

Economic Integration Agreements and the Margins of International Trade*

Scott L. Baier[†], Jeffrey H. Bergstrand[‡], Michael Feng[§]

November 21, 2011

Abstract

One of the main potential welfare gains from trade liberalization is an increase in the variety of goods consumed by importing countries from various exporting countries. Combining methodologies established in Hummels and Klenow (2005) and Baier and Bergstrand (2007), this paper: (1) provides the first evidence using gravity equations of *both* the intensive and extensive (goods) margins being affected by economic integration agreements (EIAs) employing a panel data set with a large number of country-pairs, product categories, and EIAs from 1962-2000, (2) provides the first evidence of the differential average treatment effects (ATEs) of various “types” of EIAs on these intensive and extensive margins of trade, and (3) finds a novel differential “timing” of the two margins’ ATEs with intensive-margin effects occurring sooner than extensive-margin effects, consistent with recent theoretical predictions. These results are robust to correcting for Heckman sample-selection bias and for heterogeneous firms.

Key words: Free Trade Agreements; International Trade; Extensive Margins; Intensive Margins

JEL classification: F1; F15

* *Acknowledgements:* Acknowledgements will be added later.

[†] *Affiliation:* John E. Walker Department of Economics, Clemson University, Clemson, SC 29634 and Federal Reserve Bank of Atlanta, Atlanta, GA, USA. *E-mail:* sbaier@clemson.edu.

[‡] *Affiliation:* Department of Finance, Mendoza College of Business, and Kellogg Institute for International Studies, University of Notre Dame, Notre Dame, IN 46556 USA and CESifo, Munich, Germany. *E-mail:* bergstrand.1@nd.edu.

[§] *Affiliation:* John E. Walker Department of Economics, Clemson University, Clemson, SC 29634 USA. *E-mail:* fdayu@clemson.edu.

1 Introduction

The gravity equation has long dominated the international trade literature as the main econometric approach toward estimating *ex post* the “partial” average treatment effects (ATEs) of economic integration agreements and other natural and policy-based bilateral trade costs on aggregate bilateral trade flows.¹ Economic integration agreements (EIAs) refer broadly to preferential trade agreements, free trade agreements, customs unions, common markets, and economic unions.² Recently, Baier and Bergstrand (2007) demonstrated that estimation (*ex post*) of the partial effects of EIAs suffered from endogeneity bias, mainly due to self-selection. They showed that – after accounting for such bias using panel techniques – EIAs had much larger ATEs on trade flows than revealed in the earlier gravity equation literature and these estimates were more precise. Anderson and Yotov (2011) confirmed these findings using panel data also. Such results followed in the footsteps of empirical trade studies such as Trebler (1993) and Lee and Swagel (1997) that showed that previous estimates of trade-policy liberalizations on imports were underestimated considerably due to endogeneity bias.

While such positive ATE estimates for EIA dummy variables were interpreted in the context of either Armington or Krugman models as EIAs increasing trade volumes of existing homogeneous firms (the “intensive margin”), consideration of zeros in bilateral trade, fixed export costs, and firm heterogeneity have led researchers more recently to examine various “extensive margins” of trade. Such extensive margins fall under three general categories: country, goods, and firm. The existence of zeros in aggregate bilateral trade flows among many country-pairs has led some researchers to explore the probability that a pair of countries trades at all; to the extent that an EIA affects this probability, this changes the *country* extensive margin of trade and potentially economic welfare.

However, the few empirical studies to date using gravity equations for a large number of country-pairs and EIAs have led to two puzzling results. First, two studies have used cross sections for a single year. Helpman, Melitz, and Rubinstein (2008), or HMR, found evidence using a cross-section and a two-stage estimator that EIAs influenced the country extensive

¹Partial (or direct) effects refer to the absence of general-equilibrium (or indirect) effects; see Anderson and van Wincoop (2003, 2004) and Baier and Bergstrand (2009) on partial versus general equilibrium trade effects of trade-cost changes. “Aggregate” refers to all “goods” (or industries or product categories) in which country-pairs have trade flows.

²In this study, we use the term “preferential trade agreement” to denote one with only partial liberalization (not free trade).

margin, but had *no* significant effect on the intensive margin of trade (for existing firms). Egger, Larch, Staub, and Winkelmann (2011) explored the country extensive and intensive margins also using a cross-section and a two-stage estimator and found in contrast a significant positive effect of EIAs on the intensive margin in their preferred specification, but no extensive margin effect.³ The absence of country extensive-margin effects from an EIA suggests that trade liberalization does *not* lead to increases in varieties of goods from new trade partners, a potentially important source of welfare gains. The opposite EIA effect findings of the two papers is a puzzle.

A second margin is known as the “goods” margin of trade. Hummels and Klenow (2005), or HK, introduced this notion by examining zeros in bilateral trade flows at *highly disaggregated product-category levels*.⁴ The motivation for HK was to explore in a cross-section a fundamental question: Do large economies export more because they export larger quantities of a given good (intensive margin) *or* a wider set of goods (extensive margin)?⁵ They found in their cross-section that about 60 percent of larger exports of large economies was attributable to the extensive margin of “goods” trade; specifically, as the exporter country’s economic size grew, it exported a larger number of product categories (or “goods”) to more markets. The finding also that larger economies import more goods from more partners is important because it suggests an *improvement in welfare* owing to the consumption of more varieties; yet, 40 percent of the increased trade was still explained by the intensive margin in this cross-section. However, HK did not explore the relationship between trade liberalizations and the intensive and extensive goods margins of trade.

Only three empirical studies have explored the effect of trade liberalizations – and, in particular, EIAs – on the intensive and extensive goods margins of trade using the HK methodology. The earliest study using the HK decomposition to explore this issue is Hillberry and McDaniel (2002), focusing solely on the North American Free Trade Agreement (NAFTA). Although they do not attempt to establish causal effects from NAFTA to trade increases, they provide a decomposition of post-NAFTA trade among the three partners into goods intensive

³They also evaluated whether their results were biased by omitting firm-heterogeneity, but concluded that firm-heterogeneity had no significant effect (also in contrast with HMR). Their preferred specification accounted for endogeneity of EIAs.

⁴By contrast, both HMR and Egger, Larch, Staub, and Winkelmann (2011) used only bilateral *aggregate* trade flows to determine zeros in trade.

⁵Each “good” was a 6-digit SITC category. They also explored the effects of country size and per capita GDP on the quality of goods exported, as well as the two margins.

and extensive margins using 4-digit SITC data. They find evidence of both margins changing between 1993-2001. Kehoe and Ruhl (2009) examined NAFTA, the earlier Canada-US FTA trade liberalization, and some structural transformations using a modified-version of the HK decomposition methodology and applied to a series of cross-sections. Similar to Hillberry and McDaniel (2002), they do not conduct an econometric analysis trying to explain the effect of NAFTA (or the Canada-US FTA) on trade flows conditional on other variables. They decompose actual goods extensive- and intensive-margin changes post-agreement also using 4-digit SITC data for goods categories from Feenstra et al. (2005). They find significant evidence of both extensive and intensive margin changes using their modified HK decomposition methodology. Both studies' evidence of goods intensive and extensive margins of trade expanding following the signing of NAFTA suggests the *need* for a comprehensive *econometric* analysis (conditional on other covariates) of the effects of EIAs in general on the goods intensive and extensive margins of trade, in the spirit of HK's original analysis of the effect of country size and per capita GDP on the two goods' margins.⁶

Our paper extends this literature by offering three potential empirical contributions. First, we extend Baier and Bergstrand (2007)'s panel econometric methodology for the (partial) effects of EIAs on aggregate trade flows using a gravity equation to examine in a large country-pair setting the effects of virtually all EIAs on the extensive and intensive goods margins, using the HK trade-margin-decomposition methodology. In the context of an econometric analysis, we are the first to find economically and statistically significant EIA effects on both the intensive *and* extensive (goods) margins in the context of a large number of country-pairs, EIAs, and years, in contrast to HMR and Egger, Larch, Staub, and Winkelmann (2011).

Second, we allow for various *types* of EIAs – one-way preferential trade agreements (OWPTAs), two-way preferential trade agreements (TWPTAs), free trade agreements (FTAs), and a variable for customs unions, common markets and economic unions (CUCMECU) – and we decompose trade flows into extensive and intensive margins using the HK methodology (which is based on Feenstra, 1994).⁷ While two recent studies have adapted the Baier-Bergstrand methodology for estimating the effect of differing “types” of EIAs on bilateral aggregate trade

⁶A third study will be discussed shortly. Using a methodology similar to HK for estimating the goods extensive margin, Feenstra and Kee (2007) provided an econometric analysis of the effect of NAFTA on the extensive margin of Mexico's exports to the United States; they found a significant effect of NAFTA's reduction in tariff rates on this margin.

⁷Due to few observations on common markets and economic unions, we combine these two types of “deeper” EIAs with customs unions to form the variable CUCMECU, representing “deep” EIAs.

flows, *no econometric studies* have examined the effect of various types of EIAs on the (goods) extensive and intensive margins of trade using a large number of country-pairs and EIAs.⁸ Neither HMR nor Egger, Larch, Staub, and Winkelmann (2011) distinguished among various types of EIAs in their analyses of country intensive and extensive margins. We find not only that deeper EIAs have larger trade effects than FTAs, and the latter have larger effects than (partial) two-way and one-way PTAs, but we distinguish between these various trade effects at the extensive and intensive margins using a panel of (disaggregate) bilateral trade flows from 1962-2000 covering 98 percent of world exports.

Third, Bernard, Jensen, Redding and Schott (2009) is likely the only empirical study to date to explore the “timing” of extensive and intensive margin responses to shocks. Using cross-sectional variation to examine long-run aspects, Bernard, Jensen, Redding and Schott (2009) find that variation in trade flows across country pairs is explained largely by the extensive margin, using firm-level data (the “firm” margin); this result is consistent with HK using their “goods” margin. But using *time-series* variation, Bernard, Jensen, Redding and Schott (2009) find that a larger proportion of trade variation can be explained by the intensive margin at short (five-year) time intervals. They show that, following the Asian financial crisis of 1997, virtually all of the variation in trade flows within 2-3 years could be explained by the *intensive* margin. This finding is consistent with two recent theoretical studies arguing that the low trade-cost elasticity found in macroeconomic analyses of business cycles should be associated with the intensive margin of trade compared with the relatively higher trade-cost elasticity found in international trade, which reflects extensive margin effects.⁹ In this paper, we allow for differential “timing” of EIA effects using panel data. We find the first comprehensive empirical evidence that the shorter-term effects of EIAs on trade flows are more at the (goods) intensive margin and longer-term effects are more at the extensive margin (the latter entailing either fixed export costs or staggered “Calvo” pricing by consumers), consistent with intuition and results in Bernard, Jensen, Redding and Schott (2009).

The only study to our knowledge that like us uses a data set for a large number of country-pairs and years, a large number of EIAs, and the HK methodology is Foster, Poeschl, and

⁸The two studies that extended the Baier-Bergstrand framework to differing types of EIAs are Magee (2008) and Roy (2010); both found that customs unions had larger trade effects than FTAs. However, neither study examined extensive versus intensive margin issues.

⁹Ruhl (2008) explains the delayed effect of the extensive-margin effects to fixed export costs on the supply side, while Arkolakis, Eaton, and Kortum (2011) explain the delayed effect of the extensive-margin effects to “Calvo-pricing” by consumers on the demand side.

Stehrer (2011). However, the average treatment effect they found of an EIA on the goods extensive margin was an economically insignificant 10 percent, and they found no effect of EIAs on the intensive margin. The latter result is a second puzzle because it is existing exporters and importers that seek EIAs and theoretical studies just noted suggest that the shorter-term effect of EIAs should be on the intensive margin. Yet, we note several differences between our study and theirs. First, conceptually their estimation was based upon a traditional gravity-equation specification ignoring recent theoretical developments that emphasize the importance of relative price or “multilateral resistance” terms. Our paper is based upon state-of-the-art gravity-equation specifications, such as discussed in Arkolakis, Costinot, and Rodriguez-Clare (2011). Second, Foster, Poeschl, and Stehrer (2011) use a short three-year window on both sides of the EIA formation, and consequently can only capture *short-term* EIA effects; this likely explains their economically small treatment effects but does not explain finding only an extensive-margin effect. As shown in Baier and Bergstrand (2007), EIAs can take 10-15 years to have their full impact on aggregate bilateral trade flows. Moreover, by allowing longer lags, we can distinguish between short-term vs. longer-term effects. Third, Foster, Poeschl, and Stehrer (2011) examine the impact of EIAs using a single dummy variable; we use multiple EIA variables to distinguish the effects of one-way PTAs, two-way PTAs, FTAs, and deeper EIAs on aggregate trade flows, extensive margins, and intensive margins. Even disregarding the potential endogeneity biases introduced by their ignoring relative prices, their study did not distinguish between various “types” of EIAs and did not distinguish between the “timing” of intensive- and extensive-margin effects. Indeed, Foster, Poeschl, and Stehrer (2011) suggest in their concluding paragraph that examination of the shorter-run versus longer-run effects and accounting for the differing “depth and breadth” of EIAs would be useful extensions.

Finally, evidence over the last decade indicating that firms are *heterogeneous* in productivities suggests that trade liberalizations may potentially increase both the volume of trade of existing exporting firms as well as the number of exporting firms (the “firm extensive margin”), cf., Melitz (2003). Melitz (2003) showed that such extensive-margin effects are potentially important because they suggest another channel for the “gains from trade,” specifically, the increase of average productivities within countries.¹⁰ While there have been numerous country-specific trade liberalizations analyzing firm-level intensive- and extensive-margin effects using

¹⁰See Feenstra (2009) for a thorough discussion of the potential gains from trade in this class of models.

firm-level data sets, as noted above HMR and Egger, Larch, Staub and Winkelmann (2011) used aggregate trade level data and theory to estimate both the country *and firm* intensive-and-extensive-margin effects of trade liberalizations for a large number of country pairs and EIAs for a single cross-section. HMR use a two-stage instrumental-variables estimation procedure to account for non-random selection of countries into exporting (the country extensive margin) and allow for firm heterogeneity in productivities.¹¹ The purpose of the first stage is to employ bilateral *aggregate* trade flow observations to construct probit estimates of the probability of a positive bilateral aggregate trade flow; the probit estimates are used to generate two variables for the second stage that capture Heckman selection bias and firm heterogeneity. As noted earlier, the estimated coefficients in HMR for their Heckman selection and firm-heterogeneity variables and for their EIA dummy suggest that virtually *all* of the change in trade from an EIA is associated with the extensive margin and virtually none is associated with the firm intensive margin. Since our main results for goods extensive and intensive margins are potentially biased by omitting corrections for sample-selection bias and a measure of firm heterogeneity, in a robustness analysis we control for both using two-stage estimates. Our main results are robust to correcting for sample selection and firm heterogeneity.

The remainder of this paper is as follows. Section 2 discusses our methodology, based on the HK linear trade-margins-decomposition method and the Baier and Bergstrand (2007) approach for estimating ATEs of EIAs on trade flows in gravity frameworks. Section 3 discusses data and measurement issues. Section 4 provides the empirical results and discusses the three potential contributions. Section 5 concludes.

2 Methodology

2.1 Accounting for Endogenous EIAs: The Baier-Bergstrand Methodology

Baier and Bergstrand (2007), or BB, re-evaluated usage of the gravity equation econometrically for estimating partial ATEs of EIAs on pairs of countries' trade flows.¹² The first of two

¹¹As noted above, Egger, Larch, Staub, and Winkelmann (2011) use a similar methodology.

¹²Partial (or direct) effects ignore general-equilibrium (or indirect) effects, as discussed earlier. While techniques exist for estimating the indirect effects, such estimation is beyond this paper's scope, cf., Anderson (2011), Bergstrand and Egger (2011), and Egger, Larch, Staub, and Winkelmann (2011) on these issues.

main contributions was that self-selection of country-pairs into EIAs (cf., Baier and Bergstrand, 2004a) likely created a significant endogeneity bias in previous gravity-equation estimates of the ATEs of EIAs on trade flows. This is precisely the concern raised in Arkolakis, Costinot, and Rodriguez-Clare (2011, section V) for gravity-equation estimates of trade elasticities; the observed variable trade cost measure may be correlated with unobservable trade costs hidden in the gravity equation’s error term. The second main contribution of BB was that – given the slow-moving nature of EIAs’ determinations – gravity equation estimation could use panel techniques and data to avoid endogeneity bias and also capture lagged influences, incorporating bilateral fixed effects to capture time-invariant bilateral unobservable RHS variables as well as exporter-time and importer-time effects to capture time-varying unobservable “multilateral price/resistance” terms of the exporter and importer. BB showed that EIAs on average increased two members’ bilateral trade by approximately 100 percent after 10-15 years. Such a panel approach allows estimates of the “timing” of EIAs’ effects on trade flows between short-run and longer-run, as well as offers an alternative approach to instrumental variables using cross-sectional data (and potentially avoids possible shortcomings of the latter approach).

Motivated by the theoretical framework in Anderson and van Wincoop (2003), or AvW, which allows only intensive-margin trade impacts of trade-cost changes, BB studied the gravity equation:

$$\ln \left(\frac{X_{ijt}}{Y_{it}Y_{jt}} \right) = \beta_0 + \beta_1(\ln DIST_{ij}) + \beta_2(ADJ_{ij}) + \beta_3(LANG_{ij}) + \beta_4(EIA_{ijt}) \quad (1)$$

$$- \ln \Pi_{it}^{1-\sigma} - \ln P_{jt}^{1-\sigma} + \epsilon_{ijt},$$

where X_{ijt} is the value of the aggregate trade flow from country i to country j in year t , Y_{it} (Y_{jt}) is gross domestic product, or GDP, in country i (j) in year t , $DIST_{ij}$ is the bilateral distance between the economic centers of i and j , ADJ_{ij} is a dummy variable assuming the value 1 if the two countries share a common land border (and 0 otherwise), $LANG_{ij}$ is a dummy assuming the value 1 if the two countries share a common language, EIA_{ijt} is a dummy assuming the value 1 if the two countries have an EIA in year t , and $\ln \Pi_{it}^{1-\sigma}$ ($\ln P_{jt}^{1-\sigma}$) is exporter i ’s (importer j ’s) nonlinear and unobservable multilateral price/resistance term (as defined in A-vW, with σ the elasticity of substitution in consumption) – a function of all countries’ unobservable producers’ prices – and ϵ_{ijt} is a normally distributed error term.

BB raised the following concerns about estimating the ATEs of EIAs (β_4) in gravity equa-

tions such as equation (1), where the effect of an EIA between i and j is to cause a reduction in unobservable bilateral trade costs. Suppose that bilateral trade costs are a function of numerous bilateral time-invariant variables and a dummy variable representing the presence or absence of an EIA (EIA_{ijt}). A standard problem in empirical work is the potential endogeneity of RHS variables. If any of the RHS variables in this linear equation (1) is correlated with the error term ϵ_{ijt} , that variable is considered econometrically endogenous and ordinary least squares (OLS) may yield biased and inconsistent coefficient estimates. While omitted variables, simultaneity, and measurement error are all potential sources of endogeneity bias, BB argued that the most important source of endogeneity bias was self-selection of country pairs into EIAs. For instance, error term ϵ_{ijt} in equation (1) may be representing unobservable (to the econometrician) policy-related barriers – tending to reduce X_{ijt} – that are not accounted for in the RHS variables in equation (1) but may be correlated with the decision to form an EIA. As an example, suppose two countries i and j have extensive unmeasurable domestic regulations (e.g., internal shipping regulations) that inhibit trade (causing ϵ_{ijt} to be negative). The likelihood of the two countries’ governments selecting into an EIA may be high if there is a large expected welfare gain from potential bilateral trade creation if the EIA deepens liberalization beyond tariff barriers into domestic regulations (and other non-tariff barriers). Thus, EIA_{ijt} and the intensity of domestic regulations may be positively correlated in a cross-section of data, but the gravity equation error term ϵ_{ijt} and the intensity of regulations may be negatively correlated. This reason suggests that EIA_{ijt} and ϵ_{ijt} are negatively correlated, and the EIA coefficient will tend to be underestimated. Numerous authors have noted that one of the major benefits of regionalism is the potential for deeper integration. Lawrence (1996) notes that “Once tariffs are removed, complex problems remain because of differing regulatory policies among nations” (p. 7). He argues that “agreements are also meant to achieve deeper integration” (p. 7). Gilpin (2000) notes “... the inability to agree on international rules or to increase international cooperation in this area has contributed to the development of both managed trade *and regional agreements*” (p. 108; italics added).¹³

BB also discussed that there are primarily two methods to address the potential endogeneity bias using parametric techniques, instrumental variables (IV) using cross-section data and alternatively panel techniques. We address each in turn. Cross-section estimation, such as in

¹³For discussions on how omitted variables, simultaneity, and measurement error can bias the EIA coefficient estimate, see BB, section 3.

HMR and Egger, Larch, Staub, and Winkelmann (2011), allows estimates of long-run effects.¹⁴ As discussed in Baier and Bergstrand (2002, 2004b, 2007), in the case of endogenous EIAs, the first stage of the estimation procedure would involve estimating a binary response model to generate predicted probabilities, Φ^P , of *EIA*. The second stage is a linear regression of the *EIA* binary variable on a constant, Φ^P , a vector of gravity variables \mathbf{q} such as in (1) (excluding *EIA*), and a vector of identifying variables \mathbf{z} which are not in \mathbf{q} . The third stage is the (instrumental-variables, or IV) estimation of gravity equation (1) using the predicted values from the second-stage regression. This IV estimator is consistent and asymptotically efficient, and the usual two-stage least squares standard errors and test statistics are asymptotically valid, cf., Wooldridge (2002).

As argued in BB, the problem with using cross-section data and consequently having to employ IV techniques to account for *EIA* selection bias is the inability practically of satisfying the “exclusion restriction” with confidence. As noted in Baier and Bergstrand (2004b, 2007), estimating average treatment effects (ATEs) of EIAs using IV or Heckman control functions yield unstable findings. Most variables that influence trade flows also explain selection into EIAs, and it is difficult to find a variable that explains EIAs (in \mathbf{z}) that does not also explain trade flows (in \mathbf{q}).¹⁵

Given the problems associated with accounting for endogeneity of EIAs using instrumental variables and cross-section data, BB argued that a better approach to eliminate endogeneity bias of EIAs is to use panel techniques. In the context of the theory and endogenous self-selection of country-pairs into EIAs, BB argue that one method to obtain consistent estimates of the ATE of EIAs is by fixed effects estimation of:

$$\ln X_{ijt} = \beta_0 + \beta_1 EIA_{ijt} + \eta_{ij} + \delta_{it} + \psi_{jt} + \epsilon_{ijt} \quad (2)$$

where η_{ij} is a country-pair fixed effect to capture all time-invariant unobservable bilateral

¹⁴We ignore momentarily the “zeros” issue raised in both studies, which introduces another layer in the estimation.

¹⁵Egger, Larch, Staub and Winkelmann (2011) used the IV approach to account for the endogeneity of EIAs in their single cross-section, also allowing for selection into zeros trade (i.e., a bivariate probit model). They found using an approach similar to HMR, except also allowing for endogenous EIAs, that EIAs predominantly affected trade at the intensive margin. However, it is important to note that they used bivariate probits on aggregate bilateral trade flows to estimate first the country extensive margin, and then used predicted probits to adjust for Heckman selection and heterogeneous firms in the second-stage gravity equation. Hence, their approach differs from ours in terms of which extensive margin was explored and using a cross-section approach, as in HMR. We address the zeros issue and firm-heterogeneity later in a sensitivity analysis.

factors influencing nominal trade flows and δ_{it} and ψ_{jt} are exporter-time and importer-time fixed effects, respectively, to capture time-varying exporter and importer GDPs as well as all other time-varying country-specific unobservables in i and j influencing trade, *including* the exporter’s and importers’ “multilateral price/resistance” terms (cf., Anderson and Wincoop, 2003). We refer to this as the Fixed-Effects (FE) specification. It is important to note that, in most gravity-equation applications using a comprehensive set of RHS variables, the vast bulk of “bilateral” trade-cost variables are time-invariant, such as bilateral distance, common border, common language, etc. For instance, in HMR the only time-varying bilateral trade-cost variables are their EIA dummy and a dummy for the presence or absence of a currency union. As explained in BB, using panel data ATEs of EIAs (represented by β_1) estimated using equation (2) are likely to be consistent and precise. BB showed that the ATE of the typical EIA on nominal trade flows was about 0.76, implying that the typical EIA increased bilateral trade by about 114 percent after 10-15 years.

BB also employed an alternative specification using first-differencing:

$$\Delta \ln X_{ij,t-(t-5)} = \beta_0 + \beta_1 \Delta EIA_{ij,t-(t-5)} + \delta_{i,t-(t-5)} + \psi_{j,t-(t-5)} + v_{ij,t-(t-5)} \quad (3)$$

We refer to this as the First-Difference (FD) specification. Note that the bilateral country-pair fixed effects are eliminated; however, the exporter-time and importer-time fixed effects are retained to capture time-varying exporter and importer GDPs and multilateral price terms. The latter effects were ignored in Foster, Poeschl, and Stehrer (2011), creating potential omitted variables bias.

Standard econometric discussions of treating endogeneity bias using panel data focus on a choice between estimation using the FE and FD specifications, cf., Wooldridge (2002, Ch. 10). As Wooldridge notes, when the number of time periods (T) exceeds two, the FE estimator is more efficient under the assumption of serially uncorrelated error terms ϵ_{ijt} . The FD estimator is more efficient (when $T > 2$) under the assumption that the error term ϵ_{ijt} follows a random walk (i.e., that the error term $v_{ij,t-(t-5)} = \epsilon_{ijt} - \epsilon_{ij,t-5}$ is white noise).¹⁶

First-differencing the panel data yields some potential advantages over fixed effects. First, it is quite plausible that the unobserved factors influencing the likelihood of an EIA (say, trade

¹⁶When the number of time periods is exactly two ($T=2$), estimation with FE and FD produce identical estimates and inferences; then, FD is easier to estimate. When $T > 2$, the choice depends upon the assumption one wants to make about the distribution of the error term ϵ_{ijt} .

below its “natural” level) are likely slow moving and hence serially correlated. If the ϵ_{ijt} are highly serially correlated, the inefficiency of FE is exacerbated as T gets large. This suggests that differencing the data will increase estimation efficiency for our large- T panel. Second, aggregate trade flow data and real GDP data are likely “close to” unit-root processes. Using FE is equivalent to differencing data around the mean (in our sample, year 1980); this may create a problem since T is large in our panel. As Wooldridge (2000, p. 447) notes, if the data follow unit-root processes and T is large, the “spurious regression problem” can arise in a panel using FE. FD yields data that deviates from the previous period of our panel, and thus is closer to a unit-root process. In the following, we estimate using both the FD and FE approaches, but for the reasons just stated we place emphasis on the FD specifications and place most of the FE estimates in our appendix.

One of the other potential contribution’s of BB’s panel methodology was to show that the full impact of EIAs on trade flows took 10-15 years. One reason is that most EIAs are “phased-in” over 5-10 years. The second reason is the lagged effect of the trade-cost changes (such as terms-of-trade changes) on trade flows. As in BB, using a panel allows for differentiating the shorter-terms effects (5 years) from the longer-term effects (5-15 years). In the context of the recent developments in the trade literature emphasizing intensive versus extensive margin effects, our panel approach allows for differential timing of these effects. In reality, one would expect that the intensive margin would be affected by a trade-cost change sooner than the extensive margin, because intensive margin changes in volumes do not require any startup costs. Such costs – critical to the extensive margin – may delay the entry of new firms into exporting, and thus we should expect the intensive margin to be influenced in the shorter term and the extensive margin in the longer term, as the results in Bernard, Jensen, Redding, and Schott (2009) show. Our panel data approach allows for evaluating this hypothesis.¹⁷

BB did not estimate differential effects of various types of EIAs (in terms of depth of integration) on trade flows. Magee (2008) and Roy (2010) using the methodology of BB found

¹⁷These differential timing effects were ignored in Foster, Poeschl, and Stehrer (2011). As discussed earlier, two recent theoretical papers suggest a reason for the low trade-cost elasticity of trade flows in macroeconomic analyses using time-series data and the relatively higher trade-cost elasticities of trade in cross-sectional trade analyses. Ruhl (2008) explains this puzzle by noting that the macroeconomic time-series approach is estimating the intensive-margin effect of trade, whereas the trade literature’s cross-sectional approach is capturing the extensive-margin effect, due to export fixed costs for new producers delaying trade effects and entry. In a complementary approach, Arkolakis, Eaton, and Kortum (2011) present a demand-oriented staggered-adjustment “Calvo-pricing” approach to explain the lower time-series elasticity in terms of solely an intensive-margin effect, and the higher long-run cross-section trade-cost elasticity capturing the longer-term extensive-margin elasticity.

that trade flows were impacted by larger amounts for customs unions relative to FTAs. However, *no empirical study* has examined the differential impact of FTAs relative to deeper EIAs on the extensive *versus* intensive margins – much less the differential *timing* of such effects; these are goals of this paper.¹⁸ The next section discusses how we decompose data into the two margins.

2.2 The Hummels-Klenow Margin-Decomposition Methodology

Hummels and Klenow (2005), or HK, was the first paper to highlight a tractable method for decomposing transparently the extensive and intensive *goods* margins of trade for a large set of countries’ bilateral trade flows using publicly available disaggregate trade data.¹⁹

Let X_{ijt} denote the value of country i ’s exports to country j in year t . Following HK, the extensive margin of goods exported from i to j in any year t is defined as:

$$EM_{ijt} = \frac{\sum_{m \in M_{ijt}} X_{Wjt}^m}{\sum_{m \in M_{Wjt}} X_{Wjt}^m} \quad (4)$$

where X_{Wjt}^m is the value of country j ’s imports from the world in product m in year t , M_{Wjt} is the set of all products exported by the world to j in year t , and M_{ijt} is the subset of all products exported from i to j in year t . Hence, EM_{ijt} is a measure of the fraction of all products that are exported from i to j in year t , where the each product is weighted by the importance of that product in world exports to j in year t .²⁰

¹⁸It is useful to note here a parallel literature examining the effect of GATT and/or WTO membership on trade flows. For brevity, we note that there now appears little convincing evidence of substantive GATT/WTO effects on trade, once one accounts for EIA dummies, multilateral resistance, and unobserved country-pair fixed effects (as we do here). This is the conclusion of Eicher and Henn (2011) (though they found a non-trivial WTO “terms-of-trade” effect) and of Felbermayr and Kohler (2010) who examined possible extensive-margin effects; Eicher and Henn (2011) ignored extensive- versus intensive-margin effects. We also note an issue raised in Martin and Ng (2004), which is the role of multilateral tariff reductions under the GATT/WTO. Most-Favored-Nation (MFN) tariff cuts could also be affecting results. However, such MFN tariff cuts by country would be accounted for by the exporter-time and importer-time fixed effects.

¹⁹Studies have also used country-specific data on individual plants (or firms) to study the extensive and intensive *firm* margins of trade liberalization, but such studies have necessarily been confined to *particular countries* because such data is widely known to be much more costly to access and such data sets have not been concurred for international comparisons, as noted in HMR. See Eaton, Kortum, and Kramarz (2008) for a study of French firms, Treffer (2004) for a study of Canada and the United States, and Pavcnik (2002) for a study of Indian firms. Another relevant theoretical and empirical piece with similar overtones is Arkolakis, Demidov, Klenow, and Rodriguez-Clare (2008).

²⁰Alternatively, one could use an unweighted average, which would then be simply the fraction of all products exported from i to j . However, HK – as well as researchers since then – use the weighted average. A weighted average seems more appropriate since cars and pencils do not have the same values in trade. Also, since we will use a time series of cross sections, we will consider later two alternative methods for fixing the trade-share weights over time.

HK define the intensive margin of goods exported from i to j as:

$$IM_{ijt} = \frac{\sum_{m \in M_{ijt}} X_{ijt}^m}{\sum_{m \in M_{ijt}} X_{Wjt}^m} \quad (5)$$

where X_{ijt}^m is the value of exports from i to j in product m in year t . Thus, IM_{ijt} represents the market share of country i in country j 's imports from the world within the set of products that i exports to j in year t .

One of the notable properties of the HK decomposition methodology is that the product of the two margins equals the ratio of exports from i to j relative to country j total imports:

$$EM_{ijt}IM_{ijt} = \frac{\sum_{m \in M_{ijt}} X_{ijmt}}{\sum_{m \in M_{Wjt}} X_{Wjmt}} = X_{ijt}/X_{jt} \quad (6)$$

where X_{jt} denotes j 's imports from the world. Taking the natural logs of equation (6) and some algebra yields:

$$\ln X_{ijt} = \ln EM_{ijt} + \ln IM_{ijt} + \ln X_{jt}. \quad (7)$$

Consequently, the HK decomposition methodology yields that the log of the value of the trade flow from i to j in any year t can be decomposed *linearly* into (logs of) an extensive margin, an intensive margin, and the value of j 's imports from the world. We note three issues regarding the HK methodology. First, since we will apply empirically FE and FD specifications similar to equations (2) and (3), the term $\ln X_{jt}$ will be subsumed in the importer-time fixed effect ψ_{jt} . Second, HK applied their methodology to only a cross section. By contrast, we are applying it to a time series of cross sections. Consequently, the trade weights used in constructing EX_{ijt} and IM_{ijt} will likely vary from year to year. To address this, we will also consider later in a sensitivity analysis fixed-year trade-share weights and also a chain-weighting technique. Third, there are numerous zeros in the variables in equation (7). Consequently, we will also examine whether the results are sensitive to sample selection bias, in the spirit of HMR.²¹

Because of the various extensive ‘‘margins’’ that have surfaced in the trade literature in

²¹Besedes and Prusa (2011) emphasize that extensive-margin (intensive-margin) effects may be overstated (understated) in examining effects of liberalizations using panel data, due to ‘‘survival’’ issues. Addressing this issue is beyond the scope of this particular paper. However, our results may not be biased excessively by this issue since we find material intensive-margin effects from EIAs, unlike Foster et al. (2011).

the last decade, it is important to distinguish between them. In particular, we want to be clear about the distinction between the “goods” (or HK) margin that we focus on versus the widely discussed “firm” margin, as addressed in Melitz (2003), HMR, and Redding (2011); we will also refer to the firm margin as the HMR margin. HK themselves noted that the relationship between the goods extensive and intensive margins with economic size and per capita incomes could likely be interpreted as influenced by the interaction of export fixed costs and firm heterogeneity. To fix ideas, we focus on two points. First, as Eaton and Kortum (2002), Arkolakis, Costinot, and Rodriguez-Clare (2011), and others have noted, there is a wide class of quantitative trade models (Armington, Krugman, Ricardian, Melitz) that yield iso-morphic gravity equations. Following Arkolakis, Costinot, and Rodriguez-clare (2011), under very few assumptions one can generate a standard gravity equation:

$$X_{ijt}^m = N_{it}^m Y_{jt}^m \left(\frac{(a_{Lit}^m)^{-\gamma^m} w_{it}^{-\gamma^m} \tau_{ijt}^{-\gamma^m} f_{ijt}^{-[\gamma^m/(\sigma^m-1)-1]}}{\sum_{k=1}^K N_{kt}^m (a_{Lkt}^m)^{-\gamma^m} w_{kt}^{-\gamma^m} \tau_{kjt}^{-\gamma^m} f_{kjt}^{-[\gamma^m/(\sigma^m-1)-1]}} \right) \quad (8)$$

where X_{ijt}^m is the trade flow from i to j in year t in “good” m , N_{it}^m is the number of firms in i (exporting and non-exporting) that produce output in good m , Y_{jt}^m is the expenditure in j on good m , a_{Lit}^m is the lower bound of the Pareto distribution of productivities in m in i , γ^m is an index of productivity heterogeneity among firms in good m , w_{it} is the wage rate in i , τ_{ijt} is variable trade costs of i products into j , f_{ijt} is fixed export costs from i to j , σ^m is the elasticity of substitution in consumption, and $\gamma^m > \sigma^m - 1$. Note that the relative price term in large parentheses is a standard representation of relative prices in the gravity equation, but now also reflecting productivity heterogeneity (through γ^m) and fixed exporting costs (f_{ijt}), cf., Melitz (2003), Chaney (2008), Redding (2011), and Arkolakis, Costinot, and Rodriguez-Clare (2011). In the context of these models, variable trade costs, τ_{ijt} , affect X_{ijt}^m via both the intensive and extensive margins. As Chaney (2008) demonstrates in his Melitz-type model, $\gamma^m = (\sigma^m - 1) + [\gamma^m - (\sigma^m - 1)]$. $\sigma^m - 1$ represents the intensive-margin elasticity of variable trade costs whereas $\gamma^m - (\sigma^m - 1)$ is the extensive-margin elasticity of variable trade costs. For finite means and variances in the theory, $\gamma^m/(\sigma^m - 1)$ must lie between 1 and 2. Hence, these models suggest that the intensive-margin variable-trade-cost elasticity should be larger than the extensive-margin variable-trade-cost elasticity. However, f_{ijt} works entirely through the extensive margin, so that there is no clear theoretical hypothesis for the relative sizes of

intensive- and extensive-margin effects of a given EIA formation.²²

Second, HMR (p. 451) helps to clarify changes that affect the bilateral goods extensive and intensive margins versus the firm extensive and intensive margins. It is possible for no firm in good m in country i to be productive enough to export to country j ; consequently, there would be zero trade in good m from i to j . In the HK extensive margin, an EIA may cause τ_{ijt} and/or f_{ijt} to fall sufficiently so that at least one firm's productivity is sufficiently high that positive exports occur from i to j in m following the trade liberalization, tending to increase the HK extensive margin. The HK intensive margin may increase as well. However, suppose that trade from i to j in m was already positive. The fall in variable and fixed trade costs from the EIA may cause the number of firms in country i in good m to expand because they cross the productivity threshold for exporting. This would tend to increase the "firm" extensive margin, but not increase the "good" extensive margin. In this case, changes in the HK intensive margin could be composed of increases in trade from existing exporting firms *as well as* increases in the number of (heterogeneous) firms exporting from i to j in good m . Thus, HK intensive margin changes can incorporate firm-level extensive-margin changes alongside firm-level intensive margin changes. We discuss these issues in more analytical detail in the Appendix, noting now that in our empirical sensitivity analysis later we will account for heterogeneous firms.

Finally, we are distinguishing between various types of EIAs and allow for lagged effects. Hence, the actual specification for the FD versions of our model with no lags is:

$$\begin{aligned} \Delta \ln X_{ij,t-(t-5)} &= \beta_0 + \beta_1 \Delta CUCMECU_{ij,t-(t-5)} + \beta_2 \Delta FTA_{ij,t-(t-5)} \\ &+ \beta_3 \Delta TWPTA_{ij,t-(t-5)} + \beta_4 \Delta OWPTA_{ij,t-(t-5)} + \delta_{i,t-(t-5)} + \psi_{j,t-(t-5)} + v_{ij,t-(t-5)} \end{aligned} \quad (9)$$

$$\begin{aligned} \Delta \ln EM_{ij,t-(t-5)} &= \theta_0 + \theta_1 \Delta CUCMECU_{ij,t-(t-5)} + \theta_2 \Delta FTA_{ij,t-(t-5)} \\ &+ \theta_3 \Delta TWPTA_{ij,t-(t-5)} + \theta_4 \Delta OWPTA_{ij,t-(t-5)} + \delta_{i,t-(t-5)} + \psi_{j,t-(t-5)} + v_{ij,t-(t-5)} \end{aligned} \quad (10)$$

²²The key assumption pertains to the effect of an EIA on fixed costs versus variable costs. Suppose $\ln \tau_{ijt} = \rho_0 - \rho_1 \ln DIST_{ij} - \rho_2 EIA_{ijt} + \mu_{ijt}$, as is conventional (μ_{ijt} denoting a random error term). Suppose also $\ln f_{ijt} = \delta_0 - \delta_1 \ln DIST_{ij} - \delta_2 EIA_{ijt} + \phi_{ijt}$ (ϕ_{ijt} denoting a random error term). If $\delta_2 > \rho_2$ by a large amount, the effect of f_{ijt} changes on the extensive margin may be sufficient to cause the extensive margin trade elasticity to exceed the intensive margin trade elasticity. To date, to our knowledge, there are no firm estimates of δ_2 because there is no data on $\ln f_{ijt}$.

$$\begin{aligned} \Delta \ln IM_{ij,t-(t-5)} &= \lambda_0 + \lambda_1 \Delta CUCMECU_{ij,t-(t-5)} + \lambda_2 \Delta FTA_{ij,t-(t-5)} \\ &+ \lambda_3 \Delta TWPTA_{ij,t-(t-5)} + \lambda_4 \Delta OWPTA_{ij,t-(t-5)} + \delta_{i,t-(t-5)} + \psi_{j,t-(t-5)} + v_{ij,t-(t-5)} \end{aligned} \quad (11)$$

Analogous specifications hold for the FE specifications and for specifications with lags, omitted here for brevity. Consequently, the literature to date suggests that the endogeneity of EIAs in typical gravity equations may bias the ATEs of such EIAs. We pursue the panel approach in BB to address this issue. In this paper, we follow the methodology established in HK to decompose trade flows (log-)linearly into extensive and intensive margins. In the next section, we use the methodology in BB to estimate equations such as (9)-(11) above to determine if both intensive and extensive (goods) margins are affected by EIAs, the differential effects of various “types” of EIAs on these extensive and intensive margins of trade, and the differential “timing” of the ATEs of EIAs on the two margins associated with the relative sizes of the current and lagged changes of the EIA dummies.

3 Data

The two key variables for our empirical analysis are disaggregate bilateral trade flows and a multichotomous index of the level of economic integration agreement (EIA) between a large number of country-pairs for a large number of years; a few other peripheral variables will be discussed shortly.

First, while several earlier gravity-equation analyses have used dummy variables indicating the presence or absence of an EIA between country-pairs for numerous years, such as Rose (2004), there are few publicly available systematic data sets that have multichotomous indexes of EIAs for a large number of country-pairs and number of years (a panel). We use the data set constructed by Scott Baier and Jeffrey Bergstrand and provided at Jeffrey Bergstrand’s website, www.nd.edu/~jbergstr/.²³ The index is defined as: no EIA (0), one-way preferential trade agreement, or *OWPTA* (1), two-way preferential trade agreement *TWPTA* (2), free trade agreement, or *FTA* (3), customs union (4), common market (5), and economic union (6). The definitions are conventional, based upon Frankel (1997), and are defined explicitly in the data

²³This data set was constructed under National Science Foundation grants SES-0351018 and SES-0351154 and includes annually from 1960-2005 for the pairings of 195 countries an index ranging from 0-6 of the level of any EIA between the pair.

set. Because of the small number of “deeper EIAs,” we combined customs unions (4), common markets (5), and economic unions (6) into one variable, *CUCMECU*. One of the strengths of the Baier-Bergstrand EIA panel is, for 98.6 percent of the cells where the EIA status of the country-pair changes (from 0 to 1, 0 to 3, 2 to 3, etc.), there exists a hyperlink to a copy (PDF format) of the original treaty. Because of the extensive number of bilateral and plurilateral EIAs of differing degrees of depth, the difficulty of listing all such agreements precludes a table; we recommend downloading the zip file from the web site.

Second, annual bilateral trade flows for 1962-2000 are from the NBER-United Nations trade data set at www.nber.org/data and documented in Feestra et al. (2005). The data are organized by 4-digit Standard Industrial Trade Classification (SITC), Revision 2. It covers trade flows reported by 149 countries and covers 98 percent of world exports. This is the most disaggregated publicly available data set for bilateral trade flows for a large number of years and a large number of country pairs, constructed on a consistent basis, necessary for the analysis at hand. This 4-digit SITC level data was also used in Hillberry and McDaniel (2002), Kehoe and Ruhl (2009), and Foster, Poeschl, and Stehrer (2011). In 1962, this resulted in 969 categories of “goods”; in 2000, this resulted in 1,289 categories of goods. One concern is that the level of disaggregation is not high enough, biasing results toward the intensive margin, as discussed transparently in HK; as one aggregates up, there become fewer categories in which a country does not export. Other studies have used higher levels of disaggregation to calculate trade flows in a good to avoid this (intensive-margin) aggregation bias. For instance, Broda and Weinstein (2006) used 7-digit data for the years 1972-1988 and then 10-digit data for the years 1990-2001. However, this was *U.S.* import data only, and for two short panels. Constraining ourselves to *U.S.* import trade flows only is problematic because the United States has only a small number of FTAs and has *no deep EIAs*, precluding evaluating our hypotheses. Since our decomposition of the extensive and intensive margins is based upon HK, we compare our data set to theirs. First, HK used only a cross-section of UNCTAD TRAINS data for the year 1995, but for a large number of country pairs like here.²⁴ Using a cross-section is problematic because we are relying upon panel techniques to avoid endogeneity bias. Second, HK used data at the Harmonized System 6-digit classification code; this yielded 5,017 goods categories, five times our number of categories. However, HK examined the levels of extensive margins

²⁴For robustness, HK also examined a cross-section of *U.S.* imports only at a higher level of disaggregation. But this raises the same problem as in Broda and Weinstein (2006) of *U.S.* data only.

and the correlations between the extensive margin and factors influencing the extensive margin in cross-sectional data – GDP, employment (L), GDP/L – at various levels of aggregation (1-digit, 2-digit, ..., 6-digit) and found two interesting results. As expected, extensive margins were much lower, and the correlation between the extensive margin of trade and determinants of it (such as GDP) were lower, as data became more aggregated. For example, at the 6-digit level, 62 percent of the GDP elasticity of exports could be explained by the extensive margin, whereas at the 1-digit level only 11 percent of the elasticity could be explained by the extensive margin. However, interestingly, at the 4-digit level, a sizable 54 percent of this elasticity could *still* be explained by the extensive margin. For the per capita GDP elasticity of exports, the 4-digit and 6-digit shares of the extensive margin were 62 and 64 percent, respectively. These results suggest that the (intensive-margin) aggregation bias may not be that severe in using 4-digit categories, rather than the 6-digit categories in HK.

While the focus of our empirical work will be on the FE and FD specifications described earlier, we will initially report some standard gravity equation results not allowing for country-pair fixed effects (nor first-differencing). These specifications will include standard time-invariant bilateral variables found in typical gravity equations without country-pair fixed effects, to illustrate the bias introduced by ignoring such fixed effects (or first-differencing). These variables are the natural log of bilateral distance between the country-pair’s economic centers (*LDIST*), a dummy variable for a pair of countries sharing a common land border (*ADJACENCY*), and a dummy variable for a pair of countries sharing a common official language (*LANGUAGE*), with all these data obtained from the CEPII website (as is standard). We will use data for every five years, beginning in 1965 and ending in 2000.

4 Empirical Results

The empirical results are presented in three tables associated with the text; two tables with additional FE results are provided in the Appendix. The first table presents results using fixed-effects (FE) specifications in section 4.1. In Set 1, we include only exporter-time and importer-time fixed effects, but not time-invariant country-pair effects. Set 2 also includes country-pair fixed effects. (Appendix Table 1A includes specifications with lagged values of the RHS variables.) In section 4.2, Table 2 reports the results using instead the econometrically

preferred first-difference (FD) specifications, with and without lags. In section 4.3, Table 3 reports alternative FD results using a technique similar to that in HMR to correct for sample-selection bias and firm heterogeneity and then also includes similar estimates but incorporating an alternative chain-weighted technique for weights due to the estimation of a time-series of cross-sections.

4.1 Fixed-Effects Results

Table 1 reports the results of two sets of panel regression results using alternative assumptions about fixed effects, labeled Set 1 and Set 2. Within each set are the results of running the same specification for three alternative LHS variables. TRADE refers to the aggregate bilateral trade flow between a pair of countries (or X_{ijt} in equation (7)). EM refers to the extensive margin (or EM_{ijt} in equation (7)). IM refers to the intensive margin (or IM_{ijt} in equation (7)). As we know from equation (7), once changes in $\ln X_{jt}$ are controlled for, the sum of variations in the extensive and intensive margins must equal the variation in the aggregate trade flow. This allows us to use the empirical results to infer the extensive and intensive margin elasticities to a given trade liberalization. Since both sets of results Set 1 and Set 2 include an importer-year (jt) fixed effect (as well as an exporter-year it fixed effect), this controls for $\ln X_{jt}$. Note that since both Set 1 and Set 2 results include exporter-year and importer-year fixed effects, we account for variation in exporter and importer GDPs as well as their multilateral resistance terms. The results in Set 1 are distinguished from those in Set 2 because the former exclude country-pair (ij) fixed effects. In Set 1, we use typical time-invariant bilateral gravity variables, consistent with equation (1). In Set 2, we include country-pair fixed effects in the FE equation.

We discuss Set 1 results first. In column (1a), we note that the coefficient estimates for $LDIST_{ij}$, $ADJACENCY_{ij}$, and $LANGUAGE_{ij}$ all have values similar to those found in the gravity-equation literature for aggregate trade flows and are statistically significant. However, several estimated EIA coefficient values in column (1a) tend to go against intuition. We have ordered vertically the listing of the EIAs by order of their degree of “depth”: $CUCMECU_{ijt}$ represents (the presence or absence of) a customs union, common market, or economic union, FTA_{ijt} represents a free trade agreement, $TWPTA_{ijt}$ represents a two-way preferential (or partial) trade agreement, and $OWPTA_{ijt}$ represents a one-way preferential trade agreement. Based upon the EIAs’ treaties and implementations, one would expect that the degree of trade

liberalization would be deepest in $CUCMECU_{ijt}$, next deepest in FTA_{ijt} , and most shallow in either $TWPTA_{ijt}$ or $OWPTA_{ijt}$.²⁵ However, $CUCMECU_{ijt}$ has a smaller positive effect on aggregate $TRADE_{ijt}$ than does $TWPTA_{ijt}$ or $OWPTA_{ijt}$. Also, FTA_{ijt} has the smallest influence on $TRADE_{ijt}$ than any of the other EIA dummies. Two-way and one-way PTAs have economically and statistically significant effects. Similar seemingly counterintuitive results hold for the two trade margins, EM_{ijt} and IM_{ijt} . Not only do two-way and one-way PTAs have economically and statistically significant positive effects on the extensive margin, but deeper EIAs have economically and statistically significant *negative* effects on the extensive margin. Such counterintuitive empirical results warrant further investigation.

It is likely that the EIA dummy variables are capturing some unobservable time-invariant bilateral factors that are influencing trade, but are correlated with the error term, creating endogeneity bias. As argued earlier and in BB, the use of country-pair ij fixed effects will eliminate the influence of these unobserved time-invariant bilateral factors. Set 2 of Table 1 reports the results of estimating the FE equation adding country-pair bilateral fixed effects which replace $LDIST_{ij}$, $ADJACENCY_{ij}$, and $LANGUAGE_{ij}$. There is a major difference in the estimated coefficient estimates for the EIA dummies in Set 2 relative to Set 1. In specification (2a) for aggregate trade, $CUCMECU_{ijt}$ now has the largest impact as expected and it is positive and both economically and statistically significant. The effect of common membership in an FTA is also positive as well as economically and statistically significant, and it is smaller than that for $CUCMECU_{ijt}$. These results are consistent with the findings in Magee (2008) and Roy (2010), the only two other studies that have examined ATEs of different types of agreements (specifically, FTAs versus customs unions) in a theoretically-motivated properly specified gravity model. The only counterintuitive result in column (2a) is that one-way PTAs have a negative and significant effect on trade flows.

The remaining interesting findings in Table 1 are the estimates in Set 2 for the extensive and intensive margin effects. First, consider the deep EIAs. $CUCMECU_{ijt}$ has a positive, economically significant, and statistically significant effect on *both margins*. Also, the intensive margin effect is larger than the extensive effect, as expected since there are no lags. Second, FTA_{ijt} also has a positive, economically significant, and statistically significant effect on the extensive margin, but one which is smaller than that for deep EIAs. However, the intensive-margin effect

²⁵However, many one-way PTAs include *fully* liberalized trade in one direction, so $OWPTA$ coefficients could be larger than $TWPTA$ coefficients.

is smaller than the extensive-margin effect, suggesting further investigation. Third, one-way PTAs have negative and significant effects on the extensive margin of trade in Set 2 results, also warranting further investigation. However, as discussed earlier given that trade is close to a unit-root process, it is likely that the FE results are causing spurious regression. Moreover, since the gravity-equation error terms are likely serially correlated due to sluggish trade-policy decisions, the FE estimates also are inefficient. Because of these econometric shortcomings, we move on in the next section to the FD results. Nevertheless, we do provide in the Appendix the FE specifications including lagged values of the EIA dummy variables, as well as a robustness analysis for one of these specifications for sample-selection bias and for controlling for firm-heterogeneity. On net, adding lagged values does not much influence the FE outcomes; for example, coefficient estimates of the lagged values of the one-way PTA dummies are negative and statistically significant, similar to the FE results without lagged values. However, we provide in the Appendix a brief analysis of those results.

4.2 First-Difference Results

This section reports the main empirical results using instead the FD equation rather than the FE specification in the previous section. As discussed in section 2, first-differencing the panel data yields two potential advantages over fixed effects. First, the unobserved factors influencing the likelihood of an EIA – for instance, trade is below its “natural” level – are likely slow moving and hence serially correlated. If ϵ_{ijt} in equation (2) are highly serially correlated, the inefficiency of FE is exacerbated as T gets large. First-differencing the data will increase estimation efficiency for our large- T panel. Second, aggregate trade flow data are likely “close to” unit-root processes. Using FE is equivalent to differencing data around the mean (in our sample, year 1980); this may create a problem since T is large in our panel. As Wooldridge (2000, p. 447) notes, if the data follow unit-root processes and T is large, the “spurious regression problem” can arise in a panel using FE. FD yields data that deviates from the previous period of our panel, and thus is closer to a unit-root process.

Table 2 reports the results using FD equations (9)-(11) with no lag in Set 3 results and with one lag in Set 4 results. Our first set of results in Set 3 are promising; we highlight the key issues. First, we note in column (3a) that deeper EIAs have larger effects on aggregate trade flows than FTAs, and the latter have larger effects on trade than two-way or one-way

PTAs, as expected. The coefficient estimates for $\Delta CUCMECU_{ij,t-(t-5)}$ and for $\Delta FTA_{ij,t-(t-5)}$ are economically and statistically significant. Second, for deep EIAs, FTAs, and one-way PTAs, the intensive-margin effects are economically and statistically significant. The effect is largest for $\Delta CUCMECU_{ij,t-(t-5)}$, next largest for $\Delta OWPTA_{ij,t-(t-5)}$, and slightly smaller for $\Delta FTA_{ij,t-(t-5)}$; the $\Delta TWPTA_{ij,t-(t-5)}$ effect is trivially small. Third, the extensive margin effects are statistically significant for deep EIAs and FTAs; moreover, the extensive margin effect for $\Delta CUCMECU_{ij,t-(t-5)}$ is larger than that for FTAs, as expected. Fourth, we note that – in the absence of lagged effects – for these three types of EIAs the intensive margin effect is always larger than its corresponding extensive margin effect. Thus, the Set 3 results using the FD specifications (9)-(11) provide substantive support for our three major hypotheses. However, by adding a lagged change in the EIAs, we can even provide even sharper evidence of the third hypothesis, extensive (intensive) margin effects should become relatively more (less) important with time.

In Table 2, even more interesting results are in Set 4 where we allow current and lagged changes, consequently capturing the effects on trade and the margins of 10 years of EIA changes. As discussed earlier and in BB, EIAs are likely to have delayed impacts on trade flows. One reason is that EIAs are typically “phased-in” over 5 to 10 years, delaying the full implementation of liberalization. Second, the impact on trade flows of EIAs works largely through terms-of-trade effects. It is well known that terms-of-trade changes can also have a delayed impact on trade flows. BB found that most of the impact of EIAs on aggregate trade flows was captured using panel techniques in a period of 10 years. Consider deep EIAs in columns (4a)-(4c). First, for aggregate trade, consistent with earlier results common membership in a deep EIA increases two members’ bilateral trade flow by a plausible 101 percent ($e^{0.387+0.309} = 2.01$). Second, we find that in the short-term, the effect on trade comes primarily through the intensive margin, as expected. The intensive-margin effect (0.281) dominates the extensive-margin effect (0.106) for $\Delta CUCMECU_{ij,t-(t-5)}$, with the latter effect statistically significant. However, for the lagged change the extensive margin effect (0.131) is statistically significant and larger than the current-period extensive-margin effect. Also, while the lagged intensive-margin effect (0.179) is statistically significant, it is smaller than the current-period effect (0.281). Third, when we sum the current period and lagged effects for each margin, we find an extensive-margin effect of 0.237 and an intensive-margin effect of 0.460. Also, we note that a second additional lag

$((t - 10) - (t - 15))$ only had a positive and statistically significant effect on $\Delta EM_{ij,t-(t-5)}$; for brevity, we omit presenting those results.

Consider now FTAs. First, for aggregate trade, common membership in an FTA increases two members' bilateral trade flow by 60 percent ($e^{0.242+0.228} = 1.60$). However, in the case of FTAs, the intensive-margin effect dominates the extensive-margin effect in both current period and lagged period; however, in the robustness analysis later, we find the relative size of the extensive-margin effect increasing in the lag. Second, when we sum the current period and lagged effects for each margin, we find an extensive margin effect of 0.169 and an intensive margin effect of 0.302. Hence, the ratio of these effects is similar to that for *CUCMECU*. Also, we note that a second additional lag $((t - 10) - (t - 15))$ had a statistically significant effect only on $\Delta EM_{ij,t-(t-5)}$, and hence we omit presenting those results.

For two-way PTAs, we found in Set 4 only a statistically significant effect of the lagged *TWPTA* change on $\Delta Trade_{ij,t-(t-5)}$. However, we found small positive but not statistically significant effects on the two margins. Nevertheless, the relative effects for $\Delta TWPTA_{ij,t-(t-5)}$ conformed to earlier results. For the short term, the relatively larger effect was at the intensive margin (0.046 for intensive relative to 0.023 for extensive). For the lagged change, the relatively larger effect was at the extensive margin (0.071 for extensive relative to 0.041 for intensive).

For one-way PTAs, we found in the current period, the intensive-margin effect (0.159) was economically and statistically significant and dominated the statistically insignificant extensive-margin effect (-0.043). For the lagged effect, the extensive-margin effect dominated the intensive-margin effect as expected. In fact, the extensive margin effect of 0.171 was statistically significant and the intensive margin effect of 0.114 was also statistically significant.

Our evidence indicates that the three hypotheses posed generally hold: both extensive and intensive margins are affected by EIAs, deeper EIAs tend to have larger impacts on trade, the extensive margin, and the intensive margin than shallower agreements, and intensive-margin effects tend to occur sooner than extensive-margin effects. In the next section, we evaluate the robustness of these results. First, there are “zeros” in our data set of aggregate trade flows, as well as in measures of EM_{ijt} and IM_{ijt} . To account for possible sample-selection bias and the possibility that the results are biased by omission of a control for “firm-heterogeneity,” in section 4.3 we employ the methodology in HMR to correct for both potential biases and show that the results are largely insensitive to either issue, owing to the presence in our second-stage

gravity-equation of either country-pair fixed effects or the use of first-differencing (as well as exporter-year and importer-year fixed effects) that control for both issues. Second, unlike HK which used a cross-section, we have a panel of trade flows, extensive margins, and intensive margins. The latter two variables incorporate trade weights, which are changing from year to year along with the (unweighted) margins. In section 4.3, we consider an alternative technique that employs the concept of “chain-weighting” that holds constant the trade-share weights for adjoining periods.

4.3 Sensitivity Analysis

First, as is well known, most bilateral trade flows among a large number of country-pairs include numerous observations with “zeros.” As noted in both HMR and Egger, Larch, Staub, and Winkelmann (2011), such zeros can be associated with selection bias (often referred to as “selection into exporting”); this can arise from the existence of fixed exporting costs and may be associated with firm heterogeneity in productivities (but does not require it). Our results may be sensitive to sample-selection bias and to the omission of a control variable for firm heterogeneity. Following HMR, we consider for sensitivity an alternative two-stage approach to control for selection bias and firm heterogeneity.²⁶

The procedure is easiest understand in the context of the FE specification. Following HMR for aggregate trade flows, the two-stage methodology entails estimating first a probit equation to determine the probability of a positive observation between a country-pair in each of the 8 cross-sections of our sample for each of $Trade_{ijt}$, EM_{ijt} , and IM_{ijt} . The probit estimates are then used to construct inverse Mills’ ratios (denoted IMR_{ijt}) to capture selection bias and variables Z_{ijt} and Z_{ijt}^2 to control for heterogeneous productivities. IMR_{ijt} , Z_{ijt} , and Z_{ijt}^2 are then used as additional regressors in the second-stage gravity-equation specification.

While HMR and Egger, Larch, Staub, and Winkelmann (2011) used single cross sections, we have a time series of cross sections. While the probit estimates are constructed by individual cross sections in the first stage, we choose a pooled specification in log-levels for the second-stage gravity equation. We conducted the sensitivity analysis for the FE specification including two lags, which are reported in Appendix Table A2. For the FD specifications, we used the predicted probits to construct the first-differences of the inverse Mills’ ratio, IMR_{ijt} , and of

²⁶This issue was ignored in Foster, Poeschl, and Stehrer (2011) and Eicher and Henn (2011).

Z_{ijt} and Z_{ijt}^2 .

To anticipate the results, we note the following intuition. In the cross-sectional context of HMR and Egger, Larch, Staub, and Winkelmann (2011), predicted probit values explaining the probability of exports from i to j are generated in the first stage. The predicted probit values are then used to construct the inverse Mill' ratio and the firm-heterogeneity measure, which are employed as controls in the second-stage gravity equation. The difference of our study relative to HMR and Egger, Larch, Staub, and Winkelmann (2011) is that our second-stage gravity equations either include country-pair ij fixed effects or use first-differences (as well as exporter-year and importer-year fixed effects). Since most of the variation in the predicted probit values is likely cross-sectional, the impact of variables constructed from these values – IMR_{ijt} , Z_{ijt} , and Z_{ijt}^2 – is reduced dramatically; in other words, omitted variables bias IMR_{ijt} , Z_{ijt} , and Z_{ijt}^2 is potentially eliminated largely by the ij fixed effects.²⁷ Analogously, first-differencing all the data removes the time-invariant pair-specific factors influencing IMR_{ijt} , Z_{ijt} , and Z_{ijt}^2 and thus controls for a major portion of the influences of selection and firm heterogeneity.

The first three columns (Set 5) in Table 3 report the results for the FD specification analogous to Set 4. The main finding from a comparison of the Set 5 results with the Set 4 results is that there is no material change in the results after correcting for sample-selection bias and firm heterogeneity. To emphasize, this does not imply that selection bias and firm heterogeneity are absent in the data; our results imply that such biases are largely eliminated due to the first-differencing in our second-stage regressions. For $CUCMECU_{ijt}$, the coefficient estimates are approximately the same between Sets 4 and 5; however, the effect of the concurrent $CUCMECU$ change on the extensive margin becomes statistically insignificant. For FTA_{ijt} , the coefficient estimates for the intensive-margin (extensive-margin) effect of the current FTA change is slightly larger (smaller), and the estimate for the extensive-margin effect becomes statistically insignificant.²⁸ For two-way PTAs, the coefficient estimates for aggregate trade, extensive margin, and intensive margin are insignificant in Sets 4 and 5 (with the exception of the statistically significant trade-flow change effect of $TWPPTA_{ij,(t-5)-(t-10)}$). For one-way

²⁷In the context of HMR (p. 453), this is shown clearly by reference to HMR's equation (9). The potential bias in this second-stage gravity from ignoring firm heterogeneity is omission of their term w_{ij} and from ignoring sample-selection surfaces in their error term u_{ij} . (Note, like us, they include exporter and importer fixed effects as well.) Our country-pair ij fixed effects removes the influence of their w_{ij} and the selection bias embodied in their u_{ij} .

²⁸However, the decreasing (increasing) importance of the intensive (extensive) margin effect is clearer in Set 5. The lagged extensive-margin effect remains statistically significant at the 10 percent significance level.

PTAs, the coefficient estimates are virtually the same across Sets 4 and 5. The intensive-margin effect is economically and statistically significant and the extensive-margin effect is trivial in the current period. For the lagged *OWPTA* change, the coefficient estimate for the intensive-margin effect is lower than in the current period but still significant; the extensive margin effect for the lagged *OWPTA* change is economically and statistically significant, consistent with expectations.²⁹

Second, the original HK analysis of extensive and intensive margins was conducted using a cross section for a particular year (using two alternative data sets). Our analysis uses a time series of cross sections. Consequently, the variables EM_{ijt} and IM_{ijt} defined in equations (4) and (5), respectively, include trade “weights” that change over time. Consequently, we consider two alternative methods for holding constant the trade weights, X_{Wjmt} , over time.

The first alternative measure of the extensive margin, denoted EM_{ijt}^* , uses values for X_{Wjmt} set to a “base year”; we chose 1995. Hence, EM_{ijt}^* is defined as:

$$EM_{ijt}^* = \frac{\sum_{m \in M_{ijt}} X_{Wjm1995}}{\sum_{m \in M_{Wjt}} X_{Wjm1995}} \quad (12)$$

where $X_{Wjm1995}$ is the value of country j 's imports from the world in product m in year 1995. M_{Wjt} and M_{ijt} are defined as before. In order to ensure that the log-linear decomposition holds analogous to equations (6) and (7), we define the intensive margin of goods exported from i to j as:

$$IM_{ijt}^* = \frac{\sum_{m \in M_{ijt}} X_{ijmt}}{\sum_{m \in M_{ijt}} X_{Wjm1995}}. \quad (13)$$

Consequently, the product of EM_{ijt}^* and IM_{ijt}^* results in:

$$\ln X_{ijt} = \ln EM_{ijt} + \ln IM_{ijt} + \ln X_{j1995}. \quad (14)$$

In estimation (using log-levels), the last term on the RHS is subsumed in the constant term.

One problem with the fixed-year trade weights is that the particular year chosen may bias the results, especially given that our data set spans 1965-2000. To address this, we considered another trade-weight approach. The second alternative measure for our FD specification uses

²⁹The Appendix presents and discusses the HMR results for the FE specifications in Appendix Table A2.

a “chain-weighted” approach. Since the LHS variables are 5-year differences, we used the same chain-weighting technique as used in National Income Accounts. For the X_{Wjmt} s, we used for t the geometric average of the trade flows for the corresponding years.

Table 3 presents in the last three columns (Set 6) the results of using the chain-weighted trade weights in the two-stage HMR approach; consequently, the results for Set 6 can be compared readily to the previous identical Set 5 specifications (which use the time-varying weights).³⁰ For brevity, we report only the result for the FD specification including one lag. It is readily seen that there is no material difference between Set 5 and Set 6 results.

5 Conclusions

Recent developments in estimating *ex post* the average treatment effects (ATEs) of economic integration agreements (EIAs) on international trade flows to account for the endogeneity of such agreements have led to larger and more precise estimates than found earlier in the gravity-equation literature. Separately, recent developments in international trade theory allowing for firm heterogeneity and fixed export costs have allowed researchers to explore the effects of trade-policy liberalizations on various extensive and intensive “margins of trade,” but most have used simulations. Only one study has explored econometrically the effect of EIAs on the intensive and extensive *goods* margin of trade with a large number of country-pairs’ trade flows, product-categories, years, and EIAs, and *no study* has explored the relative effects on the margins of trade of alternative *types* of EIAs – partial trade agreements, free trade agreements, customs unions, etc. – much less the relative *timing* of such effects. Systematic evidence of an increase in varieties of goods exported by various countries to numerous importing countries as a result of trade liberalizations is important for convincing policymakers of the welfare gains from EIAs.

Extending methodology established in Hummels and Klenow (2005) and Baier and Bergstrand (2007), this paper provided the first evidence using gravity equations of *both* the intensive and extensive (goods) margins being affected by economic integration agreements (EIAs) employing a panel data set with a large number of country-pairs, product-categories, and EIAs from 1962-2000. We found the first evidence of the differential average treatment effects (ATEs) of various “types” of EIAs on these intensive and extensive margins of trade with deeper in-

³⁰The 1995 base-period approach yielded results not materially different.

tegration agreements having larger impacts on aggregate trade flows, extensive margins, and intensive margins than shallower agreements. Finally, we found the first evidence of a novel differential “timing” of the two margins’ ATEs, with intensive-margin effects occurring sooner for deep EIAs, FTAs, and one-way PTAs than extensive-margin effects consistent with two recent theoretical studies. These results were robust to correcting for Heckman sample-selection bias and heterogeneous firms as in HMR.

6 Appendix

6.1 Distinguishing the (HK) Goods Margin vs. the (HMR) Firm Margin

It is important to distinguish clearly between – what we term – the “HK” (or goods) extensive and intensive margins versus the “HMR” (or firm) extensive and intensive margins. In a multi-sector context (or, in HK terms, “categories”), the HMR model yields a gravity equation (adapting HMR’s equation (6) for our notation):

$$X_{ijt}^m = (N_{it}^m V_{ijt}^m) \left(\frac{Y_{jt}^m w_{it}^{1-\sigma^m} \tau_{ijt}^{1-\sigma^m}}{\sum_{k=1}^K n_{kt}^m V_{kjt}^m w_{kt}^{1-\sigma^m} \tau_{kjt}^{1-\sigma^m}} \right) \quad (15)$$

where X_{ijt}^m is exports from country i to country j in sector m , N_{it}^m is the number of firms (non-exporting and exporting) in i in sector m , V_{ijt}^m is a monotonic function (defined shortly) of the fraction of i ’s firms that export to j in sector m (based upon firms’ relative productivities and fixed export costs to j in m), Y_{jt}^m is expenditures in j in sector m , w_{it} is the wage rate in i , τ_{ijt} is the variable trade cost in products exported from i to j expressed as 1 plus the *ad valorem* equivalent rate for shipping and/or policy barriers such as tariff rates³¹, and σ^m is the elasticity of substitution in consumption among products in sector m . Formally, V_{ijt}^m is defined in HMR as:

$$V_{ijt}^m = \int_{a_L^m}^{a_{ijt}^m} (a^m)^{1-\sigma^m} dG(a^m) \quad (16)$$

³¹Formally, for every 1 unit shipped from i , $1/(1 + \tau_{ijt})$ arrives at destination j . τ includes both natural and policy-based trade barriers.

for $a_{ijt}^m \geq a_L^m$, and $V_{ijt}^m = 0$ otherwise, where a_{ijt}^m is the amount of labor needed per unit of output for a firm in i to export to j and earn zero economic profits in sector m , a_L^m is the Pareto distribution support level for which a_{ijt}^m needs to be above for positive exports in sector m from i to j , and $G(\cdot)$ is the Pareto cumulative distribution function (cdf).

In the context of the HMR model, a lowering of trade costs as a result of formation of an EIA affects X_{ijt}^m potentially via both the intensive and extensive margins. The more traditional channel is the intensive margin. A fall in τ_{ijt} lowers the relative price of country i 's products of existing exporters in m raising relative demand in j by the variable trade-cost's elasticity, $\sigma^m - 1$. However, with heterogeneous firms, formation of an EIA may lower both variable trade costs (τ_{ijt}) as well as fixed export costs (typically denoted f_{ijt}) allowing less productive firms in sector m in country i to now profitably export from i to j , due to the rise in a_{ijt}^m . The latter raises V_{ijt}^m .

The ‘‘HMR extensive margin’’ influences, but is not the same as, the ‘‘HK extensive margin.’’ The HK extensive margin discussed above could also be written as:

$$EM_{ijt} = \frac{\sum_{m \in M_{ijt}} I_{ijt}^m X_{Wjt}^m}{\sum_{m \in M_{Wjt}} X_{Wjt}^m} \quad (17)$$

where I_{ijt}^m is an indicator variable equal to 1 if $X_{ijt}^m > 0$. Thus, the numerator of EM_{ijt} will increase as the HK extensive margin (number of sectors of positive ij exports) increases. In the context of the HMR model, a trade-cost decline will cause the HK extensive margin to increase if X_{ijt}^m becomes positive, which requires a_{ijt}^m – previously below a_L^m – to rise above a_L^m , allowing V_{ijt}^m to become positive. The HMR extensive margin increases with a fall in trade costs when a_{ijt}^m rises relative to a_L^m (assuming *either* $a_{ijt}^m \geq a_L^m$ initially *or* $a_{ijt}^m - a_L^m$ becomes positive) whereas the HK extensive margin increases when a_{ijt}^m – previously below a_L^m – rises above a_L^m . In this sense, the HK extensive-margin effect of a trade-cost decline tends to be smaller than the HMR extensive-margin effect.

The intensive margin in HMR is also similar but not identical to the HK intensive margin. The HMR intensive margin measures only the effect on volumes of existing exporters from i to j in sector m of a fall in trade costs, but not a change in the number of exporters. The intensive margin in HK allows for an increase in exports of existing exporters from i to j in sector m , but will also reflect an increase in the number of (profitable) exporters from i to j in m *if and only if* there were already profitable exporters from i to j in m (so that initially

$X_{ijt}^m > 0$). Consequently, the HK intensive-margin effect will tend to be larger than the HMR intensive-margin effect. Hence, extensive-margin effects of EIA formations (in the HMR sense) will tend to be understated using the HK extensive margin whereas intensive-margin effects (in the HMR sense) will tend to be overstated using the HK intensive margin. Of course, further disaggregation will tend to reduce these biases.

6.2 FE Specifications with Lags

Appendix Table A1 reports the results of including lagged values of EIA dummies in the FE specifications. Set A1 results add to the FE equation a five-year lag of the EIA dummies. Regarding aggregate trade in specification (A1a), we note for deep EIAs that $CUCMECU_{ijt}$ and $CUCMECU_{ij,t-5}$ both have economically and statistically significant positive effects, with the lagged effect even larger than the current period effect. The combination of coefficient estimates in (A1a) is 0.768, implying that the presence of a deep EIA between a country-pair increases their aggregate trade by 116 percent ($e^{0.768} = 2.16$), other things constant. By contrast, common membership in an FTA increased trade by 43 percent using the statistically significant lagged FTA coefficient in (A1a) ($e^{0.357} = 1.43$), and membership in either a two-way or one-way PTA had statistically significant *negative* effects, as was found in Table 1.

However, the distinguishing feature of our analysis is that we can differentiate extensive-margin from intensive-margin effects *and* we can differentiate the “timing” of these two margins’ effects. For brevity, we focus on $CUCMECU$ and FTA effects. First, consider deep EIAs in Set A1. In the current period t , or “short-run,” common membership increased bilateral trade at the intensive margin by 15 percent ($e^{0.139}$) and at the extensive margin by 16 percent ($e^{0.145}$); both coefficient estimates were economically and statistically significant. In addition, when we account for the lagged effect of $CUCMECU_{ij,t-5}$ on EM_{ijt} , we also have an additional prominent effect of 32 percent ($e^{0.276}$). Since the extensive-margin effect is reflecting in principle new firm entrants, it is not surprising that the extensive margin effect is even larger than the intensive margin effect with a lag.³² Second, consider the FTA_{ijt} and $FTA_{ij,t-5}$ coefficient estimates in column (A1a). We note that the lagged effect of an FTA on aggregate trade (0.357) is larger than the current period effect (0.016). In columns (A1b) and (A1c), both lagged FTA effects are significant, with the lagged extensive-margin effect dominating the intensive-margin

³²The lagged FTA intensive-margin effect also increases, but by a lesser percentage.

effect. For *TWPTA* and *OWPTA*, there are few statistically significant coefficient estimates in columns (A1a)-(A1c). The only statistically significant effects on margins are counter-intuitive negative lagged extensive-margin effects for *TWPTA* and *OWPTAs*, as was found in Table 1, Set 2.

For brevity, we provide only a few comments on the Set A2 results, which add to the previous specifications another lag – a 10-year lag. For *CUCMECU*, the additional 10-year lagged coefficient estimate for the extensive margin specification (A2b) is positive and statistically significant. For *FTA*, the additional 10-year lagged coefficient estimate for the extensive margin is also economically and statistically significant. However, the notable finding in this set is the number of negative and statistically significant coefficient estimates for extensive and intensive margin effects, which tends to confirm the econometric concern raised earlier about FE specifications. With large T , likely serially correlated errors, and unit-root processes for most RHS variables, spurious regression results can surface. For two-way and one-way PTAs, the 10-year lagged coefficient estimates for the aggregate trade, extensive margin, and intensive margin effects are negative, economically significant, and statistically significant. Evidence of spurious regression problems suggested the need for exploring the more appropriate FD specifications in the text of the paper.

Also, none of the issues just raised are alleviated by using the HMR procedure, adapted to panel, to account for sample-selection bias or firm heterogeneity. Appendix Table A2 provides the results in Set A3 of re-estimating the specifications in Set A2 but now using the HMR two-stage estimation procedure. As for the FD specifications, there was no material differences between the two sets of results.

References

- Anderson, James E., 2011. The gravity model, *Annals of Economic Research* **3**, 133-160.
- Anderson, James E., and Eric van Wincoop, 2003. Gravity with gravitas: A solution to the border puzzle, *American Economic Review* **93** (1), 170-192.
- Anderson, James E., and Eric van Wincoop, 2004. Trade costs, *Journal of Economic Literature* **42** (3), 691-751.

- Anderson, James E. and Yoto Yotov, 2011. Terms of trade and global efficiency effects of free trade agreements, 1990-2000, National Bureau of Economic Research Working Paper No. 17003.
- Arkolakis, Costas, Arnaud Costinot, and Andres Rodriguez-Clare, 2011. New trade models, same old gains? *American Economic Review* forthcoming.
- Arkolakis, Costas, Svetlana Demidova, Peter J. Klenow, and Andres Rodriguez-Clare, 2008. Endogenous variety and the gains from trade, *American Economic Review* **98** (2), May, 444-450.
- Arkolakis, Costas, Jonathan Eaton, and Samuel Kortum, 2011. Staggered adjustment and trade dynamics, working paper.
- Baier, Scott L., and Jeffrey H. Bergstrand, 2002. On the endogeneity of international trade flows and free trade agreements, manuscript, August.
- Baier, Scott L., and Jeffrey H. Bergstrand, 2004a. Economic determinants of free trade agreements, *Journal of International Economics* **64**, 29-63.
- Baier, Scott L., and Jeffrey H. Bergstrand, 2004b. Do free trade agreements actually increase members' international trade? Manuscript, October.
- Baier, Scott L., and Jeffrey H. Bergstrand, 2007. Do free trade agreements actually increase members' international trade? *Journal of International Economics* **71**, 72-95.
- Baier, Scott L., and Jeffrey H. Bergstrand, 2009. *Bonus vetus* OLS: A simple method for approximating international trade-cost effects using the gravity equation, *Journal of International Economics* **77** (1), 77-85.
- Bergstrand, Jeffrey H., and Peter Egger, 2011. Gravity equations and economic frictions in the world economy, in D. Bernhofen, R. Falvey, D. Greenaway, and U. Kreickemeier (eds.), *Palgrave Handbook of International Trade* forthcoming.
- Bernard, Andrew B., J. Bradford Jensen, Stephen J. Redding, and Peter K. Schott, 2009. The margins of U.S. trade, *American Economic Review* **99** (2), May, 487-493.

- Besedes, Tibor, and Thomas J. Prusa, 2011. The role of extensive and intensive margins and export growth, *Journal of Development Economics* forthcoming.
- Broda, Christian, and David E. Weinstein, 2006. Globalization and the gains from variety, *Quarterly Journal of Economics* **121** (2), 541-585.
- Chaney, Thomas, 2008. Distorted gravity: The intensive and extensive margins of international trade, *American Economic Review* **98** (4), September, 1707-1721.
- Eaton, Jonathan, and Samuel Kortum, 2002. Technology, geography, and trade, *Econometrica* **70** (5), 1741-1779.
- Eaton, Jonathan, Samuel Kortum, and Francis Kramarz, 2008. An anatomy of international trade: Evidence from French firms, National Bureau of Economic Research Working Paper No. 14610, December.
- Egger, Peter, Mario Larch, Kevin E. Staub, and Rainer Winkelmann, 2011. The trade effects of endogenous preferential trade agreements, *American Economic Journal: Economic Policy* **3** (3), August, 113-143.
- Eicher, Theo, and Christian Henn, 2011. In search of WTO trade effects: Preferential trade agreements promote trade strongly, but unevenly, *Journal of International Economics* **83**, 137-153.
- Feenstra, Robert C., 2009. Measuring the gains from trade under monopolistic competition, National Bureau of Economic Research Working Paper No. 15593, December.
- Feenstra, Robert C., and Hiau Looi Kee, 2007. Trade liberalization and export variety: A comparison of Mexico and China, *The World Economy* **30** (1), January, 5-21.
- Feenstra, Robert C., Robert E. Lipsey, Haiyan Deng, Alyson C. Ma, and Hengyong Mo, 2005. World trade flows: 1962-2000, National Bureau of Economic Research Working Paper No. 11040, January.
- Felbermayr, Gabriel, and Wilhelm Kohler, 2010. Modelling the extensive margin of world trade: New evidence on GATT and WTO membership, *The World Economy* **33** (11), November, 1430-1469.

- Foster, Neil, Johannes Poeschl, and Robert Stehrer, 2011. The impact of preferential trade agreements on the margins of international trade, *Economic Systems* **35**, 84-97.
- Frankel, Jeffrey, 1997. *Regional Trading Blocs*, Washington, DC: Institute for International Economics.
- Gilpin, Robert, 2000. *The Challenge of Global Capitalism*, Princeton, NJ: Princeton University Press.
- Greenaway, David, and Richard Kneller, 2007. Firm heterogeneity, exporting, and foreign direct investment, *Economic Journal* **117**, February, F134-F161.
- Helpman, Elhanan, Marc Melitz, and Yona Rubinstein, 2008. Estimating trade flows: Trading partners and trading volumes, *Quarterly Journal of Economics* **123** (2), May, 441-487.
- Hillberry, Russell, and Christine McDaniel, 2002. A decomposition of North American trade growth since NAFTA, United States International Trade Commission Working Paper 2002-12-A, July.
- Hummels, David, and Peter J. Klenow, 2005. The variety and quality of a nation's exports, *American Economic Review* **95** (3), 704-723.
- Kehoe, Timothy J. and Kim J. Ruhl, 2009. How important is the new goods margin in international trade? Federal Reserve Bank of Minneapolis Staff Report 324, June.
- Lawrence, Robert Z., 1996. *Regionalism, Multilateralism, and Deeper Integration*, Washington, DC: Brookings Institution.
- Lee, Jong-Wha, and Philip Swagel, 1997. Trade barriers and trade flows across countries and industries, *Review of Economics and Statistics* **79**, 372- 382.
- Magee, Christopher S.P., 2008. New measures of trade creation and trade diversion, *Journal of International Economics* **75**, 349-362.
- Martin, Will and Francis Ng, 2004. A note on sources of tariff reductions in developing countries, 1983-2003, World Bank, October.
- Melitz, Marc J., 2003. The impact of trade on intra-industry reallocations and aggregate industry profitability, *Econometrica* **71** (6), November, 1695-1725.

- Pavcnik, Nina, 2002. Trade liberalization, exit, and productivity improvements: evidence from Chilean plants, *Review of Economic Studies* **69**, 245-276.
- Redding, Stephen J., 2011. Theories of heterogeneous firms and trade, *Annals of Economic Research* **3**, 77-105.
- Rose, Andrew, 2004. Do we really know that the WTO increases trade? *American Economic Review* **94** (1), March, 98-114.
- Roy, Jayjit, 2010. Do customs union members engage in more bilateral trade than free-trade agreement members?, *Review of International Economics* **18** (4), 663-681.
- Ruhl, Kim, 2008. The international elasticity puzzle, working paper.
- Trefler, Daniel, 1993. Trade liberalization and the theory of endogenous protection: An econometric study of U.S. import policy, *Journal of Political Economy* **101** (1), February, 138-160.
- Trefler, Daniel, 2004. The long and short of the Canada-U.S. free trade agreement, *American Economic Review* **94** (4), 870-895.
- Wooldridge, Jeffrey M., 2000. *Introductory Econometrics*, Chicago, IL: South-Western Publishing.
- Wooldridge, Jeffrey M., 2002. *Econometric Analysis of Cross-Section and Panel Data*, Cambridge, MA: MIT Press.

Table 1

Variables	Set 1			Set 2		
	(1a) TRADE _{ijt}	(1b) EM _{ijt}	(1c) IM _{ijt}	(2a) TRADE _{ijt}	(2b) EM _{ijt}	(2c) IM _{ijt}
LDIST _{ij}	-1.184*** (0.014)	-0.810*** (0.011)	-0.374*** (0.011)			
ADJACENCY _{ij}	0.403*** (0.056)	0.283*** (0.044)	0.120*** (0.038)			
LANGUAGE _{ij}	0.746*** (0.026)	0.649*** (0.020)	0.097*** (0.020)			
CUCMECU _{ijt}	0.193*** (0.065)	-0.480*** (0.059)	0.673*** (0.041)	0.558*** (0.059)	0.252*** (0.046)	0.306*** (0.049)
FTA _{ijt}	0.147*** (0.042)	-0.076** (0.037)	0.223*** (0.030)	0.143*** (0.045)	0.094*** (0.035)	0.049* (0.037)
TWPTA _{ijt}	0.621*** (0.042)	0.591*** (0.032)	0.031 (0.030)	-0.0003 (0.052)	0.035 (0.043)	-0.035 (0.045)
OWPTA _{ijt}	0.711*** (0.037)	0.579*** (0.029)	0.132*** (0.030)	-0.222*** (0.052)	-0.171*** (0.038)	-0.051 (0.046)
Constant	-1.092*** (0.120)	-0.454*** (0.091)	-0.622*** (0.089)	-6.637*** (0.276)	-4.930*** (0.238)	-1.707*** (0.234)
<u>Fixed Effects:</u>						
Exporter-Year (it)	Yes	Yes	Yes	Yes	Yes	Yes
Importer-Year (jt)	Yes	Yes	Yes	Yes	Yes	Yes
Country-Pair (ij)	No	No	No	Yes	Yes	Yes
R ²	0.675	0.613	0.520	0.860	0.815	0.740
No. of observations	65,292	65,292	65,292	65,292	65,292	65,292

Notes: Robust standard errors in parentheses. *, **, and *** denote statistical significance at the 10, 5, and 1 percent levels, respectively, in one-tailed t-tests.

Table 2

Variables	Set 3			Set 4		
	(3a) $\Delta\text{TRADE}_{ij,t-(t-5)}$	(3b) $\Delta\text{EM}_{ij,t-(t-5)}$	(3c) $\Delta\text{IM}_{ij,t-(t-5)}$	(4a) $\Delta\text{TRADE}_{ij,t-(t-5)}$	(4b) $\Delta\text{EM}_{ij,t-(t-5)}$	(4c) $\Delta\text{IM}_{ij,t-(t-5)}$
$\Delta\text{CUCMECU}_{ij,t-(t-5)}$	0.329*** (0.072)	0.100* (0.062)	0.229*** (0.065)	0.387*** (0.077)	0.106* (0.067)	0.281*** (0.069)
$\Delta\text{CUCMECU}_{ij,(t-5)-(t-10)}$				0.309*** (0.076)	0.131** (0.065)	0.179*** (0.072)
$\Delta\text{FTA}_{ij,t-(t-5)}$	0.192*** (0.053)	0.074* (0.043)	0.118*** (0.046)	0.242*** (0.056)	0.085** (0.046)	0.157*** (0.049)
$\Delta\text{FTA}_{ij,(t-5)-(t-10)}$				0.228*** (0.055)	0.084** (0.047)	0.145*** (0.049)
$\Delta\text{TWPTA}_{ij,t-(t-5)}$	-0.001 (0.072)	0.012 (0.055)	-0.013 (0.063)	0.069 (0.079)	0.023 (0.063)	0.046 (0.068)
$\Delta\text{TWPTA}_{ij,(t-5)-(t-10)}$				0.111* (0.071)	0.071 (0.063)	0.041 (0.068)
$\Delta\text{OWPTA}_{ij,t-(t-5)}$	0.072 (0.065)	-0.064 (0.053)	0.135** (0.062)	0.116** (0.067)	-0.043 (0.055)	0.159*** (0.063)
$\Delta\text{OWPTA}_{ij,(t-5)-(t-10)}$				0.285*** (0.069)	0.171*** (0.055)	0.114** (0.067)
Constant	0.413*** (0.120)	0.167* (0.130)	0.247* (0.182)	0.608*** (0.247)	-0.089 (0.274)	0.697** (0.342)
<u>Fixed Effects:</u>						
Exporter-Year (i,t-(t-5))	Yes	Yes	Yes	Yes	Yes	Yes
Importer-Year (j,t-(t-5))	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.294	0.365	0.357	0.330	0.402	0.403
No. of observations	48,619	48,619	48,619	41,767	41,767	41,767

Notes: Robust standard errors in parentheses. *, **, and *** denote statistical significance at the 10, 5, and 1 percent levels, respectively, in one-tailed t-tests.

Table 3

Variables	Set 5			Set 6		
	(5a) $\Delta\text{TRADE}_{ij,t-(t-5)}$	(5b) $\Delta\text{EM}_{ij,t-(t-5)}$	(5c) $\Delta\text{IM}_{ij,t-(t-5)}$	(6a) $\Delta\text{TRADE}_{ij,t-(t-5)}$	(6b) $\Delta\text{EM}_{ij,t-(t-5)}$	(6c) $\Delta\text{IM}_{ij,t-(t-5)}$
$\Delta\text{CUCMECU}_{ij,t-(t-5)}$	0.366*** (0.084)	0.032 (0.0719)	0.334*** (0.076)	0.366*** (0.0837)	0.042 (0.073)	0.324*** (0.078)
$\Delta\text{CUCMECU}_{ij,(t-5)-(t-10)}$	0.316*** (0.076)	0.130** (0.065)	0.185*** (0.071)	0.316*** (0.076)	0.123** (0.067)	0.192*** (0.073)
$\Delta\text{FTA}_{ij,t-(t-5)}$	0.226*** (0.065)	0.021 (0.054)	0.205*** (0.059)	0.226*** (0.065)	0.027 (0.055)	0.200*** (0.060)
$\Delta\text{FTA}_{ij,(t-5)-(t-10)}$	0.224*** (0.055)	0.074* (0.047)	0.150*** (0.049)	0.224*** (0.055)	0.066* (0.048)	0.157*** (0.049)
$\Delta\text{TWPTA}_{ij,t-(t-5)}$	0.088 (0.084)	0.077 (0.068)	0.011 (0.073)	0.088 (0.084)	0.079 (0.069)	0.009 (0.074)
$\Delta\text{TWPTA}_{ij,(t-5)-(t-10)}$	0.110* (0.071)	0.076 (0.062)	0.034 (0.068)	0.110* (0.071)	0.058 (0.064)	0.052 (0.069)
$\Delta\text{OWPTA}_{ij,t-(t-5)}$	0.143** (0.067)	0.003 (0.056)	0.140** (0.064)	0.143** (0.067)	-0.006 (0.057)	0.149** (0.065)
$\Delta\text{OWPTA}_{ij,(t-5)-(t-10)}$	0.285*** (0.069)	0.172*** (0.059)	0.113** (0.067)	0.285*** (0.069)	0.169*** (0.060)	0.116** (0.068)
ΔIMR_{ijt}	-0.198 (0.296)	0.419** (0.234)	-0.616** (0.256)	-0.198 (0.296)	0.440** (0.240)	-0.637** (0.260)
ΔZ_{ijt}	0.070 (0.236)	0.521*** (0.195)	-0.452** (0.215)	0.070 (0.236)	0.537*** (0.198)	-0.467** (0.218)
ΔZ^2_{ijt}	-0.009 (0.031)	-0.059** (0.026)	0.050** (0.028)	-0.009 (0.031)	-0.061*** (0.026)	0.052** (0.028)
Constant	0.555** (0.247)	-0.005 (0.284)	0.560* (0.365)	0.555** (0.247)	0.012 (0.286)	0.548* (0.369)
<u>Fixed Effects:</u>						
Exporter-Year (i,t-(t-5))	Yes	Yes	Yes	Yes	Yes	Yes
Importer-Year (j,t-(t-5))	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.330	0.402	0.403	0.330	0.402	0.403
No. of observations	41,767	41,767	41,767	41,767	41,767	41,767

Notes: Robust standard errors in parentheses. *, **, and *** denote statistical significance at the 10, 5, and 1 percent levels, respectively, in one-tailed t-tests.

Appendix Table A1

Variables	Set A1			Set A2		
	(A1a) TRADE _{ijt}	(A1b) EM _{ijt}	(A1c) IM _{ijt}	(A2a) TRADE _{ijt}	(A2b) EM _{ijt}	(A2c) IM _{ijt}
CUCMECU _{ijt}	0.284*** (0.071)	0.145*** (0.057)	0.139** (0.061)	0.236*** (0.076)	0.191*** (0.062)	0.046 (0.067)
CUCMECU _{ij,t-5}	0.484*** (0.072)	0.276*** (0.057)	0.209*** (0.061)	0.419*** (0.082)	0.204*** (0.067)	0.215*** (0.071)
CUCMECU _{ij,t-10}				0.154** (0.080)	0.153*** (0.064)	0.0003 (0.069)
FTA _{ijt}	0.016 (0.052)	0.010 (0.042)	0.006 (0.043)	0.040 (0.056)	0.047 (0.044)	-0.007 (0.048)
FTA _{ij,t-5}	0.357*** (0.053)	0.247*** (0.042)	0.110*** (0.045)	0.333*** (0.061)	0.181*** (0.048)	0.152*** (0.052)
FTA _{ij,t-10}				0.113** (0.060)	0.206*** (0.047)	-0.094 (0.052)
TWPTA _{ij,t}	0.031 (0.063)	0.055 (0.050)	-0.024 (0.055)	-0.089 (0.070)	0.031 (0.057)	-0.120** (0.060)
TWPTA _{ij,t-5}	-0.112* (0.064)	-0.100** (0.051)	-0.011 (0.057)	0.019 (0.071)	-0.047 (0.062)	0.067 (0.065)
TWPTA _{ij,t-10}				-0.358*** (0.070)	-0.177*** (0.057)	-0.181*** (0.062)
OWPTA _{ij,t}	-0.127* (0.066)	-0.046 (0.051)	-0.081 (0.060)	-0.253*** (0.079)	-0.059 (0.062)	-0.194*** (0.074)
OWPTA _{ij,t-5}	-0.095 (0.060)	-0.093** (0.047)	-0.002 (0.054)	0.145** (0.072)	0.079 (0.056)	0.066 (0.066)
OWPTA _{ij,t-10}				-0.439*** (0.065)	-0.286*** (0.049)	-0.153** (0.059)
Constant	-6.807*** (0.072)	-3.669*** (0.062)	-3.138*** (0.074)	-7.051*** (0.117)	-3.516*** (0.084)	-3.535*** (0.128)
Fixed Effects:						
Exporter-Year (it)	Yes	Yes	Yes	Yes	Yes	Yes
Importer-Year (jt)	Yes	Yes	Yes	Yes	Yes	Yes
Country-Pair (ij)	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.872	0.827	0.760	0.888	0.842	0.790
No. of observations	57,742	57,742	57,742	48,914	48,914	48,914

Notes: Robust standard errors in parentheses. *, **, and *** denote statistical significance at the 10, 5, and 1 percent levels, respectively, in one-tailed t-tests.

Appendix Table A2

Set A3	(A3a) TRADE _{ijt}	(A3b) EM _{ijt}	(A3c) IM _{ijt}
CUCMECU _{ijt}	0.256*** (0.092)	0.185*** (0.075)	0.071 (0.080)
CUCMECU _{ij,t-5}	0.419*** (0.083)	0.213*** (0.067)	0.206*** (0.071)
CUCMECU _{ij,t-10}	0.155** (0.080)	0.164*** (0.063)	-0.009 (0.069)
FTA _{ijt}	0.066 (0.071)	0.020 (0.059)	0.046 (0.061)
FTA _{ij,t-5}	0.333*** (0.061)	0.197*** (0.048)	0.136*** (0.051)
FTA _{ij,t-10}	0.119** (0.059)	0.213*** (0.047)	-0.095** (0.052)
TWPTA _{ij,t}	-0.113* (0.077)	0.049 (0.064)	-0.162*** (0.066)
TWPTA _{ij,t-5}	0.019 (0.071)	-0.047 (0.062)	0.066 (0.065)
TWPTA _{ij,t-10}	-0.355*** (0.070)	-0.180*** (0.057)	-0.175*** (0.062)
OWPTA _{ij,t}	-0.292*** (0.084)	-0.069 (0.067)	-0.223*** (0.0771)
OWPTA _{ij,t-5}	0.143** (0.072)	0.079 (0.056)	0.064 (0.066)
OWPTA _{ij,t-10}	-0.436*** (0.065)	-0.281*** (0.049)	-0.156*** (0.059)
IMR _{ijt}	0.236 (0.299)	0.690*** (0.244)	-0.455 (0.283)
Z _{ijt}	-0.047 (0.217)	0.541*** (0.179)	-0.587*** (0.190)
Z ² _{ijt}	0.005 (0.028)	-0.068*** (0.022)	0.073*** (0.024)

Constant	-6.859*** (0.369)	-4.524*** (0.307)	-2.335*** (0.338)
Importer-Year	Yes	Yes	Yes
Exporter-Year	Yes	Yes	Yes
Country-Pair	Yes	Yes	Yes
Observations	48,914	48,914	48,914
R ²	0.888	0.842	0.790

Notes: Robust standard errors in parentheses. *, **, and *** denote statistical significance at the 10, 5, and 1 percent levels, respectively, in one-tailed t-tests