the conflict intensity after the World War II.

¹⁸We use term “standard” logit to distinguish it from the duration analysis, which can also use logit for discrete-time data. For clear comparison with duration analysis and easy understanding of the coefficients in terms of odds ratio, logit rather than probit is used.

Table 1. Descriptive Statistics, 2000

Variables	Full Sample					FTA=1					FTA=0				
	Obs	Mean	S.D.	Min	Max	Obs	Mean	S.D.	Min	Max	Obs	Mean	S.D.	Min	Max
FTA dummy	5151	0.09	0.29	0	1	480	1	0	1	1	4671	0	0	0	0
Sum of log(GDP)	5151	22.32	2.54	14.16	31.39	480	23.42	2.46	14.16	29.71	4671	22.21	2.52	14.74	31.39
Diff. in log(GDP)	5151	2.06	1.53	0.00	9.28	480	1.47	1.09	0.00	4.92	4671	2.12	1.56	0.00	9.28
log(distance)	5151	7.74	0.77	3.89	8.94	480	6.52	0.79	3.89	8.47	4671	7.86	0.64	4.07	8.94
Land Adjacency	5151	0.02	0.16	0	1	480	0.13	0.34	0	1	4671	0.01	0.12	0	1
Landlock	5151	0.47	0.60	0	2	480	0.40	0.60	0	2	4671	0.48	0.60	0	2
Islands	5151	0.18	0.40	0	2	480	0.07	0.26	0	2	4671	0.19	0.41	0	2
Diff. in Polity	5151	5.82	5.02	0	17	480	3.34	4.82	0	16	4671	6.08	4.97	0	17
Diff. in Child Labor	5151	14.69	14.85	0	51.13	480	2.51	5.25	0	39.13	4671	15.94	14.96	0	51.13
Diff. in CO2	5151	1.88	1.43	0	7.37	480	0.71	0.61	0	4.21	4671	2.00	1.44	0	7.37
Common Colony	5151	0.10	0.30	0	1	480	0.11	0.31	0	1	4671	0.10	0.30	0	1
Hostility	5151	0.01	0.09	0	2.43	480	0.02	0.11	0	1.18	4671	0.01	0.09	0	2.43
Alliance	5151	0.08	0.27	0	1	480	0.42	0.49	0	1	4671	0.05	0.21	0	1
Remoteness	5151	4.13	0.05	4.01	4.27	480	4.07	0.06	4.01	4.27	4671	4.13	0.05	4.02	4.27
Diff. in log(k)	5151	1.58	1.22	0.00	5.41	480	0.78	0.65	0.00	3.20	4671	1.67	1.24	0.00	5.41

Notes:

1. The data used in this table are associated to the regressions in columns (2) and (3) of Table 2;
2. "Diff." means absolute difference.

Table 2. Standard logit regression results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Y=2000	Y=2000	Y=2000	Pooled	Pooled	Panel	Panel
Sum of log(GDP)	0.263*** (0.029)	0.189*** (0.037)	0.177*** (0.037)	-0.037 (0.039)	-0.037 (0.039)	0.153*** (0.021)	0.153*** (0.021)
Diff. in log(GDP)	-0.216*** (0.051)	-0.204*** (0.060)	-0.207*** (0.060)	-0.075 (0.061)	-0.078 (0.061)	-0.346*** (0.036)	-0.346*** (0.036)
log(distance)	-2.025*** (0.10)	-1.728*** (0.127)	-1.703*** (0.125)	-1.611*** (0.110)	-1.602*** (0.110)	-3.110*** (0.072)	-3.093*** (0.071)
Diff. in log(k)	-0.416*** (0.078)	0.965*** (0.158)	0.969*** (0.158)	0.103 (0.152)	0.102 (0.150)	0.196*** (0.065)	0.199*** (0.065)
GINI _i *[log(k _i)-log(k _j)]		-0.018*** (0.006)		-0.001 (0.005)		-0.006** (0.003)	
GINI _j *[log(k _i)-log(k _j)]		0.019*** (0.006)		-0.002 (0.005)		-0.000 (0.003)	
Q3 _i *[log(k _i)-log(k _j)]			0.073*** (0.027)		-0.019 (0.021)		-0.013 (0.011)
Q3 _j *[log(k _i)-log(k _j)]			-0.073*** (0.028)		0.009 (0.021)		-0.004 (0.011)
Land Adjacency	-0.73*** (0.27)	-0.087 (0.333)	-0.049 (0.333)	-0.466 (0.286)	-0.454 (0.285)	-0.976*** (0.170)	-0.966*** (0.170)
Landlock	-0.315*** (0.115)	-0.304** (0.140)	-0.310** (0.140)	-0.060 (0.133)	-0.071 (0.133)	-0.381*** (0.084)	-0.390*** (0.084)
Islands	-0.67*** (0.24)	-0.708** (0.278)	-0.705** (0.277)	0.067 (0.261)	0.063 (0.262)	-0.206 (0.130)	-0.206 (0.131)
Diff. in Polity		-0.087*** (0.016)	-0.089*** (0.016)	-0.042*** (0.013)	-0.041*** (0.013)	-0.058*** (0.006)	-0.059*** (0.006)
Diff. in Child Labor		-0.074*** (0.014)	-0.073*** (0.014)	0.011 (0.009)	0.012 (0.009)	-0.008 (0.005)	-0.007 (0.005)
Diff. in CO2		-0.941*** (0.134)	-0.948*** (0.133)	-0.632*** (0.096)	-0.631*** (0.097)	-0.787*** (0.059)	-0.784*** (0.059)
Common Colony		-0.028 (0.273)	-0.029 (0.273)	0.185 (0.250)	0.190 (0.250)	0.680*** (0.129)	0.686*** (0.129)
Hostility		-2.566*** (0.803)	-2.573*** (0.808)	-1.524*** (0.414)	-1.524*** (0.410)	-3.200*** (0.401)	-3.185*** (0.400)
Alliance		2.589*** (0.22)	2.608*** (0.22)	1.995*** (0.165)	2.014*** (0.165)	2.213*** (0.093)	2.230*** (0.093)
Remoteness	-9.16*** (1.348)	-14.85*** (2.104)	-15.40*** (2.098)	-17.53*** (1.662)	-17.77*** (1.675)	-15.80*** (0.941)	-16.03*** (0.939)
Year Dummies				Yes	Yes	Yes	Yes
Country Pair RE						Yes	Yes
Observations	5886	5151	5151	169787	169787	169787	169787
Log Likelihood	-939	-671	-672	-12054	-12031	-5700	-5700
Pseudo R2	0.46	0.58	0.58	0.56	0.56		
rho						0.60	0.60
% predict (1)	45	59	59	46	46	36	36
% predict (0)	98	98	98	99	99	99	99
% predict (all)	93	94	94	97	97	97	97

Notes:

1. Standard errors in parentheses (clustered by countries pairs in regressions (4) and (5));

2. *significant at 10%; **significant at 5%; ***significant at 1%;

3. “% predict(0)” is the percentage of country pairs correctly predicted with no FTAs; “% predict(1)” is the percentage of pairs correctly predicted with FTAs; and “% predict(all)” is the percentage of pairs correctly predicted with or without FTAs.

A. Standard logit results

Table 2 shows the standard logit results. If we believe that FTA decisions are mainly based on cross sectional variations of the covariates, we can use cross sectional data and apply standard probit or logit, as in Baier and Bergstrand (2004). The first three columns use cross sectional data for year 2000 only (the most recent year with capital per worker data available). Column (1) uses the baseline specification, including some economic and geographic variables often used in the literature. In column (2), we add the key interaction terms between *Gini* ($Q3$) and the difference in the logarithms of capital per worker and some other political covariates. As reflected by the sign flips of the coefficients of the key interaction terms in column (2), the median voter model is strongly supported. The Pseudo R^2 , measured as one minus the ratio of the log-likelihood value for the estimated model to that of the intercept only model, increases from 0.46 to 0.58. In terms of prediction power, the percentages of correctly predicted negative outcomes are both as high as 98%, but the percentage of correctly predicted positive outcomes (FTA=1) increases from 45% in column (1) to 59% in column (2).¹⁹ Overall the increase in goodness-of-fit by adding the political economy variables is fairly large. This is different from the finding by Baier and Bergstrand (2004), who claim that the political economy variables are not important in FTA formation. It is also interesting to note that the coefficient of the absolute difference in $\log(k)$ is negative and significant in column (1),²⁰ but turns positive and significant in column (2). The prediction by the Heckscher-Ohlin model is supported only after the other relevant political economy covariates are controlled for. This again implies the importance of controlling for other relevant political economy variables. In addition, the variety effect is also supported as reflected by the coefficients on the sum and absolute difference of $\log(\text{GDP})$; and most of the other covariates have the expected signs.

¹⁹A positive outcome is correctly predicted if its predicted probability is larger than 0.5; and a zero outcome is correctly predicted if its predicted probability is less than 0.5. Although the number of FTAs has been increasing rapidly in recent years, FTA is still a rare event as only about 9% of country pairs in 2000 had positive outcomes (FTA=1). For a rare event, it is always much easier to predict zero outcomes (FTA=0) than positive outcomes. A constant only logit regression would result in a predicted probability of FTA of 0.09 for every observation, hence the percentage of correctly predict zero outcomes would be $1-9\%=91\%$. This is why we do not rely on the percentage of pairs correctly predicted with zero outcomes [% predict(0) in Table 2] or the overall prediction power [% predict(all) in Table 2]. The low prediction power for positive outcomes is not surprising, given that FTA is a rare event.

²⁰This is consistent to the unconditional means of the absolute difference in $\log(k)$ in Table 1 for the sub-sample with FTA=1 and the sub-sample with FTA=0 (i.e. $0.78 < 1.67$).

Although the cross sectional analyses already provides support for the median voter model, these results are not perfect as they do not use all the information in time-varying variables and are often contaminated by omitted variable bias. It is naturally the next step to extend the cross sectional analyses to panel analyses. Columns (4)-(7) in Table 2 show the standard logit results using pooled data with year dummies and panel data with both year dummies and country pair random effects. Most of the coefficients of the interaction terms, however, become insignificant. In terms of the prediction power for the positive outcomes (FTA=1), the random effect regression is worse than the pooled data regression (36% vs. 46%), and both are worse than the cross section regression (59%). In addition, the country pair random effects regressions are highly sensitive to the number of points used for quadrature.²¹ This suggests that these results are not reliable. More importantly, with the violation of the conditional independence assumption, the results from random effects logit are biased and misleading. To address this problem, we turn to duration analysis.

B. Duration Analysis

Table 3 shows the results from the duration analysis (cloglog). The duration data include all the years from 1960 to 2000, but were amended by dropping all but the first positive outcomes for each country pair. The duration dependence is captured by a sixth-order polynomial of the time counter, which provides the best fit to the data.²² The first two columns show the results using a baseline specification. Even in the baseline regressions, the median voter model, including the variety effect, is strongly supported. Similar to the standard logit results, the coefficient of the absolute difference in $\log(k)$ is still negative and significant without controlling for other political economy variables. Columns (3) and (4) use the full specification with other political economy and geographic variables. Now the coefficient on the absolute difference in $\log(k)$ is positive and significant, as predicted by the Heckscher-Ohlin model. Although the variety effect is still supported, the magnitudes of the coefficients on the sum and absolute difference of $\log(\text{GDP})$ are greatly reduced. On the contrary, the magnitudes of the coefficients on the two key interaction terms almost double, implying even stronger support for the median

²¹The random effects in logit can not be integrated out in the likelihood function, so have to be approximated by quadrature.

²²Plotting the estimated duration dependence against time (year) yields a smooth upward sloping curve.

voter model. The last two columns in Table 3 report the results from country pair random effects cloglog regressions. The unobserved heterogeneity accounts for 17% of the overall variations in the error term, as shown by the “*rho*” in the last

Table 3. Duration analysis results (complementary log-log)

	(1)	(2)	(3)	(4)	(5)	(6)
	Gini	Q3	Gini	Q3	Gini	Q3
Sum of log(GDP)	0.228*** (0.019)	0.228*** (0.019)	0.071*** (0.026)	0.067*** (0.026)	0.085*** (0.028)	0.080*** (0.027)
Diff. in log(GDP)	-0.206*** (0.043)	-0.203*** (0.043)	-0.086** (0.042)	-0.082** (0.042)	-0.104** (0.046)	-0.100** (0.046)
log(distance)	-1.613*** (0.075)	-1.607*** (0.075)	-1.381*** (0.082)	-1.363*** (0.082)	-1.500*** (0.106)	-1.476*** (0.105)
Diff. in log(k)	-0.409*** (0.061)	-0.410*** (0.061)	0.366*** (0.096)	0.374*** (0.097)	0.352*** (0.099)	0.365*** (0.099)
GINI _i *[log(k _i)-log(k _j)]	-0.008*** (0.003)		-0.014*** (0.004)		-0.015*** (0.004)	
GINI _j *[log(k _i)-log(k _j)]	0.009*** (0.003)		0.013*** (0.003)		0.014*** (0.004)	
Q3 _i *[log(k _i)-log(k _j)]		0.025** (0.011)		0.041*** (0.015)		0.043*** (0.016)
Q3 _j *[log(k _i)-log(k _j)]		-0.024** (0.011)		-0.043*** (0.015)		-0.045*** (0.016)
Land Adjacency			-0.480** (0.215)	-0.453** (0.215)	-0.436** (0.199)	-0.407** (0.199)
Landlock			-0.147 (0.108)	-0.153 (0.108)	-0.127 (0.098)	-0.135 (0.099)
Islands			-0.544*** (0.193)	-0.547*** (0.193)	-0.525*** (0.199)	-0.527*** (0.198)
Diff. in Polity			-0.061*** (0.011)	-0.063*** (0.011)	-0.066*** (0.010)	-0.068*** (0.010)
Diff. in Child Labor			-0.038*** (0.009)	-0.038*** (0.009)	-0.038*** (0.008)	-0.037*** (0.008)
Diff. in CO2			-0.572*** (0.090)	-0.563*** (0.090)	-0.583*** (0.088)	-0.572*** (0.087)
Common Colony			0.205 (0.191)	0.208 (0.191)	0.279 (0.180)	0.284 (0.180)
Hostility			-0.880** (0.354)	-0.879** (0.355)	-1.061** (0.412)	-1.055** (0.410)
Alliance			0.826*** (0.168)	0.831*** (0.168)	0.899*** (0.139)	0.904*** (0.139)
Remoteness			-7.951*** (1.266)	-8.179*** (1.277)	-8.339*** (1.233)	-8.600*** (1.232)
Duration Dependence	Yes	Yes	Yes	Yes	Yes	Yes
Country Pair RE					Yes	Yes
Observations	196971	196971	163506	163506	163506	163506
Log Likelihood	-2529	-2530	-2210	-2213	-2201	-2211
rho					0.167	0.166

Notes: Standard errors in parentheses (clustered by countries pairs in regressions (1)-(4));

*significant at 10%; **significant at 5%; ***significant at 1%;

All the regressions use a six-order polynomial of the time counter to capture duration dependence.

row of Table 3. The support for the median voter model is slightly stronger than in Columns (3) and (4), but using random effects does not change the results much.²³ It is worth noting that we do not attempt to predict the timing of FTA formation in duration analysis because positive outcomes of FTA variable (“1”s) only account for 0.3% of the observations after the repeated “1”s are dropped. If an event is rare, it is always difficult to predict.

Taking the results in columns (5) and (6) in Table 3 as preferred, we can quantify the magnitude of the coefficients of the interaction terms. Taking derivative of $Prob(FTA)$ with respect to $GINI_i$ yields

$$\frac{\partial Prob(FTA)}{\partial GINI_i} = \beta_5^* [\log(k_i) - \log(k_j)] \tag{7}$$

Based on the mean of the absolute difference in $\log(k)$ (=1.58 in Table 1)²⁴, the above partial derivative equals to 0.023. This says that one unit increase in *Gini* coefficient in the rich (poor) country decreases (increases) the odds of signing an FTA by 2.3%, *ceteris paribus*. By a similar calculation, the results in column (6) of Table 3 imply that one unit increase in *Q3* in the rich (poor) country increases (decreases) the odds ratio by about 6.95%, *ceteris paribus*. The magnitude of the coefficient of *Q3* is much bigger than that of *Gini* because the mean values of *Gini* and *Q3* are very different (40 and 15 respectively). In terms of elasticity, the two magnitudes are very similar (about 0.9 for *Gini* and 1.0 for *Q3*).²⁵ This says that one percent change in *Gini* (*Q3*) can change the odds ratio of an FTA by 0.9 (1.0) percent. These magnitudes are significant economically, compared to the elasticity of geographic distance (1.5) – the most important and robust determinant of FTAs, and the elasticities of the product and ratio of GDPs (0.08 and -0.1 respectively).²⁶ For example, if *Gini* increases by 33% from 45 (e.g. Peru, 1995) to 60 (e.g. Brazil, 1995), the odds ratio of an FTA will, on average, change by $0.9 \times 33\% = 30\%$, *ceteris paribus*.

Most of the coefficients of other covariates have the expected signs and magnitudes.

²³As robustness checks, we also consider the gamma mixture frailty proposed by Meyer (1990) and the Heckman and Singer-type nonparametric representations of frailty (Heckman and Singer, 1984). The results are very similar to those in the last two columns of Table 3.

²⁴The mean is still the same (1.58) if we consider separately the cases when $k_i > k_j$ and $k_i < k_j$.

²⁵The elasticity of *Gini* is calculated as $0.023 / (1/40) = 0.9$; the elasticity of *Q3* is calculated as $0.0695 / (1/15) = 1$.

²⁶The coefficients of distance and the product and ratio of GDPs can be directly taken as the elasticities because they are in logarithms.

Countries close in distance are more likely to form FTAs. Conditioned on distance, adjacent countries are surprisingly less likely to form FTAs, which is the same as the finding by Magee (2004). One possible explanation is that, conditional on geographic distance, adjacent countries are more likely to suffer from conflicts, such as many Middle East and African countries. Landlocked and island countries are less likely to form FTAs. Countries with similar polity scores and labor/environment standards are more likely to form FTAs. Countries ever colonized by the same parent country are more likely to form FTAs, but not statistically significant. Hostility prevents countries from forming FTAs; while formal alliance relationship helps FTA formation. Remoteness surprisingly reduces the likelihood of FTAs, which is different from what theories predict, but not uncommon in gravity model literature.²⁷

We also test for the natural trading partner hypothesis by including bilateral trade variable as a covariate. This hypothesis says that countries that trade more with each other are more likely to sign an FTA.²⁸ The coefficient of the bilateral trade variable is always positive and significant; hence it supports the natural trading partner hypothesis. To alleviate endogeneity problem of bilateral trade, we also tried one year and five years lagged bilateral trade and got similar results. The bilateral trade variable, however, is not included in previous regressions because it is highly correlated with the other covariates such as distance and the sum of GDPs. With or without bilateral trade variable in the duration analysis, the median voter model is always supported.

VII. Robustness Checks and Related Complexities

A. *Democracy vs. autocracy?*

²⁷For example, Frankel, Stein and Wei (1997, p. 143) and Dalgin, Mitra and Trindade (2004), among many others, also find negative coefficient on remoteness from some gravity regressions for bilateral trade flows. Baier and Bergstrand (2004) find positive coefficient on remoteness for FTAs, but their remoteness measure is multiplied by the dummy variable indicating the same continent. We do get positive coefficient when using the same definition as Baier and Bergstrand (2004), but the explanations will be different. The measure of remoteness does not affect the support for the median voter model.

²⁸“Natural” FTAs are often associated with more trade creation and less trade diversion (Krugman, 1995, p. 182). Krishna (2003), however, does not find empirical support for the welfare-improving effects from US data. The estimation strategy of our paper is not based on a structure model and bilateral trade simply enters the regression as a control variable. Due to many differences in the data and methods, the results of Krishna (2003) and ours are not really comparable. But we believe that the endogeneity problem of trade flows can be better addressed in a general equilibrium framework as in Krishna (2003).

If the support for the median voter model is the evidence of electoral competition (or any policies seeking popular support) or governments' inequality concern, it is likely that the support will be stronger in democratic countries than in autocratic countries. This is because democratic governments, to ensure their re-election, are more likely to adopt policies that benefit the majority. In this subsection, we run the regressions on the sub-samples of the data according to the degrees of partner countries' democracies. The data on countries' democracies come from the Polity IV Project (Marshall and Jaggers, 2002). "Polity" score data range from -10 to 10, with higher values representing more democratic regimes. Countries with positive "polity" scores are considered as democratic countries; and countries with zero or negative scores are taken as non-democratic countries (neutral or autocratic).

Table 4 reports the duration analysis results using the sub-samples.²⁹ Columns (1) and (2) use the sub-sample with both countries in a pair being democratic; while columns (3) and (4) use the sub-sample with neither country in a pair being democratic. As reflected by the larger magnitudes of the coefficients on two key interaction terms in the first two columns, the support for the median voter model is indeed stronger compared to the previous results using the full sample. By contrast, in columns (3) and (4), the key interaction terms are always insignificant when both partners are non-democratic. These results are consistent to the fact that democratic governments are more concerned about getting popular support due to electoral competition compared to autocratic governments.

B. Partial observability of FTA formation

For the FTA dummy, we observe "1" when both partners want to sign an FTA, and zero when either country does not want to sign the FTA. When the FTA dummy equals to zero, we only observe the product of two partners' decisions, rather than their individual preferences. This is a problem of "partial observability". With this problem, an increase in one country's willingness of an FTA, due to changes in its own inequality, might not necessarily imply an increase in the probability of signing the agreement if the other country is unwilling to sign the FTA. This problem, however, only makes our decision in favor of the null hypothesis that inequality does not affect the probability of signing an FTA. Given that we have already found strong support for the median voter model, the support

²⁹Because using country pair random effects make little difference in most of the sub-samples, unobserved heterogeneity is ignored in these regressions.

Table 4. Duration analysis with “Polity” sub-samples

	(1)	(2)	(3)	(4)
	Gini_D	Q3_D	Gini_ND	Q3_ND
Sum of log(GDP)	0.158*** (0.031)	0.154*** (0.030)	-0.329*** (0.083)	-0.334*** (0.085)
Diff. in log(GDP)	-0.034 (0.052)	-0.040 (0.051)	0.143 (0.136)	0.138 (0.135)
log(distance)	-1.341*** (0.101)	-1.315*** (0.103)	-1.788*** (0.214)	-1.802*** (0.212)
Diff. in log(k)	0.520*** (0.128)	0.529*** (0.127)	-0.311 (0.247)	-0.294 (0.240)
GINI _i *[log(k _i)-log(k _j)]	-0.019*** (0.006)		0.011 (0.010)	
GINI _j *[log(k _i)-log(k _j)]	0.021*** (0.005)		-0.008 (0.010)	
Q3 _i *[log(k _i)-log(k _j)]		0.066*** (0.025)		-0.028 (0.039)
Q3 _j *[log(k _i)-log(k _j)]		-0.066** (0.026)		0.037 (0.038)
Land Adjacency	-0.503* (0.257)	-0.462* (0.256)	-0.370 (0.494)	-0.387 (0.492)
Landlock	-0.264** (0.132)	-0.277** (0.132)	0.427 (0.296)	0.438 (0.292)
Islands	-0.641*** (0.216)	-0.642*** (0.215)		
Diff. in Polity	-0.247*** (0.047)	-0.245*** (0.046)	0.053 (0.056)	0.055 (0.056)
Diff. in Child Labor	-0.060*** (0.011)	-0.058*** (0.012)	0.013 (0.017)	0.011 (0.017)
Diff. in CO2	-0.853*** (0.126)	-0.847*** (0.126)	-0.072 (0.167)	-0.066 (0.172)
Common Colony	0.543** (0.254)	0.526** (0.254)	-0.046 (0.339)	-0.047 (0.339)
Hostility	-0.605 (0.425)	-0.599 (0.431)	-0.788 (1.293)	-0.869 (1.301)
Alliance	0.600*** (0.178)	0.598*** (0.178)	0.795* (0.429)	0.799* (0.425)
Remoteness	-7.525*** (2.045)	-7.837*** (2.055)	-17.868*** (3.406)	-17.736*** (3.390)
Duration Dependence	Yes	Yes	Yes	Yes
Observations	51467	51467	29972	29972
Log Likelihood	-1228	-1231	-343	-343

Notes:

1. Robust standard errors (clustered by country pairs) in parentheses;

2. *significant at 10%; ** significant at 5%; *** significant at 1%;

3. The regressions use up to six-order polynomial of the time counter to capture duration dependence;

4. Regressions (1) and (2) use the observations with both countries in a pair being democratic (D);

5. Regressions (3) and (4) use the observations with both countries in a pair being “non-democratic (ND)” (i.e. either “neutral” or “autocratic”);

6. “Islands” variable is dropped in some regressions because it perfectly predicts the negative outcomes.

would be even stronger were we able to solve the partial observability problem. Moreover, the possible bias in the coefficients due to the partial observability problem is alleviated by the fact that reciprocity and concessions are common practice in FTA formation. One country can request from or offer to its partner some “benefits” through international transfers or side-payments if one of them is unwilling to sign the agreement. For example, in order to conclude on the NAFTA, Mexico had to sign the side agreements on labor and environment standards proposed by the US.³⁰ In this sense, an increase in the desirability of an FTA to one country alone can increase the probability of signing the agreement through international transfers, even if the other country does not want to sign it initially. Therefore, the previous duration analysis is still valid by testing the effects of each partner’s inequality on joint FTA decisions.

To address the partial observability problem formally, we follow the early literature on partial observability (e.g., Poirier, 1980; Abowd and Farber, 1982) by estimating the probability of signing an FTA as the product of the desirabilities of the FTA to each partner. The log likelihood function can be written as:

$$\ln L = \sum_{FTA = 1} \ln [F(\beta'_R x_R) F(\beta'_P x_P)] + \sum_{FTA = 0} \ln [1 - F(\beta'_R x_R) F(\beta'_P x_P)] \quad (8)$$

where $prob(FTA = 1) = F(z) = e^z / (1 + e^z)$; x_R and x_P are the factors underlying the decisions of the rich and poor partners respectively. A minimum requirement for identification is that x_R should contain a regressor not in x_P and vice versa. Two different sets of variables are needed for rich and poor partners. The estimation of the likelihood function (8) will produce two sets of coefficients, β_R and β_P , for rich and poor partners respectively.

Different from previous analyses, we do not have to use the absolute values of the differences in GDP, polity scores, child labor and CO2 per capita (unidirectional differences). Rather, the directional differences in these variables can be used. For example, we include “dGDP(RP)” [=log(GDP(R))-log(GDP(P))] and “dGDP(PR)” [=log(GDP(P))-log(GDP(R))] as separate variables for the rich and poor partners respectively. We expect that the coefficients of dGDP(RP) and dGDP(PR) are both negative. In other words, a rich (poor) partner is less likely to sign an FTA if its

³⁰Theoretical discussions on the international transfers and side-payments under customs unions, free trade areas and the WTO can be found in Kemp and Wan (1976), Aghion, Antras and Helpman (2004) and Bagwell and Staiger (2004), among others.

poor (rich) partner's GDP is small. The differences of all the other variables are constructed in the same way. "Remoteness" variable is also separated as "remoteness(R)" and "remoteness(P)" for rich and poor countries.³¹ According to the theory, we assume that inequality of a country does not affect the desirability of FTA of the other country in a pair. Hence we include Gini(R) or Q3(R) in x_R , and Gini(P) or Q3(P) in x_P .

The results from logit regressions with partial observability are shown in Table 5. The first two columns are from the regressions using Gini coefficient, while the last two columns use Q3 measure. The signs of the coefficients of Gini and Q3 still strongly support the median voter model. As expected, the magnitudes on average are even bigger than what we get from duration analyses in Table 3. Many other covariates are less significant, compared to previous duration analyses. It is not surprising because the log likelihood functions of partial observability model are not globally concave and are much more difficult to estimate.³² To save space, we do not discuss here in details the results for other covariates.³³

C. Left and right censorings of duration spells

In order to use duration analysis, researchers must know when the spell before an FTA starts and which countries are in the risk set. It is natural to think that all of the countries started to face the "risk" of FTAs since the formation of the first FTA, that is, European Economic Community (EEC, Treaty of Rome) in 1958. Even if we believe that the starting point is 1958 or even earlier, the choice of year 1958 as the first year does not cause much problem. Equally left censoring is irrelevant for the estimation of coefficients if no FTA formed before 1958, because the discrete hazard is a conditional probability and all the probabilities of the earlier years are conditioned out of the likelihood function (Jenkins, 1995, and Sueyoshi, 1995). Because many covariates in the regressions have data available only since 1960, year 1960 rather than 1958 is taken as the first year in our analysis, at the cost of losing the country pairs between the original signatories of the Treaty of Rome.³⁴

³¹See Data section for the definition of "remoteness" for a single country.

³²We use "BHHH" method with "difficult" option.

³³We do not consider unobserved heterogeneity. LIMDEP8.0 develops a routine for panel data, but the random parameters model is difficult to implement, even for simulated data (Greene, 2002, E17.5). We are not aware of any application of this panel data procedure in the literature.

³⁴They are Belgium, France, Germany, Italy, Luxembourg and Netherlands.

Table 5. Partial observability of FTA formation

	(1)		(2)	
	Rich Partner	Poor Partner	Rich Partner	Poor Partner
Sum of log(GDP)	-0.151*	0.450**	-0.104	0.256***
	(0.092)	(0.228)	(0.096)	(0.081)
Diff. in log(GDP)	-0.042	-0.024	-0.192**	-0.230*
	(0.099)	(0.137)	(0.095)	(0.122)
log(distance)	-1.629***	-0.213	-1.801***	-0.111
	(0.340)	(0.853)	(0.229)	(0.313)
Diff. in log(k)	-0.065	-0.332	-1.228***	1.440***
	(0.199)	(0.327)	(0.407)	(0.334)
GINI _R *[log(k _R)-log(k _P)]	-0.012***			
	(0.004)			
GINI _P *[log(k _R)-log(k _P)]		0.035***		
		(0.009)		
Q3 _R *[log(k _R)-log(k _P)]			0.090***	
			(0.027)	
Q3 _P *[log(k _R)-log(k _P)]				-0.087***
				(0.031)
Land Adjacency	-0.339	-0.321	-1.188**	0.812
	(0.468)	(0.862)	(0.521)	(0.501)
LandlockR	-0.678**	0.454	-0.828**	0.424
	(0.334)	(0.490)	(0.326)	(0.389)
LandlockP	0.377	-0.598	0.958*	-0.941**
	(0.277)	(0.482)	(0.507)	(0.464)
IslandR	0.740	-1.225	1.770***	-1.883***
	(0.505)	(1.084)	(0.626)	(0.507)
IslandP	0.259	-3.351***	-1.518	-0.102
	(0.506)	(0.890)	(1.021)	(1.327)
Diff. in Polity	-0.019	0.036	0.044	0.090***
	(0.020)	(0.024)	(0.027)	(0.025)
Diff. in Child Labor	-0.009	-0.059**	0.062***	0.067***
	(0.015)	(0.028)	(0.019)	(0.022)
Diff. in CO2	-0.159	0.437***	-1.347***	-0.835**
	(0.148)	(0.175)	(0.354)	(0.444)
Common Colony	0.780**	-0.723	1.625**	-1.210**
	(0.345)	(0.933)	(0.715)	(0.585)
Hostility	-1.362**	-0.298	-1.660	-0.364
	(0.610)	(1.213)	(1.360)	(1.321)
Alliance	0.247	2.953*	1.653***	-0.339
	(0.335)	(1.653)	(0.367)	(0.586)
RemoteR	-17.271	-10.242	14.774*	-40.258***
	(13.438)	(23.886)	(7.918)	(8.961)
RemoteP	-4.477	-17.308	-32.050***	29.743***
	(14.552)	(30.847)	(6.373)	(7.192)
Observations	163506	163506	163506	163506
Log-likelihood	-2047.559	-2047.559	-2091.083	-2091.083

Notes:

1. Robust standard errors (clustered by country pairs) in parentheses;
2. * significant at 10%; ** significant at 5%; *** significant at 1%;
3. The regressions use five-order polynomial of the time counter to capture duration dependence,
4. The variables in difference (i.e., "Diff.") are directional, rather than absolute values (see the text for more details). To save space, the directional differences for both rich and poor partners share the same variable name in the table.

The subsequent accessions of other European Community members after 1960, however, are still kept in the analysis. Another problem is associated to those countries that enter the sample after 1960, such as the newly-independent countries from former colonies, the USSR or Yugoslavia. Most of the new states from formal USSR and Yugoslavia, however, are not covered in the sample due to missing covariate data.

Spells are “right-censored” if some country pairs have never signed FTAs until the end of the sample period. This is not a problem because these country pairs simply contribute a string of zeros, with no final one, to the likelihood.

D. Multi-country FTAs

Many FTAs have more than two member countries. When a country decides to join an FTA, it considers all the members together. We could consider multi-country FTAs by including the variables measuring the bloc sizes. This method, however, suffers from endogeneity problem because the bloc size measure uses the information from the dependent variable (FTA membership). This problem is especially serious in our duration analysis with few positive outcomes (0.3%). Please note that, even for these multi-country FTAs, many decisions were actually made on a bilateral basis. For example, the admission of new members into the EU requires unanimous approval from all the current EU members, which is an important aspect of equal right of the members. Moreover, a great deal of effort was expended to take into account the fact that different countries might join the same agreement at different years. Hence treating multi-country FTAs as bilateral ones, we may still obtain unbiased point estimates of the coefficients.³⁵

VIII. Conclusions

Using duration analyses with a large FTA panel data, this paper finds that the predictions from the median voter model are strongly supported. The support for the median voter model implies that governments’ concern for general interests on average dominates the concern for special interests in FTA formation. Despite its simplicity (direct democracy, uni-dimensional policy space and single-peaked individual preference), median voter model has received compelling empirical

³⁵The correlations between the pairs within multi-country FTAs can contaminate the estimates of standard errors. Clustered standard errors can not be obtained because one country might have several FTAs and clusters can not be mutual-exclusively defined.

support in the literature.³⁶ But the support for the median voter model should not be taken too literally. It instead should be taken as a short-hand way to capture electoral competition, general policy orientation in seeking popular support or governments' inequality concern. Moreover, the support for the median voter model does not mean that special interest groups or lobbies are absent. The exclusion of some sensitive sectors from trade agreements, the prolonged phase out period of tariffs and the exploitation of rules of origin are all probably due to the pressures from interest groups.

The findings in this paper are important for the understanding of the conflicting distributional forces in an unequal world and the underlying political economy rationales for FTA formation. Understanding the interrelationship between FTA formation and these political economy factors can help us design more effective policies to achieve many other economic goals, such as growth, inequality and poverty reduction. In addition, this paper also offers important policy implications for both the static welfare analysis and the dynamic time path issues of FTAs. The support for the median voter model implies less trade diversion from FTAs, hence sends an encouraging message to the current movement of regionalism.

Acknowledgements

I am extremely grateful to Devashish Mitra, Dan Black, Kishore Gawande, Mary Lovely, David Richardson and Shang-Jin Wei for their encouragement and guidance. I also thank the anonymous referees and participants in the seminars at Syracuse University and the International Monetary Fund for helpful comments and suggestions. All the remaining errors are mine.

Received 21 August 2007, Accepted 1 February 2008

References

Abowd, J., Farber, H. (1982) Job queues and the union status of workers. *Industrial and Labor Relations Review* 35(3), 354-367.

³⁶For example, Poole and Daniels (1985) find that 80-90% of all the recorded voters in the US Congress can be explained with one dimension space. In public finance, Congleton and Shughart (1990) find that the median voter model provides a better explanation for large scale public programs than lobbying models. See Persson and Tabellini (2002, chapter 3 and 6) for more discussions. Dutt and Mitra (2002) discuss the evidence in favor of the median voter model in the trade literature from both cross country and micro-level survey analyses.

- Alesina, A., Rodrik, D. (1994) Distributive politics and economic growth. *Quarterly Journal of Economics* 109(2), 465-490.
- Anderson, J., van Wincoop, E. (2003) Gravity with gravitas: a solution of the Land Adjacency puzzle. *American Economic Review* 93(1), 170-192
- Bagwell, K., Staiger, R.W. (2004) Backward stealing and forward manipulation in the WTO. NBER Working Paper No. 10420.
- Baier, S.L., Bergstrand, J.H. (2004) Economic determinants of free trade agreements. *Journal of International Economics* 64(1), 29-63.
- Baier, S.L., Bergstrand, J.H. (2006) Bonus Vetus OLS: A Simple Approach for Addressing the "Land Adjacency Puzzle" and Other Gravity-Equation Issues. Draft.
- Baier, S.L., Dwyer, G.P., Tamura, R. (2006) How Important Are Capital and Total Factor Productivity for Economic Growth? *Economic Inquiry* 44(1), 23-49.
- Baldwin, R. (1996) A domino theory of regionalism. NBER Working Paper No. 4465.
- Beck, N., Katz, J.N., Tucker, R. (1998) Taking time seriously: time-series-cross-section analysis with a binary dependent variable. *American Journal of Political Science* 42(4), 1260-1288.
- Bhagwati, J.N. (1993) "Regionalism and Multilateralism: An Overview," in Jaime De Melo and Arvind Panagariya (Ed.) *New Dimensions in Regional Integration*. Cambridge University Press.
- Congleton, R.D., Shughart, W.F. (1990) The Growth of Social Security: Electoral Push or Political Pull? *Economic Inquiry* 28:109-32.
- Dalgin, M., Mitra, D., Trindade, V. (2004) Inequality, nonhomothetic preferences, and trade: a gravity approach. NBER Working Paper No. 10800.
- Dollar, D., Kraay, A. (2002) Growth is good for the poor. *Journal of Economic Growth* 7(3): 195-225.
- Dutt, P., Mitra, D. (2002) Endogenous trade policy through majority voting: an empirical investigation. *Journal of International Economics* 58(1): 107-133.
- Frankel, J.A., Stein, E., Wei, S.-J. (1997) *Regional Trading Blocs*. Institute for International Economics, Washington, DC.
- Gawande, K., Krishna, P. (2003) Political economy of U.S. trade policy: empirical approaches, in James Harrigan and E. Kwan Choi (Ed.), *Handbook of International Economics*. Basil Blackwell. pp.213-250.
- Ghosn, F., Palmer, G. (2003) Codebook for the Militarized Interstate Dispute Data, Version 3.0. Online: <http://cow2.la.psu.edu>
- Gibler, D.M., Sarkees, M. (2004) Measuring Alliances: the Correlates of War Formal Interstate Alliance Dataset, 1816-2000. *Journal of Peace Research*, 41(2), 211-222.
- Greene, W.H. (2002) *LIMDEP Version 8.0 User's Manual*. Plainview, NY: Econometric Software.
- Grossman, G.M., Helpman, E. (1995) The politics of free-trade agreements. *American Economic Review* 85(4), 667-690.
- Heckman, J.J., Singer, B. (1984) A Method for minimizing the impact of distributional assumptions in econometric models for duration data, *Econometrica* 52(2), 271-320.

- Jenkins, S.P. (1995) Easy ways to estimate discrete time duration models. *Oxford Bulletin of Economics and Statistics* 57(1), 129-138.
- Jenkins S. (2004) Survival Analysis. <http://www.iser.essex.ac.uk/teaching/degree/stephenj/ec968/pdfs/ec968lnotesv5.pdf>
- Kemp, M.C., Wan Jr, H.Y. (1976) An elementary proposition concerning the formation of customs unions. *Journal of International Economics* 6(1), 95-97.
- Krishna, P. (1998) Regionalism and Multilateralism: A Political Economy Approach. *Quarterly Journal of Economics* 113 (February), 227-251.
- Krishna, P. (2003) Are regional trading partners 'natural'? *Journal of Political Economy* 111(1), 202-226.
- Krugman, P. (1995) The move toward free trade zones, in: P. King (Eds.), *International Economics and International Economic Policy: A Reader*, McGraw-Hill: New York, pp. 163-182.
- Lancaster, T. (1990) *The Econometric Analysis of Transition Data*. Cambridge University Press, Cambridge.
- Levy, P. (1997) A political-economic analysis of free-trade agreements. *American Economic Review* 87(4), 506-519.
- Limao, N., Panagariya, A. (2006) Inequality and Endogenous Trade Policy Outcomes. *Journal of International Economics*, forthcoming.
- Liu, X., (2007) A political-economic analysis of free-trade agreements: Comment. AER Internet Comment at <http://www.aeaweb.org/aer/contents/icomments.html>
- Magee, C. (2003) Endogenous preferential trade agreements: an empirical analysis. *Contributions to Economic Analysis & Policy* 2(1), Article 15.
- Maggi, G., Rodriguez-Clare, A. (2005) A Political-Economy Theory of Trade Agreements. CEPR Discussion Paper No. 5321.
- Marshall, M.G., Jagers, K. (2002) Polity IV Project: Political Regime Characteristics and Transitions, 1800-2002.
- Mayer, W. (1984) Endogenous tariff formation. *American Economic Review* 74(5), 970-985.
- Meyer, B.D. (1990) Unemployment insurance and unemployment spells. *Econometrica* 58(4), 757-782.
- Panagariya, A. (1994) East Asian and the New Regionalism. *The World Economy* 17(6), 817-839.
- Persson, T., Tabellini, G. (2002) *Political Economics: Explaining Economic Policy*. The MIT Press. Cambridge, Massachusetts.
- Poirier, D.J. (1980) Partial observability in bivariate probit models. *Journal of Econometrics* 12, 209-217.
- Poole, K.T., Daniels, R.S. (1985) Ideology, party and voting in the US Congress, 1959-1980. *American Political Science Review* 79, 373-99.
- Prentice, R., Gloeckler, L. (1978) Regression analysis of grouped survival data with application to breast cancer data. *Biometrics* 34(1), 57-67.
- Rodrik, D. (1986) Tariffs, subsidies, and welfare with endogenous policy. *Journal of International Economics* 21(3/4), 285-299.

Schiff, M., Winters, A. (2003) *Regional Integration and Development*. The World Bank and Oxford University Press.

Sueyoshi, G.T. (1995) A class of binary response models for grouped duration data. *Journal of Applied Econometrics* 10(4), 411-431.

Viner, J. (1950) *The Customs Union Issue*. Carnegie Endowment for International Peace, New York.

Appendix 1: List of FTAs covered in this paper (1960-2000)

Agreement	Date into force	Agreement	Date into force
EEC, EC, EU	1-Jan-58	EC--Bulgaria	31-Dec-93
EFTA	3-May-60	Chile--Colombia	1-Jan-94
CACM	12-Oct-61	Chile--Venezuela	1-Jan-94
CACM--Costa Rica	1-Jan-62	NAFTA	1-Jan-94
Chana--Burkina Faso	1-Jan-62	Chile--Ecuador	1-Jan-95
ACM	1-Jan-64	G3 (Group of Three)	1-Jan-95
EC--Switzerland/Liechtenstein	1-Jan-73	Mexico--Bolivia	1-Jan-95
EC--Norway	1-Jul-73	Mexico--Costa Rica	1-Jan-95
CARICOM	1-Aug-73	EC--Turkey	1-Jan-96
MRU	1-Jan-74	MERCOSUR--Chile	1-Oct-96
ECOWAS	1-Jan-75	Canada--Israel	1-Jan-97
CEPGL	1-Jan-76	MERCOSUR--Bolivia	2-Mar-97
EC--Algeria	1-Jul-76	Turkey--Israel	1-May-97
PATCRA	1-Feb-77	Canada--Chile	5-Jul-97
EC--Egypt	1-Jul-77	Czech Republic--Israel	1-Dec-97
CER	1-Jan-83	GAFTA	1-Jan-98
ECCAS	1-Jan-85	Mexico--Nicaragua	1-Jan-98
United States--Israel	19-Aug-85	Hungary--Israel	1-Feb-98
EFTA -- Finland	1-Jan-86	EC--Tunisia	1-Mar-98
CAN (CU after 1995)	25-May-88	Poland--Israel	1-Mar-98
AMU	1-Jan-89	Hungary--Turkey	1-Apr-98
United States-- Canada	1-Jan-89	Czech Republic--Turkey	1-Sep-98
MERCOSUR	29-Nov-91	Bulgaria--Turkey	1-Jan-99
EC--Czech Republic	1-Mar-92	CACM--Chile	1-Jan-99
EC--Hungary	1-Mar-92	Chile--Mexico	1-Aug-99
EC--Poland	1-Mar-92	EFTA--Morocco	1-Dec-99
EFTA--Turkey	1-Apr-92	EC--South Africa	1-Jan-00
EFTA--Czech Republic	1-Jul-92	EC--Morocco	1-Mar-00
EFTA--Israel	1-Jan-93	Poland--Turkey	1-May-00
CEFTA	1-Mar-93	EC--Israel	1-Jun-00
EFTA--Bulgaria	1-Jul-93	EC--Mexico	1-Jul-00
EFTA--Hungary	1-Oct-93	Mexico--Israel	1-Jul-00
EFTA--Poland	15-Nov-93		

Notes: All the subsequent accessions of other members to EEC, EC and EU by 2000 are covered in our analysis, although they are not shown individually in the table. The service agreements are not treated as separate agreements in this paper because they were usually signed together or after the corresponding FTAs in goods.

Data Sources:

1. WTO: http://www.wto.org/english/tratop_e/region_e/region_e.htm (accessed in 2004)
2. WTO Archive, *WTO*, Geneva, Switzerland.
3. Frankel, Stein and Wei (1997) and Schiff and Winters (2003)
4. Foreign Trade Information System: <http://www.sice.oas.org/TRADEE.ASP>