Beyond Tariffs: Quantifying Heterogeneity in the Effects of Free Trade Agreements *

Thomas Zylkin†

March, 2016

Abstract

Using NAFTA as an illustrating example, I examine heterogeneity in the effects of free trade agreements (FTAs), both within and across agreements. Most notably, the asymmetries I observe for NAFTA's effects—and for NAFTA's effects on Mexican trade in particular—disagree strongly with expectations based on pre-NAFTA tariffs. I also show these overlooked sources of heterogeneity have important first-order implications for evaluating the incidence of NAFTA's welfare benefits. Using tariffs to project welfare not only greatly underestimates the overall welfare gains for all three NAFTA countries (Mexico's especially), but also overstates the benefits for U.S. producers. Additional evidence from other FTAs signed during the 1990s suggests NAFTA is not an outlier in this context: FTAs between countries with different levels of economic development generally favored the exports of the less developed country.

JEL Classification Codes: F13, F14, F15

Keywords: Free Trade Agreements, International Trade, Gravity

*I am grateful to the members of my dissertation committee—Costas Syropoulos, Yoto Yotov, Eric Bond, Vibhas Madan, and Irina Murtazashvili—for their comments and guidance. I am also grateful for research support from the NUS Strategic Research Grant (WBS: R-109-000-183-646) awarded to the Global Production Networks Centre (GPN@NUS) for the project titled "Global Production Networks, Global Value Chains, and East Asian Development". Other acknowledgments will be added later. All errors are mine.

†Contact information: Department of Economics, National University of Singapore. Website: www.tomzylkin.com.

1 Introduction

"In all the time I've been in Congress, I've never seen a trade bill that benefits the American producer or the American worker," U.S. Congresswoman Louise Slaughter (D-NY) recently declared. "People are sick and tired of the one way trade deal."[1] Such complaints are standard in political debates surrounding major trade agreements. Despite many well-established arguments for the idea that countries generally "gain from trade", policymakers supplying crucial votes on trade bills continually want to know how a given free trade agreement (FTA) will affect their country's exports vs. its imports.[2] The idea that a potential major agreement—such as the forthcoming Trans-Atlantic or Trans-Pacific trade partnerships—could be a "one-way trade deal" for one or more countries involved is invariably a sensitive topic. Furthermore, as evidenced by Rep. Slaughter's comments, "one way" perceptions of past deals weigh heavily on the prospects for future trade integration.

This paper works towards addressing these issues by offering a tractable methodology by which general "heterogeneity" in the effects of FTAs, both within and across agreements, can be identified and analyzed. Using a structural estimation of changes in trade costs over time, I am able to infer directly from the trade data, *ex post*, what effects a particular FTA (in this case, NAFTA) has had on trade barriers for each of its member countries. Similar *ex post* approaches, starting with seminal empirical work by Baier & Bergstrand (2007), have illustrated many new facts about FTAs in recent years, with an increasing focus on identifying sources of heterogeneity. These findings have included, among other things: how average FTA effects differ across industries (Anderson & Yotov, 2016), how the average effect differs based on the type of agreement (Baier, Bergstrand, & Feng, 2014), what FTAs mean for different "margins" of trade (Egger, Larch, Staub, & Winkelmann, 2011; Baier, Bergstrand, & Feng, 2014), and how FTAs with different provisions may affect trade differently (Kohl, Brakman, & Garretsen, 2013).

This literature has not yet considered the substantial *directional* heterogeneity that may exist within the very same agreement, however—i.e., cases where one member country receives more access to another country's market than it grants in return. Using the NAFTA as my illustrating example, I show that this latter source of heterogeneity in particular may not be well understood. Even within a given FTA, there can exist significant asymmetries in trade barrier reductions which have important implications for the gains from trade.

Specifically, I use the case of NAFTA to ask two broad questions about the implications of allowing for asymmetries in FTA effects. First, *how much might directional heterogeneity in the effects of free trade agreements differ from what we might expect to observe?* It turns out that, in the specific example of NAFTA,

[1] "Specter of NAFTA hurts Obama's trade dreams", *Politico*, February 19th, 2014.

[2] Viner (1950) originates the current literature on the welfare analysis of trade integration. This literature is large and continually growing. Bhagwati, Krishna, & Panagariya (1999) summarize many of the major theoretical developments since Viner. Current approaches to quantitative welfare analysis have roots in two major papers: Eaton & Kortum (2002) and Anderson & van Wincoop (2003). Head & Mayer (2014) and Costinot & Rodríguez-Clare (2014) survey more recent developments.

the answer is "a lot". The idea that a free trade agreement may have different effect on trade barriers for different countries is not necessarily surprising; we know offhand that some countries can often have very different trade barriers *ex ante* and thus we would expect them to be affected differently by an agreement. The typical perspective in the literature for analyzing the effects of individual agreements, and the effects of NAFTA in particular, has been to assume these differences can be summarized by differences in *ex ante* tariff levels and other observable trade policy barriers.[3] I show, however, that the *ex post* realization of effects associated with NAFTA differs substantially from expectations based on North American tariffs *circa* 1990; I also observe some suggestive evidence from the broader data on FTAs that a plausible, under-appreciated source of asymmetries within NAFTA may be the difference in the level of development between the U.S. and Mexico.

Second, I also wish to know: *how important are these asymmetries for the gains from trade integration?* Using a series of simple simulation experiments, I show that prevailing approaches to identifying the welfare benefits of FTAs (and the welfare benefits of NAFTA in particular) would tend to mischaracterize NAFTA's benefits in large, important ways.

My estimation approach builds on previous panel implementations of structural gravity by Baier & Bergstrand (2007) and Anderson & Yotov (2016). The key for my purposes is that structural gravity provides a clear logic for identifying the role of trade barriers in determining trade between countries, which generalizes across a very wide range of trade models. In this context, my innovation is that I allow the effect of an FTA to vary both at the level of the individual agreement (such that NAFTA affects trade differently from other FTAs) and at the level of the direction of trade (such that NAFTA does not, for example, affect U.S. imports from Mexico the same way it does U.S. exports to Mexico). Cipollina & Salvatici (2010) and Kohl (2014) have each previously explored heterogeneity in FTA effects at the level of the individual agreement for aggregate trade. However, the gravity literature has yet to examine the question of asymmetries in FTA effects within agreements and whether they truly correspond to observable differences in trade policy barriers.

My empirical findings for NAFTA confirm the presence of substantial heterogeneity in FTA effects, both between NAFTA and other FTAs and within NAFTA itself. For example, I find that NAFTA has promoted trade significantly more than other FTAs have.[4] Furthermore, I establish that NAFTA has had strongly asymmetric effects on trade barriers, both for aggregate trade and within individual industries.[5] To name two examples: Mexico's Metals producers have received substantially more access to Canadian and U.S. import markets than U.S. and Canadian Metals producers have received in return; whereas in the Food sector, the

[3] Shikher (2012*a*) documents that the available data on nontariff barriers leads to very similar qualitative predictions about NAFTA's effects as the data on tariffs.

[4] Cipollina & Salvatici (2010) and Kohl (2014) each document similar findings for NAFTA versus other FTAs. My innovation on this point is that I allow for general differences between NAFTA and other FTAs to differ across industries.

[5] It is important to note that the use of exporter-by-time and importer-by-time fixed effects in the estimation cleanly distinguishes these asymmetries from notions of comparative advantage or any other tendency towards trade not directly related to trade costs. This includes changes in the valuation of a country's currency, an important consideration in the case of Mexico.

opposite has occurred.

Most notably, I document that these asymmetries differ significantly from what would have been expected to happen based on pre-NAFTA trade policies. Whereas projections based on tariffs alone would have suggested that U.S. and Canadian producers should have gained much more access to Mexican import markets than they received in return, the *ex post* changes in trade frictions I identify tell a different story. Overall, NAFTA coincided with sizable 63.1% decrease in barriers to U.S. import markets for Mexico's exporters, compared with a 45.9% decrease in barriers against U.S. products headed in the other direction. The trade barrier reductions between Canada and Mexico are relatively more symmetric, but still sit at odds with the fact that Mexico's pre-NAFTA tariffs were two-and-a-half times those of Canada.

Intriguingly, additional evidence from other FTAs signed during the 1990s strongly suggests that NAFTA is not a unique outlier in this context. Again contrary to expectations, FTA pairings between highly developed and less developed economies (such as the U.S. and Mexico in the case of NAFTA) generally favor the exports of the less developed partner, despite the fact that less developed countries tend to have higher initial tariffs. Section 5 discusses some possible explanations for these puzzling findings (e.g., the role that FTAs may play in reducing the costs of offshoring).

The under-appreciated effect of NAFTA on imports from Mexico may be contributing to current skepticism towards free trade deals in the U.S. My results do not support Rep. Slaughter's complaints of "one way trade" *per se*, but they do strongly suggest that the realization of NAFTA has differed substantially from the prevailing narratives available at the time. Even still, one important caveat is in order. An extension of the main analysis allowing for prior trends shows that, if anything, NAFTA's main effect on U.S./Mexico trade barriers was to reinforce existing trends. The same cannot be said, however, for trade barrier reductions between Mexico and Canada, which only became increasingly asymmetric in favor of Mexico's exports after the signing of NAFTA.

For the welfare analysis, I use a highly general modeling perspective. I assume a simple multi-sector endowment economy setting, such that all welfare effects occur strictly through first-order effects on prices for both buyers and sellers. These first-order price effects I obtain in turn have direct implications for additional welfare channels in many more elaborate trade models that share the same essential structure.[6] The endowment economy setting reveals the following: i) implied welfare gains for all three NAFTA members are 50% to 100% as large as would be implied by an average FTA effect and two to three times as large as would be found using tariff reductions; ii) the size of Mexico's implied gains from NAFTA in particular (5.07% vs. 12.32%) are underestimated using tariffs; iii) tariff changes tend to obscure the nature of the U.S.'s gains from NAFTA: the U.S. has benefited mostly from an increased ability to buy goods from

[6]Table 1 in Head & Mayer (2014) provides a (non-exhaustive) overview of these "more elaborate" models. These frameworks include allowances for monopolistic competition and increasing returns (Krugman, 1980), endogenous firm-level export entry (Melitz, 2003), trade in intermediates (Eaton & Kortum, 2002; Caliendo & Parro, 2015), and variable firm mark-ups (Melitz & Ottaviano, 2008; Behrens, Mion, Murata, & Südekum, 2014). Costinot & Rodríguez-Clare (2014) demonstrate how many of these elements can be combined within a single framework.

Canada and Mexico (i.e., lower buyer prices), not from being better able to sell its own goods to the other two markets (i.e., higher producer prices).[7]

Because NAFTA has remained controversial in North American politics, it has motivated a substantial literature dedicated to analyzing its effects on trade, welfare, and other outcome variables. The following section will touch specifically on how this paper fits within and informs this literature. In subsequent sections, I present the empirical analysis, offer some discussion of how to explain the patterns of asymmetry I observe in NAFTA's effects, and demonstrate basic implications for welfare analysis. The last section adds concluding remarks.

2 Looking Back at NAFTA and Looking Ahead to Future FTAs

Due to my focus on NAFTA in particular, my findings resonate with the recent literature that has tried to re-evaluate the policy impact of NAFTA using progressively more updated tools for analyzing gains from trade. Caliendo & Parro (2015), for example, simulate the effects of NAFTA in a calibrated model with cross-sectoral input linkages that explains a significant portion of the aggregate changes in country-level exports and imports that occurred post-NAFTA. Their work builds on the analyses performed by Brown, Deardorff, & Stern (1992), Anderson & van Wincoop (2001), Romalis (2007), Shikher (2012*a*), and others to try to characterize NAFTA's imprint on each member's welfare using analytical models. Of these, Brown, Deardorff, & Stern (1992) and Shikher (2012*a*) are notable in that they compute effects for removing non-tariff barriers as well as for removing tariffs. Otherwise, however, these papers generally do not look at how bilateral frictions between members may have been affected along other dimensions besides tariffs.[8]

One feature of the trade data that cannot easily be explained based on tariffs alone is how the U.S. and Canadian trade balances with respect to Mexico changed during the mid-1990s. For example, Agama & McDaniel (2002) document that, despite the presence of already-existing U.S. tariff preferences on imports from Mexico (on the order of about 3% pre-NAFTA), the large growth in U.S.-Mexico trade in the mid-90s and early-2000s was mostly driven by growth in U.S. imports from Mexico (190%) rather than U.S. exports to Mexico (a still substantial 93%). Meanwhile, Canadian imports from Mexico grew by 366% over the same period, versus 198% for its exports to Mexico (Trembley, 2013). The surprisingly large increase in U.S. and Canadian imports from Mexico in particular invites consideration of how NAFTA may have affected trade barriers beyond tariffs.

[7] As is often discussed in the trade literature (see Arkolakis, Costinot, & Rodríguez-Clare, 2012), the magnitudes obtained for "gains from trade" are highly sensitive to the assumptions used for the "trade elasticity" (i.e., the sensitivity of trade to changes in trade costs). I use values from Broda, Greenfield, & Weinstein (2006) for these elasticities. Using other elasticity values would affect absolute magnitudes of welfare gains but would not significantly affect relative magnitudes across the different ways of measuring NAFTA's effects. I stand on the latter as my main result.

[8] The estimates for nontariff barriers in Shikher (2012*a*) are from the Nicita & Olarreaga (2007) "Trade, Production, and Protection" database. These estimates still show much higher initial liberalization for Mexican imports than for Mexican exports, which is still contrary to the results I obtain for NAFTA's trade impact *ex post*.

At the same time, there are valid reasons why simulations of FTA effects have continued to limit their attention to tariffs. Tariff reductions associated with FTAs are known *ex ante* and easily observable *ex post.* Pure econometric estimates of changes in trade cannot necessarily disentangle the direct effects of an FTA from other concurrent factors that may have affected trade barriers. In the case of NAFTA, Krueger (1999) and Agama & McDaniel (2002) each cite Mexico's peso devaluation in the mid-1990s as an important difficulty in identifying NAFTA's effects *ex post*. Notably, Krueger (1999) provides a useful reasoning by which these two competing narratives can be evaluated. In the discussion following my results, I describe how the fixed effects estimation implied by structural gravity incorporates this reasoning.

Looking beyond the set of assumptions underlying the structural gravity model reveals some possible explanations for the asymmetries I observe in my NAFTA estimates. In addition to the discussion of currency volatility, I also discuss how the model accounts for trade in intermediates along global supply chains. The extent that the actual behavior of individual firms within the supply chain may differ from the assumed behavior implied by this class of models may help explain the asymmetries in NAFTA's effects (as well as in the effects of other agreements between developed/less developed pairs).

The choice to focus on NAFTA reflects the usefulness of NAFTA for illustrating issues that are important for characterizing the impact of modern FTAs. Since NAFTA, FTAs have been becoming increasingly expansive with regards to hard-to-quantify non-tariff provisions. An appraisal of "NAFTA at 20" in 2014 by the U.S. Congressional Research Service (Villareal & Fergusson, 2014) reserves special praise for the lasting influence NAFTA has left on a "new generation of trade agreements" as a model for how to incorporate, among other things, guarantees on freedom of investment, intellectual property rights, and investment property rights, as well as increased cooperation on the setting of regulatory standards and customs procedures. Acknowledging the opaqueness of these issues, Baier, Bergstrand, & Clance (2015) have recently shown that using *ex post* estimates of FTA effects offers a promising way of predicting the *ex ante* effects of future agreements for pairs of countries with similar configurations. What my findings suggest in relation to theirs is that, even within the same agreement, there exists important, under-appreciated variation in how the same agreement may affect different countries asymmetrically.

3 Estimation Approach and Data

This section describes the techniques and data sources used in the analyses that follow. The methodology draws on Anderson & Yotov (2016) as a basic starting point, but goes beyond their approach in allowing for heterogeneous effects within and across FTAs, as I will discuss.

3.1 Estimation Approach

The "structural gravity" equation, as generalized by Head & Mayer (2014), naturally motivates a panel fixed effects estimation strategy for identifying the impact of free trade agreements on trade. Fixed effects gravity models of this type have been widely used in the literature for estimating the effects of FTAs. The general framework assumes an R country world with K sectors and costly trade in differentiated goods within each sector. Exports from i to j in sector k (X_{ij}^k) then can be expressed via the following gravity equation, which will explicitly motivate the estimation that follows:

$$X_{ij}^k = \frac{Y_i^k}{\Omega_i^k} \cdot \frac{E_j^k}{\Phi_j^k} \cdot \phi_{ij}^k. \tag{1}$$

E_j^k and Y_i^k here are, respectively, j's expenditure on industry k and the value of i's production in k. The longstanding logic of "gravity", which dates back to Tinbergen (1962), is that trade is increasing in the size of the two countries (i.e., E and Y) and decreasing in the trade costs between them, which here would be reflected in the bilateral parameter ϕ_{ij}^k. To complete the analogy to Newtonian gravity, ϕ_{ij}^k can be said to vary inversely with how far apart i and j are geographically.

Also playing an important role in (1), however, are Φ_j^k and Ω_i^k, which themselves have the following structural interpretations:

$$\Phi_j^k = \sum_i \frac{\phi_{ij}^k}{\Omega_i^k} \cdot Y_i^k \tag{2}$$

$$\Omega_i^k = \sum_j \frac{\phi_{ij}^k}{\Phi_j^k} \cdot E_j^k. \tag{3}$$

Intuitively, these so-called "multilateral resistance" terms index the total incidence of trade costs on an individual country's ability to access world markets, both on the buyer side (in the case of Φ_j^k) and on the seller side (in the case of Ω_i^k). The more easily a producer in i is able to sell to world markets in general, the less inclined he will be to sell to any one particular destination j for any given level of bilateral trade costs ϕ_{ij}^k. A similar logic applies for buyers: better access to sellers around the world all else equal makes them less likely to buy from any one particular exporter i. These structural terms were originally introduced in Anderson & van Wincoop (2003), but are common to a surprisingly wide class of models that fall under the heading "structural gravity" and have different structural interpretations in each case.[9]

The advantage of presenting the gravity model in this way is that it is very general; thus it will allow me to make claims about both identification and welfare implications that will generalize across many different theoretical settings. In particular, I will use a multi-sector "Armington" model (with nationally differentiated

[9] The expression for Ω_i^k, (3), is usually not shown in the presentation of these other models. Nonetheless, Head & Mayer (2014) show it is a general result that follows from an accounting identify for any model where (1) and (2) already hold.

products) in order to simulate general equilibrium outcomes. However, the system defined by (1)-(3) can also be used to describe other gravity models founded on (for instance) within-industry comparative advantage (Eaton & Kortum, 2002), monopolistic competition (Krugman, 1980), or variable firm productivity and endogenous export entry (Melitz, 2003).[10]

For empirical purposes, my main parameter of interest is ϕ_{ij}^{k}, the parameter reflecting how trade costs directly affect trade between i and j. Without loss of generality, $\phi_{ij}^{k} \in (0,1)$ can be thought of as the amount of "market access" that sellers in i enjoy when attempting to sell their variety of good k to import market j: when trade integration lowers trade barriers, ϕ_{ij}^{k} increases and trade in turn increases proportionately with the change in market access, all else equal. The empirical question I am looking to examine is how bilateral market access depends on the presence of a free trade agreement between i and j and how these market access effects may vary within the same agreement. I will refer continually in this paper to $\%\Delta\phi_{ij}^{k} - 1$ as the "amount by which market access increased" in industry k as the result of an agreement. Without loss of generality, I can also call $1 - \%\Delta\left(1/\phi_{ij}^{k}\right)$ as the "amount by which barriers to market access fell".

One last point to clarify about estimating ϕ_{ij}^{k} in this setting is that the combined system (1)-(3) is "modular" (or "separable"). That is, even though the values for production and expenditure for each sector Y_i^k and E_j^k depend on what occurs across all sectors in general equilibrium, if I simply take these terms as given, it follows from (1)-(3) that I can treat the ϕ_{ij}^{k}'s in each individual industry k as an independent set of parameters to be estimated separately, with no cross-equation restrictions across industries.

Let $\phi_{ij,t}^{k}$ indicate the value of ϕ_{ij}^{k} at time t. Following Baier & Bergstrand (2007), I assume $\phi_{ij,t}^{k}$ can be specified in the following manner for each sector:

$$\phi_{ij,t}^{k} = e^{\delta^k Z_{ij} + \beta^k FTA_{ij,t}}, \tag{4}$$

where $FTA_{ij,t}$ is an indicator variable (or set of indicator variables) reflecting whether i and j have an FTA at time t and Z_{ij} is a set of controls for inherent bilateral characteristics assumed to have some effect on trade (e.g., the distance between i and j, whether they share a common language, whether they have a prior colonial relationship, presence of a common border, etc.).

Together, (1) and (4) specify my baseline estimating equation for trade in each industry k:

$$X_{ij,t}^{k} = \exp\left(\xi_{i,t}^{k} + \psi_{j,t}^{k} + \eta_{ij}^{k} + \beta^k FTA_{ij,t}\right) + \varepsilon_{ij,t}^{k}. \tag{5}$$

Note how the identification of β^k is driven by the particular fixed effects structure used in (5). $\xi_{i,t}^{k}$ and $\psi_{j,t}^{k}$ are *time-varying* exporter and importer fixed effects which absorb the Y_i^k, E_j^k, Ω_i^k and Φ_j^k terms in (1) (which themselves are endogenous to trade), leaving only the trade cost term ϕ_{ij}^{k}. By extension, these terms also absorb the effects of each country's "comparative advantage" (by controlling for changes in production

[10] The derivation of "gravity" from the Melitz (2003) theoretical model is generally credited to Chaney (2008).

patterns across industries) as well as the effects of a change in exchange rates (by controlling for the value of each individual country's currency).

Accordingly, the (symmetric) pair-specific fixed effect term η_{ij}^{k} is meant to capture all *time-invariant* bilateral relationships between i and j that influence trade (effectively absorbing $\delta^{k} Z_{ij}^{k}$ in (4)).[11] In panel estimation terms, due to the presence of η_{ij}^{k}, β^{k} essentially serves as a "within" fixed effects estimator for the effect of an FTA on exports from i to j for goods of type k.[12] As Baier & Bergstrand (2007) demonstrate, the use of pair-specific fixed effects in a panel gravity setting is a simple-to-apply procedure for identifying the average treatment effect of FTAs and this approach has become standard in the literature. In accordance with Santos Silva & Tenreyro's (2006) recommendations for minimizing bias in gravity estimations, I will use the Poisson Pseudo-Maximum Likelihood (PPML) estimator to estimate (5).[13]

Allowing for heterogeneity. For my main specifications, I break with the Baier & Bergstrand approach by allowing $FTA_{ij,t}$ to vary by agreement and, subsequently, by the direction of trade flows. Specifically, I focus on NAFTA as a suitable example to show that the effect of an FTA can be conditional on the direction of trade. To do this, I split the single $FTA_{ij,t}$ term in (5) into a set of variables, as shown below:

$$X_{ij,t}^{k} = \exp\left(\xi_{i,t}^{k} + \psi_{j,t}^{k} + \eta_{ij}^{k} + \beta^{0,k} FTA_{ij,t}^{0} + \beta^{N,k} NAFTA_{ij,t}\right) + \varepsilon_{ij,t}^{k}. \tag{6}$$

Here, the superscript "0" on $FTA_{ij,t}^{0}$ is meant to indicate that $\beta^{0,k}$ is now measuring the average effect of all FTAs aside from NAFTA. $\beta^{N,k}$ is then measuring all FTA effects that are associated specifically with NAFTA. I can then split the $NAFTA_{ij,t}$ variable even further in order to isolate directional effects. For example, I will allow $NAFTA_{CANMEX,t}$ to be a single dummy for post-NAFTA exports from Canada to Mexico and $NAFTA_{MEXCAN,t}$ to be a separate dummy for post-NAFTA flows in the other direction. NAFTA is a three country agreement, so there will be 6 directional NAFTA effects to measure in all for each sector, plus the $FTA_{ij,t}^{0}$ term to control for the average effect of all other FTAs in effect. The directional asymmetries in trade barrier reductions these $\vec{\beta}^{N,k}$ terms reveal can then be compared with what one might expect to observe based on tariff reductions associated with NAFTA.

In this last case, I need to be careful. The pair-wise fixed effect η_{ij}^{k} is intended to identify the average effect of an FTA on average trade barriers in industry k for a given pair. However, when I allow NAFTA's effects to be directional I am now (for example) interested in the specific effect of NAFTA on trade frictions

[11] Baier & Bergstrand specifically motivate η_{ij}^{k} as controlling for the propensity of i and j to "select endogenously" into an FTA. The key point, however, is that cross-sectional estimates of FTA effects are biased because i and j's propensity to form an FTA may be correlated with unobservable aspects of the trade costs between them. η_{ij}^{k} resolves the endogeneity problem by explicitly controlling for these unobservable trade costs.

[12] By construction, these pair fixed effects are symmetric. I relax this restriction later when I introduce direction-specific effects.

[13] A more natural approach would seemingly be to estimate (5) via OLS. Santos Silva & Tenreyro show that log-linearizing (5) introduces an important source of bias due to heteroskedasticity in ε_{ij} and measurement error in trade flows, which PPML estimation helps to minimize. Furthermore, as shown in Fally (2014), the PPML estimator is especially suitable for "structural" gravity estimation in particular, since using PPML to estimate (5) implicitly imposes that the structural equations (2) and (3) hold with equality.

for Canadian Food producers trying to sell in US markets, and vice versa. If trade barriers for Canadian Food producers selling in the US are different than those faced by US producers trying to sell in Canada, then $\beta^{N,Food}_{USCAN}$ and $\beta^{N,Food}_{CANUS}$ will in part reflect this initial difference in trade barriers, rather than identifying the differences in how the FTA played out.

As such, I write down this last empirical model as follows,

$$X^k_{ij,t} = \exp\left(\xi^k_{i,t} + \psi^k_{j,t} + \eta^k_{ij} + \vec{\eta}^k_{ij} + \beta^{0,k} FTA^0_{ij,t} + \vec{\beta}^{N,k} NAF\vec{T}A_{ij,t}\right) + \varepsilon^k_{ij,t}, \tag{7}$$

where the "arrow" superscript indicates a set of effects that is allowed to vary by direction. The additional (asymmetric) pair effects $\vec{\eta}^k_{ij}$ are only in play for flows between the US, Canada, and Mexico. $\vec{\beta}^{N,k}$, the set of directional NAFTA effects, by definition also varies with the direction of flows.[14]

This same simple procedure could be easily repeated for any FTA or set of FTAs in order to identify direction-specific effects. NAFTA is an especially useful illustrating example for my purposes, however, not just because of its continuing notoriety in current trade policy debates, but also because its three country structure will offer the opportunity to make unique inferences about the observable patterns of FTA effects within industries. Furthermore, as we will see, the availability of tariff data surrounding NAFTA makes NAFTA a useful example for demonstrating how asymmetric effects within FTAs may differ from what we might expect to observe.

One last point needs to be addressed regarding how the specification used in (7) distinguishes the effects of NAFTA from that of Mexico's peso devaluation, which occurred in December of 1994. As noted above, the inclusion of time-varying exporter- and importer- fixed effects effectively controls for any country's exchange rate relative to a baseline currency (e.g., the U.S. dollar). This specification thus closely resembles the strategy used by Krueger (1999) for distinguishing the effects of NAFTA from that of the devaluation. In both cases, time-variation in Mexico's trade to NAFTA partners relative to its trade with other partners identifies the effects of NAFTA specifically.

3.2 Data

The data used here builds on the data set used in Anderson & Yotov (2016). This data set spans the period 1990 until 2002 for a sample of 40 individual countries plus an aggregate "Rest of the World", for a total of 41 trading regions in all. The main source for trade flows is the CEPII "TradeProd" data base, supplemented with data on exports from UN COMTRADE accessed using the WITS World Bank trade service. The original data uses observed trade flows from the years 1990, 1994, 1998, and 2002 only. I then add additional data—referring to the original sources and construction methods—for the years 1992, 1996, and 2000, such

[14] Estimates are very similar if I use a full set of directional pair dummies as well as if I estimate different individual effects for every single FTA. I assume symmetry for non-NAFTA pairs for computational convenience and because it provides more efficiency in estimating the (symmetric) "other FTA effect", $\beta^{0,k}$ and generally speeds up the estimation.

that the full data set is for every two years.[15] The number of observations for each of the main results shown then is 11,765—i.e., 1,681 trading pairs (the square of the number of regions) times 7, the number of years in the data.

The level of aggregation for the sectoral results is the ISIC (Revision 2) 2 digit level, which is comprised of 9 2 digit manufacturing industry classifications: 31. Food and Beverages, 32. Textiles, 33. Wood Products, 34. Paper Products, 35. Chemicals, 36. Minerals, 37. Metals, 38. Machinery, and 39. Other Manufacturing. However, since some countries report some Machinery products under Other Manufacturing and vice versa in their output data, these two sectors are combined into a single "Manufacturing" category in the final trade data.

A key feature of this data is the inclusion of internal trade flows (i.e., "domestic sales"). Related studies by Dai, Yotov, & Zylkin (2014), Bergstrand, Larch, & Yotov (2015), and Anderson & Yotov (2016) each emphasize that, because international trade barriers are measured relative to internal trade barriers, changes in internal trade should be taken into account in obtaining "theory-consistent" estimates of FTA effects. In addition, one cannot perform a true general equilibrium analysis without some form of accounting for domestic sales. These flows are constructed as the difference between total sectoral output and total sectoral exports to all trading partners. Because exports are measured on a "gross" (rather than value-added) basis, the data likewise uses gross output data for these purposes. Like the trade data, the output data is mainly taken from TradeProd and then supplemented with another source, in this case the United Nations UNIDO Industrial Statistics ("IndStat") database. Missing internal trade values have been extrapolated by comparing the share of internal trade with respect to output in non-missing sectors and non-missing years.[16] Each country's total expenditure on a given industry, which plays an important role in the welfare analysis, can then be calculated by adding together internal trade and total imports.

The data on FTAs is taken mainly from Baier & Bergstrand (2007) and updated with additional data from the WTO's web site. Because the trade data begins in 1990, only FTAs that entered into effect after that year are coded. NAFTA, which went into effect in 1994, is obviously included, but the Canada-U.S. Free Trade Agreement of 1989 (CUSFTA, which preceded NAFTA) is not.[17] Overall, there are 252 country pairs in the data that entered into either a free trade agreement or customs union during the period under study.[18]

[15] The reason why I do not include every year is because, as Cheng & Wall (2005) point out, performing fixed effects gravity estimations over consecutive years may fail to address the fact that trade patterns may not adjust right away to changes in trade costs. For this reason, it is worth mentioning that the results I show in this paper are robust to using four year intervals.

[16] Anderson & Yotov (2010) describe how to impute missing internal trade values.

[17] This is regrettable, since CUSFTA went into effect only 5 years before NAFTA. Accordingly, I place more emphasis on analyzing the estimates for NAFTA pairs involving Mexico.

[18] Customs Unions are included as FTAs for these purposes. Preferential Trade Agreements (PTAs) are not considered.

Table 1: Industry-Level Results: NAFTA vs. Other FTAs

	Food	Textile	Wood	Paper	Chemicals	Minerals	Metals	Machinery
A. Sectoral FTA Estimates								
All FTAs	0.451	0.713	0.006	-0.017	0.228	0.192	0.447	0.467
	(0.076)**	(0.120)**	(0.072)	(0.057)	(0.041)**	(0.065)**	(0.062)**	(0.118)**
B. Individual FTA Estimates (NAFTA vs. All Other FTAs)								
NAFTA	0.504	1.166	0.140	0.371	0.462	0.604	0.339	0.613
	(0.067)**	(0.063)**	(0.095)	(0.032)**	(0.030)**	(0.051)**	(0.136)*	(0.144)**
All Other FTAs	0.465	0.646	-0.038	-0.050	0.183	0.133	0.405	0.371
	(0.082)**	(0.118)**	(0.068)	(0.052)	(0.032)**	(0.056)*	(0.060)**	(0.082)**
C. Significance Tests (NAFTA vs. the Average FTA)								
NAFTA vs. Average	0.040	0.520	0.177	0.422	0.280	0.471	-0.067	0.242
	(0.103)	(0.107)**	(0.117)	(0.062)**	(0.038)**	(0.076)**	(0.150)	(0.125)+

Sectoral trade for 41 trading regions from 1990-2002, every 2 years ($41^2 \cdot 7 = 11,765$ observations per sector.)
Robust standard errors, clustered by pair, are reported in parentheses. + $p < 0.10$, * $p < .05$, ** $p < .01$.
Estimates obtained using Poisson PML, w/ exporter-by-time, importer-by-time, and exporter-by-importer FEs.

4 Illustrating Heterogeneity: The Case of NAFTA

In this section, I apply the empirical techniques and data described in Section 3 to structurally estimate the effects of NAFTA, in the context of a comparison with the average FTA effect. I examine these effects at the industry-level first, in order to illustrate broad patterns of heterogeneity both across FTAs as well as within NAFTA. Section 4.2 then focuses more generally on the overall asymmetries within NAFTA and the question of "one-way" trade effects.

4.1 Industry-Level Analysis

I present my main evidence in Tables 1 and 2. First, in Table 1, Panel A, I estimate average sectoral FTA effects using equation (5) in order to document how the effects of FTAs generally vary significantly across industries.[19] These initial results serve as a useful baseline for establishing how much variation is missed out on by looking solely at average FTA effects.

Panels B and C then explore the implications of allowing for more specific FTA effects. In Panel B, following equation (6), I separate out the individual effect of NAFTA from all other FTAs. The estimates from Panel B reveal that the effects of NAFTA are quite different from both the average sectoral FTA estimates from Panel A and from the estimates of effects of all other FTAs, excluding NAFTA (the "All Other FTAs" term in Panel B). For example, I find that NAFTA has led to a significant increase in trade in Paper products among the three NAFTA members, while the effect of all other FTAs is still negative, small, and marginally significant. Panel C confirms the significance of these differences. These results support the hypothesis that

[19]Qualitatively, these results are virtually identical to Anderson & Yotov's (2016) results for this same specification, and differ only because I include additional years in the data.

Table 2: Industry-Level Results: Directional NAFTA Effects

	Food	Textile	Wood	Paper	Chemicals	Minerals	Metals	Machinery
Directional FTA Estimates (NAFTA)								
NAFTA Can-Mex	1.614	1.725	1.850	0.517	1.163	1.806	0.587	1.400
	(0.165)**	(0.125)**	(0.211)**	(0.116)**	(0.091)**	(0.163)**	(0.114)**	(0.133)**
NAFTA Mex.-Can.	-0.249	1.544	0.754	1.769	1.931	1.472	1.674	1.291
	(0.164)	(0.115)**	(0.231)**	(0.122)**	(0.088)**	(0.164)**	(0.115)**	(0.200)**
NAFTA Can-U.S.	0.595	1.348	0.209	0.081	0.412	0.643	0.330	0.441
	(0.069)**	(0.108)**	(0.181)	(0.080)	(0.086)**	(0.100)**	(0.098)**	(0.064)**
NAFTA U.S.-Can	0.510	0.832	-0.197	0.767	0.457	0.667	-0.028	0.481
	(0.067)**	(0.104)**	(0.182)	(0.079)**	(0.078)**	(0.101)**	(0.095)	(0.045)**
NAFTA Mex.-U.S.	-0.445	1.592	0.768	0.371	0.536	0.620	1.098	1.235
	(0.162)**	(0.128)**	(0.176)**	(0.104)**	(0.064)**	(0.153)**	(0.104)**	(0.207)**
NAFTA U.S.-Mex	1.000	0.938	0.123	0.293	0.512	0.395	0.415	0.727
	(0.163)**	(0.132)**	(0.148)	(0.099)**	(0.060)**	(0.152)**	(0.102)**	(0.119)**
All Other FTAs	0.478	0.645	-0.041	-0.050	0.186	0.130	0.405	0.377
	(0.081)**	(0.118)**	(0.068)	(0.052)	(0.032)**	(0.056)*	(0.060)**	(0.079)**

Sectoral trade for 41 trading regions from 1990-2002, every 2 years ($41^2 \cdot 7 = 11,765$ observations per sector.)
Robust standard errors, clustered by pair, are reported in parentheses. + $p < 0.10$, * $p < .05$, ** $p < .01$.
Estimates obtained using Poisson PML, w/ exporter-by-time, importer-by-time, and exporter-by-importer FEs.

individual FTAs can have very different effects across industries.

Then, in Table 2, I allow for country-specific and directional differences in the effects of NAFTA, as in (7). These latter results confirm that FTA effects can indeed be strongly directional. For example, my estimates for Food reveal that NAFTA had strong positive effects on Canadian exports to Mexico (NAFTA Can-Mex= 1.614, std.err 0.165) and on U.S. exports to Mexico (NAFTA U.S.-Mex=1.000, std.err 0.163). However, I do not obtain statistically significant estimates of the effects of NAFTA on Mexico's exports to Canada; Furthermore, Mexico's exports to the U.S. actually seem to have been negatively affected.[20]

Estimates from several other sectors exhibit similar regularities, with different orderings. Mexican Paper exports to Canada increase a prodigious amount, whereas other countries' exports expand more modestly. We also see this same kind of pattern where one country seems to generally gain the most access to the other two markets in Metals and Machinery (again Mexico in both cases).

One might be tempted to interpret these asymmetries in how NAFTA has affected trade in different industries as reflecting some measure of "comparative advantage". Perhaps these results are due to the U.S. and Canada having comparative advantage in Food, Mexico having comparative advantage in Machinery, and so on, the reasoning might go. Again, however, I call attention to the presence of time-varying exporter and importer fixed effects in my empirical specification. There is little question that trade liberalization impacts relative prices across industries and that changes in relative prices in turn cause factors of production to be

[20]The negative sign for NAFTA's effect on Mexico's exports to the U.S. is hard to interpret. One possible explanation is that NAFTA required relatively non-committal co-ordination on food safety standards (Hooker, 1999).

re-allocated. But the structural model in (1) explicitly controls for the endogeneity of production and prices with respect to trade costs. These estimates I am showing are indeed changes in market access as defined in the structural model, controlling for all endogenous responses to trade within each country. This includes notions of comparative advantage.[21]

4.2 A "One Way" Trade Deal?

What can we say then about the possibility of "one way" effects in general? The industry-level data is suggestive of certain patterns at a high level—Mexico seems to have had especially strong gains as an exporter for instance (especially in the Textiles, Wood, Metals, and Machinery sectors). However, to really make concrete statements one way or the other, it is necessary to examine results for aggregate trade. In Table 3, I show a number of specifications meant to highlight the broad heterogeneity in FTA effects both between NAFTA and other FTAs and within NAFTA itself.

In column 1, I document that the average effect of FTAs on trade in my sample, across all industries combined is 0.401, which corresponds to an average increase in market access of $e^{0.401} - 1 = 49.3\%$. Column 2 then replicates the specification from Table 1 Panel B. Unsurprisingly, NAFTA appears to have promoted trade much more overall than the other FTAs in my sample—about twice as much in fact.[22] Furthermore, as I show explicitly in Column 3, the difference between NAFTA and the average FTA is itself significant at the 5% level.

Columns 4 to 6 then isolate the directional impact of NAFTA on aggregate trade flows. The patterns seen here for combined trade feature less pronounced asymmetries than we saw in the industry-level results. While there is some definite variation across pairs (trade between Mexico and Canada vs. trade between the U.S. and Canada), the asymmetry seen within pairs is muted. In columns 5 and 6, we see for instance that each country generally tended to receive more or less the same amount of access to its import markets that it offered to its partners in return. One exception here—where we are able to say there have been some asymmetric effects—is the effect of NAFTA on trade between the U.S. and Mexico. The results in column 4 show that Mexico overall has received a $e^{0.999} - 1 = 171\%$ increase in access to U.S. markets, whereas the U.S. has received an $e^{0.615} - 1 = 85.0\%$ increase in return.[23] This last result is particularly interesting because it goes directly against what would have been predicted to happen based on tariffs. As was well-publicized by the pro-NAFTA campaign at the time, Mexico's tariffs on U.S. products were about two-and-a-half times

[21] In theory, a similar argument also should apply to trade in intermediate goods. As an example, consider the many-country trade model with cross-sectoral input-output linkages from Caliendo & Parro 2015. In their model, a shock to trade barriers in one industry affects the price levels of inputs across all industries—an importer-side effect—which leads to a general expansion of production in industries with more favorable access to inputs. By assumption, this general expansion of production raises exports to all destinations; thus, it is simply captured as an exporter-side effect. In practice, however, FTAs may encourage the formation and expansion of pair-specific supply chain relationships. Section 5 discusses how this dynamic may lead to asymmetries in the resulting estimates for FTA effects, particularly in the context of NAFTA.

[22] $e^{0.580} - 1 = 78.6\%$; $e^{0.319} - 1 = 37.6\%$.

[23] These increases in market access can also be described as a 63.2% decrease in US import barriers for Mexican products vs. a 45.9% decrease in Mexican import barriers for U.S. products, as stated in the opening remarks.

Table 3: NAFTA Effects, Aggregate Trade

	(1)	(2)	(3)	(4)	(5)	(6)
Average FTA Effect	0.401		0.320			
	(0.065)**		(0.042)**			
All Other FTAs		0.319		0.318	0.317	0.321
		(0.042)**		(0.042)**	(0.042)**	(0.042)**
NAFTA		0.580	0.261			
		(0.095)**	(0.100)**			
NAFTA Can-Mex				1.182		
				(0.073)**		
NAFTA Mex.-Can.				1.099		
				(0.074)**		
NAFTA U.S.-Can				0.470		
				(0.044)**		
NAFTA Can-U.S.				0.451		
				(0.044)**		
NAFTA Mex.-U.S.				0.999		
				(0.071)**		
NAFTA U.S.-Mex				0.615		
				(0.069)**		
NAFTA (Canadian exports)					0.443	
					(0.042)**	
NAFTA (Mexican exports)					1.076	
					(0.066)**	
NAFTA (U.S. exports)					0.507	
					(0.044)**	
NAFTA (Canadian imports)						0.417
						(0.064)**
NAFTA (Mexican imports)						0.845
						(0.103)**
NAFTA (U.S. imports)						0.586
						(0.080)**

Aggregate manuf. trade for 41 trading regions from 1990-2002, every 2 years ($41^2 \cdot 7 = 11,765$ observations.)
Robust standard errors, clustered by pair, are reported in parentheses. + $p < 0.10$, * $p < .05$, ** $p < .01$.
Estimates obtained using Poisson PML, w/ exporter-by-time, importer-by-time, and exporter-by-importer FEs.

those on products shipped in the other direction.[24] It would seem then that focusing on tariff provisions does not necessarily do a good job of explaining how much market access will change in response to the signing of a free trade agreement; I return to this theme in the ensuing discussion in section 5.

4.3 Issues Related to Timing

Another relevant finding from Baier & Bergstrand (2007) is the fact that the effects of FTAs tend to "phase in" (i.e., accumulate) over time. Using 5- and 10-year lags on their FTA variable, they show that the measured effect of an FTA on trade doubles when one accounts for how trade adjusts over time to the signing of an FTA. This finding bears on my analysis of NAFTA for two main reasons. First, Canada and the U.S. already had an FTA in place at the time of NAFTA—the 1989 Canada-U.S. Free Trade Agreement ("CUSFTA"). If the effects of CUSFTA were still being phased in in the early 90s, it admittedly may not be clear how to disentangle CUSFTA's phasing in effects from NAFTA's independent effect on Canada-U.S. Trade. Second, Mexico alone formed 35 FTA pairs during the sample period; if these FTAs also exhibited accumulation effects, then the econometric model may not be accurately identifying the determinants of relative trade flows for Mexico in a given year.

To deal with these general issues, I start by offering some further analysis of aggregate trade flows. The experiments in Table 4 explore how to account for the possible phasing-in effects of CUSFTA, as well as more general sources of trend that may be present in the data. In column 1, I set up a baseline, where all directional NAFTA effects are measured in relation to the average FTA effect. These results embed the asymmetries documented in Table 3 and also show that each of 6 directional effects of NAFTA seem to be large relative to the effect of the average FTA. In column 2, I introduce a (2 year) lead dummy for U.S.-Canada trade ("Can-U.S. pre-NAFTA"), which turns out to be significant. I also note that including this dummy cuts directly into the observed effect of NAFTA on US-Canada trade. While it is not clear whether the Can-U.S. lead-in dummy captures phasing-in due to CUSFTA or possible anticipation effects for NAFTA, it does cast some doubt on the measured effect of NAFTA on U.S.-Canada trade.

To take a stand on this issue, I suppose that the possible phase-in pattern of CUSFTA resembles that of the average FTA. For columns 3-5, I introduce 4- and 8-year lagged dummies for the existing presence of a signed FTA (including for NAFTA).[25] Thus, the "Average FTA Effect (after phase-in)" coefficient is meant to report the effect of an FTA after allowing for 8 years of phase-in. In turn, the results for the directional NAFTA dummies are now measured relative to that of a phased-in FTA. I also use the 4- year lag for CUSFTA to apply in the year 1992, a way of accounting for the phasing-in of CUSFTA that may have occurred between 1988/1989 and 1992.[26] Indeed, in the presence of these phase-in dummies, the

[24] See for instance "NAFTA: Good for America", *Washington Post*, July 4th, 1993.

[25] It is possible to use other combinations of lagged dummies to obtain similar results. Using 4- and 8- year lags generally yields significant estimates for each of the lagged terms (other combinations do not) and is mostly consistent with Anderson & Yotov (2016) (who also use 4- and 8- year lags) and Baier & Bergstrand (2007) (who use 5- and 10- year lags).

[26] The timing of CUSFTA is somewhat unusual in that it was signed in January, 1988 and not implemented until a full year later,

Table 4: Accounting for FTA Phase-in and Lead-in Effects, Aggregate Trade

	Baseline		FTA effects "phase in" over time		
Effects of NAFTA on Directional Trade Flows (in Deviations from the Average FTA Effect)					
NAFTA Can-Mex	0.863	0.856	0.749	0.817	0.787
	(0.079)**	(0.078)**	(0.078)**	(0.097)**	(0.099)**
NAFTA Mex.-Can.	0.780	0.780	0.685	0.751	0.711
	(0.079)**	(0.079)**	(0.084)**	(0.100)**	(0.102)**
NAFTA U.S.-Can	0.151	0.038	0.121	0.168	0.132
	(0.058)**	(0.059)	(0.055)*	(0.082)*	(0.084)
NAFTA Can-U.S.	0.132	0.015	0.099	0.142	0.106
	(0.057)*	(0.058)	(0.055)+	(0.078)+	(0.080)
NAFTA Mex.-U.S.	0.679	0.679	0.580	0.647	0.607
	(0.077)**	(0.077)**	(0.081)**	(0.097)**	(0.099)**
NAFTA U.S.-Mex	0.296	0.296	0.186	0.261	0.231
	(0.075)**	(0.075)**	(0.075)*	(0.098)**	(0.099)*
Average FTA Effect	0.319	0.319			
	(0.048)**	(0.048)**			
Average FTA Effect			0.620	0.614	0.651
(after phase-in)			(0.069)**	(0.070)**	(0.073)**
Can-U.S. pre-NAFTA		0.227	-0.031		
		(0.011)**	(0.045)		
$NAFTA_{t+1}$				0.114	0.018
				(0.101)	(0.103)
All $FTAs_{t+1}$					0.095
					(0.024)**

Aggregate manuf. trade for 41 trading regions from 1990-2002, every 2 years ($41^2 \cdot 7 = 11,765$ observations.)
Robust standard errors, clustered by pair, are reported in parentheses. $^{+}$ $p < 0.10$, * $p < .05$, ** $p < .01$.
Columns 3—5 incorporate 4- and 8-year lagged FTA dummies, such that all FTA effects "phase in" over time.
These FTA lags also account for pre-existing trends due to pre-1990 FTAs, such as the 1988 Canada-U.S. FTA.
All estimates obtained using Poisson PML, w/ exporter-by-time, importer-by-time, and exporter-by-importer FEs

Table 5: Industry-Level and Aggregate Trade with Linear Time Trends

	Food	Textile	Wood	Paper	Chemicals	Minerals	Metals	Machinery	Aggregate
Directional FTA Estimates (NAFTA)									
NAFTA Can-Mex	-0.052	0.309	0.929	0.366	0.477	2.012	-0.531	0.015	0.139
	(0.153)	(0.236)	(0.209)**	(0.091)**	(0.205)*	(0.090)**	(0.192)**	(0.122)	(0.076)+
NAFTA Mex.-Can.	-0.587	-0.408	0.944	0.067	1.373	0.515	0.950	0.316	0.254
	(0.154)**	(0.236)+	(0.226)**	(0.119)	(0.185)**	(0.093)**	(0.211)**	(0.142)*	(0.083)**
NAFTA Can-U.S.	0.060	0.436	0.396	0.449	0.485	0.640	0.024	0.418	0.425
	(0.084)	(0.119)**	(0.115)**	(0.054)**	(0.184)**	(0.064)**	(0.182)	(0.098)**	(0.061)**
NAFTA U.S.-Can	0.173	0.306	0.410	-0.065	0.195	0.827	-0.036	-0.257	0.066
	(0.085)*	(0.132)*	(0.115)**	(0.051)	(0.207)	(0.067)**	(0.184)	(0.091)**	(0.060)
NAFTA Mex.-U.S.	-0.507	-0.052	0.670	-0.300	0.275	-0.130	0.429	0.255	0.131
	(0.139)**	(0.237)	(0.202)**	(0.090)**	(0.058)**	(0.096)	(0.153)**	(0.112)*	(0.070)+
NAFTA U.S.-Mex	0.025	0.398	0.123	0.287	0.299	0.339	0.052	0.414	0.335
	(0.138)	(0.217)+	(0.188)	(0.087)**	(0.056)**	(0.092)**	(0.140)	(0.083)**	(0.056)**
All Other FTAs	0.269	0.174	-0.057	0.021	0.034	-0.001	0.180	0.052	0.110
	(0.047)**	(0.063)**	(0.056)	(0.034)	(0.026)	(0.034)	(0.042)**	(0.033)	(0.024)**

Sectoral trade for 41 trading regions from 1990-2002, every 2 years ($41^2 \cdot 7 = 11,765$ observations per sector.)
Robust standard errors, clustered by pair, are reported in parentheses. + $p < 0.10$, * $p < .05$, ** $p < .01$.
Estimates obtained using Poisson PML, w/ exporter-by-time, importer-by-time, and exporter-by-importer FEs.

estimate for the lead dummy for U.S.-Canada trade vanishes to nearly zero. In addition, while the Average FTA Effect nearly doubles when we account for phase-in (consistent with the original finding from Baier & Bergstrand, 2007), the difference between NAFTA's effects and the average FTA effect remains significant, if slightly diminished. Furthermore, the observed asymmetries within NAFTA remain qualitatively the same as in column 1. It follows that using average phase-in effects seems to be a reasonable means for drawing a distinction between NAFTA's effects and the possible lingering effects of CUSFTA. It also follows that allowing for phase-in of outside FTAs does not alter the main results.

Column 4 then generalizes the analysis of pre-NAFTA trade by introducing a general NAFTA lead dummy, NAFTA_{t+1}. The estimated coefficient for this lead dummy is positive, if not statistically significant ($p = .260$). To establish a baseline for comparison, in column 5 I examine how the observed lead-in effect associated with NAFTA differs from the average lead-in effect for all FTAs ("All FTAs_{t+1}"). Interestingly, it seems to be a fairly general feature of the data that agents within countries are able to anticipate FTAs before they go into effect and consistently begin investing in trade opportunities just before the actual signing occurs.[27] It also seems that, to the extent the same anticipation effect may have occurred in the case of NAFTA, it is no different than other FTAs; the independent effect observed for NAFTA_{t+1} shrinks to basically zero.

Lastly, in Table 5, I apply a more rigorous approach to accounting for general trends in trade barriers

in 1989. The assumption that CUSFTA's effects phased in from 1988 is for simplicity. Computing 3- year lags for all FTAs leads to similar results.

[27]This finding that trade tends to increase slightly 2 years before the actual signing of an FTA is consistent with detailed analysis of the timing of FTA effects performed by Baier, Bergstrand, & Feng (2014) (in their supplementary appendix).

by adding linear time trends for all NAFTA trading pairs, as well as for all outside pairs. Noting that FTAs themselves are a major source of trends in the trade data (via their phase-in effects), it is unsurprising that all FTA effects are measured to be significantly smaller in this last specification. Interestingly, we also see some reversals. Allowing for pre-existing trends switches the sign of the asymmetry for U.S./Mexico trade, both for aggregate trade and for several industries (most notably, Machinery, the largest industry in the data). On the other hand, the measured effect of NAFTA on Mexico/Canada trade is now more puzzling: whereas before (in Table 3), we observed a relatively symmetric effect for Mexico/Canada, the results with time trends in 5 suggest a much more asymmetric effect favoring Mexico's exports.

As we know from Table 5, it is not unusual for FTAs to have prior effects just before they go into effect; however, it is not immediately clear how to interpret the reversal of the sign for NAFTA's effect on U.S./Mexico trade. Obviously, we cannot discount the possibility that asymmetries in trade barrier reductions between the U.S. and Mexico during the 1990s are entirely explained by the persistence of pre-NAFTA trends, rather than directly due to NAFTA itself. Another possibility, however, is that NAFTA actually served to reinforce these earlier trends. Indeed, as Tornell & Hernández (1997) and Whalley (1998) have argued, the primary benefit of NAFTA from Mexico's perspective on Mexico/U.S. trade may have been for Mexico to demonstrate to foreign investors its commitment to "locking in" its own prior reforms towards liberalization (and liberalization of FDI from the U.S. in particular).

Nonetheless, in broad terms, the results shown here remain very similar to those obtained previously: the average FTA Effect is positive and significant and we still observe that NAFTA has promoted trade more strongly than the average FTA.[28] Furthermore, we still see a strong asymmetry between NAFTA's effects on Mexico's exports to Canada versus its imports from Canada, which is again strongly contradictory to expectations based on *ex ante* tariffs.

5 A NAFTA Tariff Puzzle?

From a quantitative modeling perspective, a standard expectation is that asymmetries in FTA effects should reflect differences in *ex ante* tariffs and other policy barriers to trade. This section reviews the data on industry-level tariffs for the three NAFTA countries. Even at the industry level, I find no meaningful correlation between the *ex post* estimates of NAFTA's effects and the effects that would have been implied by *ex ante* protection, especially for trade to and from Mexico.

To dig deeper into this discrepancy, I consider various possible explanations for why one might observe similar asymmetries in the effects of FTAs (and in the effects of NAFTA in particular). A particularly important consideration is the peso devaluation that occurred within Mexico during this period. The question of what role trade in intermediate inputs may play in the analysis also needs to be examined. Intriguingly, other

[28] The *p*-value for the test that NAFTA as a whole has promoted trade more as compared to the average FTA in this last specification is 0.027.

Table 6: Pre-NAFTA Tariff Protection Levels

Exporter	Importer	Food	Textiles	Wood	Paper	Chemicals	Minerals	Metals	Machinery	Aggregate
A. Pre-NAFTA Import Tariffs (*ad valorem*, Effectively Applied)										
Canada	Mexico*	10.02	15.29	14.95	5.17	10.09	13.72	7.14	14.08	9.62
Canada	U.S.	2.40	7.52	1.11	0.28	3.79	1.56	1.93	0.38	0.96
Mexico	Canada	5.32	13.67	13.18	3.14	6.39	3.32	0.71	4.14	3.88
Mexico	U.S.	6.20	13.16	4.05	3.57	5.07	5.83	3.38	3.81	4.31
U.S.	Canada	5.24	16.5	6.98	3.19	7.50	6.44	4.31	5.40	5.55
U.S.	Mexico*	10.81	16.98	14.91	6.46	10.13	14.44	10.03	13.99	11.74

Weighted avg. "effectively applied" tariffs from UNCTAD TRAINS tariff database, accessed via WITS

*Mexican tariffs only available for 1991 in Panel A. All other Panel A tariffs from 1989.

FTAs between developed/less-developed pairs generally exhibit the same (puzzling) pattern of asymmetry observed in NAFTA.

5.1 Comparing NAFTA Effects with pre-NAFTA Tariffs

Table 6, Panel A, shows industry-level *ad valorem* "effectively applied" tariff levels for each country pair before NAFTA, using data from the UNCTAD TRAINS tariff database. All values shown are from the year 1989, with the exception of Mexican tariffs, which use 1991 because of missing data for 1989 and 1990. For simplicity, I focus on which country had the highest pre-NAFTA tariff in a given industry. While it is not appropriate to use raw tariffs to compare implications for trade across sectors—because different sectors have different trade elasticities—within a given sector, one might expect larger tariff reductions to correspond with larger trade increases.[29]

The first point to document concerns the aggregate tariff levels in the right-most column of Panel A. Based on these tariffs, we would have expected, if anything, trade barriers to fall the most for Mexico's import markets, since Mexico had by far the highest tariffs to start with and also reduced its tariffs by the largest amount amount overall. Clearly, however, this expectation is not borne out in the estimates from Section 4.

Turning to the industry-level tariffs, one can find some cases where tariffs line up more closely with the estimates reported above in Table 2. In the Wood sector, for example, the U.S. only had a 1.11% tariff on Canadian Wood products before NAFTA—whereas Mexico's was 14.95%—and indeed Canada's Wood exports to Mexico increased by more than its exports to the U.S. did. Estimates in the Food sector in Table 2 do indeed correspond to asymmetries in initial Food tariffs (as well as the NTBs on Food products shown in Panel B). Generally speaking, however, there is major variation in the effects of NAFTA that cannot be explained based on tariffs. In the Metals sector, for instance, we cannot rationalize Mexico's large gains

[29] Admittedly, focusing on initial tariffs only may overstate the magnitude of each country's short- and medium-term tariff cuts. I have also computed the effects of tariff *changes* (as opposed to complete tariff removal) using tariff data from 1999 and 2002. Mexico does indeed lower its tariffs more slowly than its NAFTA partners, but the overall implications remain the same. Using estimates of (tariff equivalent) non-tariff barrier levels from Shikher (2012*a*) also leads to the same insights.

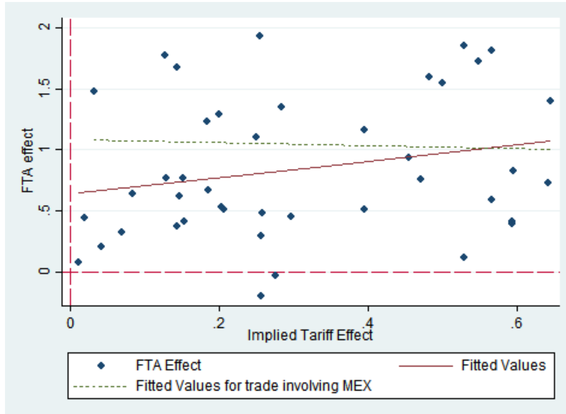

Figure 1: Comparing NAFTA Estimates vs. Implied Tariff Effects

in market access as an exporter of Metals ($\beta^{N,Metals}_{MEX,CAN} = 1.674$ in Table 2) based on tariffs alone. There is likewise no clear reason to be gleaned from Table 6 why U.S. and Canadian import markets for Paper, Chemicals, Minerals, and Machinery opened up as much as they did.

To summarize these comparisons more succinctly, Figure 1 compares the industry-level estimates of NAFTA's effects from Table 2 with the implied FTA effects from removing initial industry-level tariffs. To enable these comparisons, the "implied effect" of removing tariffs in a given industry is calculated using industry-level import elasticities from Broda, Greenfield, & Weinstein (2006). While there is a slight positive relationship between tariff effects and FTA effects, the fit is nowhere near the one-to-one relationship one might expect. Furthermore, this relationship is especially weak for trade to and from Mexico, again reflecting the surprising asymmetry found in the gravity estimates.

5.2 Reviewing Possible Explanations

Exchange rate volatility. An important factor that comes to mind in the case of NAFTA is the sharp devaluation of the Mexican peso that began in December 1994. How can NAFTA's effect be separated from the dramatic change in U.S.-Mexico exchange rate?

Krueger (1999) also considers this question in her own *ex post* analysis of trade creation under NAFTA. By comparing changes in Mexican *export shares* to U.S. and Canadian markets vs. changes in its share as exporter to non-NAFTA import markets, she observes that NAFTA did not significantly promote Mexican exports to the U.S. moreso than it did its exports to non-NAFTA destinations; she concludes that Mexico's currency devaluation was the main contributing cause of the increase in Mexico-U.S. exports during the period. She does not, however, specifically consider the question of whether NAFTA may have had asymmetric effects on trade.

In my approach, which directly incorporates the reasoning of Krueger (1999), the use of time-varying exporter- and importer- fixed effects directly controls for the effect of fluctuations in each country's currency on its exports and imports. It is undeniable that overall Mexican exports grew strongly during the period, likely in no small part due to the peso. Other countries' currencies likewise fluctuated in value during the period. The time-varying country fixed effects in the specification should be absorbing how these fluctuations would generally change each country's disposition toward trade, both as an importer and as an exporter (and across the different industries for the industry-level analysis).

Trade in intermediate inputs. Even at the industry-level, the data used in this study does not provide substantial information regarding trade in "intermediate inputs" versus trade in "final output". Given the well-known production-sharing arrangements between the U.S. and Mexico (and Canada for that matter) across certain industries (automobiles for instance), it is important to consider how the segmentation of production is accounted for by the model.

Structural gravity frameworks which incorporate trade in intermediates—e.g., Eaton & Kortum (2002); Shikher (2012*b*); Caliendo & Parro (2015)—can be used to account for how different types of trade may analyzed together based on aggregated data. The structure of these models (which is embedded in the empirical design described above) dictates that cheaper access to intermediate inputs, regardless of their origin, should generally lower production costs for importing firms and therefore promote more exports to *all* destinations, not just the original supplier of their inputs.

Admittedly, however, firm-to-firm relationships in reality may not always adhere to the structure imposed on aggregated trade flows by current modeling approaches. Firms who operate within well-defined bilateral production-sharing relationships likely respond very differently to liberalization of imported inputs than other firms who serve as general export platforms.

Under this alternative perspective, a symmetric reduction in tariffs could (seemingly paradoxically) show up in the data as promoting a country's exports back to its original supplier moreso than imports from that supplier. This logic would follow because the processed goods being returned to the original supplier of intermediates are presumably higher in the value chain and therefore represent a larger nominal quantity of trade in the data. An FTA's effect on "offshoring costs" (perhaps via investment provisions or settlement dispute provisions) could have a similar effect. Further consideration of this interesting possibility using more disaggregated data is reserved for future work.

5.3 Evidence from Other FTAs

This analysis thus far suggests an interesting puzzle behind NAFTA's effects, but how general is it? To briefly widen the scope of the analysis, I estimate one final specification for the average FTA effect along the lines of (5), only with an added interaction term between the $FTA_{ij,t}$ dummy and the difference in development levels for the exporter and importer countries (i.e., the difference in their GDP per capita). The basic hypothesis is that that FTAs between highly developed and less developed countries may generally exhibit the same asymmetry seen in NAFTA.[30] To add a more explicitly economic motivation, we may think of FTAs as making it easier for the more advanced country to offshore assembly tasks to the less developed country, in order to take advantage of lower production wages.

The results of this basic test are shown in Table 7, column 1. Strikingly, the results for all FTAs in

Table 7: FTA Effects for Developed/Less Developed Pairs

Average FTA Effects on Aggregate Trade (PPML)	
FTA dummy	0.412
	(0.069)**
FTA dummy * ($\ln y_i - \ln y_j$)	-0.103
	(0.018)**
N obs.	11,765
Fixed Effects	(i,t); (j,t); $(\vec{i,j})$
Standard Errors	robust; clustered by pair

y_i: exp. GDP/capita; y_j: imp. GDP/capita
+ $p < 0.10$, * $p < .05$, ** $p < .01$.

my sample are very much consistent with my results for NAFTA.[31] The negative and (strongly) significant coefficient on the interaction term confirms that FTAs between developed/less-developed pairs tend to favor exports more for less-developed partners. Furthermore, it is notable that this result again contradicts what we would expect based on the logic of tariffs, since the degree of tariff protection is known to vary inversely with the degree of development.

This basic analysis obviously could be extended to explore richer interactions between institutional indicators and specific features of FTAs (how Investor Settlement Dispute mechanisms may affect the enforcement of contracts within supply chains, for example). My intent here though is to document some suggestive evidence for why the effects I observe in NAFTA differ from expectations based on quantifiable policy measures; sharper characterizations of how FTAs interact with institutions to produce asymmetric effects are reserved for future work.

6 Welfare Analysis

In this section, I illustrate that these large directional differences in direct FTA effects in turn can also have important implications for general equilibrium welfare analysis of FTA formation. The simulation structure I use here is the same as in Anderson & Yotov (2016), where welfare effects occur strictly through changes in buyer and producer prices and accordingly may be thought of as "terms of trade effects". This

[30] GDP per capita data is taken from the CEPII "gravity" data set made available in conjunction with Head & Mayer (2014). Aside from simply "wages", GDP per capita could also proxy for the quality of a country's institutions. Waugh (2010) has also shown that developing countries generally face higher trade barriers than developed countries in reaching developed country import markets. I remain agnostic on the exact interpretation.

[31] Results are similar with or without NAFTA. Results are also similar with linear time trends, as in Table 5.

particular approach to simulating counterfactuals is especially suitable for demonstration purposes here for two main reasons: (i) because the impact of trade agreements on domestic producers specifically is often a key issue in trade policy debates, and (ii) because the first-order price effects it identifies would hold important implications for the additional welfare channels addressed in other quantitative models.

6.1 Model

I start by specifying a multi-sector Armington (1969) trade model where all products are differentiated based on the country of origin. Each industry in each country possesses an endowment of a differentiated good q_i^k, which it sells to the world at factory gate price p_i^{*k}. Consumers have common CES preferences over the set of nationally-differentiated varieties within each industry, with elasticity parameter σ^k. It follows that nominal exports from i to j of goods of type k can be written as

$$X_{ij}^k = \left(\frac{\mu_i^k p_i^{*k} t_{ij}^k}{P_j^k} \right)^{1-\sigma^k} \cdot E_j^k, \tag{8}$$

with t_{ij}^k the standard "iceberg cost" of transporting goods of type k from i to j and $P_j^k \equiv \sum_i (\mu_i^k p_i^{*k} t_{ij}^k)^{1-\sigma^k}$ the CES price index for buyers in j. μ_i^k is the standard "Armington" demand shifter for products from a certain origin.[32]

It is straightforward to re-write the CES Demand function in (8) in the form of a structural gravity equation as in (1) by adjusting the notation. Let $\phi_{ij}^k \equiv (t_{ij}^k)^{1-\sigma^k}$ serve as the parameter measuring bilateral market access and let $\Phi_j^k \equiv (P_j^k)^{1-\sigma^k}$. It follows then from an accounting identity that $(\mu_i^k p_i^{*k})^{1-\sigma^k} = Y_i^k / \Omega_i^k$, as in (1), with Ω_i^k still defined as in (3) as a measure of access to world markets for sellers in i.[33] The structural interpretation for Φ_j^k in (2) likewise follows from the same accounting identity.

Furthermore, using the fact $(\mu_i^k p_i^{*k})^{1-\sigma^k} = Y_i^k / \Omega_i^k$—with $Y_i^k = p_i^k q_i^k$—enables me to write:

$$p_i^{*k} \;=\; \widetilde{\mu}_i^k \left(\Omega_i^k \right)^{\frac{1}{\sigma^k}}, \tag{9}$$

where $\widetilde{\mu}_i^k = (\mu_i^k)^{1-\sigma^k/\sigma^k} (q_i^k)^{-1/\sigma^k}$ is a combined parameter describing how prices reflect the quantity and quality of i's endowment in industry k, varying negatively with the size of its overall supply, q_i^k, but positively with its embedded demand shifter, $(\mu_i^k)^{1-\sigma^k}$.

Note that, by the duality between Φ_j^k and the buyer price index, and by the dependence of seller prices on

[32] μ_i^k is typically interpreted as (an inverse measure of) the "embedded quality" of i's production technology for products of type k.

[33] Specifically, if $Y_i = \sum X_{ij} = S_i \left(\sum_j \phi_{ij} E_j / \Phi_j \right)$, then $S_i = Y_i / \Omega_i$ follows directly.

Ω_i^k in (9), we have the following intuitive relationships between price changes and access to world markets:

$$\Delta \ln p_i^{*k} = \frac{1}{\sigma^k} \Delta \ln \Omega_i^k \qquad\qquad \Delta \ln P_j^k = \frac{1}{1-\sigma^k} \Delta \ln \Phi_j^k$$

That is, sellers face higher aggregate demand for their products when they are more able to reach world markets (higher Ω_i^k) and, by $\sigma^k > 1$, buyers likewise enjoy lower prices when there is more competition in their import markets (higher Φ_j^k). Regarded separately, these two effects allow for a simple decomposition of how trade integration affects buyers vs. producers in each economy. Regarded together, the combined effect provides a useful notion of the "terms of trade" effects from liberalization.

Adapting Anderson & Yotov's method, I generate a solvable R-by-K system of equations for factory gate prices $\{p_i^{*k}\}$ by summing (8) over j and dividing by world output. The resulting expression is

$$\frac{Y_i^k}{Y^k} = \sum_j \frac{\left(\mu_i^k p_i^k\right)^{1-\sigma^k} \phi_{ij}^k}{\Phi_j^k} \cdot \frac{E_j^k}{Y^k}, \tag{10}$$

where Y_i^k, E_j^k, Y^k, and Φ_j^k are each themselves functions of prices. To close the model, I assume that each country's expenditure on manufacturing goods is a constant share δ_j of its total manufacturing income $\sum_k Y_i^k$ and that buyers in each country allocate expenditure across sectors according to a Cobb-Douglas function with common share parameters $\{\alpha^k\}$.

To simulate general equilibrium FTA effects within this structure, I impose the normalization that all initial seller prices p_i^{*k} are equal to one, such that all endowments $\{q_i^k\}$ are given by initial 1990 output levels; that is, $q_i^k = Y_{i,1990}^k$. By then introducing a new system of ϕ_{ij}^k's reflecting NAFTA, I can recover changes in buyer and producer prices, terms of trade, and welfare using the linkages implied by the structural model.[34] Specifically, I solve (10) for $\{p_i^{*k}\}$ subject to the following:

$$\frac{Y_i^k}{Y^k} = \frac{p_i^{*k} q_i^k}{\sum_k p_i^k q_i^k} \tag{11}$$

$$\frac{E_j^k}{Y^k} = \frac{\delta_j \sum_k p_i^{*k} q_i^k}{\sum_j \delta_j \sum_k p_j^k q_j^k} \tag{12}$$

$$\Phi_j^k = \sum_j \left(\mu_i^k p_i^{*k}\right)^{1-\sigma^k} \phi_{ij}^k. \tag{13}$$

In practice, I solve (10)-(13) twice: once with 1990 trade costs and all $p_i^{*k} = 1$ in order to recover initial values for the combined $\{(\mu_i^k)^{1-\sigma^k}\}$ terms, and then again for new p_i^{*k}'s using new post-NAFTA ϕ_{ij}^k's. In addition, I take values for elasticity parameters $\{\sigma^k\}$ from data compiled by Broda, Greenfield, & Weinstein

[34] For the initial ϕ_{ij}^k's, I follow Anderson & Yotov (2016) in using the pair fixed effects from gravity estimation (η_{ij}^k's). Specifically, I use the η_{ij}^k's recovered from estimating equation (7).

(2006). I also construct expenditure share parameters $\{\delta_j\}$ and $\{\alpha^k\}$ from 1990 trade and output data.[35]

Finally, I construct aggregate measures for changes in buyer prices, seller prices, and welfare in each country i. Aggregate buyer prices in i ("$\mathbf{P}_i$") are simply a function of aggregate access to world markets in each individual industry, aggregated by expenditure shares. That is, $\mathbf{P}_i \equiv \prod_k (\Phi_i^k)^{\alpha^k/1-\sigma^k}$. Country i's welfare, $\mathbf{W}_i \equiv \mathbf{P}_i^{-1} \sum_k p_i^{*k} q_i^k$, is then directly analogous to its real income. Intuitively, national welfare increases with seller prices (as the price of one's own endowments increases) and decreases with buyer prices (as access to other countries' endowments improves). To compute aggregate changes in national supplier prices ("$\mathbf{p}_i$"), I simply compare $\sum_k q_i^k$ with $\sum_k p_i^{*k} q_i^k$.[36] The overall change in welfare can thus be neatly decomposed into buyer- and seller-side effects on domestic prices: $\Delta \ln \mathbf{W}_i = \Delta \ln \mathbf{p}_i - \Delta \ln \mathbf{P}_i$.

6.2 NAFTA's Terms of Trade Effects

The first three panels of Table 8 then show the simulation results for the effects of NAFTA on price indices and welfare for each of the three specifications shown in Tables 1 and 2 for each NAFTA country and for an aggregate non-NAFTA group. All changes shown are relative to the 1990 baseline year. Price and welfare changes for the non-NAFTA group are aggregated by expenditure shares. To establish a baseline, I first note the main features of Panel A, which assumes NAFTA has been no different than any other FTA. Notably, Mexico and Canada show larger net positive benefits from NAFTA than the U.S. This result is intuitive because Canada and Mexico are both relatively small countries compared with the U.S., with Mexico seeing the largest gains because it is the smallest of the three. The Rest of the World, which is also very large, only suffers very mildly from trade diversion.

When I allow for NAFTA to have its own distinct effect on trade costs, however (Panel B), we see favorable differences for both buyers and sellers in all three NAFTA countries, with each country realizing roughly 50% larger implied gains than in the baseline case. These results not only re-affirm the message of Panel B in Table 1—that NAFTA has been significantly more effective than the average FTA in promoting trade—but also quantify that message in terms of real income effects. What this approach shows is that using the simplifying assumption of a single average FTA effect across all agreements can greatly distort the true welfare gains from trade liberalization for individual countries, even large ones. NAFTA, of course, is just one example; focusing on other agreements instead would likely reveal further large differences, both positive and negative.

Panel C displays price and welfare results implied by the asymmetric NAFTA effects from Table 2. Noting that Mexico had the largest gains in both exports and imports in Table 3, it is unsurprising that the differences in welfare changes between Panels B and C heavily favor Mexico, and that these gains are

[35]Note that all α^k terms cancel out of (12). However, these terms will still be needed for constructing aggregate buyer prices and welfare.

[36]One last normalization is needed when we endogenize Y and E because the system in (10) is homogeneous of degree zero. So I impose $\sum_i p_i^k q_i^k = Y_{1990}^k$ both pre- and post-NAFTA, such that total nominal production in each sector stays the same over time.

Table 8: Initial Terms of Trade Effects (NAFTA only)

	A. Average FTA effect			B. Average NAFTA effect		
	%ΔSupplier Prices	%ΔBuyer Prices	%ΔWelfare	%ΔSupplier Prices	%ΔBuyer Prices	%ΔWelfare
Canada	1.54	-0.69	2.24	2.08	-1.30	3.42
Mexico	3.15	-1.96	5.21	3.86	-3.87	8.05
U.S.	0.08	-0.15	0.23	0.12	-0.22	0.34
ROW	-0.10	-0.09	-0.01	-0.14	-0.12	-0.02
	C. Directional NAFTA effects			D. Using Tariffs		
	%ΔSupplier Prices	%ΔBuyer Prices	%ΔWelfare	%ΔSupplier Prices	%ΔBuyer Prices	%ΔWelfare
Canada	1.62	-1.09	2.74	-0.50	-1.38	0.89
Mexico	7.47	-4.32	12.32	0.57	-4.29	5.07
U.S.	0.12	-0.21	0.33	0.20	0.08	0.12
ROW	-0.15	-0.13	-0.02	-0.06	-0.05	-0.01

Supplier prices and welfare for the Rest of World ("ROW") region aggregated by output shares.

Buyer prices for ROW aggregated by expenditure shares.

All changes relative to 1990 price and welfare levels, assuming fixed supply quantities.

especially large for Mexican sellers in particular. Clearly, Mexico has benefited the most from NAFTA out of the three countries. It is also reasonable that these gains have come largely at the expense of Canada, who has had the smallest overall increase in market integration from NAFTA, again going by columns 5 and 6 of Table 3.

Outcomes for the U.S. and the Rest of the World, however, are much less sensitive to these asymmetries. In part, this lack of an effect is simply due to the fact that the U.S. and ROW are both very big regions. But given the politicization of NAFTA as a supposed "one way" trade deal, the lack of an apparent difference across Panels B and C for either U.S. buyers and U.S. sellers overall is notable. Clearly, as one can infer from Table 2, asymmetries within NAFTA work in favor of U.S. seller prices in some industries (e.g., Food, Paper, Chemicals), but fall for others (e.g., Textiles, Wood, Metals). That the differences in the overall price indices turn out to be negligible reflects the observation made in Section 4.2 that analyzing asymmetries at the aggregate level understates the heterogeneity in FTA effects that occurs across different industries.

The more interesting comparison, though, is with what would have been expected to happen based on tariffs. The last panel of Table 8 shows the counterfactual for prices and welfare if changes in trade costs were driven solely by the tariff changes shown in Table 6.

As we can see, estimating welfare effects based on tariff changes not only greatly underestimates the the magnitudes of the overall gains from NAFTA for all three countries, it also tends to mischaracterizes the composition of each country's gains. As Panel D shows, estimates based on tariffs alone would suggest U.S. producers benefited from NAFTA at the expense of producers in the other two countries, as well as at the expense of U.S. buyers. Based on Table 6, the reason for the large difference in the results between Panel D and the other three panels (particularly Panels B and C) is straightforward: projections based on tariffs alone would suggest that, contrary to the econometric estimates, the U.S. sellers should have gained a relatively large increase in access to the other two NAFTA markets (not just Mexico's import market, but Canada's as well).

Thus, welfare benefits from NAFTA, and the attribution of those benefits, unsurprisingly turn out to be very sensitive to the underlying trade cost changes used to estimate them. The asymmetries I have identified in NAFTA's effects have generally favored Mexico's terms of trade at the expense of Canada's, while not necessarily having much of an affect (on average) on either buyers or sellers in the U.S. A particularly dramatic difference can be seen between the price and welfare effects implied by the tariff changes associated with NAFTA versus those implied by *ex post* estimates. It follows that tariffs may do a poor job of predicting how FTAs shape the gains from trade and that more work should be done understanding other means by which FTAs can affect trade.

7 Closing Remarks

The question of how much free trade agreements actually increase trade between countries remains an important, actively studied topic in the trade literature. While most empirical work has focused on identifying how different types of agreements affect trade on average, my refined approach makes it plain that some FTAs increase trade more than others and, furthermore, that the effects of FTAs are not necessarily symmetric between partners within a particular sector. Accounting for NAFTA's particular effects on trade in a general equilibrium simulation setting reveals quantifiably large implications for welfare: ignoring these variations both understates the overall gains from trade and also misses important differences in the incidence of price effects for buyers vs. sellers in each country.

A particularly notable empirical finding is that the patterns of market liberalization I identify within NAFTA do not correspond with what would have been expected to happen based on tariffs. This finding turns out to be quite general; differences in the level of development between partner countries have played an important, overlooked role in determining asymmetric effects within FTAs. It is not surprising then that basing welfare predictions on tariff reductions seems to fundamentally mischaracterize NAFTA's first-order effects on prices. In particular, I show that U.S. sellers did not realize the benefits that would be predicted based on tariffs; these predicted benefits for the U.S. instead accrued (and then some) to U.S. buyers.

I envision there being significant potential for future work both expanding on the approach I have pursued in this paper and examining the reasons for the differences I obtain. My findings for trade within NAFTA motivate a larger-scale project geared towards characterizing how FTAs with particular provisions tend to affect trade differently for countries with different observable characteristics. For example, further work may consider how FTAs with strong investment provisions may affect trade barriers differently depending on differences in the level of financial development across member countries. It is also worth knowing whether some countries simply tend to gain more market access as a result of FTAs than others, all else equal. These unsettled issues aside, the main messages that this paper intends to communicate are: (i) directional heterogeneity in FTA effects may differ substantially from what we might expect based on tariffs alone; and (ii) accounting for heterogeneity in FTA effects, especially directional heterogeneity within FTAs, can have important implications for how we quantify the gains from trade.

References

Agama, L.-A. & McDaniel, C. A. (2002), "The NAFTA Preference and US-Mexico Trade", Working Paper 2002-10-A, USITC Office of Economics.

Anderson, J. E. & van Wincoop, E. (2001), "Borders, Trade and Welfare", Working Paper 8515, National Bureau of Economic Research.

Anderson, J. E. & van Wincoop, E. (2003), "Gravity with Gravitas: A Solution to the Border Puzzle", *American Economic Review* **93**(1), 170–192.

Anderson, J. E. & Yotov, Y. V. (2010), "Specialization: Pro- and Anti-globalizing, 1990-2002", Working Paper 16301, National Bureau of Economic Research.

Anderson, J. E. & Yotov, Y. V. (2016), "Terms of Trade and Global Efficiency Effects of Free Trade Agreements, 1990-2002", *Journal of International Economics* **(in press)**.

Arkolakis, C., Costinot, A., & Rodríguez-Clare, A. (2012), "New Trade Models, Same Old Gains?", *The American Economic Review* **102**(1), 94–130.

Armington, P. S. (1969), "A Theory of Demand for Products Distinguished by Place of Production", *Staff Papers (International Monetary Fund)* **16**(1), 159–178.

Baier, S. L. & Bergstrand, J. H. (2007), "Do free trade agreements actually increase members' international trade?", *Journal of International Economics* **71**(1), 72–95.

Baier, S. L., Bergstrand, J. H., & Clance, M. (2015), "Heterogeneous Economic Integration Agreements", mimeo.

Baier, S. L., Bergstrand, J. H., & Feng, M. (2014), "Economic integration agreements and the margins of international trade", *Journal of International Economics* **93**(2), 339–350.

Behrens, K., Mion, G., Murata, Y., & Südekum, J. (2014), "Trade, Wages, and Productivity", *International Economic Review* **55**(4), 1305–1348.

Bergstrand, J. H., Larch, M., & Yotov, Y. V. (2015), "Economic Integration Agreements, Border Effects, and Distance Elasticities in the Gravity Equation", *European Economic Review* **78**.

Bhagwati, J. N., Krishna, P., & Panagariya, A. (1999), *Trading Blocs: Alternative Approaches to Analyzing Preferential Trade Agreements*, MIT Press.

Broda, C., Greenfield, J., & Weinstein, D. (2006), "From Groundnuts to Globalization: A Structural Estimate of Trade and Growth", Working Paper 12512, National Bureau of Economic Research.

Brown, D. K., Deardorff, A. V., & Stern, R. M. (1992), "North American Integration", *The Economic Journal* **102**(415), 1507–1518.

Caliendo, L. & Parro, F. (2015), "Estimates of the Trade and Welfare Effects of NAFTA", *Review of Economic Studies* **82**(1), 1–44.

Chaney, T. (2008), "Distorted Gravity: The Intensive and Extensive Margins of International Trade", *The American Economic Review* **98**(4), 1707–1721.

Cheng, I. H. & Wall, H. J. (2005), "Controlling for Heterogeneity in gravity models of trade", *Federal Reserve Bank of St. Louis Review* **87**(1), 49–63.

Cipollina, M. & Salvatici, L. (2010), "Reciprocal Trade Agreements in Gravity Models: A Meta-Analysis", *Review of International Economics* **18**(1), 63–80.

Costinot, A. & Rodríguez-Clare, A. (2014), "Trade Theory with Numbers: Quantifying the Consequences of Globalization", *Handbook of International Economics* **4**, 197–261.

Dai, M., Yotov, Y. V., & Zylkin, T. (2014), "On the trade-diversion effects of free trade agreements", *Economics Letters* **122**(2), 321–325.

Eaton, J. & Kortum, S. (2002), "Technology, Geography, and Trade", *Econometrica* **70**(5), 1741–1779.

Egger, P., Larch, M., Staub, K. E., & Winkelmann, R. (2011), "The Trade Effects of Endogenous Preferential Trade Agreements", *American Economic Journal: Economic Policy* **3**(3), 113–143.

Fally, T. (2014), "Structural Gravity and Fixed Effects", Mimeo.

Head, K. & Mayer, T. (2014), "Gravity equations: Workhorse, toolkit, and cookbook", *Handbook of International Economics* **4**, 131–196.

Hooker, N. H. (1999), "Food safety regulation and trade in food products", *Food Policy* **24**(6), 653–668.

Kohl, T. (2014), "Do we really know that trade agreements increase trade?", *Review of World Economics* **150**(3), 443–469.

Kohl, T., Brakman, S., & Garretsen, H. (2013), "Do trade agreements stimulate international trade differently? Evidence from 296 trade agreements", Working Paper 4243, CESifo.

Krueger, A. O. (1999), “Trade creation and trade diversion under NAFTA”, Working Paper 7429, National Bureau of Economic Research.

Krugman, P. (1980), “Scale Economies, Product Differentiation, and the Pattern of Trade”, *The American Economic Review* **70**(5), 950–959.

Mayer, T., Paillacar, R., & Zignago, S. (2008), “Tradeprod. the cepii trade, production and bilateral protection database: Explanatory notes”, .

Melitz, M. J. (2003), “The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity”, *Econometrica* **71**, 1695–1725.

Melitz, M. J. & Ottaviano, G. I. P. (2008), “Market Size, Trade, and Productivity”, *Review of Economic Studies* **75**(1), 295–316.

Nicita, A. & Olarreaga, M. (2007), “Trade, Production, and Protection Database, 1976–2004”, *World Bank Econ Rev* **21**(1), 165–171.

Romalis, J. (2007), “NAFTA’s and CUSFTA’s Impact on International Trade”, *Review of Economics and Statistics* **89**(3), 416–435.

Santos Silva, J. M. C. & Tenreyro, S. (2006), “The Log of Gravity”, *Review of Economics and Statistics* **88**(4), 641–658.

Shikher, S. (2012*a*), “Predicting the effects of NAFTA: Now we can do it better!”, *Journal of International and Global Economic Studies* **5**(2), 32–59.

Shikher, S. (2012*b*), “Putting industries into the Eaton–Kortum model”, *The Journal of International Trade & Economic Development* **21**(6), 807–837.

Tinbergen, J. (1962), “An analysis of world trade flows”, in *Shaping the World Economy*, pp. 1–117, Twentieth Century Fund, New York.

Tornell, A. & Hernández, G. E. (1997), “The Political Economy of Mexico’s Entry into NAFTA”, in *Regionalism versus Multilateral Trade Arrangements, NBER-EASE Volume 6*, pp. 25–56, University of Chicago Press.

Trembley, P. (2013), “Canada-Mexico (Trade and Investment)”, Publication 2013-46-E, Library of Parliament (Canada), Ottawa.

Villareal, M. A. & Fergusson, I. F. (2014), “NAFTA at 20: Overview and Trade Effects”, Federal Publication, Congressional Research Service, Washington, DC.

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

Waugh, M. (2010), "International Trade and Income Differences", *The American Economic Review* **100**(5), 2093–2124.

Whalley, J. (1998), "Why Do Countries Seek Regional Trade Agreements?", in J. A. Frankel (ed.) *The Regionalization of the World Economy*, pp. 63–90, University of Chicago Press, Chicago.

Data Appendix

Trade. The data from Anderson & Yotov (2016) covers trade, production (i.e., gross output), and expenditure values for 64 countries for 8 ISIC (revision 2) 2 digit manufacturing industries for the years 1990, 1994, 1998, and 2002. The industries and countries included are shown in Table 9. Note that, for computational reasons, 24 countries which do not sign free trade agreements during the period are combined into a single aggregate "Rest of the World" region, leaving 41 trading regions. The main data sources for trade and production are the CEPII TradeProd database (Mayer, Paillacar, & Zignago 2008), UN COMTRADE, and UNIDO INDSTAT, with the latter two sources used to supplement trade and production values (respectively) that are missing from TradeProd.

"Internal trade" is then constructed as the difference between a country's gross output and total gross exports in a given industry. Any case where industry-level production is less than gross exports is treated as a missing value, in which case internal trade is interpolated using non-missing values for that country. These issues arise only in a small number of cases and, furthermore, only for non-NAFTA countries. See Anderson & Yotov (2010) for a description of the methods used to interpolate missing internal trade values. Trade data for the years 1992, 1996, and 2000 is obtained using the same data sources and procedures.

FTAs. The data on FTAs includes 252 FTA pairs that took place between the years 1990 and 2002, as well as information on prior FTA pairings. Table 10 lists the included agreements and the year each agreement went into effect. The original source for these FTAs is Baier & Bergstrand (2007), supplemented with data from the WTO's website.

Other data sources. *Tariffs* are from TRAINS. Pre-NAFTA tariffs use 1989 tariffs for the U.S. and Canada and 1991 for Mexico (since Mexico does not report 1989 tariffs). *Trade elasticities* for each ISIC 2 digit industry are from Broda, Greenfield, & Weinstein (2006). *GDP per capita* is from Head & Mayer (2014).

Table 9: Included Industries and Countries

ISIC Rev. 2 Industries included (8)	
Industry Code	Description
31	Food (includes Beverages and Tobacco)
32	Textiles (includes Leather and Footwear)
33	Wood products (includes Furniture)
34	Paper products (includes Publishing and Printing)
35	Chemicals (includes Fuels, Rubber, and Plastic products)
36	Minerals (includes Glass and other Non-metal products)
37	Metal products
38-39*	Machinery (includes Other manufacturing n.e.c.)

Countries/regions included (64 countries/41 regions)

"Main sample" (40 countries/regions): Argentina, Australia, Austria, Bulgaria, Belgium-Luxembourg, Bolivia, Brazil, Canada, Switzerland, Chile, China, Colombia, Costa Rica, Germany, Denmark, Ecuador, Spain, Finland, France, United Kingdom, Greece, Hungary, Ireland, Iceland, Israel, Italy, Japan, South Korea, Morocco, Mexico, Malta, Netherlands, Norway, Poland, Portugal, Romania, Sweden, Tunisia, Turkey, Uruguay, United States

"Rest of World" (24 countries/regions): Cameroon, Cyprus, Egypt, Hong Kong, India, Indonesia, Iran, Jordan, Kenya, Kuwait, Sri Lanka, Macau, Mauritius, Malawi, Malaysia, Malta, Niger, Nepal, Philippines, Senegal, Singapore, Trinidad & Tobago, Tanzania, South Africa.

*This industry grouping combines two different codes due to reporting and data availability issues.

Table 10: Included Agreements

Agreement	Year	Member Countries
Multilateral Trade Blocs		
EU†	1958	Belgium-Luxembourg, France, Italy, Germany, Netherlands, Denmark (1973), Ireland (1973), United Kingdom (1973), Greece (1981), Portugal (1986), Spain (1986), Austria (1995), Finland (1995), Sweden (1995),
EFTA	1960	Norway, Switzerland, Iceland (1970), Portugal (1960-1986), Austria (1960-1995), Sweden (1960-1995) Finland (1986-1995).
LAFTA/LAIA	1993	Argentina, Bolivia, Brazil, Chile, Ecuador, Mexico.
CEFTA	1993	Poland (1993), Hungary (1993), Romania (1997), Bulgaria (1998)
Andean Community†	1993	Bolivia, Colombia, Ecuador
NAFTA	1994	Canada, Mexico, U.S.
Mercosur*†	1995	Argentina, Brazil, Uruguay

EU's outside agreements: EFTA (1973/1994), Hungary (1994), Poland (1994), Bulgaria (1995), Romania (1995), Turkey (1996), Tunisia (1998), Israel (2000), Mexico (2000), Morocco (2000).

EFTA's outside agreements: Turkey (1992), Bulgaria (1993), Hungary (1993), Israel (1993), Poland (1993), Romania (1993), Mexico (2000), Morocco (2000).

Other agreements: Bulgaria-Israel (1994), Bulgaria-Turkey (1998), Canada-Chile (1997), Canada-Israel (1997), Chile-Costa Rica (2002), Chile-Mexico (1999), Colombia-Mexico (1995), Costa Rica-Mexico (1995), Mercosur-Bolivia (1996), Mercosur-Chile (1996), Hungary-Israel (1998), Hungary-Turkey (1998), Israel-Mexico (2000), Israel-Poland (1998), Israel-Romania (2001), Israel-Turkey (2001), Poland-Turkey (2000), Romania-Turkey (1998)

Earlier agreements: U.S.-Israel (1985); U.S.-Canada (1989)

This table lists the free trade agreements used in the estimation of FTA effects. Customs Unions, such as the EU and Mercosur, are coded the same as FTAs. Preferential trade agreements (or PTAs) are not considered.