CESifo Working Papers
www.cesifo.org/wp

What Determines BITs?

Jeffrey H. Bergstrand
Peter Egger

CESIFO WORKING PAPER NO. 3514
CATEGORY 8: TRADE POLICY
JULY 2011

PRESENTED AT CESIFO AREA CONFERENCE ON GLOBAL ECONOMY, FEBRUARY 2011

CESifo *Working Paper No. 3514*

What Determines BITs?

Abstract

Bilateral investment treaties (BITs) have proliferated over the past 50 years such that the number of pairs of countries with BITs is roughly as large as the number of country-pairs that belong to bilateral or regional preferential trade agreements (PTAs). The purpose of this study is to provide the first systematic empirical analysis of the economic determinants of BITs and of the likelihood of BITs between pairs of countries using a qualitative choice model, and in a manner consistent with explaining PTAs. We develop the econometric specification for explaining the two based upon a general equilibrium model of world trade and foreign direct investment with three factors, two products, and explicit natural as well as policy trade and investment costs among multiple countries in the presence of national and multinational firms. The empirical model for BITs and PTAs is bivariate in nature and supports a set of hypotheses drawn from the general equilibrium model. Using the preferred empirical model, we correctly predict approximately 85 (75) percent of all BITs (PTAs) correctly, relative to an unconditional probability of only 11 (16) percent.

JEL-Code: F100, F200.

Keywords: bilateral investment treaties, foreign direct investment, multinational firms, free trade agreements, international trade.

Jeffrey H. Bergstrand
Department of Finance
Mendoza College of Business
University of Notre Dame
USA – Notre Dame, IN 46556
bergstrand.1@nd.edu

Peter Egger
ETH Zurich
WEH E6
Weinbergstrasse 35
Switzerland – 8092 Zurich
egger@kof.ethz.ch

July 6, 2011

1 Introduction

> *"[T]he primary problem for researchers wishing to assess the impacts of policies to promote FDI is that policy adoption is endogenously determined." (Aisbett, 2009, p. 396)*
>
> *"The literature on BITs is limited, making it difficult to truly understand the determinants for signing." (Tobin and Rose-Ackerman, 2005, p. 15)*

One of the most notable economic events since World War II is the proliferation of preferential trade agreements (PTAs), including predominantly free trade agreements (FTAs) and some customs unions (CUs). However, the proliferation of bilateral investment treaties (BITs) has been *at least as significant* as that of PTAs. For instance, in 2010 the U.S. government had 40 BITs in force while it had only 17 PTAs in force. Figure 1 presents the numbers of BITs in existence in the world in every year from 1980-2007. Moreover, Table 1 indicates the numbers of country-pairs with BITs and with PTAs (including some with both) in year 2000 for 161 countries. Table 1 confirms for this sample 923 country-pairs with BITs but not PTAs, 1,478 pairs with PTAs but not BITs, and 556 pairs with both.[1]

Yet in contrast to the vast international trade literatures on the theoretical net benefits and costs of FTAs and CUs and on the empirically-estimated effects of FTAs and CUs on trade flows, the literature on BITs is not only considerably smaller but dominated by legal scholars. Moreover, the number of studies of the effects of BITs on foreign direct investment (FDI) flows is limited to a handful relative to the trade literature for PTAs (including notably gravity-equation studies). Furthermore, there is *no study* trying to systematically explain empirically the economic determinants of BITs – much less one motivated by a general equilibrium model.[2] This paper addresses this shortcoming.

We offer three potential contributions. First, building upon the theoretical and empirical literatures on economic determinants of PTAs (cf., Baier and Bergstrand, 2004), we use comparative statics from a more general version of the general equilibrium (GE) model of world trade and FDI with national and multinational firms with three factors of productions, two goods, three countries, and natural and policy-based investment and trade costs in Bergstrand and Egger (2007) to illustrate the likely

[1] We note that, since 1990, many of PTAs have introduced substantive investment provisions; such agreements are more accurately "preferential trade *and investment* agreements" (or PTIAs). We will discuss the sensitivity of our main empirical results to this issue later in the robustness analysis.

[2] The only empirical study close to ours is Swenson (2005), who provides an econometric analysis explains the number of BITs across countries in terms of per capita incomes, expropriation risk, and pre-existing levels of FDI stocks. However, the study is not motivated by a formal economic model and does not address the economic determinants emphasized in our study.

net utility gains (losses) of BITs and PTAs between country-pairs as functions of numerous economic variables. We show how such potential utility gains from BITs and PTAs are functions of economic (GDP) size, economic similarity, and (natural and policy-based) investment and trade costs – in a world with *identical* relative factor endowments. Second, we relax the assumption of identical relative factor endowments and show the relationships between net utility gains from BITs and PTAs and relative skilled labor, unskilled labor, and physical capital endowments, using a traditional Edgeworth box framework, since the vast bulk of BITS are between developed and developing countries. Third, guided by the comparative statics from the GE model, we specify and estimate a bivariate probit model of the likelihood of BITs *and* PTAs between country pairs – a bivariate version of the simple univariate probit model in Baier and Bergstrand (2004) for predicting PTAs. We find that the model has a pseudo-R^2 of 21 percent and predicts correctly approximately *66 percent* of all cases (true positives and true negatives). Specifically, the model correctly predicts 85 percent of the 1,479 BITs in year 2000 among 12,880 pairs of 161 countries and 62 percent of the remaining 11,401 pairs with no BITS *as well as* 75 percent of the 2,034 PTAs in year 2000 among the 12,880 country-pairs and 67 percent of the remaining 10,846 pairs with no PTAs; we also discuss predicted probabilities at alternative cutoffs. Moreover, we show that the results hold up to an extensive sensitivity analysis, including simultaneity between BITs and PTAs. These results suggest that the determinants of BITs are – like those for PTAs – quantifiable, and this provides potentially useful information for addressing the issue of *endogeneity* in analysis of the effects of BITs and PTAs on FDI and trade flows, as mentioned above.

2 Motivating Economic Determinants of BITs

The international trade literature has long studied the net benefits and losses from preferential trade agreements (PTAs), especially CUs and FTAs, beginning with Viner (1950) and Lipsey (1960).[3] The study of PTAs has followed fundamentally two paths, one normative and one positive. The normative path is whether or not PTAs are welfare-improving; a full discussion of this literature is beyond the scope of this paper, but see Baldwin (2007) and Freund and Ornelas (2009) for excellent recent surveys. The second path, which is "positive," examines what factors explain and predict *which* pairs of countries have formed PTAs. Building on the work of Krugman (1991a, 1991b) and Frankel (1997), Baier and Bergstrand (2004) introduced asymmetric absolute and relative factor endowments into a Krugman-Heckscher-Ohlin-type increasing-returns/monopolistic-competition model with national

[3]There is also a more recent large literature on the net gains and losses from currency unions, which is outside the scope of this study.

exporting firms in two industries to show theoretically the net utility gains from, and consequently empirical likelihood of, a bilateral PTA depend on two countries' economic size and similarity, distance, remoteness, and relative factor endowments. Egger and Larch (2008) extended this work to show empirically how PTAs of "third-country-pairs" affected subsequent PTA formations.

However, the economic literature on BITs is *considerably* smaller, with surprisingly few contributions from economists. First, unlike the literature on PTAs, discussions of the potential benefits from BITs are dominated by legal and political science scholars rather than economists, cf., Salacuse (1990), Vandevelde (1998, 2000), Tobin and Rose-Ackerman (2005), and Buthe and Milner (2009).[4] Consequently, none of these papers addresses factors explaining BITs' formations using formal theoretical economic models, and few provide econometric empirical analyses (cf., Swenson, 2005, as an exception). Second, relative to the trade and PTA literature, there are few papers – some by economists and some by legal/political scholars – that have looked systematically and econometrically at the impact of BITs on FDI. Hallward-Driemier (2003), Tobin and Rose-Ackerman (2005), Gallagher and Birch (2006), and Aisbett (2009) find little economically and statistically significant effects of BITs on FDI flows. By contrast, Egger and Pfaffermayr (2004a), Salacuse and Sullivan (2005), Neumayer and Spess (2005), and Buthe and Milner (2009) find economically and statistically significant effects.[5]

In this paper, we examine formally, theoretically and econometrically, the economic determinants of BITs – and in a manner consistent with understanding the economic determinants of PTAs. While BITs have been examined much less in the international economics literature, the motivation for such agreements for FDI is actually quite analogous to that for PTAs for trade. While "Friendship, Commerce and Navigation" treaties surfaced as early as the 18th century, BITs were effectively created in the 1970s in response to numerous expropriations of FDIs as well as the limitation of the General Agreement on Tariffs and Trade (GATT) to trade only; (West) Germany concluded the first BIT. Hence, the first BITs were intended to reduce the relative cost of FDI, such as PTAs reduce the relative cost of trade.[6]

More recently, BITs regulate FDI-related issues such as admission, treatment, expropriation, and the settlement of disputes at the bilateral level. They establish transparency about investment risk, and consequently reduce *ex ante* the risks of FDI by a home country. BITs ensure that firms have property

[4] An exceptional edited volume of recent studies on the effect of BITs (and double taxation treaties) on FDI, including many of the papers noted in this paragraph, is Sauvant and Sachs (2009).

[5] Moreover, only one paper – Aisbett (2009) – provides a formal game-theoretic model of BITs formation based upon expropriation risk. We refer to this model later for motivating the inclusion of a measure of expropriation risk in our empirical analysis. However, the model does not address how factors such as economic size and similarity, trade and investment costs, and relative factor endowments can help explain the likelihood of a BIT between a country-pair.

[6] The importance of FDI has also led many PTAs to be broadened to include FDI provisions, such as in NAFTA. However, the number of preferential trade and investment agreements (PTIAs) still are fewer than 10 percent of the number of BITs (cf., Sauvant and Sachs, 2009) and this issue will be discussed later.

rights and they are protected from expropriation. According to the Office of Investment Affairs of the U.S. Bureau of Economic Business Affairs, BITs should: (1) protect U.S. FDI stocks where U.S. investors' rights are not protected already via existing agreements; (2) encourage host countries to adopt market-oriented domestic policies that treat foreign and domestic investors equivalently; and (3) support the development of international law standards. UNCTAD (1998) summarizes the following characteristics of many existing BITs. First, most BITS facilitate FDI by guaranteeing foreign investors fair and equitable, non-discriminatory, most-favored-nation and national treatment. Second, most BITS provide legal protection of physical and intellectual properties under international law. Third, some BITS even provide more transparent conditions for foreign investors than for their national counterparts.

Since the fundamental purpose now of a BIT is to encourage FDI flows between country pairs by reducing the relative cost of FDI and that of a PTA is to encourage trade between country pairs by reducing the relative cost of trade, economic determinants of BITs may well share many similarities to those of PTAs. Since there has been no previous formal theoretical nor econometric model of the determinants of BITs, our starting point is the new literature on the economic determinants of PTAs. This literature, surveyed in Freund and Ornelas (2009), starts with Baier and Bergstrand (2004), which developed a qualitative choice econometric model of the likelihood of a pair of countries having a PTA in a given year.[7] Motivated by a general equilibrium model of world trade with two factors, two monopolistically-competitive markets with national exporting firms, and explicit intercontinental and intracontinental trade costs among multiple countries on multiple continents, the Baier-Bergstrand model suggests that country-pairs are more likely to have a PTA: (1) the closer together they are; (2) the more remote they are from other markets; (3) the larger their joint economic size; (4) the more similar their economic sizes; (5) the larger the difference in the pairs' relative factor endowments (up to a point). Baier and Bergstrand (2004) showed that all these economic factors were economically and statistically significant (with expected signs) in explaining cross-sectional variation in country-pairs' probabilities of having PTAs with a pseudo-R^2 of 73 percent. Moreover, the probit model correctly predicted 85 percent of the 286 bilateral PTAs existing in 1996 among 1,431 pairs and 97 percent of the remaining 1,145 pairs with no PTAs. Using a larger sample of 10,585 pairs in year 2000, Egger and Larch (2008) predicted correctly about 1,000 of the 1,498 pairs with PTAs (or 67 percent) and 9,131 of the 9,322 pairs with no PTAs (or 98 percent). Of course, their pseudo-R^2 was considerably lower at 27 percent (as expected) due to their much larger and less selective sample.

[7]Recent extensions of the model include, for example, Egger and Larch (2008) and Baldwin and Jaimovich (2010), who examined the role of PTAs of "third-countries" for explaining subsequent PTA formations and enlargements, and Bergstrand, Egger and Larch (2010), who examined the role of several economic variables for the "timing" of PTA formations and enlargements.

However, the economic determinants of BITs are not likely to be explained by the same econometric model, due to several considerations. First, BITs potentially influence FDI flows. Consequently, while economic size and similarity help to predict PTAs, they may not simultaneously predict BITs; as mentioned, most BITs are between developed and developing countries (and the latter tend to be economically smaller than the former). Other factors – such as bilateral trade and investment costs and relative factor endowments – are likely to have differing effects on explaining BITs relative to PTAs. Since FDI is generated by multinational enterprises (MNEs), a theoretical framework should incorporate MNEs' behavior; consequently, a simple Krugman-Heckscher-Ohlin general equilibrium (GE) model of trade as in Baier and Bergstrand (2004) is insufficient. An extension of the Baier-Bergstrand framework to include MNEs, FDI flows, and foreign affiliate sales (FAS), in the spirit of the "Knowledge-Capital" (KC) models of Markusen (2002) and Markusen and Maskus (2001, 2002), is a natural direction. Fortunately, Bergstrand and Egger (2007) have extended the 2x2x2 KC model to three factors and three countries, and provide a ready framework to address the economic determinants of BITs and PTAs, in the spirit of the multi-country model in Baier and Bergstrand (2004) for just trade flows and PTAs. Bergstrand and Egger (2007) is especially relevant because it is the first general equilibrium model to demonstrate that bilateral FDI *and* trade are maximized between countries with *identical* relative *and absolute* factor endowments, consistent with the large literature on gravity equations that explain very well both bilateral trade *and FDI* flows.[8] Thus, the first potential contribution of this paper is to use the theoretical framework in Bergstrand and Egger (2007) to generate comparative statics to show (in the absence of any relative factor endowment differences) how: (1) economic size and size similarity of two countries influence their net utility gains from a BIT and from a PTA; and (2) bilateral investment and trade costs between two countries influence their net utility gains (or losses) from a BIT and from a PTA. Note that, in general equilibrium, the net utility gains from BITs and PTAs are influenced by the behavior of multinational enterprises as well as national (exporting) firms. In the presence of MNEs and general equilibrium, the influences of these economic variables on the net utility gains from a PTA are not *necessarily* the same as those found in Baier and Bergstrand (2004), where MNEs and FDI were absent.

Second, in reality relative factor endowments are not identical across countries, and such differences matter for economic determinants of BITs because *many* BITs (not to mention PTAs) are "North-South" in nature, that is, between countries with quite different relative factor endowments. The second potential contribution of this paper is to show the relationship between relative factor

[8] Blonigen and Piger (2011), using a Bayesian moving average statistical technique, have recently shown that the most important factors for explaining FDI flows are basic gravity equation variables parent- and host-country real GDPs and bilateral distance.

endowments between two countries and the net utility gains from BITs and from PTAs. This goal is more complicated here because Baier and Bergstrand (2004) had only two factors of production. With three factors – skilled labor, unskilled labor, *and physical capital* – the relationships are more complex. However, using traditional Edgeworth boxes and recent developments in specifying properly the relationships between relative factor endowments and bilateral FAS flows in Braconier, Norback and Urban (2005), the theoretical relationships between factor shares suggest easily specified empirical counterparts to capture some of the influences of relative factor endowments on the net utility gains from BITs and PTAs.[9]

Third, guided by the theoretical comparative statics, we specify a bivariate probit model of the probabilities of BITs and PTAs existing between country pairs in year 2000. We choose to estimate a bivariate probit model because the error terms may be correlated across probabilities, and this provides more efficient coefficient estimates. To anticipate some of the results, we find the following empirical conclusions. First, as most trade is "intra-industry" and most FDI is "horizontal," one would expect that the net utility gains from a BIT and from a PTA are positively related to economic size and similarity. Such results are confirmed here. Second, our theoretical model suggests that higher bilateral investment costs should reduce the likelihood of a BIT, but increase the likelihood of a PTA. Using a measure of investment costs that influences FDI and trade oppositely, we find evidence of this "substitution effect." Third, in the presence of three factors of production, the relationships between relative factor endowments (measured, as a conventional Edgeworth box suggests, using factor shares) and net utility gains of a pair of countries from a BIT or a PTA are complex, non-linear, and non-monotonic. However, we draw upon the geometric features of an Edgeworth box to introduce a measure of similarity of factor shares that helps explain the importance of relative abundance of skilled labor in the home country for increasing the net utility gains from a BIT for a country-pair. Alongside another (more standard) measure of deviations of capital-unskilled-labor ratios from the Edgeworth box diagonal, we show empirically how relative factor endowments affect the probabilities of a BIT and of a PTA. Finally, the bivariate empirical model has a relatively high explanatory power that holds up to an extensive sensitivity analysis. Moreover, the inclusion of different relative factor endowment variables for the BIT and PTA equations allows for potential identification in a simultaneous equations system, which we explore. Finally, we correctly predict 85 percent of the 1,479 BITs in year 2000 among 12,880 pairs of countries and 62 percent of the remaining 11,401 pairs with no BITS *as well as* 75 percent of the 2,034 PTAs in year 2000 among the 12,880 country-pairs and

[9] Blonigen and Piger (2011) show that – beyond parent- and host-country real GDPs and bilateral distance – the most important variables explaining FDI flows are parent real GDP per capita, parent *physical capital per worker*, relative skilled-labor endowments, common official language, urban concentration of the host country, and remoteness of the host country from *ROW*.

67 percent of the remaining 10,846 pairs with no PTAs. The overall pseudo-R^2 of 21 percent for our bivariate probit model is only slightly less than the 27 percent found in Egger and Larch (2008) for PTAs using a univariate probit model and a comparable sample. The results provide quantitative guidance as to the determinants of BITs and PTAs simultaneously and to addressing potentially the endogeneity bias inherent in many previous empirical analyses of the effects of BITs on FDI flows.

The remainder of the paper is as follows. In section 3, we summarize the theoretical GE model of Bergstrand and Egger (2007) and the parameter values chosen for the numerical version of the GE model. In section 4, we discuss the GE comparative static results for the relationships between the net utility gains from a BIT and from a PTA with economic size and similarity, investment costs, and trade costs, assuming identical relative factor endowments across countries. In section 5, we relax the assumption of identical relative factor endowments and, using conventional Edgeworth boxes and their geometric properties, provide GE comparative statics to motivate two relative-factor-endowment variables for the empirical analysis. In section 6, we describe our data set and provide the results from the bivariate probit empirical analysis, including the robustness analysis. Section 7 concludes.

3 Theoretical Framework: A Summary of the Bergstrand and Egger (2007) Knowledge-Capital-and-FDI Model

The model we use is a more general version of the 3-country, 3-factor, 2-good model in Bergstrand and Egger (2007) by allowing for BITs and for differences in relative skilled-to-unskilled-labor (S/U) ratios *and* relative physical-capital-to-unskilled-labor (K/U) ratios, the latter consequently generating horizontal and vertical MNEs as well as national exporting enterprises (NEs) in equilibrium. Bergstrand and Egger (2007) is a 3-factor, 3-country extension of the 2x2x2 KC model in Markusen (2002) with national exporters (NEs), horizontal MNEs (HMNEs), and vertical MNEs (VMNEs).[10] However, Bergstrand and Egger (2007) assumed identical relative factor endowments to focus on the roles of economic size and similarity and provide a theoretical foundation for gravity equations of bilateral trade, FDI, and foreign affiliate sales (FAS) simultaneously; consequently, no VMNEs surfaced in that paper, except in one sensitivity analysis. VMNEs play a role here alongside HMNEs, due to differences between countries in S/U and K/U ratios.

Although the theoretical model is used only to suggest testable hypotheses, it is important to note that the determinants of BITs are likely strongly related to the determinants of FDI flows (as

[10] Since the explicit equations are provided in Bergstrand and Egger (2007), there is no need to repeat them here in the text; we provide them in a Theoretical Appendix later, which is not intended for publication. To limit the scope of this already complex model, we assume (firm) homogeneity among NEs, among HMNEs, and among VMNEs, although there is heterogeneity between these three firm types.

Baier and Bergstrand (2004) found that determinants of PTAs were similar to determinants of trade flows). Recently, Blonigen and Piger (2011) showed using a Bayesian moving average econometric analysis that the most important factors influencing FDI flows are parent- and host-countries' real GDPs, bilateral distance, parent real GDP per capita, parent capital per worker, relative skilled-labor endowments, common official language, remoteness of host country, and urban concentration of the host country. With the exception of host country urban concentration, all of these economic variables are embedded in the Bergstrand and Egger (2007) model. Moreover, Bergstrand and Egger (2011) provide empirical support as well for the economic and statistical importance of these variables (except host urban concentration).

We now summarize the theoretical model underlying our hypotheses. The demand side in the model is analogous to that in the KC model. However, the model of Bergstrand and Egger (2007) extends the KC model in two important ways. The first distinction is the addition of "FDI" and physical capital to the KC model, where the latter focuses only on "knowledge capital" and skilled labor (alongside unskilled labor). A key notion of the 2-factor KC model is that both skilled and unskilled labor are internationally immobile; however, skilled labor possesses ("non-rival") knowledge capital for which its services can be shared costlessly by an MNE at home and abroad. Knowledge capital represents the *intangible* assets of firms, the services of which – because of their "public good" nature within the firm – can be used globally within the firm without "rivalry" and can flow relatively costlessly between countries, cf. Markusen (2002).

However, there is no factor in the KC model to represent the *tangible* assets of firms; hence, physical capital has been omitted from the KC model even though Blonigen and Piger (2011) show that it is important empirically. Tangible assets are represented by the (financial) claims of the firm on *private capital* and are distinguished from knowledge capital because of their "rival" ("private good") nature within the firm and the services can flow imperfectly across countries due to FDI impediments. In reality, of course, national and multinational firms use both (rival) private capital – often measured by the financial claims to tangible assets such as physical capital at home or abroad – as well as (non-rival) knowledge capital – often associated with skilled labor. Moreover, physical capital and skilled labor tend to be complements in production, cf., Slaughter (2000).

Thus, the key difference between our model and Markusen's KC model is that we add a third factor, private physical capital (K), the *services* of which can be used at home or transferred abroad (via FDI, and not necessarily costlessly) either as a "greenfield" investment or an acquisition.[11] We assume

[11] Hence, in the KC model, skilled labor is immobile internationally, but the services of (non-rival) "knowledge capital" are (costlessly) mobile. In our model, this still holds, but additionally physical capital is immobile internationally, but the services of (rival) "financial capital" are internationally mobile (with possible investment costs) for plant setups. Our model is static with no depreciation; the real world is dynamic and depreciation occurs. In reality in a typical year (with

that all three internationally-immobile primary factors are used in the production of the differentiated good: unskilled labor (U), skilled labor (S), and private physical capital (K). Moreover, following evidence from Griliches (1969), Goldin and Katz (1998), and Slaughter (2000), we assume that skilled labor and physical capital are complements in production, which is also consistent with evidence in Bernard, Jensen, and Schott (2005) that MNEs tend to be relatively abundant in countries that are relatively abundant in both skilled labor *and* physical capital. However, for the setups of headquarters and plants, we assume that only the services of skilled labor (physical capital) are used to setup headquarters (plants). HMNEs headquartered in any country i, for example, arise endogenously, and the services of home physical capital are "used up" (due to their "rival" nature) to setup a plant in the home country and abroad to maximize firm profits (with an implied no-arbitrage incentive for rates of return on physical capital; no profits are left "on the table"), but physical capital need not actually move internationally. In the presence of imperfect international (financial) capital mobility, firms may choose to have financial claims to physical capital at home or – via FDI – abroad; the key distinction from knowledge capital is that claims to physical capital are *rival*, and FDI – due to foreign government restrictions – may not move costlessly between countries.[12]

The second distinction of our model from the KC model is to introduce a "third country," or Rest-of-World (ROW). The presence of the third country helps explain the observed complementarity of bilateral FAS and trade flows with respect to a country pair's economic size and similarity and that bilateral FDI and FAS tend empirically to be as well explained by a gravity equation as bilateral trade flows are. One implication of a third country is that (in equilibrium) both two-country HMNEs and three-country HMNEs may surface. It is unclear what the implications of the typical "two-country" assumption in the KC model are when guiding empirical work in an N-country world.

As noted, vertical MNEs played no role in explaining FAS and trade patterns in Bergstrand and Egger (2007), except in one sensitivity analysis. By contrast, VMNEs play a central role here. We

growth), MNEs choose (based upon firm profit maximization) to allocate a portion of real fixed investment expenditures to the domestic market and portions to one or more foreign markets (either as greenfield investment or mergers-acquisition; our model generalizes to both); this definition of FDI is captured in the "current-cost" method (of "investment in plant and equipment"), cf., Borga and Yorgason (2002). In official statistics, FDI outflows are also (market-value) measured as infusions of equity capital and reinvestment of retained earnings into affiliates (and acquisitions). Economically, an FDI outflow reduces a parent's equity claims on tangible assets in the home country and increases the parent's claims on tangible assets in the country where the foreign affiliate (or acquisition) exists (hence, the "rival" nature of FDI). Conceptually, the Bergstrand and Egger (2007) model is consistent with such classic papers as Caves (1971), Markusen (1983), Helpman and Razin (1983), and Neary (1995) as well as with recent MNE models such as Muendler and Becker (2011). The numeraire homogeneous good will be produced using only unskilled labor.

[12] Physical capital (K) in our model can just as easily be interpreted as financial equity claims on any tangible assets (other than knowledge). The *key issue* for our model is that there is another factor (K) that complements skilled labor (S) in production and that (rival) K is "used up" in the fixed costs of setting up a plant at home or abroad. One can also interpret physical capital as "putty" capital, with the fraction of home capital used in domestic production and the fraction that is used up in setting up plants at home and abroad as endogenously determined by the MNE's profit maximization decision.

distinguish in this paper theoretically and empirically between differences between two countries' K/U ratios and differences between two countries' S/U ratios. Even though S and K tend to be complements in production, differences in the relative factor intensity requirements of setting up headquarters (needing S) relative to setting up plants (needing K) influence the results. Since the structure and calibration of the model is described explicitly in Bergstrand and Egger (2007), to conserve space we refer the reader to that paper for description of the theoretical model and of the calibration of its numerical version.

4 Economic Determinants of BITs and PTAs

In this section, we focus on eight theoretical results (summarized in four hypotheses), assuming countries are identical in *relative* factor endowments. First, we address the relationship between economic (GDP) size and similarity for influencing the net utility gains from BITs and PTAs (four results). Second, we discuss the relationship between bilateral investment and trade costs for influencing the net utility gains from BITs and PTAs (four results).

4.1 Economic Size, Economic Similarity, and Welfare Gains from BITs and PTAs

Using the model described above (and in Appendixes A and B), we simulated the model under alternative scenarios. First, we simulated the model with and without a BIT between countries i and j at various levels of GDP size and similarity for these two countries. In the base model without the BIT, the *ad valorem* policy-based FDI barrier (g) was 10 percent; in the scenario with the BIT, this barrier was zero. Figure 2a presents the results of the utility change for countries i and j from their forming a BIT, depending upon their joint economic size (y-axis) and their similarity of economic size (x-axis). The z-axis represents the utility change for *both* i and j from their forming a BIT (Figure 2a).[13] The y-axis is the sum of the GDPs of countries i and j. The lines on the y-axis are indexed from 1 to 1.7. The y-axis indexes country-pairs' GDPs from the smallest joint GDP (line 1) to the largest joint GDP (line 1.7).[14] The x-axis is indexed from 0.45 to 0.55. Each line represents i's share of both countries' GDPs, where the center line represents 50 percent, or identical GDP shares for i and j. Second, we simulated the model with and without a PTA between countries i and j at various levels of GDP size and similarity for these two countries. In the base model without the PTA, the *ad valorem* policy-based tariff rate (τ) was 20 percent; in the scenario with the PTA, the tariff rate was

[13] We assume the existence of transfers between the two countries' governments so that the relevant consideration is the sum of the two countries' representative consumers' utilities.

[14] See Bergstrand and Egger (2007) on determination of relative GDP sizes; the dispersion is based upon empirical GDP data.

zero. Figure 2b presents the results of the utility change for countries i and j from their forming a PTA.

Hypothesis 1: The net utility gain from and likelihood of a BIT and of a PTA between i and j is a positive function of their joint economic sizes.

The economic rationale is based upon the following intuition. Consider Figure 2a first. When i and j have identical relative factor endowments (as assumed in this section), there are horizontal MNEs in equilibrium, but no VMNEs (for a large set of parameters), consistent with Bergstrand and Egger (2007). All FDI (and FAS) is intra-industry. Consequently, when two countries are larger, there will be a larger volume of horizontal FDI. The formation of a BIT will then reduce bilateral investment costs on a larger volume of FDI for two larger countries (and associated larger number of varieties produced and consumed), and consequently increase the utility gains more for both countries relative to two smaller countries. This increases the probability of a BIT between i and j, for a given ROW GDP.

The same rationale applies to PTA and economic size, as in Baier and Bergstrand (2004) in the absence of MNEs, cf., Figure 2b. When i and j have identical relative factor endowments, all bilateral trade is intra-industry. Consequently, when two countries are larger, there will be a larger volume of intra-industry trade of NEs (and associated larger number of varieties produced and consumed). The formation of a PTA will then reduce bilateral trade costs on a larger volume of trade for two larger countries, and consequently increase the utility gains more for both countries relative to two smaller countries. This increases the probability of a PTA between i and j, for a given ROW GDP.[15]

Hypothesis 2: The net utility gain from and likelihood of a BIT and of a PTA between i and j is a positive function of the similarity in their economic sizes.

The economic rationale is based upon similar intra-industry reasoning. Consider Figure 2a first. We know from the literature on gravity equations that – for a given total economic size – flows between i and j (either trade or FDI) will be maximized when the two countries have identical sizes. In a simple frictionless world:

$$Flow_{ij} = GDP_i GDP_j / GDP^W = (GDP_i + GDP_j)^2 (sh_i sh_j) / GDP^W \quad (1)$$

where $Flow_{ij}$ denotes either a trade or FDI flow from i to j, GDP^W is world GDP, sh_i denotes i's share of countries' i and j joint GDP, and sh_j is defined analogously for j. Intuitively, suppose i had all of the two countries' (i's and j's) GDP. Then there would be no reason for i to setup an affiliated plant

[15] See Bergstrand and Egger (2007) on the theoretical foundation for bilateral trade and FDI flows related positively to both economic size and similarity.

in j, and consequently there would be no economic gain from forming a BIT. However, the volume of FDI from i to j will be maximized when each country has the same economic size. Consequently, the gains from a BIT will be maximized when the two countries have identical sizes.

A similar economic rationale holds for PTAs and trade, cf., Figure 2b, as addressed in Baier and Bergstrand (2004) in the absence of MNEs. We note, however, that the theoretical relationship between similarity and gains from a BIT does not appear as strong as that for a PTA. We leave it to the empirical analysis to confirm the relative strengths of the two relationships.

4.2 Investment Costs, Trade Costs, and Welfare Gains from BITs and PTAs

Figures 3a and 3b present the results of the utility change for two identical countries (i, j) from introducing a BIT between i and j and a PTA between i and j, respectively. First, we provide a brief explanation of the axes for these two figures. The "vertical" axis (or z-axis) represents the net utility gain (or, if negative, loss) for countries i and j from introducing either a BIT (cf., Figure 3a) or a PTA between i and j (cf., Figure 3b). The y-axis is labeled from 1 to 1.13 and represents the gross bilateral natural trade cost from i to j, τ_{ij}; τ=1 implies zero natural trade cost. The x-axis is labeled from 0.4 to 0.8 and represents a bilateral natural FDI investment cost, γ_{ij}, such as "political instability." A higher value of γ_{ij} represents a higher natural cost (in *ad valorem* terms) for an MNE with a headquarters in i to invest in j.

Although the details of the calibration are presented in Appendix B, recall that initially bilateral tariffs among all countries on goods (b_{ij}) are 20 percent and all bilateral "policy" investment costs on FDI (g_{ij}) are 10 percent (the latter are approximately one-fifth of the total foreign investment costs). The BIT between i and j, or BIT_{ij}, is captured by a reduction of the policy investment cost between i and j (g) from 10 percent to zero. The PTA between i and j, or PTA_{ij}, is captured by a reduction of the tariff rate between i and j (τ) from 20 percent to zero, similar to Baier and Bergstrand (2004).

Hypothesis 3: The net utility gains from and likelihood of a BIT between i and j are a negative function of their natural bilateral investment costs and a positive function of their natural bilateral trade costs.

A BIT between i and j tends to raise economic welfare of each of the two countries, with the gains larger the lower are natural bilateral investment costs and the higher are natural bilateral trade costs, cf., Figure 3a. To understand this, consider first the more well-known negative relationship between the gains to countries i and j of PTA_{ij} and "natural" trade costs between i and j in Figure 3b. Common examples of "natural" bilateral trade costs (from the large trade gravity equation literature) are the absence of a common land border and absence of a common language (i.e., the absence of these commonalities causes a positive trade cost). As shown in Baier and Bergstrand (2004), a low value of natural trade costs (such as the presence of a common land border) between i and j – other things the same – implies a high value of bilateral trade. Consequently, a given reduction in the bilateral tariff rate between i and j will lead to a large increase in bilateral trade and a large gain in utility for i and j. Hence, larger natural bilateral trade costs reduce the net utility gains from and likelihood of PTA_{ij}, cf., Figure 3b.

Consider now the relationship between the gains for i and j from BIT_{ij} and natural bilateral investment costs, shown in Figure 3a. An example of a "natural" bilateral investment cost for FDI is political instability of the host country government. A higher degree of political instability in host country j leads to a lower level of bilateral FDI from i to j. Since FDI is lower, the gains to FDI_{ij} from BIT_{ij} are lower. Hence, the net utility gains from and likelihood of BIT_{ij} are higher the higher (lower) the degree of political stability (instability) in j.

Consider now the "cross-price" effect. A higher natural bilateral trade cost between i and j will tend to reduce bilateral trade between i and j, but increase bilateral FDI between i to j. While for most country pairs in reality, bilateral FDI flows tend to be large when the countries' bilateral trade flows are large as addressed in Bergstrand and Egger (2007), FDI and trade are substitutes with respect to relative price effects. Thus, a BIT between i and j will liberalize a larger volume of FDI the higher are the countries' bilateral trade costs, leading to a larger net utility gain for i and j, and increasing the likelihood of BIT_{ij}, cf., Figure 3a. Hence, the net utility gains from and likelihood of BIT_{ij} are larger the higher are natural bilateral trade costs.

Hypothesis 4: The net utility gains from and likelihood of a PTA between i and j are a negative function of their natural bilateral trade costs and a positive function of their natural bilateral investment costs.

A PTA between i and j tends to raise economic welfare of each of the two countries, with the gains larger the lower are natural bilateral trade costs and the higher are natural bilateral investment costs, cf., Figure 3b.[16] We described one of these channels above already. A low natural bilateral trade cost, such as a common language, increases bilateral trade for countries i and j. Thus, a given reduction in their bilateral tariff rate will lead to a large increase in bilateral trade and a large gain in utility for i and j, increasing the likelihood of PTA_{ij}.

There is a similar cross-price effect for PTAs as well. A higher natural bilateral investment cost between i and j will tend to reduce bilateral FDI between i and j, but increase bilateral trade between i and j. Thus, a PTA between i and j will liberalize a larger volume of trade the higher are natural bilateral investment costs, leading to a larger net utility gain for i and j, and increasing the

[16] PTAs are assumed here to be trade agreements only, with no FDI provisions.

likelihood of PTA_{ij}, cf., Figure 3b. Hence, the likelihood of PTA_{ij} is higher the larger are natural bilateral investment costs, such as political instability. All four hypotheses will be "tested" later in the empirical section.

5 Relative Factor Endowments and Determinants of BITS and PTAs

As noted in the BITs literature, most BITs are between developed and developing countries, motivated initially (in the 1970s) by the risk of expropriation. Consequently, relative factor endowment differences may well be influential in the likelihood of a BIT between a country-pair. We now allow relative factor endowments to vary between countries and consider the net utility gains (or losses) from either a BIT or a PTA. For tractability, we use the traditional Edgeworth box to illustrate our results. However, an Edgeworth box is designed to illustrate the impacts in a world with two countries and two factors. In our three-factor setting, we are taking a "slice" of an Edgeworth cube for the two countries. For instance, if we consider two factors, physical capital (K) and unskilled labor (U), there is a continuum of such slices for the various values of s_i – the share of i's and j's skilled labor (S) in country i. For illustration below, we will examine the relationship between the utility gains for i and j from BIT_{ij} with k_i and u_i at $s_i = 0.5$; similarly, we will examine the relationship between the utility gains for i and j from PTA_{ij} with k_i and u_i at $s_i = 0.5$.[17] Also, since we are operating in a three-country world, we are examining these relationships for a *given* endowment of factors K, S and U in the ROW. Of course, the Edgeworth box relationships are quantitatively sensitive to the economic size of and relative factor endowments in ROW.[18]

Hypothesis 5: The net utility gain from and likelihood of a BIT between i and j are increasing in the abundance of skilled labor relative to capital and unskilled labor in i or j.

Figure 4a presents the relationships between the utility gains for i from BIT_{ij} with k_i and u_i at $s_i = 0.5$. Figure 4a suggests that the net utility gains for i and j from BIT_{ij} are maximized when either i or j has a very small amount of the two countries' K and U, given $s_i = 0.5$ (that is, i or j is very skilled labor abundant). Intuitively, the benefits of BIT_{ij} will be greater the larger the FDI/FAS created by the BIT. FAS will be larger with greater numbers of vertical MNEs (VMNEs) and horizontal MNEs (HMNEs). VMNEs and HMNEs will be prominent when skilled labor is abundant relative to capital and to unskilled labor, because such a country would have a comparative advantage in setting up headquarters (which are S intensive) and a comparative disadvantage in production at home (low

[17]Empirically, for the data set used later the actual means of s_i, u_i, and k_i all range between 0.53 and 0.56, so using $s_i = 0.5$ is a feasible choice.

[18]In the simulations below, we assume for ROW that its endowments of K, S and U are exactly one-third of the world's endowments and that trade and investment costs in the benchmark equilibrium are the same between i and j as they are between either of these countries and the ROW.

K and U for plant setups and production). Hence, the benefits from a BIT between i and j should be maximized when either i or j is abundant in S relative to K and U (note the utility gains are maximized near the two countries' origins, and $s_i = s_j = 0$.[19]

Hypothesis 6: The net utility gain from and likelihood of a PTA between i and j are decreasing in differences of both countries' ratios of physical capital to unskilled labor.

It is well established that a PTA between a pair of countries should increase trade between them. However, it *may not* be the case that the welfare of the country-pair is enhanced from this PTA – especially if the pair has large differences in capital-unskilled labor endowment ratios. Baier and Bergstrand (2004) showed in a world excluding MNEs that the welfare benefits from PTA_{ij} were positive in relative factor endowment differences, up to a point, based upon traditional comparative advantage interacted with trade costs.[20] While that study's empirical results supported that result, Egger and Larch (2008) found in a much larger sample that wider relative capital-unskilled labor ratios were negatively related to the likelihood of a PTA. The model here can help to explain this latter result.

Figure 4b illustrates the utility gain to i and j from PTA_{ij}. The prominent aspect of this figure is that the two countries' utility gains are decreasing in larger differences in their K/U ratios. Our model can explain this, drawing once again on changes in the activities of NEs, HMNEs, and VMNEs. Note initially that when i and j have identical shares of all three factors, pure intra-industry trade of NEs will be maximized, cf., Figure 4c (middle of the diagram). With a large amount of bilateral intra-industry trade, a PTA causes a large increase in trade between them, consistent with net utility gains at the center of Figure 4b. Moreover, the increase in the volume of trade due to a PTA is also very large if countries i and j have very *different* K/U ratios (cf., Figure 4c), consistent with a PTA benefiting traditional Heckscher-Ohlin trade, as in Baier and Bergstrand (2004).[21]

However, in our context with MNEs also, the utility gains on net from PTA_{ij} may in fact be a negative function of relative K/U ratios for i and j because of a large loss of varieties (and consequently utility) when K/U ratios of i and j are vastly different. The number of HMNEs are maximized when countries i and j are identical. With wider relative K/U ratios, there will be fewer HMNEs in i and

[19]We can confirm using additional figures of the numbers of VMNEs and HMNEs headquartered in i and j that both countries benefit the most from a BIT when i or j is very skilled labor abundant. It is important to note that when k_i and u_i are small, k_j and u_j are large by construction (k_i= 1-k_j). However, when k_i and u_i are small, they are small relative to s_i (since s_i=s_j=0.5 in Figure 4a), implying that i is relatively abundant in skilled labor, and consequently has a comparative advantage in setting up an MNE and benefitting from a BIT. Analogously, when k_j and u_j are small, they are small relative to s_j, implying that j is relatively abundant in skilled labor, and consequently has a comparative advantage in setting up an MNE and benefitting from a BIT. Thus, the potential gains from a BIT are maximized when either i or j is abundant in skilled labor relative to physical capital and unskilled labor.

[20]Baier and Bergstrand (2004) showed a quadratic relationship, both theoretically and empirically.

[21]Hence, bilateral trade of NEs is very large when countries i and j are identical in absolute and relative factor endowments (i.e., intra-industry trade) *and* when the two countries have very different K/U ratios (i.e., inter-industry trade).

j (and demand will be met more by NEs). The loss in volume of the relatively few varieties produced by HMNEs with PTA_{ij} – see Figure 4d – causes a greater loss of utility for i or j when K/U ratios are very large than the gain in utility for trading large volumes of the varieties goods produced by NEs, see Figure 4c.

Moreover, VMNE activity cannot offset these welfare losses. There is little change in the volume of VMNE activity from PTA_{ij}.[22] Consequently, the large utility loss from the decline in production of the few HMNEs in i or j in the presence of large K/U ratio differences offsets the utility gains from trading more of NEs' outputs following a PTA, suggesting that the probability of a PTA between i and j – in the context of MNEs – may fall the greater the K/U endowment ratio differences between the two countries, as stated in **Hypothesis 6**.

Since the Edgeworth surfaces are handy, we address briefly the method for which we will capture empirically the influence of relative K/U ratios of i and j on the likelihood of PTA_{ij}, in the context of our theoretical model. First, as in Braconier, Norback, and Urban (2005), we want the measure of relative factor endowment differences to capture as precisely as possible the relationships between relative factor endowment shares as shown in the figures. Consequently, the absolute difference in the natural logs of k_i/u_i and k_j/u_j captures deviations of relative factor endowments from the NW-SE diagonal in Figure 4b (where i's origin is the SW corner); in the regressions later, we use specifically $KURatio_{ij} \equiv |ln(k_i/u_i) - ln(k_j/u_j)|$.

By contrast, measuring the difference in economic size *along* the SW-NE diagonal in Figure 4a (from i's origin to j's origin) has not been done traditionally in the international trade literature. Typically, deviations in economic size are captured by variables such as $sh_i sh_j$, discussed earlier in equation (1), where sh_i is the share of country i's GDP in the sum of countries' i's and j's GDPs. However, one cannot just use $sh_i sh_j$ to capture the difference in endowments of both K and U along the SW-NE diagonal in Figure 4a. The reason is that – when k_i and u_i are very small – s_i in Figure 4a is *still* 0.5; hence in an Edgeworth box, variation in k_i and u_i along the diagonal changes *relative* factor endowments as well. However, there is a way to capture variation along the SW-NE diagonal in Figure 4a (for a given s_i). Using the geometric properties of the Edgeworth box, variation in the diagonal is captured by a variable, $KUDiff_{ij}$:

$$KUDiff_{ij} \equiv \ln |(k_i^2 + u_i^2)^{1/2}\mu - (k_j^2 + u_j^2)^{1/2}\mu|, \qquad (2)$$

where

$$\mu \equiv \frac{1}{(k_i^2 + u_i^2)^{1/2} + (k_j^2 + u_j^2)^{1/2}}$$

[22] Omitted figure for VMNE activity confirms this.

A rise in $KUDiff_{ij}$ reflects a wider difference in k_i *and* u_i relative to k_j *and* u_j. We expect $KUDiff_{ij}$ to be positively related to the probability of a BIT.

Finally, we provide no other hypotheses regarding the effects of relative factor endowments. Examination of the analogous comparative statics in $S - U$ space and in $K - S$ space yielded no clear empirically "testable" relationships between relative factor endowments and the utility gains or losses from a BIT or PTA.[23] Finally, in order to evaluate **Hypotheses 5** and **6** holding constant s_i and $s_j = (1 - s_i)$, we include the variable $Ssim_{ij}$ defined as:

$$Ssim_{ij} \equiv lns_i + ln(1 - s_i) \tag{3}$$

as a control variable.

6 Econometric Specification, Data Description, and Empirical Results

6.1 Econometric Specification

As in Baier and Bergstrand (2004), the econometric framework employed is the qualitative choice model of McFadden (1975, 1976). A qualitative choice model can be derived from an underlying latent variable model. For instance, let the underlying latent variable be denoted $y*$. In that study, $y*$ represented the difference in utility levels from an action (formation of a PTA), where

$$y^*_{ij} = \mathbf{x}_{ij}\boldsymbol{\beta} + e_{ij}, \tag{4}$$

where $\mathbf{x}_{ij}$ denoted a vector of explanatory variables (i.e., economic characteristics) of country-pair ij including a constant, $\boldsymbol{\beta}$ was a vector of parameters, and error term e_{ij} was assumed to be independent of $\mathbf{x}_{ij}$ and to have a standard normal distribution. In the context of the model formally, $y^*_{ij} = \Delta U_i + \Delta U_j$ where ΔU_i (ΔU_j) denotes the change in utility for the representative consumer in i (j) from a BIT or PTA. We are assuming implicitly the existence of transfers between the two countries' governments so that the relevant consideration is that the sum of the utility changes of the two countries' representative consumers needs to positive for their governments to form a BIT or PTA.

Since y^*_{ij} is unobservable, following McFadden (1975, 1976) Baier and Bergstrand (2004) defined an indicator variable, PTA_{ij}, which assumed the value 1 if two countries have a PTA and 0 otherwise,

[23] It will be important later for econometric purposes that BIT is likely related to $KUDiff_{ij}$ but not to $KURatio_{ij}$, and vice-versa for PTA. This satisfies the necessary exclusion restriction for estimating a simultaneous equation system in the robustness analysis later.

with the response probability, Pr, for PTA:

$$Pr(PTA_{ij} = 1) = Pr(y^*_{ij} > 0) = G(\mathbf{x}_{ij}\boldsymbol{\beta}), \quad (5)$$

where $G(\cdot)$ is the standard normal cumulative distribution function, which ensured that $Pr(PTA_{ij} = 1)$ was between 0 and 1. As conventional to this econometric literature, $Pr(PTA_{ij} = 1) > 0.5$ "indicated" $y^*_{ij} > 0$ and $Pr(PTA_{ij} = 1) \leq 0.5$ indicated $y^*_{ij} \leq 0$.

In this study, we are concerned with the determination of two probabilities, $Pr(PTA_{ij})$ and $Pr(BIT_{ij})$. As our theoretical model suggests, common factors may influence both probabilities simultaneously. Due to the likely presence of correlated error terms, we employ a *bivariate probit* model to estimate the likelihood of BITs and PTAs *simultaneously*.[24]

6.2 Data Description

A general equilibrium model such as the one outlined earlier is mainly informative about long-run economic relationships. Therefore, as in Baier and Bergstrand (2004) we use cross section data to infer the aforementioned hypotheses. These data capture the state of BITs and PTAs as of year 2000 and use explanatory variables which are averages of the five years prior to 2000. With regard to the dependent and independent variables, we use data from the following sources.

First, information on BITs in force as of 2000 was collected from the United Nations Conference on Trade and Development (UNCTAD). We use this information to define a binary variable BIT_{ij}, which is unity if countries i and j had a BIT in force by the end of the year 2000 and zero otherwise. Second, we collected data on preferential trade agreements (customs unions, free trade areas, and other preferential trade agreements) from the World Trade Organization (WTO) and individual countries' sources. On the basis of that data, we defined an indicator variable PTA_{ij}, which is unity whenever two countries i and j had a preferential trade agreement in force (either under or outside of the auspices of the WTO).

Third, data on a number of economic fundamentals such as real GDP in US dollars (GDP_i), labor force (L_i), and gross fixed capital formation at constant US dollars of 2000 (K_i) were taken from the World Bank's World Development Indicators (2005). These variables were used to construct the following determinants for our analysis: (i) a measure of bilateral economic size, $GDPSum_{ij} = \ln(GDP_i + GDP_j)$; (ii) a measure of similarity in bilateral economic size, $GDPSim_{ij} = \ln[sh_i(1-sh_i)]$; and (iii) capital endowments of country i, K_i, and, in turn, the variable $k_i = K_i/(K_i + K_j)$.[25]

[24] In a sensitivity analysis later, we also consider a 3SLS estimation of two linear probability models with endogenous explanatory variables as well as a (recursive) simultaneous equations probit model.

[25] We calculate K_i by using the perpetual inventory method, following Leamer (1984). For this, we calculate an initial

Fourth, data on skilled workers (S) come from a new database constructed by researchers at the World Population Program of the International Institute for Applied Systems Analysis (IIASA) which establishes panel data on attained education of the average workers in a comparable way for 120 countries (see Lutz, Crespo Cuaresma, and Sanderson, 2008). These data serve to distinguish between high-skilled workers (S) and low-skilled workers (U) in four education categories. We classify workers in education categories 3 and 4 (corresponding to upper secondary and tertiary education) as highly-skilled ones and workers with lower levels of attained education (categories 1 and 2) as unskilled workers. This obtains s_i and u_i as measures of i's share of skilled and unskilled workers, respectively, between i and j. Furthermore, we use s_i to construct $Ssim_{ij}$ to hold constant relative endowments of skilled workers between two countries i and j in some of the empirical specifications later (see above for theoretical rationale).

Fifth, we use data on the distance between economic centers of countries ($Distance_{ij}$), a common land border indicator ($Adjacency_{ij}$), and a common language indicator ($Language_{ij}$). We use bilateral distance not only to measure $Distance_{ij}$ but also to construct a measure remoteness ($REMOTE_{ij}$) of country-pair ij from the rest of the world. The latter is constructed as an average distance of i and j from all countries (except i and j), as in Baier and Bergstrand (2004) and Egger and Larch (2008).

Sixth, we employ measures of political stability of nations' governments ($PolStab_{ij}$) and the lack (or the inverse) of expropriation risk ($IExpRisk_{ij}$) in countries i and j. The former is based on data about political risk as published by BERI, and the latter we define as the inverse of investment risk available from the same source. BERI provides cross-country time-series data on components of risk. We define $PolStab_{ij}$ such that a higher level measures greater political stability in the least stable of the two countries. Moreover, we define $IExpRisk_{ij}$ such that a higher level measures less expropriation risk in the riskier one of two countries i and j for investments between them.

Table 2 presents summary statistics for all the variables used in this study.

6.3 Empirical Results

Table 3 presents the results based on seemingly unrelated bivariate probit models for BITs and PTAs. These models allow for correlation of BIT_{ij} and PTA_{ij} through two sources: (i) the observed determinants of such agreements as included in the specification of the latent process determining bivariate

stock of capital for year 0 in each country i, $K_{i0} = \sum_{t=-5}^{-1} I_{it}$, where t is a time index. This provides an estimate of the initial capital stock for a chosen year 0 equivalent to the sum of gross fixed capital investments in the five years prior to that. We chose 1980 as the base year for all countries to make sure that the weight of measurement error of the initial capital stock is negligible by 2000. Then, we calculate the capital stock in year 1 as $K_{i1} = 0.87K_{i0} + I_{i1}$, where 0.87 is one minus the depreciation rate and I_{i1} are i's real gross investments in year 1, and so forth, until we obtain K_i as a measure of the capital stock in the year 2000.

binary outcomes[26] and (ii) unobserved characteristics as included in the disturbances. Table 3 has several columns presenting the results of including variables suggested by our discussion above incrementally, similar to the presentation in Baier and Bergstrand (2004) for only PTAs. We will refer below to "a" to denote a specification associated with BIT_{ij} and "b" refers to results associated with PTA_{ij}.

Specifications 1a and 1b are reported to examine the effects on the likelihood of a BIT and a PTA, respectively, of variables from section 4, in particular, measures of economic size, economic similarity, distance, trade costs, and investment costs. We consider first economic size and similarity. We find that the sum and similarity of the two countries' GDPs ($GDPSum_{ij}$ and $GDPSim_{ij}$) have positive and statistically significant impacts on the likelihood of a BIT and of a PTA, as our theoretical model suggested.[27] This is the first study to find and explain the positive association between the economic size and similarity of a pair of countries and the likelihood of their having a BIT. Other things the same, country-pairs that are economically larger and more similar in size tend to have more FDI so that these economies' welfare will tend to rise more from a BIT, increasing the likelihood of their forming one. Moreover, country-pairs that are economically larger and more similar in size tend to have more trade so that these economies' welfare will tend to rise more from a PTA, increasing the likelihood of their forming one. Hence, **Hypotheses 1** and **2** are confirmed.

We now turn to $Distance_{ij}$, $Adjacency_{ij}$ and $Language_{ij}$. As in Baier and Bergstrand (2004), $Distance_{ij}$ was included as a proxy for "natural" trade costs.[28] However, bilateral distance is well-known from the gravity-equation literature to have a negative, economically significant, and statistically significant impact both on bilateral trade flows *and* bilateral FDI flows. Consequently, one might argue that bilateral distance is not capturing "trade" costs *per se* but rather trade as well as other "information" costs.

However, although the marginal response probabilities will be provided later, there is some information content in the relative coefficient estimates for $Distance_{ij}$ from Specifications 1a and 1b, consistent with gravity equations of FDI and trade. For instance, typically in gravity equations estimated using ordinary least squares (absolute values of) coefficient estimates on $Distance_{ij}$ are smaller for FDI flows relative to trade flows, cf., Berden, Bergstrand, and van Etten (2011), even though there are very few studies that estimate both trade and FDI gravity equations in the same study using a common specification. This result is consistent with the following notion. FDI flows are motivated

[26]Following McFadden, the latent processes underlying BIT_{ij} and PTA_{ij} could be interpreted as the net gains for country-pair ij from concluding one or the other type of agreement.

[27]We discuss the economic significance of these and other coefficients later in the section on marginal response probabilities.

[28]$Distance_{ij}$ is the natural logarithm of the country-pair's bilateral distance. Baier and Bergstrand (2004) used the variable $NATURAL_{ij}$, which was simply the log of the inverse of bilateral distance.

from two sources, with "trade costs" having opposite effects. Much vertical FDI is to set up plants abroad which serve as export platforms back to the home country; trade costs (as proxied by distance) should have a negative effect on this FDI. However, horizontal FDI sets up plants abroad to "jump-over" trade costs, so such costs should have a positive effect on this FDI. The smaller negative coefficient estimate for bilateral distance in gravity equations for FDI relative to trade flows is consistent with this behavior. Since BITs are intended to enhance FDI flows, the smaller negative coefficient estimate for $Distance_{ij}$ is consistent with this explanation.[29]

Specifications 1a and 1b reveal *opposite* coefficient signs between the BITs and PTA probits for $Adjacency_{ij}$ and $Language_{ij}$. These two variables are the most commonly used dummies in gravity equations of trade flows and typically have economically and statistically significant positive effects. It is interesting to find that these two variables have economically and statistically significant *negative* coefficient estimates in the BITs equation but positive coefficient estimates in the PTA equation. These results are consistent with (the trade-cost portions of) **Hypotheses 3** and **4**. As Figure 3b suggests, the likelihood of a PTA should be greater between two countries the lower their trade costs, which implies a positive relationship between $Adjacency_{ij}$ and $Language_{ij}$ with $Pr(PTA_{ij} = 1)$. But lower trade costs discourage horizontal FDI, which suggests (as in Figure 1a) that the net utility change for a pair of countries from a BIT is a negative function of $Adjacency_{ij}$ and of $Language_{ij}$. Thus, the results are consistent with inferring that sharing a common land border and a common language reduce trade costs.

However, it would be useful to have *explicit* measures of investment costs in order to determine if the remaining (investment-cost) portions of **Hypotheses 3** and **4** are confirmed. Specifications 2a and 2b augment Specifications 1a and 1b to include a measure of "investment costs." The measure employed here is the index $PolStab_{ij}$ described above, which is increasing in the perceived degree of political *stability*. In a recent study of several governance indicators and their relationships to FDI and trade flows (cf., Berden, Bergstrand, and van Etten, 2011), the measure of political stability had a positive and statistically significant effect on FDI flows and an estimated zero effect on trade flows, suggesting a plausible measure of investment costs; moreover, this variable had an economically and statistically significant effect on FDI relative to trade. Specification 2a reveals that greater political stability in the host country, by lowering investment costs and enlarging FDI, leads country-pairs to be more likely to form a BIT. This confirms **Hypothesis 3**. Moreover, $PolStab_{ij}$ has the opposite effect on the likelihood of a PTA. An increase in political stability reduces investment costs, which decreases the net utility gains from a bilateral PTA, consistent with Figure 3b and **Hypothesis 4**.

[29] See also Egger and Pfaffermayr (2004b) on this issue. These authors find opposite effects of distance, negative for trade and positive for FDI, consistent with our interpretation.

Figures 4a and 4b and **Hypotheses 5** and **6** suggest relationships between k_i and u_i with the welfare effects (and, hence, likelihoods) of BITs and PTAs. Figure 4a suggests a positive relationship between differences in k_i and u_i with k_j and u_j ($KUDiff_{ij}$) and the utility gains to i and j from a BIT, as a higher value of $KUDiff_{ij}$ is associated with a larger relative abundance of skilled labor in either i or j, which favors more MNEs and FDI. Figure 4b suggests a negative relationship between differences in k_i/u_i relative to k_j/u_j ($KUratio_{ij}$) and the utility gains to i and j from a PTA. Wider relative K/U ratios between i and j lead to more NEs trade, but a large decline in HMNEs activity, such that on net a PTA is welfare decreasing. Specifications 3a and 3b confirm the qualitative impacts of these two variables on their respective probabilities as consistent with these two hypotheses. Higher skilled labor abundance of either i or j leads to a higher probability of a BIT between them as expected. A wider K/U ratio between i and j leads to a lower probability of a PTA between them as expected.

Thus, all six hypotheses suggested by the numerical comparative statics discussed earlier are supported empirically.

6.4 Robustness Analysis

Tables 4 and 5 report the results of six additional sets of specifications to evaluate the sensitivity of our results to numerous issues.

6.4.1 Robustness to Expropriation Risk Measure

A measure of political stability from BERI was used as an (inverse) index of natural investment costs. We showed that this index had opposite effects on the probability of a BIT versus a PTA. Berden, Bergstrand, and van Etten (2011) showed that more political stability tends to be positively correlated with FDI-relative-to-trade flows, suggesting it is an (inverse) measure of investment costs. However, it is important to note that political stability works *indirectly* through FDI on the utility gains from a BIT. Less political stability will tend to reduce FDI so that the net utility gains from a BIT are lowered.

However, a more direct measure of investment costs – in particular, an index of expropriation risk – might influence the likelihood of a BIT *directly.* That is, less expropriation risk may directly lower the probability of a BIT, since this is one of the early main motivations for a BIT. For instance, Aisbett (2009) formed a political economy model of BIT formation motivated by expropriation risk. Inspired by this work, we introduce $IEXPrisk_{ij}$ to our model in Specifications 4a and 4b, which is a measure of inverse (or lack of) expropriation risk, reported in Table 4. $IExpRisk_{ij}$ is a bilateral variable measuring the inverse of expropriation risk for the most risky member of the pair of countries i and j;

a higher value denotes less expropriation risk of the most risky member of the pair. If expropriation risk has a *direct* impact on the likelihood of a BIT, then $IEXPrisk_{ij}$ would be expected to have a negative coefficient estimate in the BITs equation. We note that this variable has the expected negative effect on the probability of BIT_{ij} in Specification 4a; with less expropriation risk in the most risky country of the pair, there is less need for a BIT to protect the investment interests of the less risky country. This result is consistent with the model in Aisbett (2009). Moreover, the other coefficient estimates are robust to its inclusion.

6.4.2 Robustness to Including GDP Similarity

As discussed earlier, once we introduced measures of relative factor endowments that – in 3-dimensional Edgeworth-box space – also influenced relative economic sizes, it seemed appropriate to eliminate the variable measuring relative economic size, $GDPSim_{ij}$. However, one may argue that some results for relative factor endowments are biased by exclusion of $GDPSim_{ij}$ as a compact measure of similarity in absolute factor endowments. Specifications 5a and 5b add $GDPSim_{ij}$ back into the bivariate probit model. We see that the results are influenced only trivially. Although the coefficient estimate for $GDPSim_{ij}$ is positive and statistically significant at the 10 percent level in the PTA_{ij} equation, it is insignificant in the BIT_{ij} equation of the bivariate probit model. In both equations, none of the other coefficient estimates changes materially. Therefore, in the subsequent analysis of marginal response probabilities and for predicting particular BITs and PTAs, we include $GDPSim_{ij}$ but all our results generalize to excluding it as well.

6.4.3 Robustness to Including REMOTE

Another variable examined in Baier and Bergstrand (2004) was the "remoteness" of a country pair. In their model, the more economically remote a pair of countries was, the greater the utility gains from and likelihood of a PTA, as more remote countries trade more all else constant. We can show in the context of this paper's model with NEs and MNEs that the utility gains from either a BIT or a PTA are increasing in the "remoteness" of the pair. (For brevity, figures are omitted.)

Specifications 6a and 6b add to the previous specification the measure of remoteness used in Baier and Bergstrand (2004). For the PTA equation of the probit model, $REMOTE_{ij}$ has a positive and statistically significant effect, as expected. By contrast, the coefficient estimate is negative for the BITs equation; however, it is not statistically significantly different from zero for that equation. Moreover, its inclusion has no material effect on the other coefficient estimates.

6.4.4 Robustness to Preferential Trade and Investment Agreements

As discussed earlier, some of the PTAs in our sample are actually preferential trade *and investment* agreements (PTIAs), such as the European Union (EU) and NAFTA. Since our dependent variable for BITs is actual bilateral investment treaties, investment liberalizations covered under "trade" agreements are not included in our BITs dependent variable. Hence, for Germany and France, PTA_{ij} is recorded as a 1 whereas BIT_{ij} is recorded as a 0. We approached this issue two alternative ways. Specifications 7a and 7b in Table 5 comprise the bivariate probit model where we include an "adjusted" variable for BIT. The new variable – $AdjBIT$ – includes also any bilateral investment agreements akin to BITs that are within country-pairs' PTIAs. We modified 108 "0's" to "1's" of the BIT variable to capture the influence of investment agreements in the EU and in NAFTA as representing BITs. A comparison of the results in Specifications 6a and 6b using BITs and PTAs with those in Specifications 7a and 7b in Table 5 using the adjusted-BITs variable and PTAs shows that the results are materially the same in the two sets of specifications.

Alternatively, we could simply remove all the PTIAs from our sample. However, we note that – if anything – the inclusion of PTIAs in our sample of PTAs would tend to bias the results against us. For instance, the coefficient estimates for *Adjacency*, *Common Language*, *Political Stability*, and *Inverse Expropriation Risk* are expected to have opposite signs for PTAs and BITs. The inclusion of PTIAs in our PTA sample would tend to bias the coefficient estimates for these variables toward similar signs for PTAs and BITs. Yet, we find economically and statistically significant opposite coefficient signs for these variables across BITs and PTAs. This suggests that it is the trade provisions that are the primary element in our PTA variable. Nevertheless, we also re-ran Specifications 6a and 6b in Table 4 deleting the 108 observations associated with the EU members and NAFTA members, and the results do not change significantly enough to be presented. The only material changes were for the BITS specification where the Adjacency and Inverse Expropriation coefficient estimates became statistically insignificant (they had been marginally significant; their coefficient values remained basically the same); there were no material changes in the PTAs specification.

6.4.5 Robustness to Recursive Simultaneous Bivariate Probit Estimation

It is possible that the latent variable influencing the probability of a PTA for two countries also influences the latent variable influencing the probability of a BIT between the pair, or vice-versa. That is, the determination of the probabilities of PTA_{ij} and BIT_{ij} may be better represented by a simultaneous equations probit model. However, there is no econometric theoretical foundation for estimating a simultaneous equations probit model for BITs and PTAs where *both* endogenous variables appear as explanatory variables (PTA affecting BIT in one equation and BIT affecting PTA in the other equation, cf., Wooldridge, 2002). However, there is econometric theory for estimating a recursive simultaneous equations model with one of the endogenous variables entering one of the equations, cf., Schmidt (1981) and Wooldridge (2002, pp. 477-478). Moreover, in order to evaluate which equation should have the endogenous explanatory variable, we turn for guidance to a three-stage least squares (3SLS) estimation of two linear probability models first.

Models 8a and 8b in Table 5 report the results of estimating the linear probability versions of BIT and PTA using 3SLS. The notable finding from these two equations is that the PTA equation has no statistically significant coefficient estimates, including that for BIT. By contrast, the 3SLS estimates of the BIT equation show that the existence of a PTA does have a statistically significant effect on the existence of a BIT and the other coefficient estimates are plausible.[30] Given there is only empirical evidence that PTA causes BIT (using the 3SLS estimates), we estimated a recursive simultaneous probit equations model of BIT and PTA, where PTA was allowed to be an endogenous explanatory variable for the BIT equation. These estimates are reported in Specifications 9a and 9b of Table 5. The main finding is that PTA does have a statistically significant positive effect on the existence of a BIT. However, importantly *none* of the other coefficient estimates of the BIT equation are materially different from those in bivariate probit Specifications 6a and 6b in Table 4. The BIT equation is basically robust to including the endogenous explanatory variable PTA. Also, the PTA equation in Specification 9b is very similar to that in Specification 6b; the PTA equation is also robust as we never find any evidence of BITs affecting PTAs.

6.4.6 Caveat for Interdependence

In reality, the net utility gains from and likelihood of a BIT or PTA between a country-pair ij are influenced by either country's BITs and PTAs with other countries k ($k \neq i, j$) or by BITs and PTAs among other country-pairs kl ($kl \neq ij$). Egger and Larch (2008) was mainly devoted to exploring the effect of other agreements (henceforth, "interdependence") on the likelihood of a country-pair ij forming a PTA. Econometrically, Egger and Larch (2008) handled the issue using "spatial lags" of other agreements; they did so using both a panel and a cross-section. In their panel analysis, the interdependence variable was constructed using a spatial lag of other agreements five years prior to the period examined.

Unfortunately, as in Baier and Bergstrand (2004), we are limited here to a cross-section and constructing a panel of BIT_{ij} is beyond this paper's scope. The problem with a cross-section, as in

[30]Note that one cannot compare the sizes of these coefficient estimates since the linear probability model generates marginal effects.

Egger and Larch (2008), is that the spatial lag introduces a potential endogeneity bias because the spatial lag would be constructed using other BITs from the concurrent year. In a univariate setting as in Egger and Larch (2008), the endogeneity bias could be avoided by a re-transformation of the spatial lag model. However, even in the univariate case, two problems arose there: multiple integrals in the likelihood function and heteroskedasticity in the error terms. Consequently, they applied a Bayesian Markov-Chain Monte Carlo model requiring estimation using Gibbs sampling. This complicated and resource-intensive estimation technique would need to be somehow adapted econometrically for our bivariate system, an extensive undertaking well beyond the scope of this paper.

That said, we note that a comparison of the spatial and non-spatial cross-section models in Egger and Larch (2008) suggested that any bias introduced from ignoring the spatial lag was not excessive. All the coefficient estimates in the two models had identical coefficient and marginal-response-probability signs, although some variables' marginal response probabilities were quantitatively different. Moreover, the overall explanatory power of the two models were similar. While exploring interdependence remains a useful extension in this context, we leave this for future research because of the scope of the undertaking needed.

6.5 Marginal Response Probabilities and Predictions

In this section, we report two sets of findings. We discuss the marginal response probabilities of one-standard-deviation changes in the right-hand-side variables. Then, we summarize the percentages of correctly predicted PTAs and BITs (ones) and the correctly predicted "No-PTAs" and "No-BITs" (zeros) from the models.

6.5.1 Marginal Response Probabilities

Probit estimates cannot reveal the quantitative effect of a change in any RHS variable on the probability of a BIT or PTA. Given the standard bivariate normal distribution, we can calculate the marginal response probabilities to unit- or one-standard-deviation changes in the RHS variables. For brevity, we report only in Table 6 the marginal response probabilities to one-standard-deviation changes in the RHS variables, although the other results are available on request. Since it is a bivariate probit model, we report both the unconditional and conditional response probabilities.

First, the variable that has the largest quantitative effect on either the probability of a BIT or a PTA is economic size. Moreover, a one-standard-deviation change in GDP size has a larger effect on the likelihood of a BIT than on that of a PTA. This result accords well with an empirical result in Bergstrand and Egger (2007) that FDI flows are more elastic with respect to changes in economic size than are trade flows; consequently, the likelihood of a BIT is more responsive to size changes than is that of a PTA.

Second, GDP similarity has an economically and statistically significant effect on PTAs, but not on BITs, consistent with results found earlier. As discussed earlier, $Ssim$ and $KUDiff$ (for BITs) may be capturing empirically the role of size similarity.

Third, distance has a much larger (negative) impact on the likelihood of PTAs than on that for BITs. This is consistent with the discussion earlier; the influence of distance on the likelihood of BITs is likely muted by the trade-cost-jumping role of horizontal FDI. Also, of *Adjacency* and *Language*, only *Adjacency* has a significant marginal response probability. Of the two variables, *Adjacency* has a clearer economic interpretation, as an inverse natural-trade-cost measure. Lower trade costs associated with adjacent economies increases trade, reduces horizontal FDI, and lowers the likelihood of a BIT.

Fourth, both measures of investment costs have economically and statistically significant marginal response probabilities in the expected direction for likelihood of BITs and for likelihood of PTAs. Interestingly, a one-standard-deviation increase in the political stability index has a much larger absolute quantitative effect on the probability of a BIT than a one-standard-deviation increase in the (inverse) expropriation risk index.

Fifth, both relative factor endowment variables have economically and statistically significant marginal response effects on their respective probabilities. Interestingly, the negative effect on the probability of a PTA of a one-standard-deviation increase in relative K/U ratios is considerably larger quantitatively than the positive effect on the probability of a BIT of a one-standard-deviation increase in the relative skilled-labor abundance of countries i and j. This actually accords well with Figures 4a and 4b, where the negative welfare effect from a PTA of widening K/U ratios was considerably larger (in absolute terms) than the positive welfare effect from a BIT of a widening of k_i and u_i relative to k_j and u_j (and consequently larger skilled-labor abundance of i or j).

6.5.2 Explanatory Power and Predicted Probabilities

One of the interesting results from the original Baier and Bergstrand (2004) study was the high pseudo-R^2 value of 73 percent in their fullest specification of explaining the likelihood of PTAs. However, this study examined only 1,145 pairings of 54 countries. In a much larger and less selective cross-section using 146 countries and 10,585 country-pairs, Egger and Larch (2008) showed that the Baier-Bergstrand results held up, but the overall explanatory power, as captured by the pseudo-R^2, was unsurprisingly much smaller at 27 percent. In this study, we use an even larger data set than Egger-Larch, with 161 countries and 12,880 observations. Moreover, explaining the likelihood of PTAs *and* BITs *simultaneously* is a larger challenge. Hence, our pseudo-R^2 of 21 percent in our fullest specification, Model 6, is a strong result, in this context.

In the spirit of Baier and Bergstrand (2004) for just PTAs, we also consider here the percent of correctly predicted, both for "True Positives" and for "True Negatives." However, in this bivariate model, we are calculating percent correctly predicted for each of the *four* outcomes: True Positives for BITs and PTAs and True Negatives for BITs and PTAs. Baier and Bergstrand (2004) conducted this statistical summary for their cross-section analysis of year 1996 for PTAs and found that their model predicted correctly 243 of 286 PTAs, or 85 percent (using a cutoff probability of 0.5). They also predicted 1,114 of the 1,145 pairs without PTAs correctly, or 97 percent. However, this sample was quite small, and Egger and Larch conducted a similar analysis using their larger cross-section of 10,585 country-pairs. Using a cutoff probability of 0.5 for year 2000, the percent PTAs correctly predicted was 62 percent and the percent of "No-PTAs" correctly predicted was 98 percent.

In this study, we considered three potential cutoff probabilities. First, for comparison with Baier-Bergstrand and Egger-Larch, we considered a cutoff probability of 0.5. Table 7 provides the percent correctly predicted for the various outcomes at a cutoff probability of 0.5 in the second column using Model 6. While the overall correct prediction rate of BITs and No-BITs was 77 percent, the model correctly predicted only 17 percent of BITs correctly (the True-Positive rate) with a True-Negative rate of 85 percent. However, our bivariate probit model also did not perform as well for PTAs as compared to the univariate model in Egger and Larch (2008). Model 6 predicted only 76 percent of PTAs and No-PTAs correctly. It predicted 29 percent of the PTAs correctly and 84 percent of the No-PTAs correctly. In summary, the overall prediction rate for PTAs in the bivariate model was much less than that in the Baier-Bergstrand and Egger-Larch univariate models, but the overall BITs prediction rate was roughly comparable to the PTAs prediction rate in this model.

Second, it is important to note, however, that a cutoff probability of 0.5 is probably not a very relevant one. The reason is that PTAs and BITs are still somewhat rare events. As Table 1 showed, of 12,880 country-pairs in year 2000, only 1,479 had BITs and only 2,034 had PTAs. Thus, the unconditional probability of having a BIT is 11 percent and that of having a PTA is 16 percent. A cutoff probability of 0.5 consequently is too extreme a cutoff. Cohen et. al. (2003) suggest using *a priori* information about the proportion of (PTA and BIT) events and non-events in our population. The third column of Table 7 shows the percent correctly predicted using the unconditional probabilities for BITs ($E(BIT) = 0.11$) and PTAs ($E(PTA) = 0.16$) for cutoffs, as has been used in the statistical literature. Using the cutoff for BITs of 0.11, the percent of BITs correctly predicted was 85 percent and the percent of No-BITs correctly predicted was 62 percent. Using the cutoff for PTAs of 0.16, the percent of PTAs correctly predicted was 75 percent and the percent of No-PTAs correctly predicted was 65 percent.

We note that both of these prediction rates are well below that of Baier and Bergstrand (2004). However, two factors need to be emphasized again: Baier and Bergstrand (2004) predicted PTAs using a univariate probit model and that study dealt with a much smaller number of (select) observations. We have both a larger sample here as well as a bivariate probit model. Nevertheless, the lower level of predictability of True Positives and True Negatives found here – in the first systematic attempt at explaining empirically the likelihood of BITs – suggests room for further research.

In the spirit of considering other approaches, note that the goodness of fit in terms of correctly predicted binary outcomes can be determined in various ways. In general, such goodness of fit in univariate nonlinear binary response models will depend on the cutoff level chosen to determine predicted positive and negative outcomes (positive referring to a unitary entry in the indicator and negative to a zero entry). Custom measures of goodness of fit are the so-called Receiver Operating Characteristic (ROC) curve (see Cameron and Trivedi, 2005) and Matthew's (1975) Correlation Coefficient (MCC). We will focus in the remainder of this section on the latter, since the former can not be generalized in a straightforward to allow for a graphical representation of goodness of fit while the latter can be.

MCC represents a correlation coefficient between observed and predicted binary outcome and takes on real numbers in the interval $MCC \in [-1; +1]$. A higher level of MCC reflects a better prediction of positive as well as negative binary outcomes (see Baldi, Brunak, Chauvin, Andersen, and Nielsen, 2000). While Matthew (1975) focused on the case of models with a single binary indicator, we face a bivariate relationship for two indicators (BIT_{ij} and PTA_{ij}). Matthew's idea was the following. Divide the unit interval for the probability support region for a single dependent binary outcome variable into segments of equal size and refer to them as $p_\ell = 0, ..., 1$. Altogether, there will be $\mathfrak{P}$ such elements p_ℓ. Then, following Matthew (1975), one may calculate a correlation coefficient based on True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) events. Predictions of treatment status depend on the chosen cutoff level p_ℓ for predicted success probabilities. For instance, with a univariate probit model for PTA membership as in Baier and Bergstrand (2004), one could state that $\widehat{PTA}_{ij}(p_\ell) = 1$ if $Pr(PTA_{ij}|X_{ij}) > p_\ell$ where X_{ij} is the vector of explanatory variables determining the latent process underlying PTA_{ij}. Then, one may define $TP_\ell = \sum_i^J \sum_j^J [\widehat{PTA}_{ij}(p_\ell) P_{ij}(p_\ell)]$, $TN_\ell = \sum_i^J \sum_j^J [(1 - \widehat{PTA}_{ij}(p_\ell))(1 - P_{ij}(p_\ell))]$, $FP_\ell = \sum_i^J \sum_j^J [\widehat{PTA}_{ij}(p_\ell)(1 - P_{ij}(p_\ell))]$, and $FN_\ell = \sum_i^J \sum_j^J [(1 - \widehat{PTA}_{ij}(p_\ell))P_{ij}(p_\ell)]$. This obtains

$$MCC_\ell = \frac{TP_\ell TN_\ell - FP_\ell FN_\ell}{\sqrt{(TP_\ell + FP_\ell)(TP_\ell + FN_\ell)(TN_\ell + FP_\ell)(TN_\ell + FN_\ell)}}.$$

With a bivariate probit model, we may consider two cutoff probabilities, say, p_ℓ and p_∂. For convenience, let us refer to p_ℓ as the one for PTAs and p_∂ as the one for BITs. Then, one can determine $TP_{A,\ell\partial}$, $TN_{A,\ell\partial}$, $FP_{A,\ell\partial}$, $FN_{A,\ell\partial}$ for each $A \in [PTA, BIT]$. Furthermore, we may calculate $TP_{\ell\partial} = \sum_A TP_{A,\ell\partial}$ and similarly for $TN_{\ell\partial}$, $FP_{\ell\partial}$, $FN_{\ell\partial}$. It is important to note that these outcomes depend on both cutoff levels. Accordingly, we can formulate

$$MCC_{\ell\partial} = \frac{TP_{\ell\partial}TN_{\ell\partial} - FP_{\ell\partial}FN_{\ell\partial}}{\sqrt{(TP_{\ell\partial} + FP_{\ell\partial})(TP_{\ell\partial} + FN_{\ell\partial})(TN_{\ell\partial} + FP_{\ell\partial})(TN_{\ell\partial} + FN_{\ell\partial})}}. \quad (6)$$

Figure 5 may be viewed as graphical illustration of the generalized MCC as in (6) in bivariate cutoff space (p_ℓ, p_∂). The optimal cutoff level lies wherever the function drawn reaches its maximum. This happens to be the case at a probability cutoff level of about 0.25 for PTAs and of about 0.24 for BITs. At those cutoffs, the correlation coefficient between predicted and actual binary outcomes is 39 percent.

7 Conclusions

The purpose of this study was to develop an econometric model that explains the "economic" determinants of BITs – at the same time as explaining PTAs – as well as predicting successfully the likelihood of pairs of countries forming BITs and FTAs. In the spirit of Baier and Bergstrand (2004), which explained PTAs in the context of a general equilibrium model of world trade with exporters, the model in this study is the first econometric model to explain BITs (along with PTAs) in the context of an explicit general equilibrium model of world production, consumption, trade, and FDI with national and multinational firms in multiple countries on multiple continents.

The main conclusions are that the potential welfare gains from and likelihood of a BIT (PTA) between a country-pair are higher: (1) the larger and more similar in GDP are the country-pair; (2) the closer in distance are the two countries; (3) if the two countries are not adjacent (are adjacent) and do not share (do share) a common language; (4) the higher (lower) the degrees of political stability and of expropriation risk of the pair; and (5) the relatively more skilled labor abundant (the wider the relative K/U ratio of) the pair. These factors have economically and statistically significant effects on the probability of a BIT (PTA).

While there exist choices of cutoff probabilities in determining the percent correctly predicted of the alternative outcomes, using the unconditional probabilities, the preferred empirical model predicts correctly 85 percent of BITs while predicting correctly 62 percent of "No-BIT" events and *simultaneously* predicts correctly 75 percent of PTAs while predicting correctly 65 percent of "No-PTA" events.

The model provides a benchmark for incorporating other economic – and especially political science and legal variables – into understanding the determinants of BITs.

References

Aisbett, Emma, (2009), Bilateral Investment Treaties and Foreign Direct Investment: Correlation versus Causation, in Karl P. Sauvant and Lisa E. Sachs (eds.), *The Effect of Treaties on Foreign Direct Investment*, Oxford University Press, Oxford, UK.

Baier, Scott L., and Jeffrey H. Bergstrand (2004), Economic Determinants of Free Trade Agreements, *Journal of International Economics*, **64**, 29-63.

Baldi, Pierre, Søren Brunak, Yves Chauvin, Claus A.F. Andersen, and Henrik Nielsen (2000), Assessing the Accuracy of Prediction Algorithms for Classification: An Overview, *Bioinformatics* **16**, 412–424.

Baldwin, Richard E. (2007), The Logic of Trade blocs as Building and Stumbling Blocks on the Road to Global Free Trade: A Survey, *unpublished manuscript*, Graduate Institute for International Studies.

Baldwin, Richard E. and Dany Jaimovich (2010), Are free trade agreements contagious?, *unpublished manuscript*, Graduate Institute for International Studies, November.

Berden, Koen, Jeffrey H. Bergstrand, and Eva van Etten (2011), Globalization and Governance, working paper.

Bergstrand, Jeffrey H., and Peter Egger (2007), A Knowledge-and-Physical-Capital Model of International Trade Flows, Foreign Direct Investment, and Multinational Enterprises, *Journal of International Economics*, **73**, 278-308.

Bergstrand, Jeffrey H., and Peter Egger (2011), International Trade, Multinational Enterprises, and Relative Factor Endowments, *unpublished manuscript*, University of Notre Dame.

Bergstrand, Jeffrey, Peter Egger, and Mario Larch (2010), Economic Determinants of the *Timing* if Preferential Trade Agreement Formations and Enlargements, *unpublished manuscript*, University of Notre Dame, September.

Bernard, Andrew B., J. Bradford Jensen, and Peter K. Schott (2005), Importers, Exporters, and Multinationals: A Portrait of Firms in the U.S. that Trade Goods, Tuck School of Business at Dartmouth University Working Paper No. 2005-17, May.

Blonigen, Bruce A., and Jeremy Piger (2011), Determinants of Foreign Direct Investment, NBER Working Paper No. 16704, January.

Borga, Maria, and Daniel R. Yorgason, Direct Investment Positions for 2001: Country and Industry Detail, *Survey of Current Business*, July, 25-35.

Braconier, Henrik, Pehr-Johan Norback, and Dieter Urban (2005), Reconciling the Evidence on the Knowledge-Capital Model, *Review of International Economics* **13** (4), 770-786.

Buthe, Tim, and Helen V. Milner (2009), Bilateral Investment Treaties and Foreign Direct Investment: A Political Analysis, in Karl P. Sauvant and Lisa E. Sachs (eds.), *The Effect of Treaties on Foreign Direct Investment: Bilateral Investment Treaties, Double Taxation Treaties, and Investment Flows*, Oxford University Press, Oxford, UK.

Cameron, A. Colin, and Pravin K. Trivedi (2005), *Microeconometrics: Methods and Applications*, Cambridge University Press, Cambridge, UK.

Caves, Richard (1971), International Corporations: The Industrial Economics of Foreign Investment, *Economica* **38**, February, 1-27.

Cohen, Jacob, Patricia Cohen, Stephen G. West, and Leona S. Aiken (2003), *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences,* Third Edition, Lawrence Erlbaum Associates, Mahwah, NJ.

Egger, Peter, and Mario Larch (2008), Interdependent Preferential Trade Agreement Memberships: An Empirical Analysis, *Journal of International Economics* **76**, 384-399.

Egger, Peter, and Michael Pfaffermayr (2004a), The Impact of Bilateral Investment Treaties on Foreign Direct Investment, *Journal of Comparative Economics* **32**, 788-804.

Egger, Peter, and Michael Pfaffermayr (2004b), Distance, Trade, and FDI: A Hausman-Taylor SUR Approach, *Journal of Applied Econometrics* **19**, 227-246.

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

Freund, Caroline, and Emanuel Ornelas (2009), Regional Trade Agreements, *unpublished manuscript*, London School of Economics.

Gallagher, Kevin P., and Melissa B.L. Birch (2006), Do Investment Agreements Attract Investment? Evidence from Latin America," *Journal of World Investment and Trade* **7** (6), 961-976.

Goldin, Claudia, and Lawrence F. Katz (1998), The Origin of Technology-Skill Complementarity, *Quarterly Journal of Economics* **113**(3), 693-732.

Griliches, Zvi (1969), Capital-skill Complementarity, *Review of Economics and Statistics* **51**(4), 465-468.

Hallward-Driemeier, Mary (2003), Do Bilateral Investment Treaties Attract FDI? Only a Bit and They Could Bite, World Bank Policy Research Paper WPS 3121.

Helpman, Elhanan, and Assaf Razin (1983), Increasing Returns, Monopolistic Competition, and Factor Movements: A Welfare Analysis, *Journal of International Economics* **14**(3/4), 263-276.

Krugman, Paul R. (1991a), Is Bilateralism Bad? in Helpman, Elhanan and Assaf Razin, A. (eds.), *International Trade and Trade Policy*, MIT Press, Cambridge, MA.

Krugman, Paul R. (1991b), The Move Toward Free Trade Zones. Policy Implications of Trade and Currency Zones, Federal Reserve Bank of Kansas City, Jackson Hole, WY.

Leamer, Edward E. (1984), *Sources of International Comparative Advantage: Theory and Evidence*, MIT Press, Cambridge, MA.

Lipsey, Robert G. (1960), The Theory of Customs Unions: A Survey, *Economic Journal*, **70**, no. 3, 496-513.

Lutz, Wolfgang , Jesus Crespo Cuaresma, and W. Sanderson (2008), The Demography of Educational Attainment and Economic Growth, *Science* **319**, 1047-1048.

Markusen, James R. (1983), Factor Movements and Commodity Trade as Complements, *Journal of International Economics* **14**(3/4), 341-350.

Markusen, James R. (2002), *Multinational Firms and the Theory of International Trade*, MIT Press, Cambridge, Massachusetts.

Markusen, James R., and Keith E. Maskus (2001), Multinational Firms: Reconciling Theory and Evidence, in Magnus Blomstron and Linda Goldberg (eds.), *Topics in Empirical International Economics: A Festschrift in Honor of Robert E. Lipsey*, University of Chicago Press, Chicago, IL.

Markusen, James R., and Keith E. Maskus (2002), Discriminating Among Alternative Theories of the Multinational Enterprise, *Review of International Economics*, **10**, no. 4, 694-707.

Matthew, Brian W. (1975), Comparison of the Predicted and Observed Secondary Structure of T4 Phage Lysozyme, *Biochemica et Biophysica Acta* **405**, 442–451.

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

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

Muendler, Marc-Andreas, and Sascha O. Becker (2011), Margins of Multinational Labor Substitution, *American Economic Review*, **100**, no. 5, December, 1999-2030.

Neary, J. Peter (1995), Factor Mobility and International Trade, *Canadian Journal of Economics*, **28**, S4-S23.

Neumayer, E., and L. Spess, (2005), Do Bilateral Investment Treaties Increase Foreign Direct Investment to Developing Countries?, *World Development*, **33**, no. 10, 1567-1585.

Salacuse, Jeswald W. (1990), BIT by BIT: The Growth of Bilateral Investment Treaties and Their Impact on Foreign Investment in Developing Countries," *HeinOnline*, **24**, 655-675.

Salacuse, Jeswald W., and Nicholas P. Sullivan (2005), Do BITs Really Work? An Evaluation of Bilateral Investment Treaties and Their Grand Bargain, *Harvard International Law Journal* **46**, no. 1.

Sauvant, Karl P., and Lisa E. Sachs (eds.) (2009), *The Effect of Treaties on Foreign Direct Investment*, Oxford University Press, New York, NY.

Schmidt, Peter (1981), Constraints on the Parameters in simultaneous Tobit and Probit Models, in Charles Manski and Daniel McFadden (eds.), *Structural Analysis of Discrete Data and Econometric Application*, MIT Press, Cambridge, Massachusetts.

Slaughter, Matthew J. (2000), Production Transfer within Multinational Enterprises and American Wages, *Journal of International Economics* **50**(2), 449-472.

Swenson, Deborah (2005), Why Do Developing Countries Sign BITs?, *U.C. Davis Journal of International Law and Policy*, **131**.

Tobin, Jennifer, and Susan Rose-Ackerman (2005), Foreign Direct Investment and the Business Environment in Developing Countries: The Impact of Bilateral Investment Treaties, Yale Law School Center for Law, Economics, and Public Policy Research paper No. 293.

United Nations Conference on Trade and Development (UNCTAD), Recent Developments in International Investment Agreements, IIA Monitor No. 3, 2007.

Vandevelde, Kenneth J. (1998), Investment Liberalization and Economic Development: The Role of Bilateral Investment Treaties, *Columbia Journal of Transnational Law*, **36**, 501-527.

Vandevelde, Kenneth J. (2000), The Economics of Bilateral Investment Treaties, *Harvard International Law Journal*, **41**, 469-502.

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

Wooldridge, Jeffrey M. (2002), *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, Massachusetts.

Figure 1

Number of BITs in the World by Year, 1980-2007

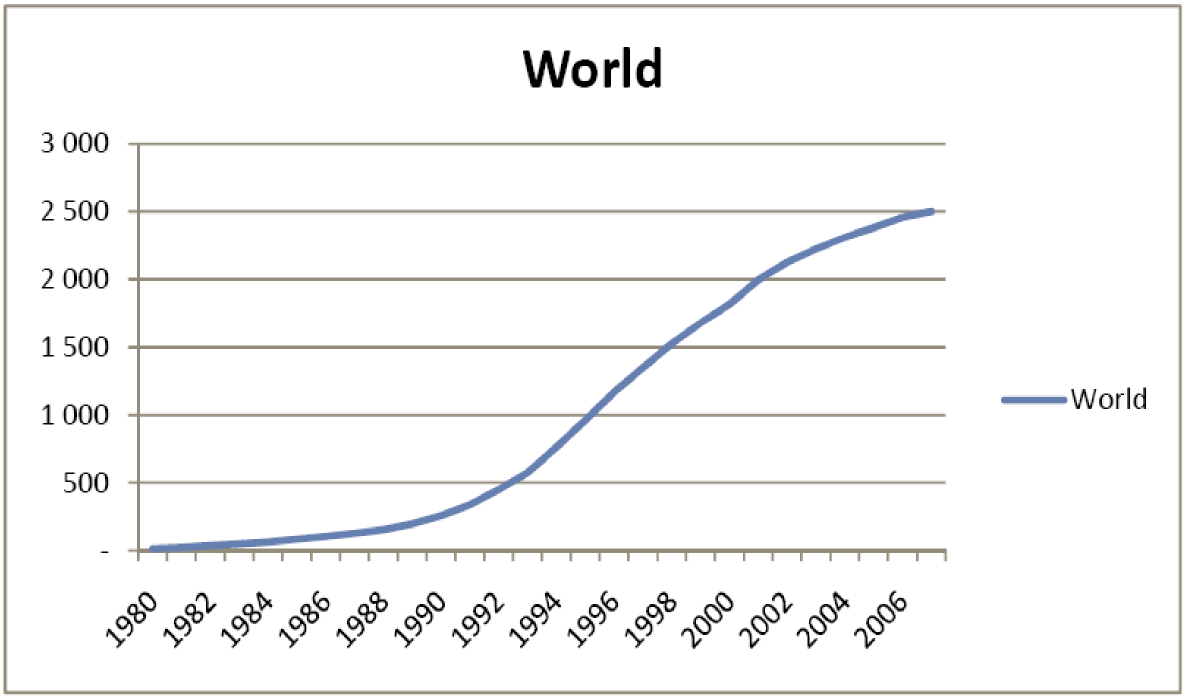


Figure 2a

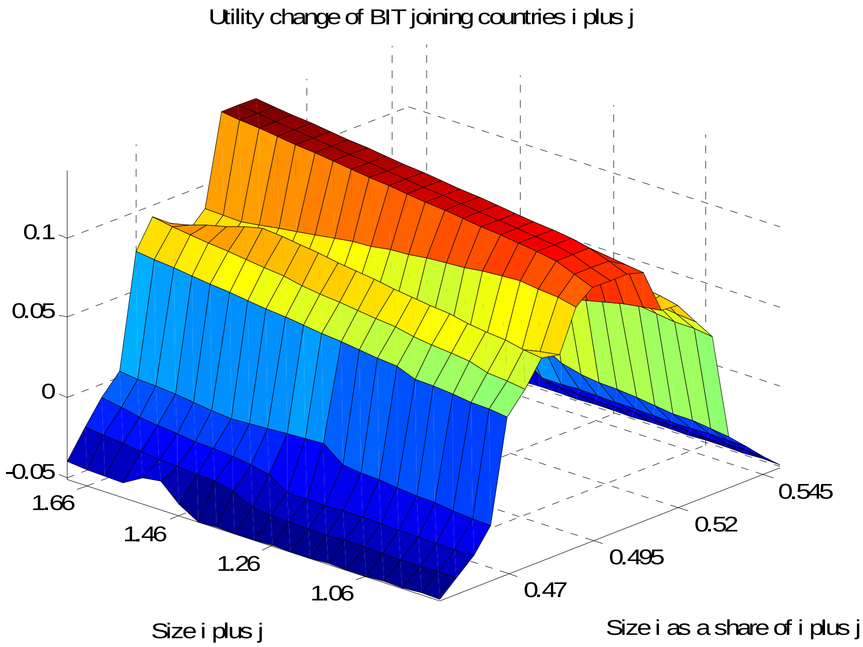


Figure 2b

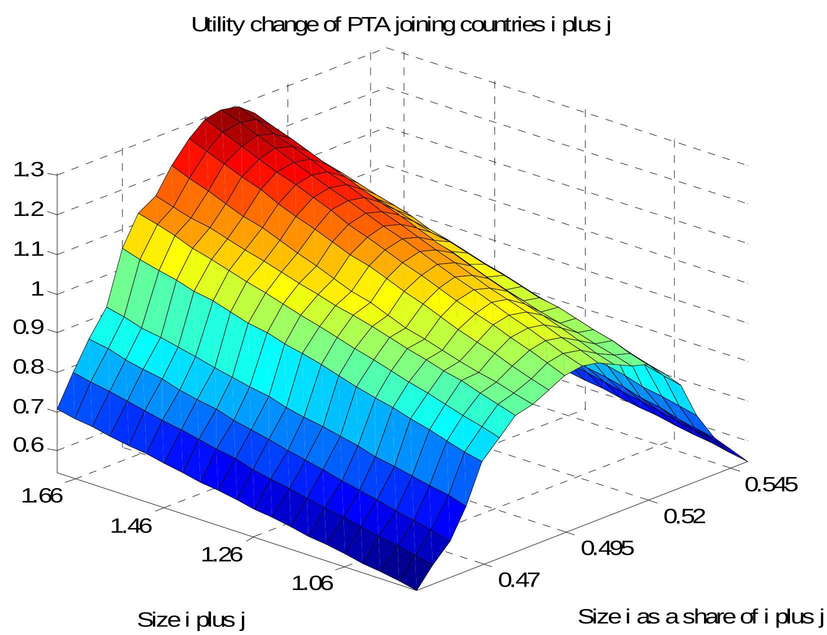


Figure 3a

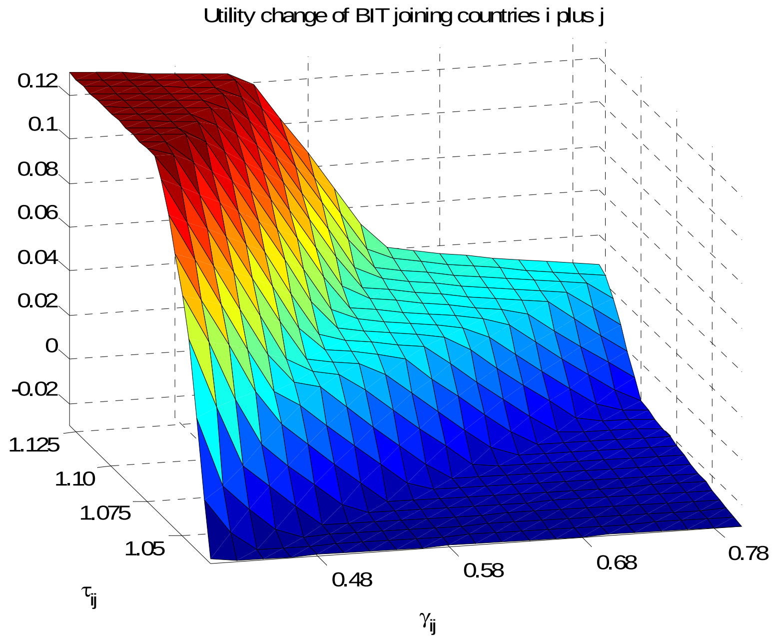


Figure 3b

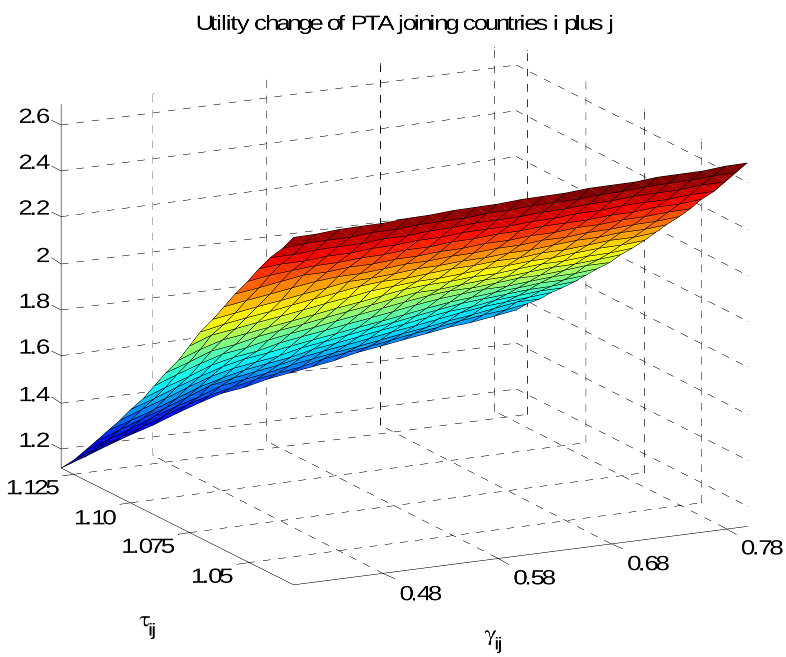


Figure 4a

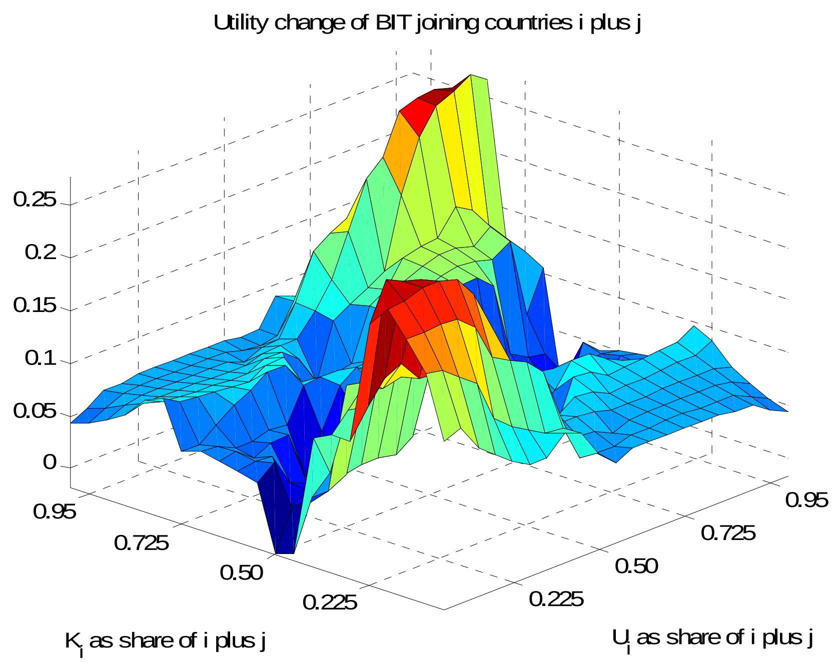


Figure 4b

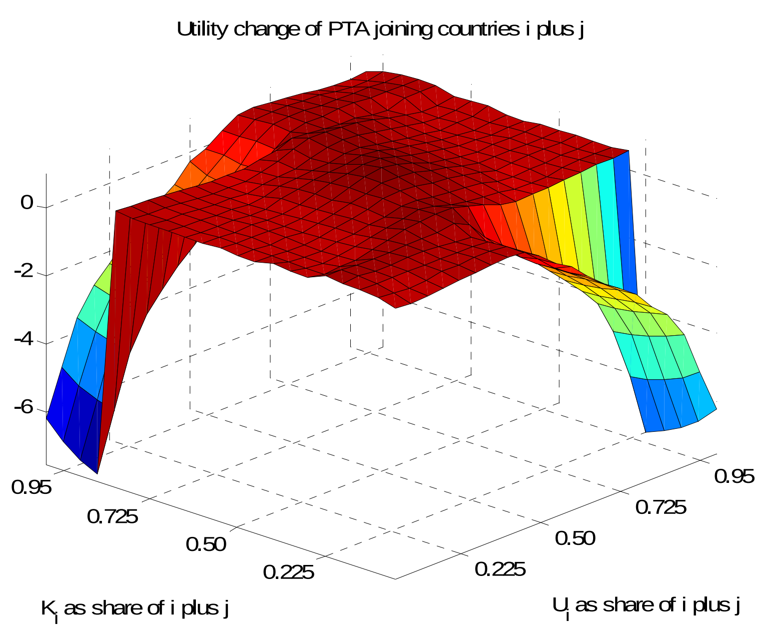


Figure 4c

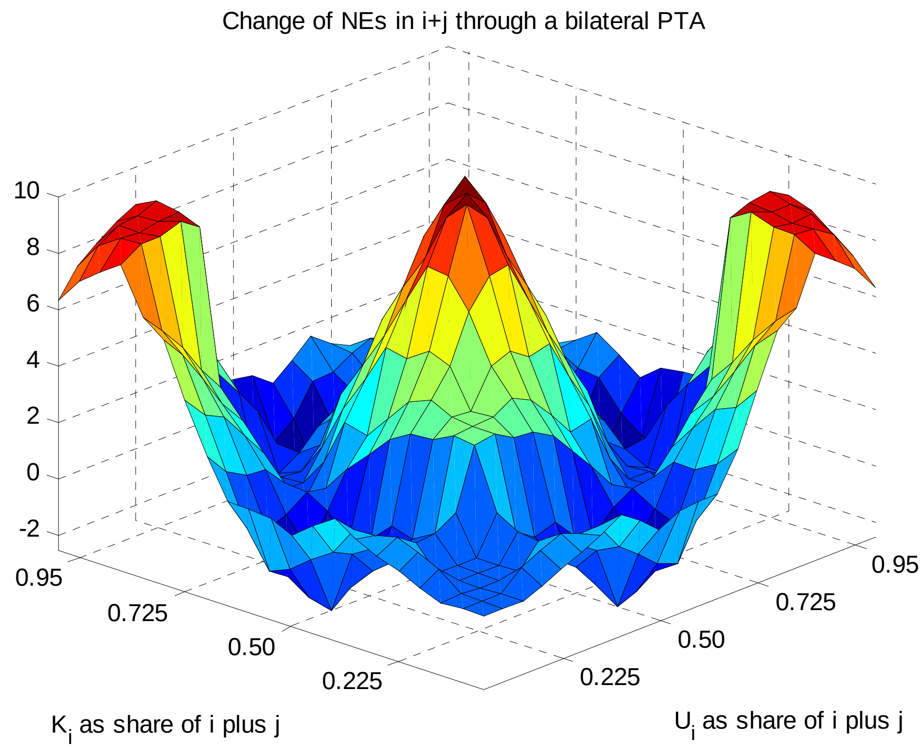


Figure 4d

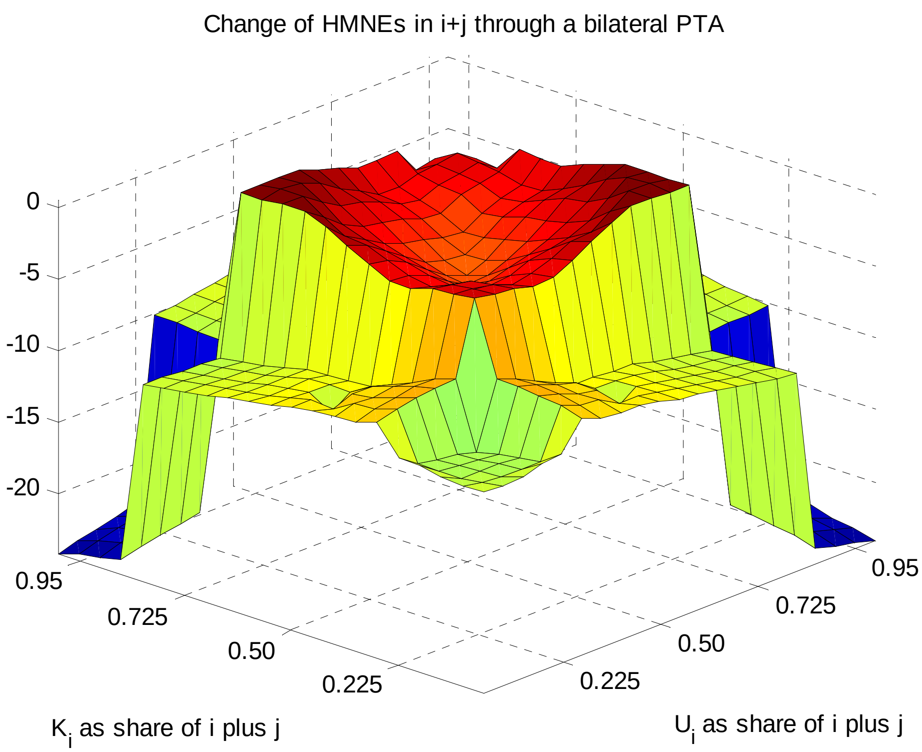
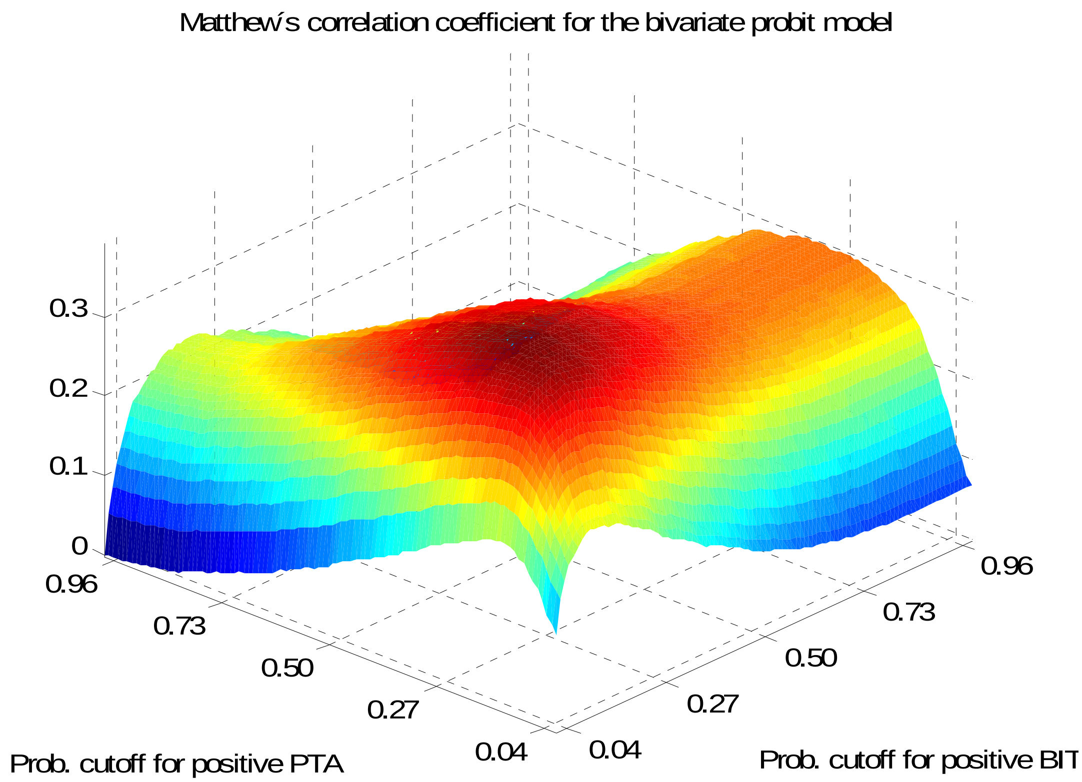


Figure 5



Matthew´s correlation coefficient for the bivariate probit model
0.3
0.2
0.1
0
0.96
0.73
0.50
0.27
0.04
0.04
0.27
0.50
0.73
0.96
Prob. cutoff for positive PTA
Prob. cutoff for positive BIT

Table 1 - Bilateral Investment Treaties (BITs) and Preferential Trade Agreements (PTAs) Across 12,880 Country-Pairs in Year 2000

		PTAs		
		Yes (1)	No (0)	Sum
BITs	Yes (1)	556	923	1479
	No (0)	1478	9923	11401
	Sum	2034	10846	12880

Notes: There are 161 countries (12,880 pairs) in the sample.

Table 2 - Summary statistics for Bilateral Investment Treaties (BITs) and Preferential Trade Agreements (PTAs) and their Key Determinants

		Unconditional		At BIT=1		At PTA=1	
Variable	Acronym	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Dependent variables							
Bilateral Investment Treaty indicator	BIT_{ij}	0.115	0.319	1.000	0.000		
Preferential Trade Agreement indicator	PTA_{ij}	0.158	0.365			1.000	0.000
Independent variables							
Log sum of i's and j's real GDP	$GDPSum_{ij}$	24.995	1.834	26.532	1.390	25.468	1.575
Log similarity of i's and j's real GDP	$GDPSim_{ij}$	-0.537	1.597	-0.793	1.790	-0.474	1.291
Log bilateral distance between i's and j's economic centers	$DIST_{ij}$	8.193	0.789	7.752	0.965	7.566	1.064
Adjacency indicator between i and j	ADJ_{ij}	0.021	0.142	0.045	0.207	0.084	0.277
Common official language indicator between i and j	$LANG_{ij}$	0.128	0.334	0.110	0.313	0.162	0.369
Political stability between i and j	$PolStab_{ij}$	72.875	11.764	80.010	9.473	73.907	11.673
Inverse expropriation risk between i and j	$IExpRisk_{ij}$	8.484	1.830	9.017	1.556	8.666	1.687
Log absolute difference in relative capital-unskilled labor ratios between i and j	$KURatio_{ij}$	2.349	1.713	2.010	1.536	1.615	1.231
Log absolute difference in capital and unskilled labor shares between i and j	$KUDiff_{ij}$	-2.727	2.388	-2.466	2.155	-2.431	2.027
Log similarity of i's and j's skilled labor endowment shares	$Ssim_{ij}$	-2.406	1.156	-2.390	1.149	-2.153	0.858
Log distance of i and j to the rest of the world	$REMOTE_{ij}$	8.436	0.153	8.385	0.163	8.403	0.180

Table 3 - The Determinants of Bilateral Investment Treaties (BITs) and Preferential Trade Agreements (PTAs) in Seemingly Unrelated Bivariate Probit Models: Main Results

		Model 1		Model 2		Model 3	
Determinants	Acronym	BITs	PTAs	BITs	PTAs	BITs	PTAs
Log sum of i's and j's real GDPs	$GDPSum_{ij}$	0.374 ***	0.147 ***	0.327 ***	0.160 ***	0.353 ***	0.203 ***
		(0.010)	(0.007)	(0.011)	(0.008)	(0.012)	(0.010)
Log similarity of i's and j's real GDPs	$GDPSim_{ij}$	0.035 ***	0.058 ***	0.029 ***	0.060 ***		
		(0.008)	(0.007)	(0.008)	(0.007)		
Log bilateral distance between i's and j's economic centers	$DIST_{ij}$	-0.529 ***	-0.617 ***	-0.509 ***	-0.614 ***	-0.475 ***	-0.542 ***
		(0.022)	(0.020)	0.022	(0.021)	(0.023)	(0.021)
Adjacency indicator between i and j	ADJ_{ij}	-0.341 ***	0.398 ***	-0.242 **	0.342 ***	-0.233 *	0.280 ***
		(0.116)	(0.099)	(0.123)	(0.100)	(0.126)	(0.101)
Common official language indicator between i and j	$LANG_{ij}$	-0.091 *	0.087 **	-0.087	0.048	-0.064	0.091 **
		(0.053)	(0.041)	(0.054)	(0.042)	(0.056)	(0.046)
Political stability between i and j	$PolStab_{ij}$	-	-	0.017 ***	-0.005 ***	0.015 ***	-0.007 ***
		-	-	(0.002)	(0.001)	(0.002)	(0.002)
Log absolute difference in capital and unskilled labor shares between i and j	$KUDiff_{ij}$	-	-	-	-	0.029 ***	-
		-	-	-	-	(0.009)	-
Log absolute difference in relative capital-unskilled labor ratios between i and j	$KURatio_{ij}$	-	-	-	-	-	-0.166 ***
		-	-	-	-	-	(0.010)
Log similarity of i's and j's skilled labor endowment shares	$Ssim_{ij}$	-	-	-	-	0.169 ***	0.226 ***
		-	-	-	-	(0.021)	(0.017)
Constant		-6.436 ***	0.296	-6.679 ***	0.334	-7.018 ***	-0.333
		(0.272)	(0.233)	(0.286)	(0.236)	(0.299)	(0.253)
Observations		12880		12561		11325	
Countries		161		159		151	
Correlation between disturbances in BITs and RTA processes		0.219		0.222		0.161	
Standard error of correlation coefficient above		0.023		0.024		0.025	
Log-likelihood of model		-8123.94		-7972.69		-7434.08	
Log-likelihood of constant-only model		-9968.14		-9968.14		-9312.86	
McFadden pseudo-R^2		0.185		0.200		0.202	

Notes: Standard errors in parentheses are robust to heteroskedasticity of unknown form. ***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively, using 2-tailed tests.

Table 4 - The Determinants of Bilateral Investment Treaties (BITs) and Preferential Trade Agreements (PTAs) in Seemingly Unrelated Bivariate Probits: A Sensitivity Analysis

		Model 4		Model 5		Model 6	
Determinants	Acronym	BITs	PTAs	BITs	PTAs	BITs	PTAs
Log sum of i's and j's real GDPs	$GDPSum_{ij}$	0.352 ***	0.205 ***	0.352 ***	0.209 ***	0.352 ***	0.210 ***
		(0.012)	(0.010)	(0.013)	(0.011)	(0.013)	(0.011)
Log similarity of i's and j's real GDPs	$GDPSim_{ij}$			0.000	0.017 *	0.000	0.022 **
				(0.009)	(0.010)	(0.009)	(0.010)
Log bilateral distance between i's and j's economic centers	$DIST_{ij}$	-0.475 ***	-0.544 ***	-0.475 ***	-0.543 ***	-0.467 ***	-0.662 ***
		(0.023)	(0.021)	(0.023)	(0.021)	(0.028)	(0.026)
Adjacency indicator between i and j	ADJ_{ij}	-0.237 *	0.278 ***	-0.237 *	0.280 ***	-0.226 *	0.139
		(0.126)	(0.101)	(0.126)	(0.101)	(0.128)	(0.104)
Common official language indicator between i and j	$LANG_{ij}$	-0.054	0.083 *	-0.054	0.084 *	-0.045	-0.009
		(0.056)	(0.046)	(0.056)	(0.046)	(0.058)	(0.049)
Political stability between i and j	$PolStab_{ij}$	0.017 ***	-0.009 ***	0.017 ***	-0.009 ***	0.017 ***	-0.010 ***
		(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Inverse expropriation risk between i and j	$IExpRisk_{ij}$	-0.024 *	0.020 *	-0.024 *	0.019 *	-0.024 *	0.024 **
		(0.013)	(0.011)	(0.013)	(0.011)	(0.013)	(0.011)
Log absolute difference in capital and unskilled labor shares between i and j	$KUDiff_{ij}$	0.031 ***	-	0.031 ***	-	0.031 ***	-
		(0.009)	-	(0.009)	-	(0.009)	-
Log absolute difference in relative capital-unskilled labor ratios between I and j	$KURatio_{ij}$	-	-0.164 ***	-	-0.162 ***	-	-0.147 ***
		-	(0.010)	-	(0.010)	-	(0.010)
Log similarity of i's and j's skilled labor endowment shares	$Ssim_{ij}$	0.172 ***	0.225 ***	0.172 ***	0.224 ***	0.172 ***	0.236 ***
		(0.021)	(0.017)	(0.021)	(0.016)	(0.021)	(0.017)
Log distance of i and j to the rest of the world	$REMOTE_{ij}$	-	-	-	-	-0.074	1.029 ***
		-	-	-	-	(0.136)	(0.122)
Constant		-6.933 ***	-0.395	-6.936 ***	-0.474 *	-6.386 ***	-8.216 ***
		(0.304)	(0.254)	(0.310)	(0.258)	(1.071)	(0.940)
Observations		11325		11325		11325	
Countries		151		151		151	
Correlation between disturbances in BITs and RTA processes		0.161		0.162		0.168	
Standard error of correlation coefficient above		0.025		0.025		0.025	
Log-likelihood of model		-7434.08		-7429.22		-7392.04	
Log-likelihood of constant-only model		-9312.86		-9312.86		-9312.86	
McFadden pseudo-R^2		0.202		0.202		0.206	

Notes: Standard errors in parentheses are robust to heteroskedasticity of unknown form. ***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively, using 2-tailed tests.

Table 5 - The Determinants of Bilateral Investment Treaties (BITs) and Preferential Trade Agreements (PTAs):
Further Sensitivity Analysis

		Model 7 (Bivariate Probit)		Model 8 (3SLS; Linear)		Model 9 (Recursive)	
Determinants	Acronym	AdjBITs	PTAs	BITs	PTAs	BITs	PTAs
Log sum of i's and j's real GDPs	$GDPSum_{ij}$	0.399 ***	0.209 ***	0.026 ***	0.266	0.311 ***	0.210 ***
		(0.013)	(0.011)	(0.005)	(0.350)	(0.013)	(0.011)
Log similarity of i's and j's real GDPs	$GDPSim_{ij}$	0.008	0.023 **	-0.004 *	0.004	-0.006	0.020 **
		(0.010)	(0.010)	(0.002)	(0.008)	(0.009)	(0.010)
Log bilateral distance between i's and j's economic centers	$DIST_{ij}$	-0.580 ***	-0.664 ***	0.054 ***	-0.508	-0.292 ***	-0.657 ***
		(0.028)	(0.026)	(0.019)	(0.507)	(0.034)	(0.026)
Adjacency indicator between i and j	ADJ_{ij}	-0.151	0.141	-0.137 *	-0.052	-0.285 **	0.166
		(0.121)	(0.104)	(0.030)	(0.284)	(0.131)	(0.109)
Common official language indicator between i and j	$LANG_{ij}$	-0.067	-0.010	-0.006 *	-0.047	-0.029	-0.010
		(0.059)	(0.049)	(-0.570)	(0.074)	(0.058)	(0.049)
Political stability between i and j	$PolStab_{ij}$	0.021 ***	-0.010 ***	0.005 ***	0.013	0.019 ***	-0.009 ***
		(0.002)	(0.002)	(0.000)	(0.022)	(0.002)	(0.002)
Inverse expropriation risk between i and j	$IExpRisk_{ij}$	-0.018	0.024 **	-0.010 **	-0.024	-0.032 **	0.026 *
		(0.013)	(0.011)	(0.003)	(0.045)	(0.013)	(0.011)
Log absolute difference in capital and unskilled labor shares between i and j	$KUDiff_{ij}$	0.037 ***	-	-0.005 *	-	0.022 **	-
		(0.009)	-	(0.002)	-	(0.009)	-
Log absolute difference in relative capital-unskilled labor ratios between i and j	$KURatio_{ij}$	-	-0.143 ***	-	-0.104	-	-0.156 ***
		-	(0.010)	-	(0.117)	-	(0.010)
Log similarity of i's and j's skilled labor endowment shares	$Ssim_{ij}$	0.199 ***	0.235 ***	-0.013 **	0.147	0.120 ***	0.240 ***
		(0.022)	(0.017)	(0.007)	(0.157)	(0.022)	(0.017)
Log distance of i and j to the rest of the world	$REMOTE_{ij}$	-0.080	1.036 ***	-0.272 ***	0.096	-0.368 ***	1.005 ***
		(0.138)	(0.121)	(0.044)	(0.322)	(0.137)	(0.122)
PTA between i and j				0.758 ***		0.765 ***	-
				(0.094)		(0.102)	-
BIT between i and j					-3.832	-	-
					5.997	-	-
Constant		-6.819	-8.264 ***	0.897 ***	-2.742	-4.668 ***	-8.036 ***
		(1.089)	(0.939)	(0.294)	(1.925)	(1.056)	(0.938)
Observations		11325		11325		11325	
Countries		151		151		151	
Correlation between disturbances in BITs and RTA processes		0.240		-		-0.243	
Standard error of correlation coefficient above		0.025		-		0.050	
Log-likelihood of model		-7264.955		-		-7379.27	
Log-likelihood of constant-only model		-9312.86		-		-9312.86	
McFadden pseudo-R^2 (in Model 7); R2 (in Models 8 and 9)		0.220		-0.286(BIT);-9.715(PTA)			

Notes: Standard errors in parentheses are robust to heteroskedasticity of unknown form. ***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively, using 2-tailed tests.

Table 6 - Impact of One-Standard-Deviation Change in the Determinants of BITs and PTAs on Marginal and Conditional Response Probabilities (Parameters are based on Model 6 in Table 4)

		Uncond. responses				Conditional responses			
Determinants	Acronym	BITs		PTAs		BITs|PTA=1		PTAs|BIT=1	
Log sum of i's and j's real GDP	$GDPSum_{ij}$	0.102	***	0.077	***	0.101	***	0.059	***
		(0.004)		(0.004)		(0.004)		(0.004)	
Log similarity of i's and j's real GDP	$GDPSim_{ij}$	0.000		0.007	**	-0.001		0.010	**
		(0.002)		(0.003)		(0.002)		(0.005)	
Log bilateral distance between i's and j's economic centers	$DIST_{ij}$	-0.058	***	-0.105	***	-0.080	***	-0.151	***
		(0.003)		(0.004)		(0.006)		(0.006)	
Adjacency indicator between i and j	ADJ_{ij}	-0.005	*	0.004		-0.013	*	0.009	
		(0.003)		(0.003)		(0.007)		(0.006)	
Common official language indicator between i and j	$LANG_{ij}$	-0.002		-0.001		-0.003		0.000	
		(0.003)		(0.003)		(0.004)		(0.004)	
Political stability between i and j	$PolStab_{ij}$	0.032	***	-0.023	***	0.043	***	-0.029	***
		(0.004)		(0.005)		(0.005)		(0.005)	
Inverse expropriation risk between i and j	$IExpRisk_{ij}$	-0.007	*	0.009	**	-0.009	**	0.011	**
		(0.004)		(0.004)		(0.004)		(0.004)	
Log absolute difference in capital and unskilled labor shares between i and j	$KUDiff_{ij}$	0.012	***			0.013	***		
		(0.003)				(0.004)			
Log absolute difference in relative capital-unskilled labor ratios between i and j	$KURatio_{ij}$			-0.051	***			-0.059	***
				(0.004)				(0.004)	
Log similarity of i's and j's skilled labor endowment shares	$Ssim_{ij}$	0.031	***	0.055	***	0.024	***	0.064	***
		(0.004)		(0.004)		(0.003)		(0.005)	
Log distance of i and j to the rest of the world	$REMOTE_{ij}$	-0.002		0.032	***	-0.007		0.044	***
		(0.003)		(0.004)		(0.005)		(0.005)	

Notes: Standard errors are in parentheses.

Table 7 - Predicting Agreements and Non-Agreements using Specification 6

Statistic (predictions) in % of true events	Chosen cutoff of predicted probabilities 0.5	E(BIT); E(PTA)
BITs if BIT=1	17	85
No BITs if BIT=0	85	62
Correct BIT successes and failures	77	65
PTAs if PTA=1	29	75
No PTAs if PTA=0	84	65
Correct PTA successes and failures	76	67