Phase Out Tariffs, Phase In Trade?*

Tibor Besedes† — Georgia Institute of Technology
Tristan Kohl‡ — University of Groningen
James Lake§ — Southern Methodist University

April 10, 2019

Abstract

What causes U.S. trade with Mexico and Canada to continue growing faster, for up to a decade, relative to countries with which the U.S. does not have a free trade agreement? Baier and Bergstrand (2007) suggest that tariff phase-out and delayed pass-through of tariffs into import prices could cause such prolonged differential import growth. We examine how tariff cuts negotiated under the Canada-US Free Trade Agreement and the North American Free Trade Agreement (NAFTA) affected U.S. import growth in 1989–2016 using detailed product-level data on tariff phase-out in the original treaties. We find essentially no evidence for the tariff phase-out or delayed pass through explanations. Rather, we find evidence for an important role played by NAFTA tariff cuts reducing the impacts of frictions at various extensive margins.

Keywords: Free Trade Agreements, CUSFTA, NAFTA, trade, phase-out, tariffs, extensive margin.

JEL classification: F1

*The authors would like to thank Adina Ardelan, Scott Baier, Jeffrey Bergstrand, Kristy Buzzard, Douglas Campbell, Shushanik Hakobyan, Benedikt Heid, Mario Larch, Maia Linask, Emanuel Ornelas, Jooyoung Park, Justin Pierce, Jayjit Roy, Charles Sawyer, Georg Schaur, Tim Schmidt-Eisenlohr, Peter Schott, Steffen Sirries, Yoto Yotov, Tom Zylkin as well as seminar and conference participants at Aarhus University, Radboud University Nijmegen, Southern Methodist University, University of Bayreuth, University of North Dakota, Bank of Mexico, Midwest Trade Meetings, ETSG, LETC, and SEA for helpful comments and discussion. Kohl gratefully acknowledges Georgia Tech and SMU's hospitality in hosting him during research visits that have contributed to this paper.

†School of Economics, Georgia Institute of Technology, 221 Bobby Dodd Way, Atlanta, GA 30332-0615, e-mail: besedes@gatech.edu

‡Faculty of Economics and Business, University of Groningen, Nettelbosje 2, 9747 AE Groningen, The Netherlands, e-mail: t.kohl@rug.nl

§Department of Economics, Southern Methodist University, 3300 Dyer Street, Suite 301, Umphrey Lee Center, Dallas, TX 75275, e-mail: jlake@smu.edu

1 Introduction

After some 18 months of politically controversial negotiations, the U.S., Canada, and Mexico recently signed a revised version of the North American Free Trade Agreement (NAFTA) and re-branded it as the United States Mexico Canada Agreement (USMCA). Signed in mid 1992, NAFTA came into effect in 1994 and incorporates the earlier U.S.-Canada Free Trade Agreement (CUSFTA) that was implemented in 1989. As one of the world's largest trade agreements, understanding the economic outcomes of NAFTA is important not only in the current political debate, but also for trade policy analysts and economists in general. Indeed, CUSFTA/NAFTA have been extensively studied to determine how Free Trade Agreements (FTAs) affect their members' trade, output, prices, welfare, and more generally the winners and losers of globalization (e.g. Trefler 2004, Romalis 2007, and Caliendo and Parro 2015).

However, a simple glance at the evolution of CUSFTA/NAFTA's trade flows reveals a well-known puzzle that the FTA literature, to the best of our knowledge, has not resolved. Figure 1 plots cumulative growth of real U.S. imports from Mexico, Canada, and the Rest of the World (ROW) as of CUSFTA's enforcement in 1989.[1] It reveals U.S. imports from Mexico started growing more rapidly, and more rapidly relative to ROW, once NAFTA came into force in 1994 and this effect does not level off until, at least, the early-mid 2000s. A similar story initially holds for U.S. imports from Canada, although the impact is much less pronounced and any differential impact disappears around 15 years after NAFTA during the great trade collapse in the late 2000s. While the phenomenon of FTAs having delayed effects on trade flows goes back to Baier and Bergstrand (2007), there is no systematic evidence in the trade literature on *why* FTAs take so long to fully impact trade flows.

Baier and Bergstrand (2007, p.89-90) suggest two hypotheses to explain these prolonged differential growth rates of real trade flows. Citing the 10-year phase-in of the original European Economic Community (EEC) agreement and NAFTA, their first hypothesis re-

[1] ROW excludes China and countries with which the U.S. formed FTAs over the sample period of 1989-2016.

volves around the observation that "... virtually every [FTA] is 'phased-in,' typically over 10 years." As such, one could naturally expect the impact of an FTA on trade flows to play out gradually over time as bilateral tariffs are slowly phased out over time.

While tariff phase-out has long been familiar to trade economists, incorporating this into empirical work has, somewhat ironically, been delayed. Baier and Bergstrand (2007) introduced "phase-in" effects for FTAs in the gravity model of international trade by including, in addition to the familiar contemporaneous binary FTA variable, one or more lagged FTA variables that span the FTA tariff phase-out period. They find that "on average, an FTA approximately doubles two members' bilateral trade after 10 years" (p. 7).

By now, inclusion of the Baier and Bergstrand (2007) lagged FTA terms has become standard in applied work (see, for example, Baier, Bergstrand, and Feng 2014 and Kohl 2014) and there is consensus that lagged FTA terms do indeed yield positive and statistically significant effects on bilateral trade for 5-10 years after the FTA enters into force. However, a striking limitation of these studies is that they do not explicitly demonstrate the causal relationship between product-level tariff phase-out and product-level trade. A key reason for this is that the product-level "staging categories" that define tariff phase-out are embedded in complicated product-level documents running hundreds of pages and, thus, cannot readily be incorporated in studies spanning multiple countries and FTAs. At best, the lagged FTA terms in aggregate studies can be assumed to capture the delayed trade growth stemming from tariff phase-out, but cannot be interpreted as evidence of a causal relationship.

The second hypothesis of Baier and Bergstrand (2007) involves changes in tariffs being passed through gradually to lower prices over time (i.e. "delayed terms-of-trade" effects). One could naturally expect the impact of an FTA on trade flows to play out over time as tariff cuts filter through to import prices. However, the literature has mixed findings on NAFTA tariff pass-through. Caliendo and Parro (2015) find that U.S. import prices fall because wages rise by less than the fall in materials' prices in Canada and Mexico. Romalis (2007) finds modest effects on U.S. import prices with prices actually *increasing* in highly protected U.S. sectors as NAFTA diverts trade away from more efficient non-members. Similarly, Kikkawa, Mei, and Santamarina (2019) show slight price increases for U.S. imports from Mexico because Mexican exporters raise their markups when NAFTA increases their U.S. market access. However, these papers do not address whether and how tariff stagings embedded in FTAs affect a country's import prices.[2]

Our paper, to the best of our knowledge, is the first to explore the relevance of either Baier and Bergstrand (2007) hypothesis—"tariff phase-out" or "delayed pass-through"–as an explanation for the delayed impact of FTAs on trade flows. To do so, we examine CUSFTA and NAFTA to determine how tariff phase-out affects trade flows and, as proxied by unit values, import prices. Specifically, we merge U.S. product-level import data and unit values (as proxies for prices) with detailed information on the tariff phase-out staging categories agreed by the U.S. in CUSFTA and NAFTA. We use a difference-in-difference-in-difference, or triple-difference (DDD) approach from the applied microeconometrics literature dating back to Gruber (1994). While a triple-difference approach has been used in the trade literature, it has not been used in the FTA literature.[3]

The DDD approach applies naturally in our tariff phase-out context. Essentially, our main empirical strategy looks at import growth from NAFTA partners vis-à-vis ROW for products where tariffs are phased out. But, to control for broader non-tariff related NAFTA effects, our empirical strategy looks at this import growth relative to import growth from NAFTA partners vis-à-vis ROW for products where tariffs are duty free both pre- and post-NAFTA.

Our central finding is that there is no evidence that tariff phase-out or delayed pass-through of tariffs to prices can help explain the delayed growth in trade flows typically observed following FTAs. We do find that imports of products receiving tariff cuts grow more, both in the short-run and long-run, than products where tariffs were already zero

[2] Anderson and Yotov (2016) provide an extensive review of the growing literature addressing the impact of an FTA's reciprocal tariff cuts on the members' terms of trade. They also estimate the terms of trade effects of FTAs in the 1990s for 40 countries.

[3] See Frazer and Van Biesebroeck (2010), Coeli (2018), and Friederich and Zator (2018).

before NAFTA. And, comfortingly, the magnitude of the effects, both within a country for products of different phase-out duration as well as across countries for products with the same phase-out duration, are broadly consistent with differences in the actual country-product specific tariff cuts embodied in NAFTA. However, the bulk of the delayed impact comes from products that *immediately* had their tariff cut to zero and, in the case of Mexico, had their product-level tariff-free access via the GSP program converted into *permanent* tariff-free access. This contrasts starkly with the expectation from the tariff phase-out hypothesis that the bulk of delayed import growth should come from products where tariffs were phased out over 5 or even 10 years. Further, we find no evidence of delayed pass-through effects.

As an alternative explanation of the delayed import growth in the wake of NAFTA, we present evidence that NAFTA tariff cuts reduce the impact of frictions related to the extensive margin of trade, especially for imports from Mexico. Our first piece of evidence is that NAFTA tariff cuts have very little effect on import growth of products continuously imported from Mexico over the entire sample period, i.e. products at the intensive margin. Thus, products at the extensive margin, especially "new products" not imported before NAFTA, drive the delayed impact of NAFTA tariff cuts on import growth in our baseline analysis. Our second piece of evidence relies on U.S. import data that reports the customs district where a product enters consumption channels. Regardless of the specific phase-out category or the intensive versus extensive margin distinction, we see a given imported product from a NAFTA partner gradually spreading out over the U.S. over time when NAFTA cuts its tariff. Of course, one could think of this expansion as a product's "spatial " extensive margin. In this light, a key takeaway from our analysis is that various dimensions of the extensive margin, including the "new products" and "spatial" extensive margins, are crucial to understanding post-NAFTA delayed import growth driven by NAFTA tariff cuts.

Our finding that effects related to the extensive margin drive the delayed growth in trade flows following NAFTA contributes to a growing trade literature emphasizing the role of the extensive margin. In the context of post-NAFTA trade growth, this dates back to at least Hillberry and McDaniel (2002). More generally in the context of FTAs, and moving in the direction of an explanation for the delayed timing of FTAs, Baier, Bergstrand, and Feng (2014, p. 339) find "intensive-margin effects occurring sooner than extensive-margin effects." However, unlike our main analysis, they do not look at the role of tariff phase-out or delayed pass-through in explaining the delayed trade impacts of FTAs.

The broader trade literature has also addressed the interplay between the extensive margin of trade and the timing of trade flows. Ruhl (2008) uses delayed impacts on the extensive margin to reconcile the large long-run trade elasticities found in applied general equilibrium models and empirical analyses of tariffs on trade flows with the small short-run trade elasticities found in international real business cycle models. Albornoz, Pardo, Corcos, and Ornelas (2012) show how firms can learn about their export profit potential over time. Very much like our "spatial extensive" margin results, they show empirically how this generates a firm-level "sequential exporting" phenomenon where successful firms expand their reach across more destination markets over time. Moreover, Arkolakis, Papageorgiou, and Timoshenko (2018) present evidence that firms learn over time about unobserved demand uncertainties in foreign markets. Our contribution to this broad literature is to say that extensive margin effects (either a new products extensive margin or a spatial extensive margin) triggered by FTA tariff cuts, rather than the original Baier and Bergstrand (2007) hypotheses of tariff phase-outs or delayed tariff pass-through to import prices, appear to drive the delayed increase in trade flows observed after FTAs.

2 Data

2.1 U.S. import data

We use the universe of product-level U.S. import data from the United States International Trade Commission (USITC).[4] These data report U.S. imports at the 10-digit HS level from

[4]Our USITC import data are the "imports for consumption" data series.

each foreign country and each import program (e.g. NAFTA, GSP, etc.) over the period 1989-2016. We aggregate these data to the exporter-year 8-digit HS level to match the product-level tariff phase-out data in the CUSFTA and NAFTA texts. After aggregation and cleaning the data for our later use of unit values, we have 4,877,858 observations.[5]

2.2 Tariff schedules

We extract the product-level tariff phase-out data from the original and publicly available CUSFTA and NAFTA treaties. Each treaty contains a tariff schedule for each member. The tariff schedules contain the product-level staging categories that govern how each member phases out tariffs on the other member(s) upon the treaty entering into force.[6] Unfortunately, the U.S. NAFTA tariff schedule often breaks a given 8-digit product into various sub-products that are identified by letters (i.e. not 10-digit HS codes) and have different staging categories. Thus, we cannot match these "Mixed" products to trade data even though they account for a non-trivial 12.7% of products in the U.S. NAFTA tariff schedule.[7]

NAFTA has five standard staging categories. Staging category A immediately cuts tariffs to zero while staging category D reflects products that were already duty free pre-NAFTA and, hence, continue duty free post-NAFTA. Starting January 1, 1994, the other three staging categories phase out tariffs in equal annual stages from the "base rate" as defined

[5]Cleaning the data entails dropping three sets of observations: (i) the 0.0007% of observations with import program "Unknown country", (ii) respectively, the 3.69% and then the 0.9% of observations where an 8-digit product is measured in different units (e.g. volume and weight) for a given exporter-year or for a given exporter, (iii) the 0.006% of observations with positive quantities despite the USITC quantity description stating that the product has no quantity dimension.

[6]The CUSFTA tariff schedules, entering into force on January 1, 1989, are introduced in Chapter 3: *Border Measures* by Article 401: *Tariff Elimination.* The schedules themselves are separately attached as Annex 401.2; the U.S. schedule is 509 pages. The NAFTA tariff schedules, entering into force on January 1, 1994, are introduced in Chapter 4: *National Treatment and Market Access for Goods* by Annex 302.2: *Tariff Elimination of NAFTA*. The schedules themselves are separately attached to Annex 302.2; the U.S. schedule is 734 pages.

[7]Often in agriculture, these "Mixed" products can be used, for example, to impose different tariffs in different months of a calendar year. For example, 0707.00.50 represents Cucumbers imported during May-June or September-November in the USHTS. But, the U.S. NAFTA tariff schedule assigns staging category C+ to 0707.00.50A (defined as imports during May or October-November) and staging category B to 0707.00.50B (defined as imports during June or September).

by the USHTS Column 1 tariff on July 1, 1991.[8] Staging category B does this over five years. Staging categories C and C+ do this over 10 and 15 years respectively. Members also have member-specific NAFTA staging categories. The U.S. tariff schedule defines two such staging categories: B6 and C10. B6 products have their tariff reduced on January 1, 1994, by "an amount equal, in percentage terms, to the base rate" and then in five equal annual stages beginning on January 1, 1995.[9] C10 products have their tariff cut non-linearly to 0% over 10 years: a 20% cut on January 1, 1994, followed by eight equal annual cuts beginning on January 1, 1996. For reference, Table 1 lists the staging categories and definitions.

Panel (a) of Figure 2 shows the distribution of NAFTA staging categories in terms of the $8,843$ products in the U.S. NAFTA schedule and in terms of bilateral imports with the relevant NAFTA partner(s) over our 1989-2016 sample period.[10] Qualitatively, these two distributions deliver very similar takeaways. Around 40-50% of the U.S. NAFTA tariff schedule relates to products that have their tariff immediately cut to 0 and another 15-20% of the schedule relates to products that continue duty free. For products with tariffs actually phased out over time, we hereafter aggregate staging categories B and B6 into a single "5-year" category and staging categories C, C10 and C+ into a single "10-year" category. While they each account for around 10% of products, the latter account for over 20% of U.S. imports from NAFTA partners but the former for less than 2% of such imports.[11,12]

Pre-NAFTA U.S. preferential arrangements with Canada and Mexico substantially impact implementation of the U.S. NAFTA staging categories just described. First, the U.S. agreed the base rate faced by Mexico would be that faced by Mexico in 1991 under the U.S. Generalized System of Preferences (GSP) and that Mexico would lose its GSP status on Jan-

[8] See General Note 2 of the U.S. tariff schedule in Annex 302.2.

[9] A product protected by a 40% base rate tariff would have the tariff reduced by 40% or 16 percentage points on January 1, 1994, with the remaining 24% tariff rate then cut in five equal installments (4.8 percentage points) until it reaches 0%.

[10] Table A.1 in the Appendix provides the data underlying Figure 2.

[11] The B6 products account for around 80% of the "5-year" products and about 60% of U.S. imports from NAFTA partners. The C products account for around 85% of the "10-year" products and over 99% of U.S. imports from NAFTA partners.

[12] The "Other" category in Figure 2 represents the "Mixed" products described above and a small number of products where the staging category is missing. See Table A.1 in the Appendix for further details.

uary 1, 1994.[13] Since Mexican GSP-eligible products enter the U.S. duty free, they continued duty free after NAFTA but this duty-free status was now permanent.[14] Because the vast majority of Mexico's GSP-eligible products are specified by the U.S. NAFTA tariff schedule to have their tariff immediately cut to zero, panel (b) of Figure 2 dramatically illustrates the reduction in the share of U.S. imports from Mexico that had their tariff immediately cut to zero. Ultimately, nearly 95% of imports from Mexico where tariffs are phased out over time fall in the 10-year staging category.

The second pre-NAFTA preferential arrangement of importance is CUSFTA. Specifically, the U.S. and Canada agreed that their CUSFTA tariff schedules would bound the NAFTA product-level tariffs they levied on each other.[15] That is, NAFTA could accelerate but not relax the CUSFTA-specified rate at which the U.S. phased out tariffs on imports from Canada. Panels (b)-(c) of Figure 2 illustrate this point. First, for NAFTA products that had their tariff immediately cut to zero, CUSFTA had already eliminated tariffs on about 45% of them by 1993. Second, for nearly all NAFTA products that had their tariff phased out over 10 years, CUSFTA had already eliminated, or started to eliminate, their tariff before 1994. These facts dramatically reduce the share of U.S. imports from Canada where tariffs are either immediately cut to zero or phased out over 10 years and, in turn, substantially increases the share that continue duty free. Ultimately, over 90% of imports from Canada where tariffs are phased out over time fall in the 5-year staging category.

Given the U.S. CUSFTA tariff concessions received by Canada, its subsequent U.S. NAFTA tariff concessions are fairly moderate, especially relative to those of Mexico. Panel (a) of Figure 3 illustrates and also shows tariff cuts, for either NAFTA partner, are actually highest for 5-year products. Nevertheless, panel (b) of Figure 3 shows, as one would expect, that the size of annual tariff cuts is highest for immediate-cut products and lowest for 10-year

[13]See https://www.cbp.gov/trade/nafta/guide-customs-procedures/effect-nafta/en-gsp and Glick (2010, p. 11).

[14]To establish Mexico's 1991 product-level GSP eligibility, we use the 1991 USITC tariff data collected by John Romalis and described in Feenstra, Romalis, and Schott (2002). This data has an 8-digit product indicator for GSP eligibility and also information on country-product specific exclusions from GSP eligibility.

[15]See NAFTA Annex 302.2(4) and 302.2(12).

products, with the 5-year products somewhere in between.[16]

2.3 Matching tariff schedules to trade data

Matching issues arise when merging the NAFTA staging categories and 8-digit USITC import data. On one hand, 91 products from the U.S. NAFTA tariff schedule do not appear in the USITC import data over our 1989-2016 sample period and are excluded from our analysis.[17]

On the other hand, 15 of the $8,690$ products imported into the U.S. in 1993, the first full year after NAFTA was signed, are not in the U.S. NAFTA tariff schedule.[18] That is, 99.83% of imported products and 98.85% of imported value in 1993 appear in the U.S. NAFTA tariff schedule. These match rates are only slightly lower in the 1989-1992 pre-NAFTA years. Ultimately, these omissions do not look like systematic attempts to exclude politically sensitive sectors from eventual tariff elimination.

Naturally, as Table B.1 in the Online Appendix illustrates, these 99% match rates fall over time. First, the World Customs Organization (WCO) periodically updates HS codes at the 6-digit level (as in 1996, 2002, 2007 and 2012). Second, based on recommendations to the President, the USITC updates 10-digit HS codes each year. In the early post-NAFTA years, these USITC changes were substantial. The 99.83% match rate of 1993 falls to 94.42% in 1994 and 82.68% in 1995. Thereafter, the match rate declines noticeably only in years of WCO HS changes (to 68.43% in 2002, 62.70% in 2007 and 59.16% in 2012). These facts motivate part of our robustness analysis that focuses on either HS codes that remain unchanged over time or concorded HS codes, i.e. "consistent codes," using our extension of the concordance of Pierce and Schott (2012).

[16] Table A.2 in the Appendix presents the data behind Figure 3.

[17] Of these products, 76 are in Chapter 98 *Special Classification Provisions* and 11 are Chapter 4 dairy products.

[18] Of the 15 products in the 1993 USITC import data that are not in the U.S. NAFTA tariff schedule, five products are not even in the initial version of the 1993 USHTS (2921.42.26, 2921.42.28, 9021.19.85, 9999.00.15 and 9999.95.00). The remaining 10 products are 0814.00.80 (citrus or melon peel); 2921.42.21, 2921.42.22 and 2922.50.11 (chemical compounds); 4418.20.40 and 4418.20.80 (wood doors), 7326.90.35 (iron or steel container); and 8521.10.30, 8521.10.60 and 8521.10.90 (video apparatus).

3 Empirical strategy

3.1 Triple difference specifications

Our aim is to identify how the U.S. phase-out of product-level tariffs under NAFTA impacts its product-level imports from NAFTA partners. Two strategies come to mind immediately. First, one could look at phase-out products and compare product-level import growth from NAFTA partners versus ROW. Intuitively, any differential import growth in this "NAFTA versus ROW" approach would reflect the tariff phase-out on NAFTA partners. Second, one could look at imports from NAFTA partners and compare product-level import growth for products where the tariff is phased out (phase-out products) versus products where the tariff is zero both pre- and post-NAFTA (continue-duty-free products). Intuitively, any differential import growth in this "phase-out versus continue-duty-free" approach would reflect the tariff phase-out. However, each of these approaches is problematic.

Both the NAFTA versus ROW and phase-out versus continue-duty-free approaches can be implemented as difference-in-difference (DD) specifications. However, the NAFTA versus ROW approach ignores the possibility that, after NAFTA, a product's NAFTA imports grow relative to its ROW imports *regardless* of whether the product's tariff is being phased out. After all, the NAFTA versus ROW approach only looks at phase-out products and ignores continue-duty-free products. Import growth from NAFTA partners relative to ROW in both phase-out and continue-duty-free products could be driven by positive supply shocks in the NAFTA partners or broad effects of NAFTA that go beyond tariff reduction.

Conversely, the phase-out versus continue-duty-free approach ignores the possibility that, after NAFTA, a phase-out product's imports grow relative to a continue-duty-free product *regardless* of the exporting country. After all, the phase-out versus continue-duty-free approach only looks at NAFTA imports and ignores ROW imports. Finding import growth of phase-out products relative to continue-duty-free products for both NAFTA imports and ROW imports could be driven, for example, by product-specific supply or demand shocks.

To avoid these problems with the intuitive NAFTA versus ROW and phase-out versus continue-duty-free approaches, we use triple difference (DDD) specifications. The simplest DDD specification is

$$
\begin{aligned}
(1) \quad \ln M_{cpt} \quad = \quad & \alpha + \beta_1 NAFTA_c + \beta_2 Phase_p + \beta_3 Post_t \\
& + \gamma_1 Phase_p \times Post_t + \gamma_2 NAFTA_c \times Post_t + \gamma_3 NAFTA_c \times Phase_p \\
& + \delta NAFTA_c \times Phase_p \times Post_t + \varepsilon_{cpt}.
\end{aligned}
$$

Here, $\ln M_{cpt}$ represents U.S. log imports from country c of product p in year t. Further, $NAFTA_c$, $Phase_p$, and $Post_t$ represent dummy variables indicating, respectively, (i) whether the exporting country c is a NAFTA partner, (ii) whether product p is a product where the tariff is phased out under NAFTA, and (iii) whether year t is in the post-NAFTA period of 1993 onwards.[19] In all our analyses, we only include either Canada or Mexico as the single NAFTA country. To avoid the phase out of tariffs across multiple U.S. FTAs simultaneously, we always exclude countries that are U.S. FTA partners at any point in time.[20]

The key coefficient of interest is the DDD coefficient δ and it has two equivalent interpretations. First, after controlling (via γ_2) for any post-NAFTA effects that impact imports from a NAFTA partner across all products, δ reflects the differential import growth of phase-out products from a NAFTA partner relative to ROW. Second, after controlling (via γ_1) for any post-NAFTA effects that impact phase-out product imports across all exporting countries, δ reflects the differential NAFTA partner import growth in phase-out products relative to continue-duty-free products.

Although improving on the intuitive NAFTA versus ROW and phase-out versus continue-duty-free approaches, equation (1) is a "no controls" DDD specification. As such, it still

[19] In practice, 1993 is neither a control year nor a treatment year. On one hand, NAFTA was signed in mid 1992 and, hence, the phase-out staging categories were known prior to 1993. But, on the other hand, tariff cuts do not start until 1994. Realistically, 1993 would be best described as a partially treated year. To this end, we treat it as the first treatment year.

[20] As a robustness check later, we also exclude countries that are FTA partners of Mexico or Canada at any point in time.

omits many potentially relevant variables.

First, the standard DDD specification in equation (1) does, by construction, control for effects that differentially impact import growth of phase-out and continue-duty-free products via γ_1. But, it does not allow such effects to vary over time at an annual frequency and does not allow the effects to vary across the various phase-out products or the various continue-duty-free products. Possible relevant factors that could vary at the product-year level include global product-level supply and demand shocks, U.S. production levels, Maquiladora production levels, import shares from U.S. FTA partners or Canada and Mexico's FTA partners (which impact the relative degree of preferential access for NAFTA partners in the US), status under the WTO Multifibre Arrangement, and Chinese import competition in the US.[21] Possible relevant factors that could vary at the country-product level include an exporter's international competitiveness or revealed comparative advantage. Nevertheless, we can control for these potentially relevant factors, among others that vary at the product-year or country-product level, by adding product-year fixed effects γ_{pt} and country-product fixed effects γ_{cp} to equation (1).[22]

Second, the standard DDD specification in equation (1) does, by construction, control for effects that differentially impact import growth from NAFTA partners versus ROW via γ_2. But, it does not allow such effects to vary over time nor across the various ROW countries. Possible relevant factors that could vary at the country-year level include GDP growth, bilateral exchange rates between the U.S. and the exporting country, and the export country's WTO status. Nevertheless, we can control for these potentially relevant factors, among others that vary at the country-year level, by adding country-year fixed effects γ_{ct} to equation (1).

Adding these three sets of fixed effects to the standard DDD specification in equation

[21]See, for example, Autor and Dorn (2013) and Bloom, Draca, and Van Reenen (2016) for the general economic importance of Chinese import competition and the WTO Multifibre Agreement.

[22]Note, these fixed effects subsume product fixed effects that would control for factors including the elasticity of substitution, the pre-NAFTA tariff, whether the good is an intermediate or primary or final good, and whether the goods is a homogeneous or differentiated good.

(1) flexibly controls for various product-year, country-product, and country-year variables, including but not limited to the ones described above. We now have the following fixed-effects, or "with controls," DDD specification:

$$(2) \qquad \ln M_{cpt} = \alpha + \delta NAFTA_c \times Phase_p \times Post_t + \gamma_{pt} + \gamma_{ct} + \gamma_{cp} + \varepsilon_{cpt}.$$

One could reasonably expect important heterogeneity in the DDD coefficient δ along two dimensions. First, as described in Section 2.2, some products are phased out over longer periods than others. Second, the effects of tariff cuts may affect import growth over time. Hence, one could reasonably expect the effects of tariff cuts to grow over time and depend on the length of a product's tariff phase-out. Thus, we augment equation (2) in two ways. First, we allow the DDD coefficient to vary over time by replacing the $Post_t$ dummy with a vector of year dummies $\mathbf{Year}_t = (1989, 1990, 1991, 1993, ..., 2016)$ with the omitted year of 1992 serving as the reference year. Second, we redefine $Phase_p$ as a vector $\mathbf{Phase}_p = (GSP_p, Immed_p, 5yr_p, 10yr_p)$ consisting of indicator variables for whether the product continues duty free because of the GSP program (GSP_p), has its tariff cut to zero immediately ($Immed_p$), has its tariff phased out over 5 or 6 years ($5yr_p$) or has its tariff phased out over at least 10 years ($10yr_p$).

Our generalized fixed-effects, or "with controls," DDD specification is:

$$(3) \qquad \ln M_{cpt} = \alpha + \boldsymbol{\delta} NAFTA_c \times \mathbf{Phase}_p \times \mathbf{Year}_t + \gamma_{pt} + \gamma_{ct} + \gamma_{cp} + \varepsilon_{cpt}.$$

Here, $\boldsymbol{\delta}$ is a vector of coefficients containing one coefficient for each year and phase-out category pair. Given the number of coefficients estimated, we present many of our results using figures that plot annual point estimates and corresponding 95% confidence intervals.

3.2 Identification

The triple-difference framework dates back to Gruber (1994) who investigated the cost pass-through to wages of married women from state-level health insurance mandates regarding maternity benefits. He observed (i) males and females and (ii) some states that did, and some that did not, implement mandates during the sample period. In turn, he states (p. 627) that the DDD identification assumption is "fairly weak: it simply requires that there be no contemporaneous shock that affects the relative outcomes of the treatment group [i.e. females relative to males] in the same state-years as the law." Translating this assumption into our setting says that there be no contemporaneous shock that affects import growth of phase-out products relative to continue-duty-free products from NAFTA partners vis-à-vis ROW.

To investigate the reasonableness of this assumption, we can look at the pre-NAFTA period to see how import growth of phase-out products relative to continue-duty-free products from NAFTA partners compares to that from ROW. Substantive differences, which would show up as statistically significant DDD estimates in the pre-NAFTA period, suggest concern that contemporaneous shocks systematically drive a wedge between these relative import growth measures and violate the identifying assumption described above. But, absence of substantive differences, which would show up as statistically insignificant DDD estimates in the pre-NAFTA period, suggests that the relative import growth in the NAFTA partners were quite similar to that in ROW in the pre-NAFTA period. As such, there are reasonable grounds to think any divergence of these growth rates in the *post*-NAFTA period could well be attributed to NAFTA itself and not contemporaneous shocks. Thus, our subsequent analysis presents the DDD coefficients in the pre-NAFTA period and, at least for Mexico, strongly support our identification assumption. Of course, this DDD exercise is the analog of investigating the parallel trends assumption in a DD setting.

In Section A.1 of the Appendix, we provide a complementary analysis that explores the empirical determinants of the U.S. NAFTA staging categories. There, we see that, to a reasonably large extent, U.S. NAFTA staging categories reflect the U.S. CUSFTA staging categories. Thus, the economic and political variables driving U.S.-Canada CUSFTA negotiations also drive the U.S. NAFTA staging categories on imports from Mexico which significantly reduces any endogeneity concerns over such staging categories. Further, the key drivers of the U.S. NAFTA staging categories are time-invariant variables (e.g. the pre-NAFTA tariff and whether a good is an intermediate good) rather than time-varying variables and, hence, are subsumed by our various product fixed effects.

4 Results: Tariff phase-out

4.1 A simple means-based approach

The standard DDD approach in equation (1) is just a comparison of mean import growth between phase-out and continue-duty-free products and between NAFTA partners and ROW. Table 2 illustrates these mechanics.

Panel A depicts the NAFTA versus ROW approach and also motivates the value of a DDD approach over a DD approach. To begin, Panel A1 shows relative import growth of phase-out products from NAFTA partners vis-à-vis ROW. While mean log imports of phase-out products from Mexico grew by 0.345 log points in the post-NAFTA period, mean log imports of phase-out products from ROW *decreased* by 0.244 log points in the post-NAFTA period. Thus, import growth of phase-out products from Mexico vis-à-vis ROW was 0.589 log points and represents a DD estimate. A very similar story holds for Canada, both qualitatively and quantitatively. Ultimately, import growth in phase-out products from Canada vis-à-vis ROW was 0.633 log points and represents a DD estimate. From these DD perspectives, NAFTA tariff cuts appear to have substantial impacts on NAFTA trade flows.

However, this DD approach overestimates the impact of NAFTA tariff cuts. Specifically, Panel A2 shows that similar, but weaker, DD effects emerge when looking at continue-duty-free products. Even though continue-duty-free products did not receive tariff cuts, import growth of continue-duty-free products from NAFTA partners was 0.326 log points for Mexico vis-à-vis ROW and 0.433 log points for Canada vis-à-vis ROW. The fact that imports from NAFTA partners grow relative to ROW even for continue-duty-free products suggests important NAFTA-specific effects on import growth that go beyond tariff cuts.

The DDD estimates take this into account by looking at the *excess* relative import growth of NAFTA partners vis-à-vis ROW in phase-out products relative to continue-duty-free products. That is, the DDD estimates are differences in DD estimates. The DDD estimates say this excess relative import growth is 0.263 log points for Mexico and 0.200 log points for Canada. On one hand, the large DD point estimates in Panel A2 show the importance of controlling for a "NAFTA effect" beyond tariff phase-outs and motivates the importance of country-year fixed effects to allow a "NAFTA effect" that varies across time and ROW partners. Nevertheless, the non-trivial DDD point estimates show that tariff cuts were an important part of the NAFTA-induced import growth.

Panel B of Table 2 depicts the phase-out products versus continue-duty-free products approach. Panel B1 shows that import growth from Mexico of phase-out products vis-à-vis continue-duty-free products was -0.069 log points: on average, imports from Mexico of phase-out products actually grew *less* than continue-duty-free products. The same story holds for for Canada with relative import growth of -0.081 log points. These relative import growth numbers are DD estimates and, by themselves, suggest that NAFTA tariff cuts may have actually reduced NAFTA trade flows.

However, these DD effects underestimate the impact of NAFTA tariff cuts. Specifically, Panel B2 shows much larger negative DD effects when looking at import growth of phase-out products relative to continue-duty-free products from ROW. Defining phase-out products based on Mexico's (Canada's) NAFTA staging categories, import growth of phase-out relative to continue-duty-free products from ROW was -0.332 (-0.281) log points.

The DDD estimates take this into account by looking at the "excess" relative import growth of phase-out products relative to continue-duty-free products for NAFTA partner imports vis-à-vis ROW imports. That is, the DDD estimates are differences in DD estimates and say this excess relative import growth is 0.263 log points for Mexico and 0.200 log points for Canada. By construction, these DDD estimates match those from the NAFTA versus ROW approach above. The very large DD point estimates in Panel B2 motivate the importance of controlling for the systematic differences in phase-out products versus continue-duty-free products that we described in Section 3.2 and, in turn, the importance of product-year and country-product fixed effects in our later analysis. Nevertheless, again, the non-trivial DDD point estimates show that tariff cuts were an important part of the NAFTA-induced trade flow growth.

4.2 Regression-based approach

While the simple means-based approach highlights the key intuition of the DDD approach, it is essentially a "no controls approach" as we described in Section 3.1. Moreover, it ignores possible heterogeneity in the DDD treatment across time and phase-out categories. Given the richness of our data, we can include country-product, country-year, and product-year fixed effects to control for a myriad of potentially confounding factors. And, we can also allow DDD estimates to vary across time and phase-out categories.

4.2.1 Mexico

We focus on Mexico as the NAFTA partner in this sub-section, with Canada being the focus of the next sub-section. Putting aside heterogeneity, Table 3 shows the impacts of moving from the "no controls" specification in equation (1) to the "with controls" specification in equation (2). Column (1) show the "no controls" DDD estimates from equation (1) and, by construction, the DDD point estimate matches that from the means-based DDD in Table 2. Columns (2)-(4), respectively, add country-product, country-year, and product-year fixed effects so that column (4) represents the "with controls" DDD specification in equation (2).

Column (2) shows that the DDD point estimate increases by about 40% upon including country-product fixed effects. This is consistent with the product composition of Mexican exports to the U.S. differing notably from that of ROW exports to the US, perhaps due to differences in comparative advantage. From the perspective of the phase-out versus continue-duty-free products approach, the smaller estimates in column (1) could reflect that, relative to their comparative advantage in continue-duty-free products, Mexico tends to have a weaker comparative advantage in phase-out products than ROW. That is, controlling for these country-product effects increases relative import growth of phase-out products (i.e. relative to continue-duty-free products) from Mexico vis-à-vis ROW.

Comparing column (4) with column (2), adding the country-year and product-year fixed effects only modestly impacts the DDD estimates. Intuitively, country-year fixed effects control for time-varying factors that are common across import growth of phase-out and continue-duty-free products from a particular exporter. And, product-year fixed effects control for time-varying factors for a particular product that are common across import growth from NAFTA members and ROW. This leaves the DDD estimates largely unchanged because, in the former case, relative import growth (phase-out versus continue-duty-free products) remains largely unchanged from a particular exporter and, in the latter case, product-level import growth remains largely unchanged from a NAFTA member vis-à-vis ROW.

Bringing in the dimensions of heterogeneity discussed above, we start with time-varying DDD estimates. Ignoring this dimension of heterogeneity, column (4) from Table 3 said the time-*invariant* DDD point estimate was 0.371 log points. But, unsurprisingly given the tariff phase-out argument from Baier and Bergstrand (2007), Figure 4 shows considerable time heterogeneity. Nevertheless, the Mexican DDD point estimate is statistically insignificant before NAFTA. As we discussed in Section 3.2, statistical insignificance of these pre-NAFTA DDD point estimates is important in terms of assuaging potential endogeneity concerns as they say that import growth from Mexico (i.e. phase-out products relative to continue-duty free products) is not diverging from that of ROW before NAFTA.

Indeed, focusing on the "with controls" specifications, the DDD point estimate only becomes statistically significant in 1997. The point estimates continue growing from 0.309 log points in 1997 to a peak, right around the 10-year mark of NAFTA, of 0.551 log points in 2004. While the effects taper off somewhat post-2004, the point estimates largely hover in the 0.4-0.45 log points in the post-2000 period and are always statistically significant.[23] Overall, Figure 4 clearly depicts the idea of a gradual and delayed increase in imports from Mexico over the first 10 years of NAFTA that stabilizes in the early to mid 2000s. Thus, Figure 4 appears to provide preliminary support for tariff phase-outs explaining delayed import growth from Mexico after NAFTA: import growth of phase-out products *relative* to continue-duty-free products from Mexico consistently exceeds that from ROW by around 0.4-0.45 log points after 2000.

Our highly disaggregated 8-digit product-level HS data allow us to investigate the tariff-phase-out story by looking at heterogeneity of the DDD estimates not only over time but also by phase-out category. Products where tariffs are immediately cut to zero should have a very different time path of import growth than products where tariffs are phased out gradually over 5 or 10 years. We expect to see a rapid increase in import growth of immediate-cut products that stabilizes quickly in contrast to a steady and gradual increase in import growth of 5-year and 10-year phase-out products that stabilizes after 5-10 years.

Figure 5 presents time-varying DDD estimates from equation (3) when splitting phase-out products into their different categories: immediate cut (A), 5-year phase-out (B and B6), 10-year phase-out (C, C10 and C+) and GSP.[24] The dashed (blue) lines illustrate our qualitative hypotheses. Products where tariffs are immediately cut to zero should see a large immediate growth in trade that remains stable thereafter. Products where tariffs are phased out over 5-6 years should see gradual trade growth that stabilizes after around 5-6 years. Products with a 10-year phase-out should see even more gradual trade growth that stabilizes after 10-

[23] Table B.2 in the Online Appendix presents the time-varying regression results for the specifications that parallel those in Table 3.

Again, product-year and country-year fixed effects matter little after country-product fixed effects are included.

[24] Table B.3 in the Online Appendix presents the regression results.

15 years. Finally, assuming NAFTA removes uncertainty about future eligibility for the GSP program, GSP products should look similar to products of which the tariff is immediately cut to zero: an immediate increase in trade that quickly stabilizes.[25] Importantly, note the statistically insignificant DDD point estimates in the pre-NAFTA period across panels (a)-(d): relative import growth (i.e. phase-out products relative to continue duty free products) from Mexico was not diverging from that of ROW before NAFTA.[26]

After NAFTA is signed in mid 1992, panel (a) in Figure 5 shows statistically significant import growth of immediate-cut products beginning in 1994 that reaches around 0.9 log points by the late 1990s and stabilizes shortly thereafter. Moreover, panel (b) shows 5-year phase-out products experience even larger import growth that eventually peaks around 1.25 log points in 2000 and stabilizes around 1.15 log points in the early 2000s. At its peak, import growth is nearly 40% larger for the 5-year phase-out products than immediate-cut products which is consistent with the 25-65% larger tariff cuts of 5-year phase-out products (B and B6) versus immediate-cut products from Table A.2. However, in contrast to the hypothesis that import growth of the immediate cut products should stabilize much sooner than the 5-year phase-out products, both seem to stabilize at similar points in time.

Indeed, this notable feature of panels (a)-(b) in Figure 5 is a robust statistical property. Table 4 shows the extent that we can detect *changes* in cumulative import growth over time. The tariff-phase-out hypothesis implies we should be able to detect cumulative import growth that increases over time but stabilizes sooner for products with shorter phase-out periods. Naturally, it is difficult to detect statistically significant changes in cumulative import growth at an annual frequency. Thus, Table 4 presents estimates of import growth over 3-year rolling windows: the year t point estimate is the difference between the DDD

[25]As part of the broader and growing literature on trade policy uncertainty (e.g. Handley 2014, Pierce and Schott 2016 and Handley and Limão 2017), Hakobyan (2019) documents the inherent legislative uncertainty surrounding GSP renewal and the adverse impact of uncertainty on import growth from beneficiary countries.

[26]The one qualification here would be the statistically significant DDD point estimate in 1989 for 5-year phase-out products in panel (b). However, the DDD point estimate is then statistically insignificant until at least 3 years after NAFTA is signed.

point estimates for cumulative import growth in year t and year $t-3$ from Figure 5.[27] A statistically significant and positive point estimate in year t says we can detect a statistically significant increase in cumulative import growth from year $t-3$ to year t.

Table 4 shows that import growth of immediate-cut products in these rolling 3-year windows is around 0.3-0.5 log points in all years between 1994 and 1999 with the caveat of not seeing import growth in 1996 relative to 1993. Statistically speaking, the post-NAFTA import growth of immediate-cut products only stabilizes from 2000. However, 5-year phase-out products grow about 0.3-0.6 log points in these rolling 3-year windows beginning in 1995 and continuing every year until 2001. Statistically speaking, the post-NAFTA import growth for 5-year phase-out products stabilizes from 2002. Import growth of immediate-cut products starts only one year earlier than 5-year phase-out products and stabilizes 7 years after NAFTA was signed, only 2 years ahead of when 5-year phase-out products stabilize. From this perspective, the dynamics of import growth for immediate-cut products is remarkably similar to that of 5-year phase-out products. This strongly undermines the tariff-phase-out hypothesis for delayed import growth.

In terms of tariff phase-out on Mexican imports, panel (b) of Figure 2 shows 5-year phase-out products account for only about 6% of imports from Mexico of which the tariffs were actually phased out over time. In contrast, 10-year phase-out products account for the remaining 94%. Thus, to the extent that tariff phase-outs help explain delayed import growth from Mexico, it should help explain the 10-year phase-out products. However, we can hardly detect import growth for these products over our 3-year rolling windows. In such windows, Table 4 shows we cannot detect increases in import growth over time.[28] Together with the 0.4-0.6 log point magnitude of import growth being much smaller than 0.9 log points for immediate-cut products despite both having average tariff cuts of around 7.5%, there is essentially no evidence that import growth of 10-year phase-out products conform with the

[27] That is, the year t point estimate in Table 4 represents the test of equality between the year t and year $t-3$ point estimates in Table B.3 of the Online Appendix that come from estimating equation (3).

[28] This inability to detect robust import growth for the 10-year phase-out products also holds when using 2-year, 4-year or 5-year rolling windows.

tariff phase-out hypothesis.

The import growth dynamics of GSP-eligible products provide more evidence against the tariff-phase-out hypothesis. NAFTA permanently removes uncertainty over Mexico's tariff-free access on GSP-eligible products by codifying these products as continue-duty-free. Thus, like immediate-cut products, we expect an initial burst of import growth with the cumulative growth impact stabilizing quickly. Yet, panel (d) of Figure 5 suggests GSP import growth after 1992 only kicks in by the late 1990s and only stabilizes by the mid 2000s with Table 4 putting this growth around 0.3-0.4 log points. Moreover, Table 4 says we can only detect 3-year rolling-window import growth in 1999 and 2001. Ultimately, the dynamics of GSP-eligible products look more like what we expect from 5-year or 10-year phase-out products than immediate-cut products. Indeed, this gradual import growth in GSP-eligible products is an important part of the overall pattern of gradual import growth from Table 4 and Figure 5 given that panel (b) of Figure 2 shows GSP-eligible products represent around 40% of the U.S. staging schedule on imports from Mexico.

4.2.2 Canada

The key reason we have focused on U.S. imports from Mexico is that the U.S. was already phasing out tariffs on Canada under CUSFTA from 1989. Thus, any effects of tariff phase-out on Canadian imports would already be present in the pre-NAFTA period. Indeed, when not distinguishing between different phase-out staging categories, Figure 4 shows such effects. And, Figure 6 illustrates this quite strongly for 5-year phase-out products which, per panel (b) of Figure 2, constitute about 90% of Canadian imports that were actually phased out over time.[29] Together, these results provide strong evidence that relative import growth from Canada (i.e. phase-out products relative to continue duty free products) was diverging from that of ROW in the pre-NAFTA period. Hence, our results on Canadian imports should be viewed with caution. As a result, we relegate much of our analysis on imports from Canada

[29]This largely follows directly from the implications of Canada's CUSFTA staging categories for their NAFTA staging categories.

to the Online Appendix.

Nevertheless, the key points emerging from our Mexican analysis also emerge for Canadian imports. Figure 6 and Table B.3 of the Online Appendix shows, post-NAFTA growth of immediate-cut products becomes statistically significant in the late 1990s, peaking around 0.25 log points in 2001 and stabilizing shortly thereafter. Indeed, Table 4 shows we can detect 3-year rolling window import growth of 0.10-0.25 log points for immediate-cut products each year during 1997-2000. The substantially smaller magnitudes for Canada than Mexico are consistent with Figure 3 showing these immediate-cut products from Canada had tariff cuts one-third as large as such Mexican products. Ultimately, as with Mexico, Canadian immediate-cut products experience the type of delayed import growth one would have expected from 5-year phase-out products.

The key set of Canadian imported products where tariffs are actually phased out over time exhibit import growth dynamics very similar to that of immediate-cut products, as was the case with Mexico. For Canada, 5-year phase-out products account for about 90% of Canadian imports where tariffs were actually phased out over time. Table 4 and Panel (b) of Figure 6 show import growth of Canada's 5-year phase-out products stabilize in 1999 at around 0.5 log points.[30] Thus, statistically speaking, immediate-cut and 5-year phase-out products both stop growing in the 1999-2000 period which contrasts starkly with the tariff-phase-out hypothesis.

4.2.3 Robustness

We now describe various robustness checks to our results presented above. First, one may be concerned about the well-documented surge of U.S. imports from China. Second, while we have excluded U.S. FTA partners from ROW throughout our analysis, one may be concerned about FTAs formed by Mexico and Canada and the associated tariff phase-outs therein. Figure 7 and Figure B.1 in the Online Appendix show these restrictions on the country

[30] This notably larger import growth compared to immediate-cut products is consistent with the substantially larger tariff cuts experienced by the 5-year phase-out products.

sample representing ROW do not affect our results.

Unfortunately, as Section 2 described, product codes change over time. In principle, this is problematic given our methodology relies on matching 8-digit HS product codes, that potentially change over time, with their NAFTA staging category. Thus, we take two approaches. First, we restrict the sample to product codes that remain unchanged over our sample period. Second, we extend the concordance from Pierce and Schott (2012) to the end of our sample period in 2016 to create "consistent codes" that correct for changes in product code definitions over time. Figure 7 and Figure B.1 in the Online Appendix show these restrictions on the sample of products do not affect our results.

5 Results: Delayed tariff pass-through

To the extent that tariffs pass-through to import prices, increases in the value of trade could come from increases in quantities or prices. Thus, we now modify equation (3) by using log real unit values as the dependent variable to proxy for import prices.[31] This analysis addresses the second hypothesis of Baier and Bergstrand (2007) that delayed pass-through of tariff cuts to import prices can explain the delayed trade flow effects of FTAs.

Figure 8 and Figure B.2 in the Online Appendix present the results where we plot coefficients from our baseline specification with controls as well as only using data on consistent codes. Quite starkly, there is no evidence of delayed pass-through effects as there is essentially no impact of tariff phase-out on unit values. In turn, the impact on trade values seen in our earlier analysis reflects growth in the quantity of trade rather than the price of imports. Given our DDD estimates measure the impact on phase-out products *relative* to the impact on continue-duty-free products, our results suggest any impact of NAFTA on U.S. import prices from NAFTA partners are driven by NAFTA effects that go beyond tariff phase-out.

[31] Real unit values are the ratio of real import value to quantity. We standardize quantity measures: count measurements are converted to singular counts of units (rather than dozens, etc.), weight measurements are converted to kilograms, length measurements are converted to meters, and volume measurements are converted to liters. As mentioned in Section 2.1, we drop products where the unit of measurement changes over time in a way that cannot be standardized.

6 Extensive margin of delayed import growth

If tariff phase-out and delayed pass-through of tariff cuts to prices do not drive delayed import growth after NAFTA, then what could? One explanation is that the extensive margin plays an essential and growing role over time. This would be consistent with the empirical finding of "intensive-margin effects occurring sooner than extensive-margin effects" from Baier, Bergstrand, and Feng (2014, p. 339) in the FTA context and theoretical (Arkolakis, Costinot, and Rodríguez-Clare 2012) and empirical (Bernard, Jensen, Redding, and Schott 2009) results in the broader trade literature. Intuitively, extensive margin effects should take longer to play out because of frictions created by fixed cost barriers that firms face when entering foreign markets or adding new products in terms of, e.g., building networks of distributors and customers. We now use two approaches to investigate this broad idea. Given our concerns over disentangling the impacts of CUSFTA from NAFTA for imports from Canada, we focus on imports from Mexico throughout this section.

We begin investigating the role of the extensive margin by thinking of the intensive margin as import growth of "continuously traded" products; that is, products imported from Mexico in every year of our sample. In turn, the extensive margin includes a "new products" extensive margin captured by import growth of products that were not imported pre-NAFTA from Mexico. But, the extensive margin also includes an "infrequent" extensive margin captured by import growth of products that were imported pre-NAFTA from Mexico but are not imported in *every* year of our sample.[32] According to the mix of products imported from Mexico, Figure 9 shows the intensive margin falls towards 50% of imported value during our sample (but is stable from the mid-late 1990s at around $20-25\%$ of products). Figure 9 also shows a growing shift away from the infrequent extensive margin

[32] Note, our new products extensive margin is actually about products that are newly traded. This is different than the new goods margin of Kehoe and Ruhl (2013). After rank ordering products to their value of trade, they define the new goods margin as the products that comprise the first 10% of trade value. Thus, in their own words, their new goods margin is a "least traded goods" margin. In the context of our subsequent analysis, we did not find any statistically robust effects when looking at their "least traded goods" margin.

towards the new products extensive margin; the latter grows towards 40% of imported value (50% of products) and the former falls towards 10% of imported value (30% of products). Thus, the new products extensive margin is an important feature of U.S. import growth from Mexico.

Since our empirical methodology relies on comparing import growth between the pre- and post-NAFTA periods, we cannot directly identify the impact of NAFTA on the new products extensive margin. But, we can identify the impact of NAFTA along the intensive and infrequent extensive margins and, as a residual, indirectly talk about the new products extensive margin.

Figure 10 shows our results with the "baseline" results being those from Section 4.2.1 that ignore heterogeneity across intensive and extensive margins. Note that Figure 10 broadens our earlier result that there is essentially no evidence that relative import growth of phase-out products from Mexico (i.e. relative to continue-duty-free products) diverge from that of ROW in the pre-NAFTA period. For the DDD estimates in the post-NAFTA period, the intensive margin effects are strikingly weaker, across all staging categories, than our baseline results and generally statistically insignificant. Since the intensive margin accounts for 50-60% of Mexican imports, this suggests another margin drives our baseline results. While the infrequent extensive margin results generate somewhat stronger impacts than our baseline analysis, this margin only accounts for about 10-15% of Mexican imports from the mid-late 1990s onwards. Our results suggests that, as a residual, the new products extensive margin plays a key role in driving the delayed import growth that we observe in our baseline results. This effect is important across staging categories and regardless of the time path of tariff cuts. In turn, this helps understand why tariff phase-out or delayed pass-through of tariffs to prices cannot explain the delayed import growth observed after NAFTA.

As the literature has suggested, fixed costs of creating networks of distributors and customers are important in explaining the empirical importance of the extensive margin and explaining how firms may gradually learn and develop these networks over time. To investigate this idea in our setting, we use Census data that record product-level imports by year, exporting country, *and* each customs district in the U.S.

The geographic U.S. consists of 42 customs districts.[33] They generally correspond to state borders, but can cover multiple states (e.g. the Boston district covers Massachusetts and Connecticut) and some states are covered by multiple districts (e.g. 5 in Texas and 3 in California). The Census uses two alternative definitions of a customs district: where the imported shipment cleared customs ("the district of unlading") and entered consumption channels ("district of entry"). Using the latter, we construct the variable D_{cpt} that represents the log number of U.S. customs districts where imports from exporting country c of product p entered consumption channels in year t. We then trace out how NAFTA impacts the spread of imports from Mexico across the U.S. over time by replacing the imports variable $\ln M_{cpt}$ in equation (3) with the districts variable $\ln D_{cpt}$.

Figure 11 shows our results. Like previous results, there is no evidence, regardless of the phase-out category or our treatment of the intensive and extensive margins, that the relative spread across the U.S. of phase-out products imported from Mexico (i.e. relative to continue-duty-free products) diverged from that imported from ROW in the pre-NAFTA period. Focusing first on the "baseline"estimates that ignore heterogeneity across the intensive and extensive margins, the post-NAFTA DDD estimates say that Mexican imports of a given phase-out product start entering more and more districts over time. And, these impacts are notable across *all* staging categories: spatial expansion across the U.S. grows steadily through the early post-NAFTA period and stabilizes around 2000 for *all* staging categories. Thus, these spatial expansion effects help explain why tariff phase-out and delayed pass-through of tariffs to prices cannot explain the delayed import growth observed after NAFTA.

Diving deeper into the spatial expansion across the U.S., we can look at heterogeneity across the intensive and infrequent extensive margins defined above and, as a residual, the

[33]This includes separate districts for Alaska and Hawaii. Additionally, two districts cover Puerto Rico and the U.S. Virgin Islands. And, three "special districts" do not conform to geographic boundaries: "vessels under their own power," "low-valued imports and exports,"and "mail shipments."

new products extensive margin. While often positive and statistically significant in the early post-NAFTA years, the intensive margin effects still remain weaker overall than the baseline results at least in terms of statistical significance in the long run. This again suggests a strong role for the extensive margin in explaining our spatial expansion results. But, the infrequent extensive margin effects largely mirror the baseline results and Figure 9 says they only account for 10-15% of U.S. imports from Mexico from the mid-1990s onward. Thus, the new products extensive margin appears to again play a very important role for our spatial expansion results, both in the short and long run.

Ultimately, NAFTA tariff cuts create important incentives for spatial expansion across the U.S. that, at least in the short run, emerge regardless of whether a product has been imported from Mexico for a long time or only starts being imported after NAFTA. Of course, one could think of a product's spatial expansion across the U.S. as a "spatial "extensive margin, regardless of whether we have labeled it as an intensive margin product, infrequent extensive margin product or a product in the new products extensive margin. In this light, a key takeaway from our analysis is that various dimensions of the extensive margin, including the "new products" and "spatial" extensive margins, are crucial to understanding post-NAFTA delayed import growth stemming from tariff cuts.

7 Conclusion

Since the seminal work of Baier and Bergstrand (2007), the literature has known that trade flows increase gradually over time following FTA formation with the rule of thumb being that trade flows stabilize after doubling over 10 years. In their paper, Baier and Bergstrand (2007) hypothesize that these effects could naturally arise because FTAs typically phase out tariffs over time and because of delayed pass-through of tariff cuts to import prices. However, to the best of our knowledge, there is no empirical evidence attempting to investigate these hypotheses. One reason for this lack of research is that there is no readily and publicly available information of the tariff phase-out embodied in FTAs. Thus, by going to the publicly available texts of the CUSFTA and NAFTA treaties, we collect the necessary data and are the first to investigate the root causes suggested by Baier and Bergstrand (2007) for the delayed import growth following FTA formation.

Our central finding is that there is no evidence to support the idea that tariff phase-outs or delayed pass-through effects can help explain the delayed growth in trade flows typically observed following FTA formation. When looking at the impact of tariff phase-outs on imports, we do find that imports of products with tariffs phased-out grow more than those where tariffs continue duty free. And, comfortingly, the magnitude of the effects that we find, both across products within a country of different phase-out duration and across countries for products with the same phase-out duration, are consistent with differences in the actual country-product specific tariff cuts embodied in NAFTA. If tariff phase-out is indeed the root cause of delayed trade flows then we would expect the bulk of delayed import growth to come from products where tariffs were phased out over 5, 10, or even 15 years. But, crucially, the bulk of the delayed import growth comes from products where tariffs were immediately cut to zero or, in the case of Mexico, had pre-NAFTA tariff-free access via the GSP program converted into permanent tariff-free access. Moreover, there is essentially no evidence of delayed pass-through effects.

We show that a more likely explanation for the delayed import growth following FTA formation is that tariff cuts reduce the impact of trade frictions associated with the extensive margin of trade. First, we present evidence that the products driving the delayed impact of NAFTA tariff cuts on U.S. import growth are products where tariffs were cut but were not imported before NAFTA or, in other words, the “new products” extensive margin. We do so by showing that there is very little delayed import growth for products where tariffs were cut and were continuously imported before and after NAFTA. Second, we present evidence highlighting the importance of a “spatial” extensive margin where, largely independent of an imported product’s length of tariff phase-out, NAFTA tariff cuts lead to such imported products spreading out gradually across more and more geographic areas of the U.S. Thus, our results suggest FTA tariff cuts reducing the impact of frictions associated with establishment of networks of distributors and consumers — either at the “new products” or “spatial” extensive margin — is crucial to understanding how trade flows respond after FTA formation.

Our findings do not invalidate the use of lagged FTA dummies in the standard gravity approach when looking at the aggregate effects of FTAs. FTAs do indeed have lagged effects on trade. However, those lagged effects are neither a consequence of tariffs being gradually reduced nor of gradual pass-through of tariffs to prices.

References

Albornoz, F., H. F. C. Pardo, G. Corcos, and E. Ornelas (2012): "Sequential exporting," *Journal of International Economics*, 88(1), 17 – 31.

Anderson, J. E., and Y. V. Yotov (2016): "Terms of trade and global efficiency effects of free trade agreements, 1990–2002," *Journal of International Economics*, 99, 279–298.

Arkolakis, C., A. Costinot, and A. Rodríguez-Clare (2012): "New trade models, same old gains?," *The American Economic Review*, 102(1), 94–130.

Arkolakis, C., T. Papageorgiou, and O. A. Timoshenko (2018): "Firm learning and growth," *Review of Economic Dynamics*, 27, 146 – 168.

Autor, D. H., and D. Dorn (2013): "The growth of low-skill service jobs and the polarization of the US labor market," *The American Economic Review*, 103(5), 1553–1597.

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

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

Bernard, A. B., J. B. Jensen, S. J. Redding, and P. K. Schott (2009): "The margins of US trade," *American Economic Review*, 99(2), 487–93.

Bloom, N., M. Draca, and J. Van Reenen (2016): "Trade induced technical change? The impact of Chinese imports on technology and employment," *Review of Economic Studies*, 83, 87–117.

Broda, C., and D. E. Weinstein (2006): "Globalization and the gains from variety," *The Quarterly Journal of Economics*, 121(2), 541–585.

Caliendo, L., and F. Parro (2015): "Estimates of the trade and welfare effects of NAFTA," *The Review of Economic Studies*, 82(1), 1–44.

Coeli, F. (2018): "Trade policy uncertainty and innovation: Evidence from China," *Mimeo*.

Feenstra, R. C., J. Romalis, and P. K. Schott (2002): "US imports, exports, and tariff data, 1989-2001," *NBER Working Paper No. 9387*.

Frazer, G., and J. Van Biesebroeck (2010): "Trade growth under the African Growth and Opportunity Act," *Review of Economics and Statistics*, 92(1), 128–144.

Friederich, B. U., and M. Zator (2018): "Adaptation to shocks and the role of capital structure: Danish firms during the cartoon crisis," *Mimeo*.

Glick, L. A. (2010): *Understanding the North American Free Trade Agreement: Legal and business consequences of NAFTA*. Kluwer Law International.

GRUBER, J. (1994): "The incidence of mandated maternity benefits," *The American Economic Review*, pp. 622–641.

HAKOBYAN, S. (2019): "GSP expiration and declining exports from developing countries," *Mimeo.*

HANDLEY, K. (2014): "Exporting under trade policy uncertainty: Theory and evidence," *Journal of International Economics*, 94(1), 50–66.

HANDLEY, K., AND N. LIMÃO (2017): "Policy uncertainty, trade, and welfare: Theory and evidence for China and the United States," *The American Economic Review*, 107(9), 2731–83.

HILLBERRY, R., AND C. MCDANIEL (2002): "A decomposition of North American trade growth since NAFTA," *US International Trade Commission Working Paper 2002-12-A*.

KEHOE, T. J., AND K. J. RUHL (2013): "How important is the new goods margin in international trade?," *Journal of Political Economy*, 121(2), 358–392.

KIKKAWA, A., Y. MEI, AND P. SANTAMARINA (2019): "The impact of NAFTA on prices and competition: Evidence from Mexican manufacturing plants," *Mimeo.*

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

PIERCE, J. R., AND P. K. SCHOTT (2012): "Concording US Harmonized System codes over time," *Journal of Official Statistics*, 28(1), 53–68.

——— (2016): "The surprisingly swift decline of US manufacturing employment," *The American Economic Review*, 106(7), 1632–1662.

RAUCH, J. E. (1999): "Networks versus markets in international trade," *Journal of International Economics*, 48(1), 7–35.

ROMALIS, J. (2007): "NAFTA's and CUSFTA's impact on international trade," *The Review of Economics and Statistics*, 89(3), 416–435.

RUHL, K. J. (2008): "The international elasticity puzzle," *Mimeo.*

TREFLER, D. (2004): "The long and short of the Canada-US Free Trade Agreement," *The American Economic Review*, 94(4), 870–895.

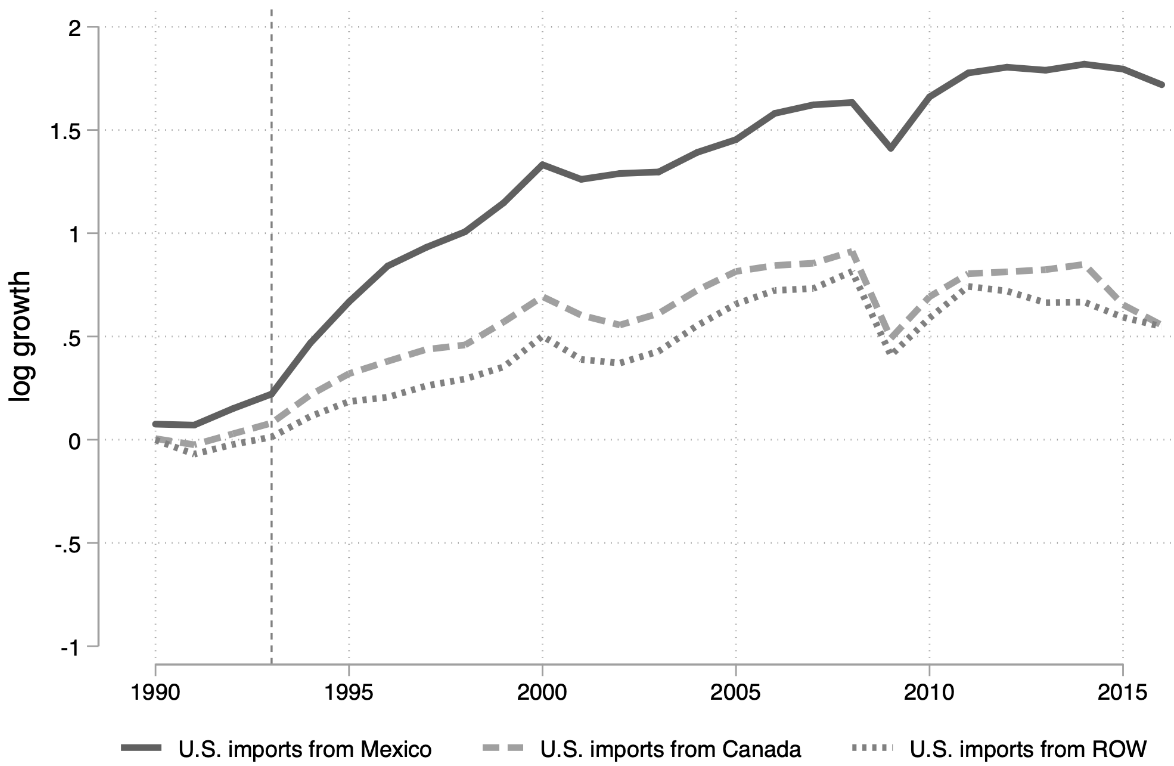
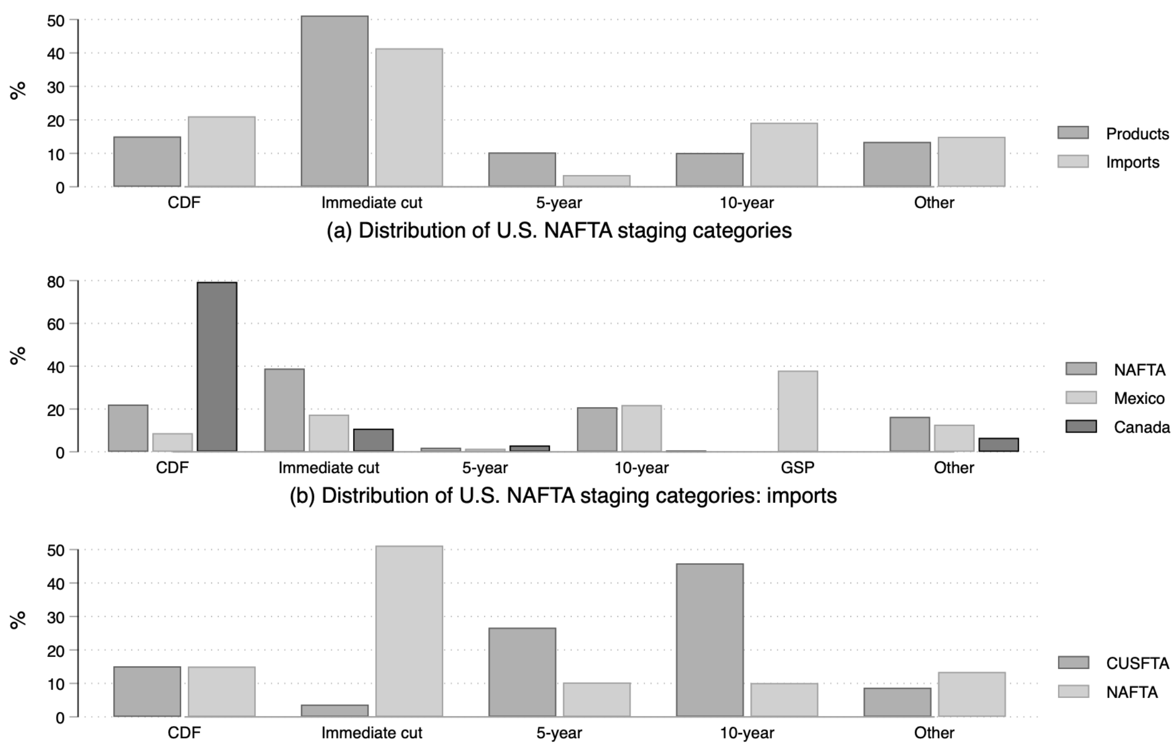


Figure 1: Cumulative real growth of U.S. imports from 1989.

Notes: Rest of the World (ROW) excludes Canada (Mexico) for Mexican imports (Canadian imports) and countries with a U.S. FTA in sample period. China excluded from ROW. Vertical line in 1993 marks year before NAFTA was implemented. Import data from USITC, GDP deflators from World Bank Development Indicators.



(a) Distribution of U.S. NAFTA staging categories

(b) Distribution of U.S. NAFTA staging categories: imports

(c) Distribution of U.S. CUSFTA and NAFTA staging categories: products

Figure 2: Distributions of U.S. NAFTA staging categories.

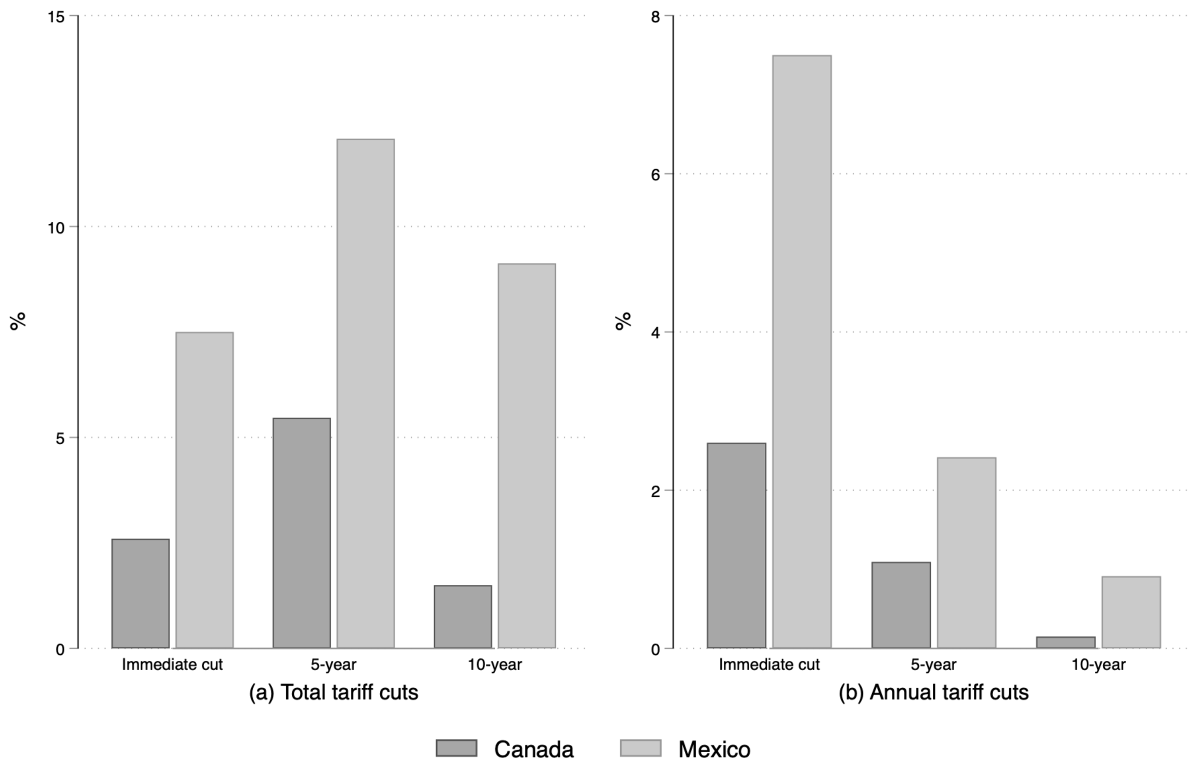


Figure 3: Tariff cuts implied by U.S. NAFTA staging categories.

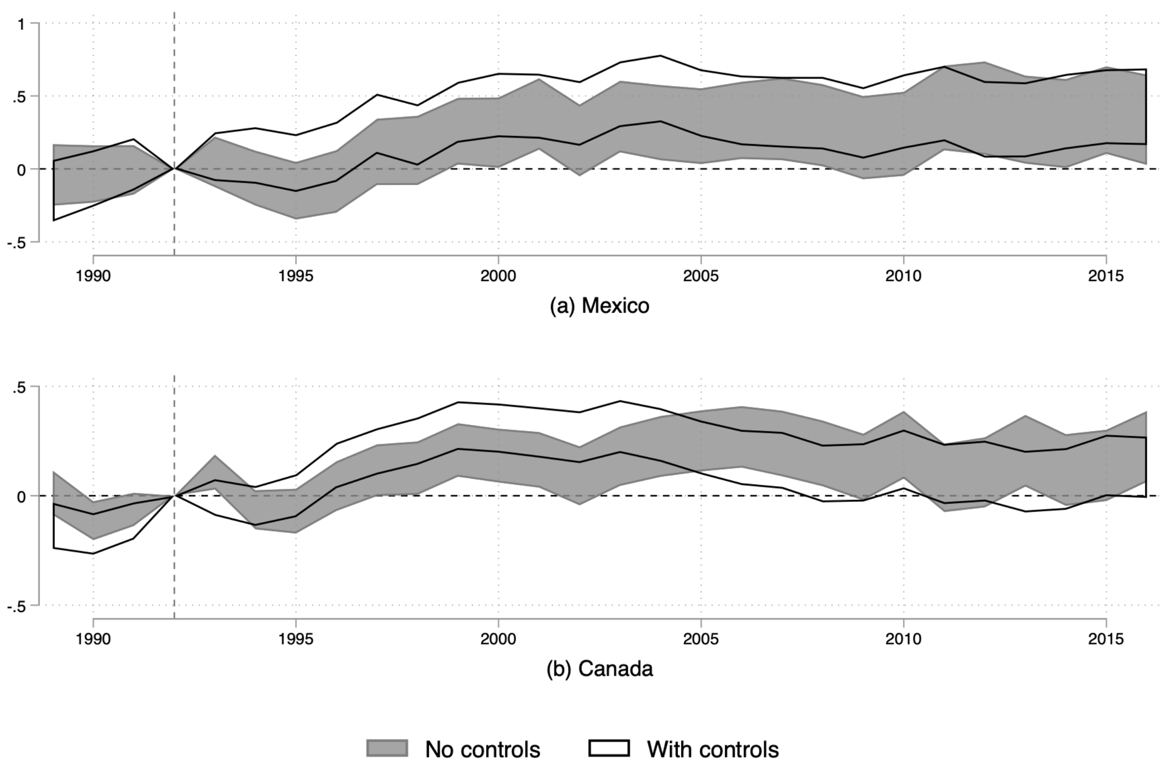


Figure 4: Time-varying homogeneous DDD estimates for Mexico and Canada.

Notes: DDD point estimates correspond to equation (1) ("No controls") and equation (2) ("With controls"), respectively. Plots represent 95% confidence intervals. Two-way clustered standard errors are used, clustering on both country-year and product-year.

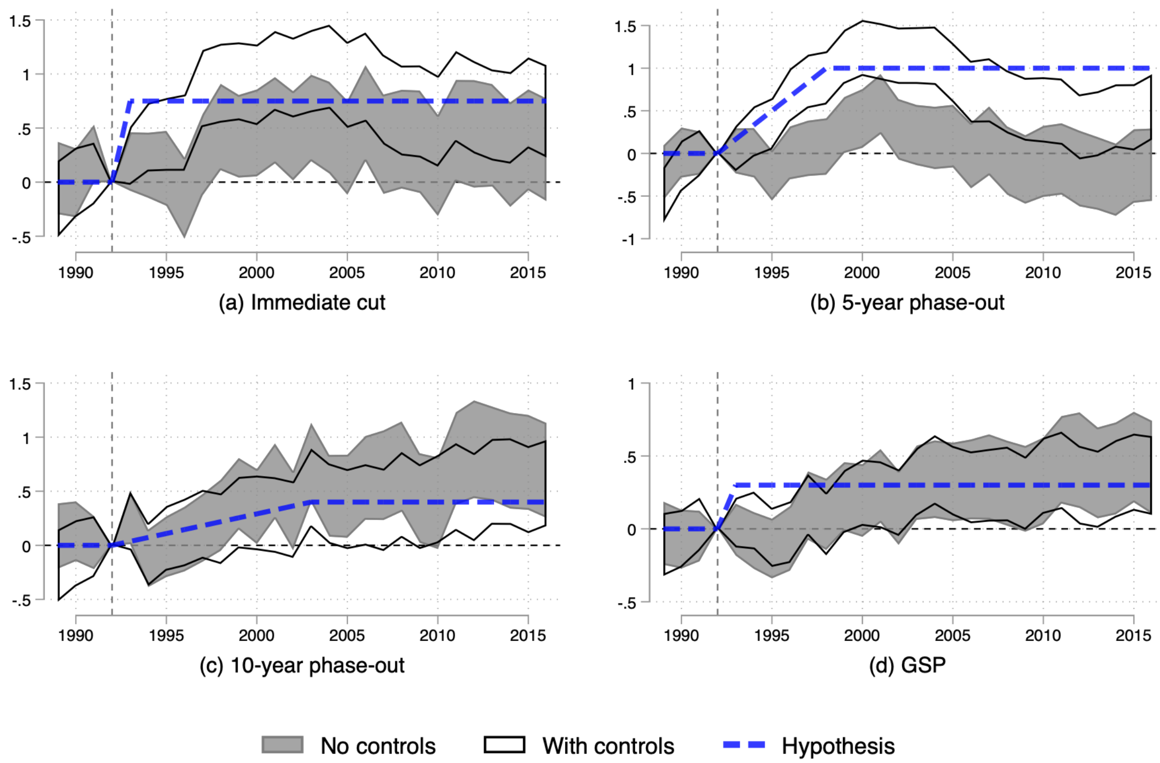


Figure 5: Time-varying heterogeneous DDD estimates for Mexico.

Notes: DDD point estimates correspond to equation (3). Plots represent 95% confidence intervals. Two-way clustered standard errors are used, clustering on both country-year and product-year.

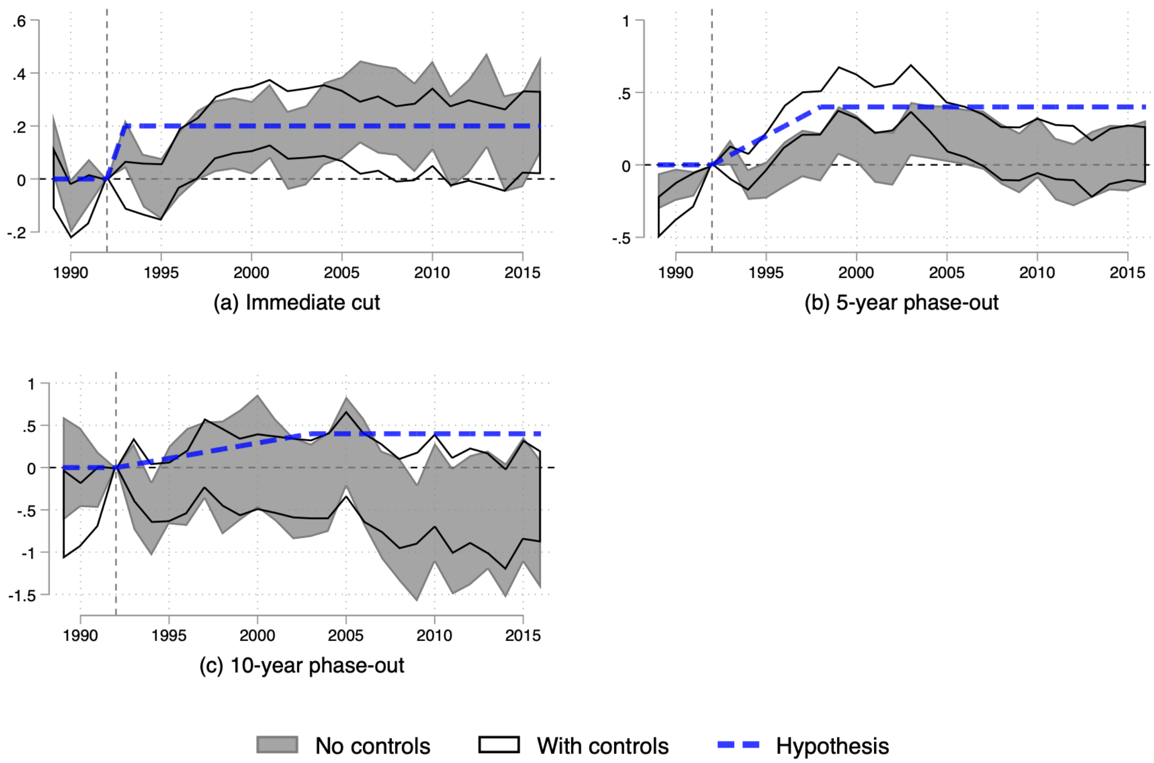


Figure 6: Time-varying heterogeneous DDD estimates for Canada.

Notes: DDD point estimates correspond to equation (3). Plots represent 95% confidence intervals. Two-way clustered standard errors are used, clustering on both country-year and product-year.

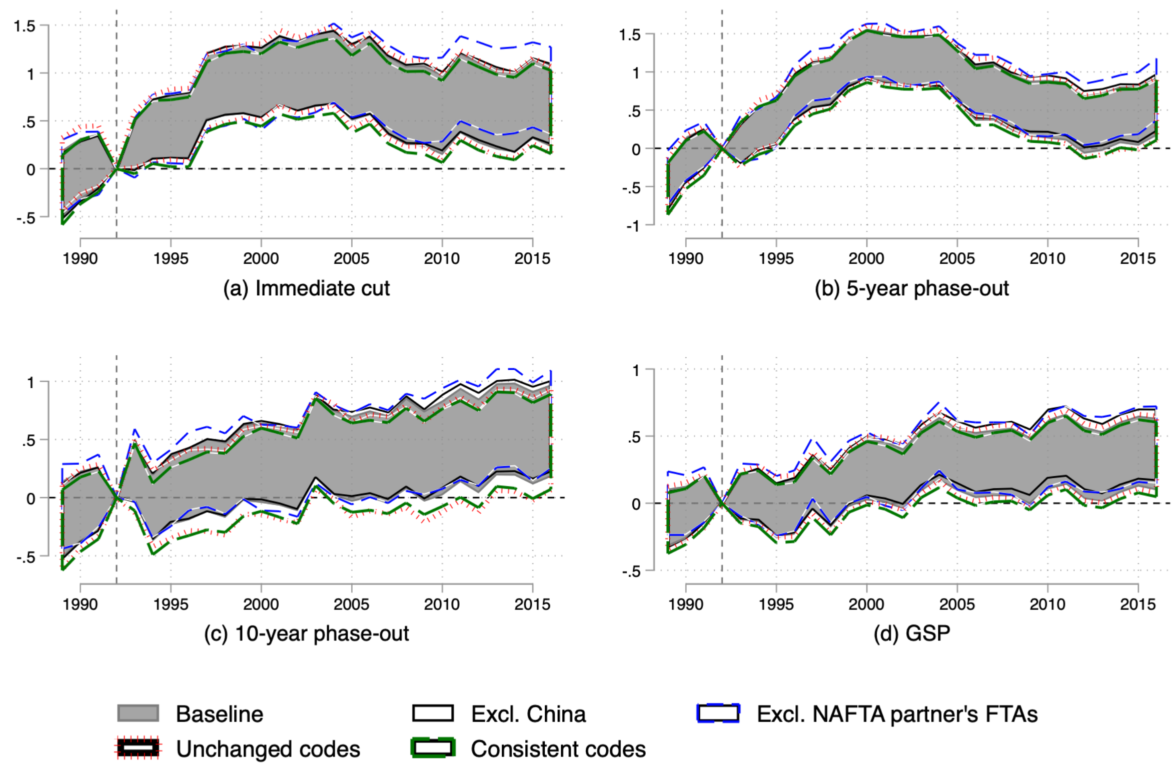


Figure 7: Time-varying heterogeneous DDD estimates for Mexico: Robustness.

Notes: DDD point estimates correspond to equation (3). Plots represent 95% confidence intervals. Two-way clustered standard errors are used, clustering on both country-year and product-year. Excl. China indicates China excluded from ROW definition; Excl. NAFTA partner's FTAs indicates that Mexico's FTA partners excluded from ROW for entire sample; Unchanged codes indicates products where product code changes over time excluded from sample; Consistent codes indicates we use extended version, through 2016, of Pierce and Schott (2012) concordance.

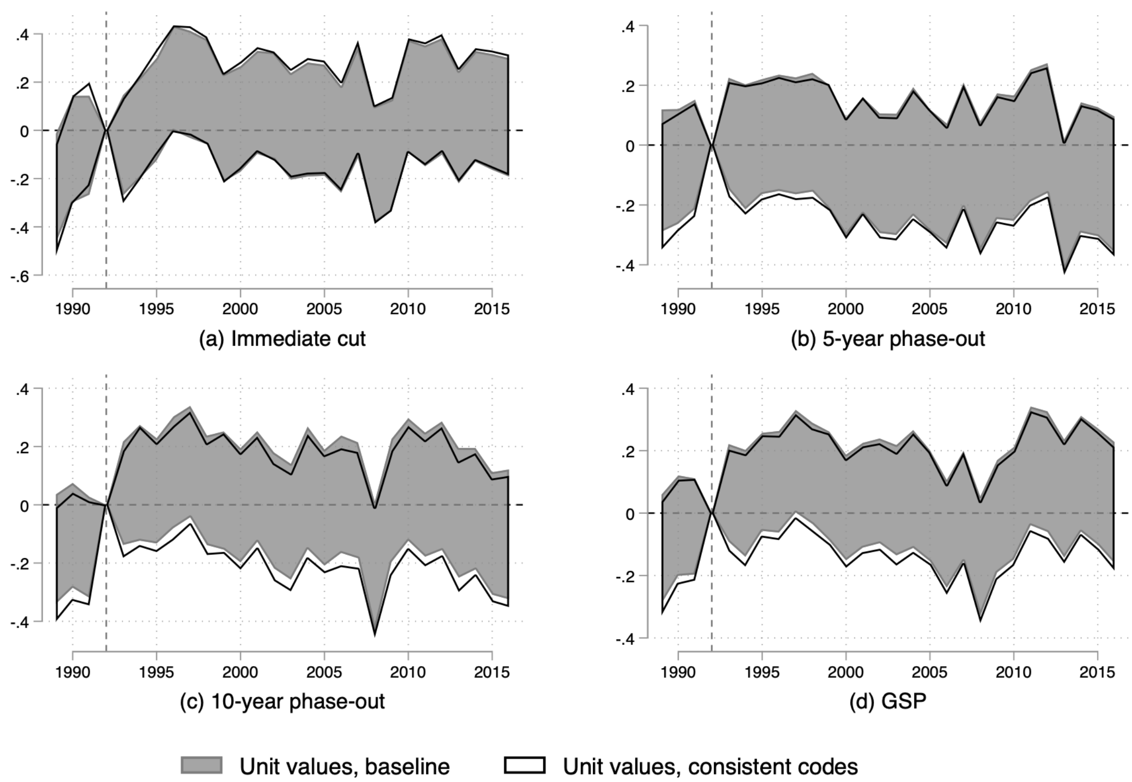


Figure 8: Time-varying heterogeneous DDD estimates for Mexico: Unit values.

Notes: DDD point estimates correspond to equation (3) with log unit values as dependent variable. Plots represent 95% confidence intervals. Two-way clustered standard errors are used, clustering on both country-year and product-year. Consistent codes indicates we use extended version, through 2016, of Pierce and Schott (2012) concordance.

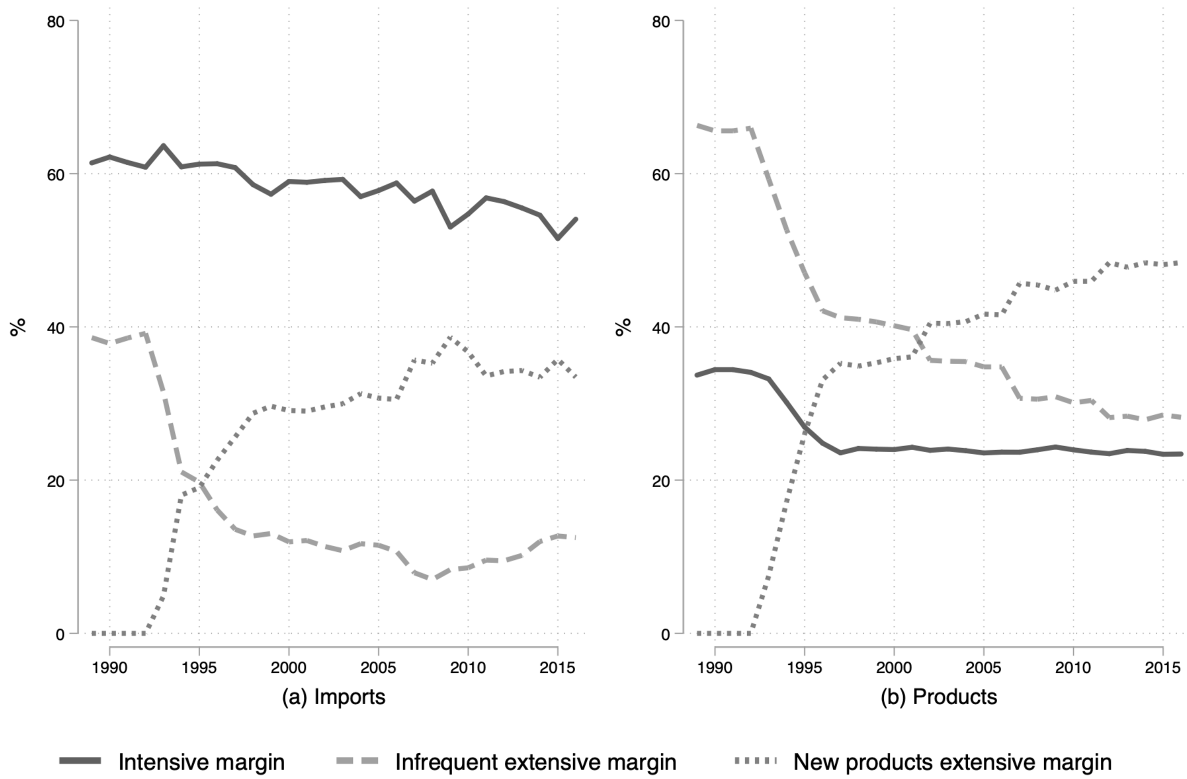


Figure 9: Imports from Mexico: Intensive margin, infrequent extensive margin, and new products extensive margin.

Notes: Intensive margin are products imported from Mexico in every year of our sample. Infrequent extensive margin are products imported from Mexico before 1993 but not imported in every year of our sample. New products extensive margin are products not imported from Mexico before 1993. Import data from USITC.

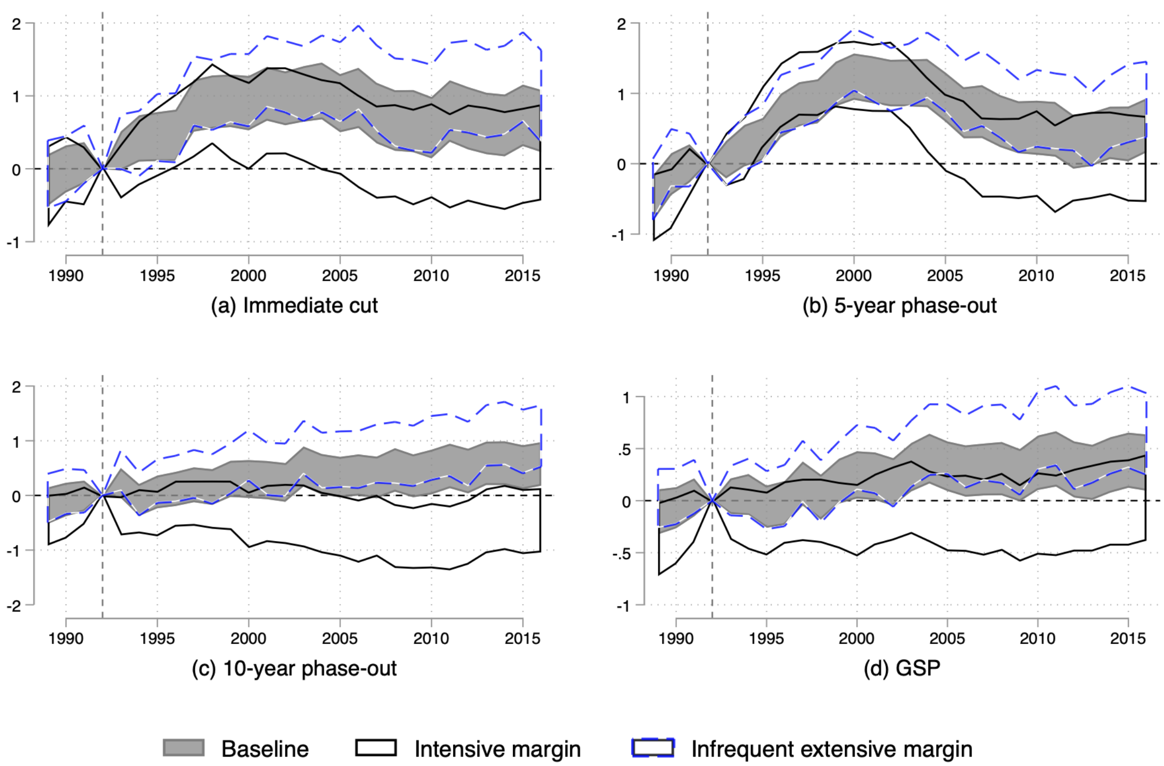


Figure 10: Time-varying heterogeneous DDD estimates for Mexico: Intensive and extensive margins.

Notes: DDD point estimates correspond to equation (3). Plots represent 95% confidence intervals. Two-way clustered standard errors are used, clustering on both country-year and product-year. Intensive margin and Infrequent extensive margin specifications restrict the sample to the respective subsample of products.

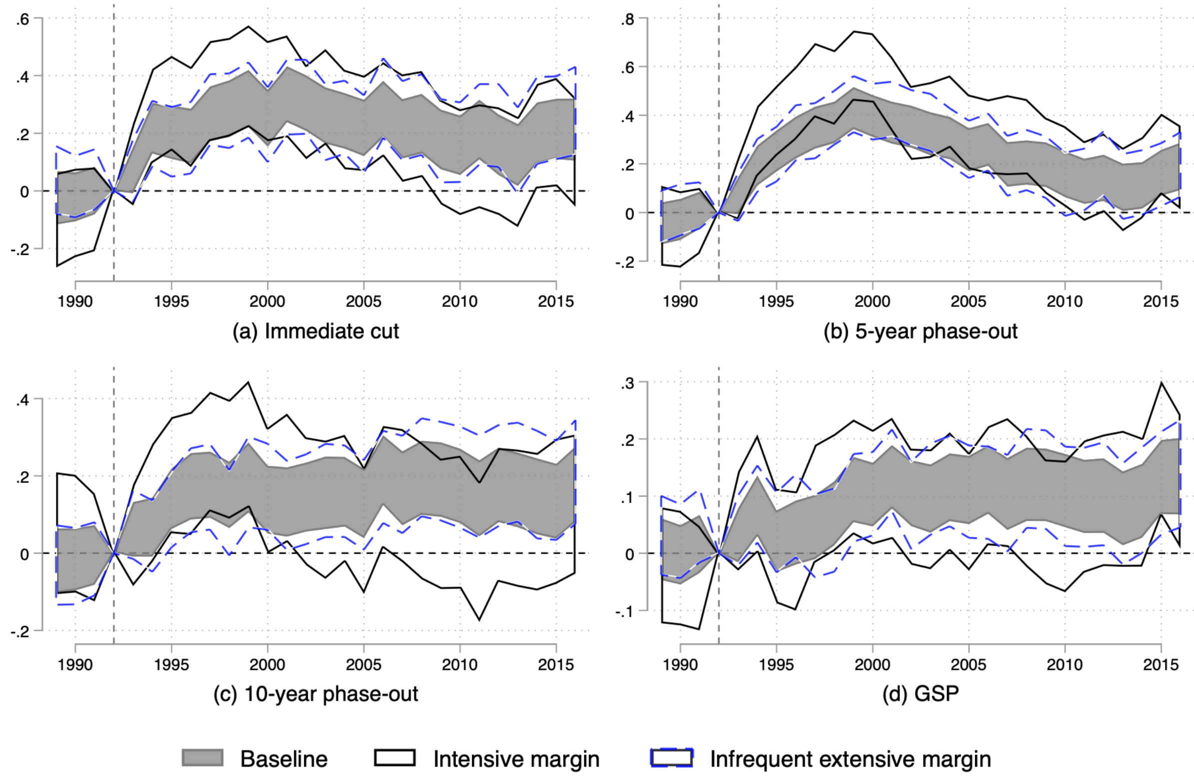


Figure 11: Time-varying heterogeneous DDD estimates for Mexico: Districts entered.

Notes: DDD point estimates correspond to equation (3) with log cumulative number of districts entered since 1989 as dependent variable. Plots represent 95% confidence intervals. Two-way clustered standard errors are used, clustering on both country-year and product-year.

Staging category	Description
A	Tariff immediately cut to 0 on January 1, 1994
B	Tariff cut to zero in 5 equal annual increments beginning January 1, 1994
B6	Tariff reduced on January 1, 1994, by "an amount equal, in percentage terms, to the base rate" and then in 5 equal annual stages beginning January 1, 1995
C	Tariff cut to zero in 10 equal annual increments beginning January 1, 1994
C10	Tariff cut non-linearly to 0 over 10 years: 20% cut on January 1, 1994, followed by 8 equal annual cuts beginning January 1, 1996
C+	Tariff cut to zero in 15 equal annual increments beginning January 1, 1994
D	Duty free before NAFTA, and continues duty free

Notes: All staging categories included in NAFTA. CUSFTA only has A, B, C and D staging categories. In the main text, we aggregate staging categories B and B6 into a "5-year" staging category and C, C10 and C+ into a "10-year" staging category.

Table 1: U.S. staging categories in CUSFTA and NAFTA

Panel A: NAFTA vs ROW approach

A1. Phase-out products

	Mexico				Canada		
	Pre-NAFTA	Post-NAFTA	Growth		Pre-NAFTA	Post-NAFTA	Growth
NAFTA partner	12.507	12.852	0.345	NAFTA partner	12.335	12.670	0.335
	(0.025)	(0.011)	(0.027)		(0.026)	(0.012)	(0.029)
	[12,690]	[79,918]			[11,612]	[61,642]	
ROW	11.603	11.359	-0.244	ROW	11.389	11.091	-0.298
	(0.005)	(0.002)	(0.005)		(0.006)	(0.002)	(0.006)
	[320,382]	[2,094,494]			[195,882]	[1,380,218]	
DD			0.589	DD			0.633
			(0.027)				(0.030)

A2. CDF-products

	Mexico				Canada		
	Pre-NAFTA	Post-NAFTA	Growth		Pre-NAFTA	Post-NAFTA	Growth
NAFTA partner	12.456	12.870	0.414	NAFTA partner	13.134	13.550	0.416
	(0.058)	(0.026)	(0.064)		(0.028)	(0.014)	(0.031)
	[2,281]	[13,524]			[10,897]	[52,998]	
ROW	11.839	11.928	0.088	ROW	11.889	11.872	-0.017
	(0.011)	(0.005)	(0.012)		(0.007)	(0.003)	(0.007)
	[52,487]	[338,494]			[151,147]	[904,627]	
DD			0.326	DD			0.433
			(.064)				(0.030)
DDD			0.263	DDD			0.200
			(0.089)				(0.051)

(continued on next page)

Table 2: Time-invariant DDD estimates of NAFTA

Panel B: Phase-out products vs CDF-products approach

B1. NAFTA partner

	Mexico				Canada		
	Pre-NAFTA	Post-NAFTA	Growth		Pre-NAFTA	Post-NAFTA	Growth
Phase-out products	12.507	12.852	0.345	Phase out products	12.335	12.670	0.335
	(0.025)	(0.011)	(0.027)		(0.026)	(0.012)	(0.029)
	[12,690]	[79,918]			[11,612]	[61,642]	
CDF-products	12.456	12.870	0.414	CDF products	13.134	13.550	0.416
	(0.058)	(.026)	(0.064)		(.028)	(0.014)	(0.031)
	[2,281]	[13,524]			[10,897]	[52,998]	
DD			-0.069	DD			-0.081
			(0.075)				(0.045)

B2. ROW

	Mexico				Canada		
	Pre-NAFTA	Post-NAFTA	Growth		Pre-NAFTA	Post-NAFTA	Growth
Phase-out products	11.603	11.359	-0.244	Phase out products	11.389	11.091	-0.298
	(0.005)	0.002)	(0.005)		(0.006)	(0.002)	(0.006)
	[320,382]	[2,094,494]			[195,882]	[1,380,218]	
CDF-products	11.839	11.928	0.088	CDF products	11.889	11.872	-0.017
	(0.001)	(0.005)	(0.012)		(0.007)	(0.003)	(0.007)
	[52,487]	[338,494]			[151,147]	[904,627]	
DD			-0.332	DD			-0.281
			(0.014)				(0.010)
DDD			0.263	DDD			0.200
			(0.089)				(0.051)

Notes: Cells contain mean log imports for the relevant group of countries, products and years. Phase-out products, CDF products, Pre-NAFTA and Post-NAFTA years defined in main text. Number of observations in square brackets. Standard errors in parentheses. For group means and growth in group means, standard errors from t-test of equivalence of group means. For difference-in-difference (DD) and triple difference (DDD) estimates, standard errors from OLS regression, clustering on both country-year and product-year. The DDD estimate in Panel A2 (B2) is the difference between the DD estimate in Panel A1 (B1) less that in Panel A2 (B2).

(cont.): Time-invariant DDD estimates of NAFTA

	Mexico				Canada			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Post	0.088 ^{b}	0.255 ^{c}			-0.017	0.243 ^{c}		
	(0.03)	(0.02)			(0.02)	(0.01)		
NAFTA	0.617 ^{c}				1.244 ^{c}			
	(0.10)				(0.05)			
Phase	-0.236 ^{c}				-0.500 ^{c}			
	(0.03)				(0.02)			
Post × NAFTA	0.326 ^{c}	0.328 ^{c}			0.433 ^{c}	0.123 ^{c}		
	(0.08)	(0.07)			(0.04)	(0.03)		
Post × Phase	-0.332 ^{c}	-0.060 ^{c}	-0.072 ^{c}		-0.281 ^{c}	-0.132 ^{c}	-0.160 ^{c}	
	(0.03)	(0.02)	(0.02)		(0.02)	(0.01)	(0.01)	
NAFTA × Phase	0.287 ^{b}				-0.299 ^{c}			
	(0.11)				(0.07)			
Post × NAFTA × Phase	0.263 ^{b}	0.363 ^{c}	0.383 ^{c}	0.371 ^{c}	0.200 ^{c}	0.298 ^{c}	0.328 ^{c}	0.260 ^{c}
	(0.09)	(0.08)	(0.08)	(0.07)	(0.05)	(0.04)	(0.04)	(0.04)
Observations	2,914,270	2,825,272	2,825,015	2,816,958	2,769,023	2,686,266	2,686,013	2,678,207
Adjusted R^2	0.013	0.726	0.740	0.763	0.033	0.730	0.744	0.766
Country × Product FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Country × Year FE	No	No	Yes	Yes	No	No	Yes	Yes
Product × Year FE	No	No	No	Yes	No	No	No	Yes

Notes: Column (1) based on equation (1) and column (4) based on equation (2). Two-way clustered standard errors in parentheses, clustering on both country-year and product-year. ^{a} $p < 0.05$, ^{b} $p < 0.01$, ^{c} $p < 0.001$.

Table 3: DDD regression: time invariant, homogeneous cumulative treatment effects

	Mexico				Canada		
	Immediate cut	5-year phase-out	10-year phase-out	GSP	Immediate cut	5-year phase-out	10-year phase-out
1990							
1991							
1992	0.157	0.492 ^{b}	0.188	0.108	-0.008	0.364 ^{c}	0.548 ^{a}
1993	0.246	0.205	0.302	0.111	0.096	0.267 ^{c}	0.528 ^{a}
1994	0.336	0.252	-0.073	0.026	0.038	0.121	0.036
1995	0.441 ^{a}	0.347 ^{a}	0.064	-0.058	-0.049	0.093	-0.287
1996	0.214	0.626 ^{c}	-0.109	-0.065	0.101	0.248 ^{c}	-0.144
1997	0.449 ^{b}	0.588 ^{c}	0.278	0.107	0.154 ^{b}	0.403 ^{c}	0.469 ^{b}
1998	0.474 ^{b}	0.539 ^{c}	0.089	0.092	0.243 ^{c}	0.265 ^{c}	0.293
1999	0.475 ^{b}	0.449 ^{c}	0.183	0.210 ^{a}	0.138 ^{b}	0.261 ^{c}	0.059
2000	0.035	0.393 ^{b}	0.105	0.083	0.111 ^{a}	0.118	-0.217
2001	0.114	0.309 ^{a}	0.127	0.200 ^{a}	0.056	0.020	-0.090
2002	0.034	0.013	-0.065	-0.008	-0.012	-0.124	-0.012
2003	0.126	-0.089	0.230	0.074	-0.016	0.055	-0.091
2004	0.039	-0.049	0.105	0.170	-0.030	0.022	-0.014
2005	-0.067	-0.199	0.098	0.151	-0.004	-0.138	0.282
2006	-0.055	-0.424 ^{c}	-0.155	-0.037	-0.055	-0.300 ^{c}	0.021
2007	-0.303 ^{a}	-0.406 ^{b}	-0.057	-0.105	-0.049	-0.231 ^{b}	-0.145
2008	-0.237	-0.342 ^{a}	0.131	-0.023	-0.067	-0.183 ^{a}	-0.584 ^{b}
2009	-0.318 ^{a}	-0.207	-0.017	-0.040	-0.016	-0.151 ^{a}	-0.243
2010	-0.201	-0.229	0.100	0.065	0.024	-0.039	0.089
2011	0.129	-0.116	0.073	0.094	-0.007	0.010	-0.021
2012	0.039	-0.206	0.088	0.056	0.006	0.006	0.028
2013	0.058	-0.163	0.160	-0.092	-0.068	-0.158 ^{a}	-0.270
2014	-0.197	-0.051	0.050	-0.058	-0.016	-0.030	-0.162
2015	0.040	0.112	0.069	0.089	0.032	0.002	0.072
2016	0.033	0.197	-0.011	0.094	0.048	0.097	0.080
Observations	2,816,958				2,678,207		
Country × Product FE	Yes				Yes		
Country × Year FE	Yes				Yes		
Product × Year FE	Yes				Yes		

Notes: Point estimate for year t coefficient is difference in year t and year $t-3$ point estimates from Figure 5, 6 and Table B.3; p-values obtained from t-test for equality of these point estimates. ^{a} $p<0.05$, ^{b} $p<0.01$, ^{c} $p<0.001$.

Table 4: DDD regression: time varying, heterogeneous cumulative treatment effects for 3-year rolling windows

A Appendix

A.1 Staging category endogeneity

A complementary way of investigating the DDD identification assumption in our setting is looking at the empirical determinants of the U.S. NAFTA staging categories. This helps address the possibility of an econometric endogeneity problem given these staging category assignments are our treatment variable but were negotiated as part of NAFTA.

In principle, the U.S. staging schedule in NAFTA need not look anything like that in CUSFTA. Indeed, panel (c) of Figure 2 starkly shows how the U.S. phased out tariffs on Canadian imports under CUSFTA for about 70% of products whereas the prima facie U.S. NAFTA tariff schedule only does this for about 20% of products. Moreover, as noted above, NAFTA also imposes that the U.S. CUSFTA tariff schedule overrides its NAFTA staging schedule for imports from Canada when the NAFTA schedule is more protectionist. Nevertheless, the U.S. NAFTA staging schedule still delivers tariff free Mexican imports in the same post-NAFTA year as predicted for Canadian imports by CUSFTA for nearly 50% of CUSFTA products. That is, economic and political economy forces driving the *CUSFTA* negotiations fundamentally shape the U.S. *NAFTA* staging schedule faced by *Mexico*. This dramatically reduces the extent to which shocks driving post-NAFTA import growth from Mexico could impact the U.S. NAFTA staging schedule faced by Mexico.

Our regression based results illustrate this point, using the ordered logit framework. When analyzing the determinants of U.S. CUSFTA staging categories for Canadian imports, the ordered outcome variables are continue duty free products (coded 0), immediate-cut products (coded 1), 5-year phase-out products (coded 2) and 10-year phase-out products (coded 3). We also run two additional regressions for Mexican and Canadian staging outcomes under NAFTA. Here, there is also the 15-year phase-out staging category (coded 4).[34] Given USITC HS data start in 1989, our CUSFTA analysis only uses 1989 data even though this is

[34]GSP, Mixed, and missing staging assignments are excluded from the analysis.

not ideal since CUSFTA was implemented on January 1, 1989. Since NAFTA negotiations concluded in mid-1992, our NAFTA analysis uses 1991 data.

In all three regressions, the independent variables are those that could drive both U.S. imports from NAFTA partners and the U.S. staging categories for imports from NAFTA partners. Specifically, we consider (i) the contemporaneous share of U.S. product-level imports and exports from Mexico or Canada and, for our NAFTA analysis, the change in this share from 1989 to 1991, (ii) the product-level log ratio of US imports to US exports with ROW, (iii) the product-level Grubel-Lloyd Index (GLI) measuring the extent of U.S. intra-industry trade with Canada or Mexico, (iv) the relevant U.S. CUSFTA or NAFTA base rate (see Online Appendix B.1), (v) the product-level elasticity of substitution, σ (from Broda and Weinstein 2006), (vi) a dummy variable for whether the product is homogeneous (as opposed to differentiated, per Rauch 1999), and (vii) dummy variables for whether the product is an intermediate good or primary good (as opposed to a final good, per the BEC classification).[35] We also include 4-digit SIC fixed effects to control for various industry-level variables including employment, value added, and the capital-labor ratio.

Table A.4 presents our results (with supporting summary statistics in Table A.3). Each coefficient measures the impact of a one-unit change in the explanatory variable on the *odds ratio* which is defined as the relative probability of a hypothetical product having at least as long phase out to having a shorter phase out. First, higher CUSFTA base rates are associated with substantially longer phase outs: a one standard deviation increase in the CUSFTA base rate increases the odds ratio by 5.42-6.19 times.[36] Second, homogeneous and intermediate or primary goods have substantially shorter phase-out periods: for homogeneous goods, the odds ratio is 0.50-0.59 times that of differentiated goods; for intermediate and primary goods respectively, these odds ratios are 0.48-0.56 and 0.06-0.10 times that of final goods. While much smaller in economic magnitude, products with higher import shares from Canada have shorter phase-out: a one standard deviation increase in the import share changes the odds

[35]Homogeeous goods are reference-priced goods or goods traded on an organized exchange.

[36]Note, $1.292^{6.599} = 5.422$. and $1.318^{6.599} = 6.185$

ratio by 0.82-0.87 times. In contrast, the other explanatory variables, including the elasticity of substitution, are not robust determinants of CUSFTA phase-out length. Nevertheless, overall, the U.S. CUSFTA staging schedule appears closely related to key economic variables.

When it comes to U.S. NAFTA staging categories, the relationships in CUSFTA are notably dampened. For example, the impact of a one standard deviation increase in the NAFTA partner base rate increases the odds ratio by 2.60-2.69 times for Canadian imports and 1.69-1.78 times for Mexican imports. Moreover, changes in the share of imports from, or exports to, NAFTA partners over 1989-1991 matter little for NAFTA phase-out length. The only exception is for Canadian imports where a rising share of imports from Canada leads to longer phase out and largely offsets the impact of a high import share.

Our econometric analysis of the factors driving the length of product-level tariff phase-out under NAFTA suggests that endogeneity of the NAFTA staging categories is not a strong concern, especially for Mexican imports. Our country-product and product-year fixed effects control for the share of imports from and exports to NAFTA partners, the base rate for each NAFTA partner, and whether goods are intermediate or primary goods among other factors.

		Panel A		Panel B						Panel C					
Staging category		CUSFTA: product level data		NAFTA: distribution of products						NAFTA: distribution of imports (tn)					
				NAFTA		Canada		Mexico		NAFTA		Canada		Mexico	
Code	Description	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
A	Immediate cut to 0	315	3.7%	4,526	51.2%	2,535	28.7%	981	11.1%	2.74	38.9%	0.48	10.8%	0.45	17.4%
B	5 equal annual cuts to 0	2,285	26.7%	179	2.0%	792	9.0%	169	1.9%	0.08	1.2%	0.11	2.5%	0.01	0.4%
B6	1 immediate cut + 5 equal annual cuts to 0			728	8.2%	728	8.2%	726	8.2%	0.05	0.7%	0.02	0.5%	0.03	1.0%
C	10 equal annual cuts to 0	3,932	45.9%	750	8.5%	94	1.1%	737	8.3%	1.46	20.7%	0.01	0.3%	0.56	21.6%
C10	Non-linear cuts to zero over 10 years			71	0.8%	0	0.0%	71	0.8%	0.01	0.1%	0.00	0.0%	0.01	0.2%
C+	15 equal annual cuts to 0			74	0.8%	3	0.0%	72	0.8%	0.00	0.1%	0.00	0.0%	0.00	0.0%
D	Continue duty free	1,295	15.1%	1,329	15.0%	3,871	43.8%	1,301	14.7%	1.56	22.1%	3.56	79.4%	0.22	8.7%
GSP								4,194	47.3%					0.98	37.9%
Mixed		745	8.7%	1,120	12.7%	755	8.5%	559	6.3%	1.11	15.7%	0.29	6.4%	0.29	11.3%
Missing		2	0.0%	66	0.7%	65	0.7%	66	0.7%	0.04	0.6%	0.01	0.1%	0.04	1.4%
Total		8,574	100%	8,843	100%	8,843	100%	8,876	100%	7.06	100%	4.48	100%	2.58	100%

Notes: Staging category data comes from CUSFTA Article 401 and Annex 401.2 and NAFTA Annex 302.2. Panels A and B describe the distribution of products in these Annexes across staging categories. Columns (5)-(6) modify the NAFTA staging categories for consistency with CUSFTA staging categories. Columns (7)-(8) modify the NAFTA staging categories for consistency with Mexico's product-level eligibility for the US Generalized System of Preferences (GSP) program. Panel C merges the NAFTA staging category data with 8-digit HS USITC data on bilateral imports from Canada and Mexico for the period 1989-2016. Panels A and B only use products that appear in this USITC import data. Imports are measured in trillions of real 2010 USD using the World Development Indicator GDP deflator. In CUSFTA, the two "Missing" products were phased out in three equal annual cuts beginning January 1, 1989. Of the 66 products listed as having a "Missing" staging category in columns (3)-(4), 37 had a non-linear phase out that was not associated with a particular staging category. For example, 0703.90.00 represents "Leeks and other alliaceous vegetables" and had its tariff cut from a base rate of 25% to 14.4% on January 1, 1993, and then had its tariff phased out over 5 equal annual cuts. A further 27 products were sets of articles (e.g. tols, textile ensembles, watch parts) where the staging category applied either to each individual item separately or the complete item specified elsewhere (e.g. 6103.22.00 representing "Men's or Boy's cotton suit ensembles".) The final two products were articles re-entering after being sent abroad for further processing or assembly out of U.S. parts. For the value of imports here, the tariff applies as if the entire article itself was imported. See Table A.2 for further details.

Table A.1: NAFTA and CUSFTA tariff schedule staging categories

Code	Description	CUSFTA tariff cuts on Canada			NAFTA tariff cuts on Canada			NAFTA tariff cuts on Mexico		
		Products	Pre-CUSFTA mean tariff	Mean annual tariff cut	Products	Pre-NAFTA mean tariff	Mean annual tariff cut	Products	Pre-NAFTA mean tariff	Mean annual tariff cut
A	Immediate cut to 0	312	3.6%	3.6%	2,508	2.6%	2.6%	961	7.5%	7.5%
B	5 equal annual cuts to 0	2,284	5.8%	1.2%	791	4.7%	0.9%	168	9.4%	1.9%
B6	1 immediate cut + 5 equal annual cuts to 0				727	6.3%	1.1%	726	12.7%	2.1%
C	10 equal annual cuts to 0	3,894	8.5%	0.9%	85	1.5%	0.2%	735	7.6%	0.8%
C10	Non-linear cuts to zero over 10 years							71	14.1%	1.4%
C+	15 equal annual cuts to 0							72	19.8%	1.3%
D	Continue duty free	1,295	N/A	N/A	3,871	N/A	N/A	1,301	N/A	N/A
GSP								4,194	N/A	N/A
Mixed		745	N/A	N/A	755	N/A	N/A	559	N/A	N/A
Missing		2	N/A	N/A	65	N/A	N/A	66	N/A	N/A
Total		8,532			8,802			8,853		

Notes: Appendix B.1 explains the construction of the pre-CUSFTA and pre-NAFTA tariffs.

Table A.2: Tariff cuts by staging category

Variable	Mean	Std. dev.	Min.	Max.	Obs.
Canada-CUSFTA					
Staging category	3.14	1.08	1	4	7,580
Import share	23.41	35.77	0	100	7,580
Export share	10.88	18.75	0	100	7,580
Import/export ratio	12.91	14.92	-32.10	36.47	7,580
GLI	.15	.26	0	1	7,580
Base rate	6.28	6.70	0	79.00	7,580
Homogeneous	.43	.50	0	1	7,580
Intermediate	.51	.50	0	1	7,580
Primary	.06	.24	0	1	7,580
σ	10.45	36.88	1.06	1,303.43	5,731
Canada-NAFTA					
Staging category	1.72	.80	1	4	7,715
Import share	23.42	35.62	0	100	7,715
Export share	12.62	21.20	0	100	7,715
Import/export ratio	12.48	14.96	-33.05	36.58	7,715
GLI	.14	.26	0	1	7,715
Base rate	1.97	3.37	0	40.44	7,715
Homogeneous	.42	.49	0	1	7,715
Intermediate	.53	.50	0	1	7,715
Primary	.06	.24	0	1	7,715
Change import share	.80	20.53	-100	100	7,715
Change export share	1.63	16.27	-100	100	7,715
σ	9.97	36.22	1.06	1,303.43	5,880
Mexico-NAFTA					
Staging category	2.33	1.15	1	5	3,829
Import share	26.63	42.05	0	100	3,829
Export share	7.14	16.88	0	100	3,829
Import/export ratio	12.75	14.83	-33.05	35.34	3,829
GLI	.09	.24	0	1	3,829
Base rate	6.80	8.20	0	80.87	3,829
Homogeneous	.45	.50	0	1	3,829
Intermediate	.49	.50	0	1	3,829
Primary	.10	.30	0	1	3,829
Change import share	1.01	19.28	-100	100	3,829
Change export share	.41	12.44	-100	100	3,829
σ	10.28	35.53	1.10	1,110.16	3,064

Notes: Descriptive statistics for variables underlying ordered logistic regression results in Table A.4.

Table A.3: Descriptive statistics

	Canada CUSFTA		Canada NAFTA		Mexico NAFTA	
	(1)	(2)	(3)	(4)	(5)	(6)
Import share	0.994[c]	0.996[c]	0.997[b]	0.995[c]	1.001	1.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Export share	0.996	0.998	0.995	0.995[a]	0.990[b]	0.991[c]
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Import/export ratio	1.005	1.007[a]	1.002	1.004	1.003	1.004
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
GLI	1.055	0.941	1.127	1.574[b]	0.840	0.851
	(0.17)	(0.13)	(0.26)	(0.28)	(0.20)	(0.16)
Base rate	1.318[c]	1.292[c]	1.340[c]	1.326[c]	1.079[c]	1.087[c]
	(0.02)	(0.02)	(0.04)	(0.03)	(0.01)	(0.01)
Homogeneous	0.497[c]	0.589[c]	0.797	0.788[a]	0.774	0.827
	(0.08)	(0.08)	(0.11)	(0.09)	(0.12)	(0.11)
Intermediate	0.482[c]	0.566[c]	0.794	0.936	0.486[b]	0.617[b]
	(0.09)	(0.07)	(0.11)	(0.09)	(0.11)	(0.11)
Primary	0.0554[c]	0.105[c]	0.126[c]	0.241[c]	0.103[c]	0.187[c]
	(0.02)	(0.03)	(0.04)	(0.06)	(0.04)	(0.05)
Change import share			1.006[b]	1.007[c]	0.995	0.997
			(0.00)	(0.00)	(0.00)	(0.00)
Change export share			1.002	1.005[a]	1.003	1.003
			(0.00)	(0.00)	(0.00)	(0.00)
σ	1.001		1.002[a]		1.001	
	(0.00)		(0.00)		(0.00)	
Observations	5,731	7,580	5,880	7,715	3,064	3,829
Possible total	7,827	7,827	8,023	8,023	4,057	4,057
SIC industry FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Ordered logistic regression results (odds ratios). Outcomes ranked from least restrictive to most restrictive staging category. Robust standard errors in parentheses, clustering at HS8 level. [a] $p < 0.05$, [b] $p < 0.01$, [c] $p < 0.001$. For each column, the possible total number of product stagings that can be predicted is calculated as the Total minus the GSP, Mixed and Missing categories in the related column in Appendix Table A.1.

Table A.4: Determinants of staging outcomes

B Online Appendix (not for publication)

B.1 Constructing pre-CUSFTA and pre-NAFTA tariffs

While we can extract staging categories from the CUSFTA and NAFTA texts, it is extremely difficult to extract base rates, i.e. pre-FTA tariffs imposed on members, from these texts. Thus, we construct pre-CUSFTA and pre-NAFTA tariffs according to the following procedure.

As a starting point for pre-CUSFTA tariffs faced by Canada, we take the 1989 U.S. MFN tariffs per John Romalis' data described in Feenstra, Romalis, and Schott (2002) (hereafter "Romalis' tariff data"). This is reasonable because adjusting these 1989 U.S. MFN tariffs by a products' CUSFTA staging category nearly always equals the 1989 preferential tariff faced by Canadian imports per Romalis' tariff data. For the 0.69% of products where the difference is more than rounding error (i.e. more than .01% points), we manually check the CUSFTA text and adjust accordingly. We also manually check the CUSFTA text for products where the tariff is immediately cut to zero and their 1989 U.S. MFN tariff is missing per Romalis' tariff data. Additionally, products 2207.10.30 and 2401.30.60 have respective ad valorem equivalent Canadian preferential tariffs per Romalis' tariff data of 673% and 97% (the next highest is 57.5%), so we treat these as outliers and exclude them for the purpose of tariff summary statistics.

Ultimately, we match $8,574$ products from the CUSFTA staging schedule to USITC import data and $7,827$ of these are not in the "Mixed" or "Missing" staging categories. Of these $7,827$ products, we have an imputed pre-CUSFTA tariff faced by Canada for $7,785$ products. Of the 42 products with missing pre-CUSFTA tariffs, 5 have specific tariffs but do not have an ad valorem equivalent tariff per Romalis' tariff data and we cannot compute one based on pre-CUSFTA imports because our USITC import data begins in 1989. The remaining 37 products have "complex" base rates that cannot be transformed into an ad valorem equivalent tariff with USITC import data.[37]

For Canada's pre-NAFTA tariff, we initially follow a two-step procedure. First, a product's pre-NAFTA tariff must be zero if its CUSFTA staging category is either A, D or B. Second, for products phased out over 10 years under CUSFTA with ad valorem tariffs, their pre-NAFTA tariff must be half of their pre-CUSFTA tariff. For remaining products, we use the 1993 Canada preferential tariff per Romalis' tariff data. If this is not available, we compute the ad valorem equivalent tariff using the CUSFTA base rate, CUSFTA staging category and the last available pre-NAFTA import level from the USITC.

Ultimately, we match $8,843$ products from the NAFTA staging schedule to USITC import data and $8,023$ of these are not in the "Mixed" or "Missing" staging categories for Canada. Of these $8,023$ products, we have an imputed pre-NAFTA tariff faced by Canada for $7,982$ products. Of the 41 products with missing pre-NAFTA tariffs, 5 have complex tariff structures and 2 are specific tariffs but we cannot compute an ad valorem equivalent because they were not imported from Canada before NAFTA per our USITC import data. A further 29 NAFTA products were not in CUSFTA and their tariff is missing per Romalis' tariff data. The final 5 products were part of a CUSFTA "mixed" product and hence we do not know its CUSFTA base rate and, in turn, cannot compute its pre-NAFTA tariff.

For Mexico's pre-NAFTA tariff, the process is much simpler. For Mexico's pre-NAFTA GSP eligible products and for NAFTA staging category D products, the pre-NAFTA tariff is zero. For other products, we first check the U.S. 1993 MFN ad valorem equivalent tariff per Romalis' tariff data.[38] For remaining products, we self-compute an ad valorem equivalent tariff using the NAFTA base rate and the last available pre-NAFTA import level from the USITC. Of the $8,876$ Mexican products that we can match from the NAFTA schedule or GSP eligibility to USITC import data, $8,251$ are not in the "Mixed" or "Missing" staging categories. Of these $8,251$ products, we have pre-NAFTA tariffs for $8,228$. Of the remaining

[37]For example, the base rate for product 2613.90.00, which is *other molybdenum ore and concentrate*, depends on the amount of molybdenum content.

[38]According to the Romalis' tariff data, the U.S. does not change any advalorem MFN tariffs between 1991 and 1993.

23 products, 19 have complex tariff structures and 4 have specific MFN tariffs but we cannot self-compute an ad valorem equivalent tariff because the product was not imported from Mexico before NAFTA per our USITC import data.

B.2 Supplemental tables

Year	Product-level data			Exporter-product-level data					
	Panel A			Panel B: Products			Panel C: Import values ($tn)		
	Trade	Stagings	Match %	Trade	Stagings	Match %	Trade	Stagings	Match %
1989	8,602	8,393	97.57%	131,048	127,390	97.21%	$0.44	$0.43	97.11%
1990	8,677	8,456	97.45%	126,447	122,960	97.24%	$0.47	$0.45	97.22%
1991	8,659	8,523	98.43%	125,963	123,708	98.21%	$0.46	$0.45	97.79%
1992	8,745	8,642	98.82%	129,326	127,600	98.67%	$0.50	$0.49	98.13%
1993	8,690	8,675	99.83%	134,926	134,541	99.71%	$0.53	$0.53	98.85%
1994	8,994	8,492	94.42%	145,319	136,326	93.81%	$0.62	$0.53	85.69%
1995	9,568	7,911	82.68%	151,752	129,641	85.43%	$0.69	$0.55	78.68%
1996	9,770	7,449	76.24%	158,050	125,800	79.60%	$0.74	$0.54	72.37%
1997	9,997	7,461	74.63%	168,033	130,389	77.60%	$0.80	$0.57	71.34%
1998	9,896	7,392	74.70%	168,495	130,903	77.69%	$0.85	$0.59	70.18%
1999	9,876	7,406	74.99%	170,030	132,860	78.14%	$0.94	$0.65	69.72%
2000	9,908	7,412	74.81%	178,080	138,807	77.95%	$1.11	$0.78	70.41%
2001	9,917	7,406	74.68%	178,476	138,543	77.63%	$1.03	$0.76	73.47%
2002	10,163	6,955	68.43%	185,114	134,846	72.84%	$1.05	$0.71	67.79%
2003	10,179	6,953	68.31%	188,279	136,934	72.73%	$1.13	$0.77	67.91%
2004	10,155	6,950	68.44%	191,986	139,445	72.63%	$1.33	$0.90	67.44%
2005	10,172	6,944	68.27%	195,741	141,474	72.28%	$1.52	$1.03	67.60%
2006	10,188	6,951	68.23%	198,368	142,945	72.06%	$1.69	$1.14	67.83%
2007	10,116	6,343	62.70%	197,675	133,373	67.47%	$1.78	$1.15	64.51%
2008	10,095	6,339	62.79%	192,709	130,455	67.70%	$1.92	$1.26	65.64%
2009	10,043	6,326	62.99%	183,535	124,129	67.63%	$1.40	$0.87	62.17%
2010	10,053	6,326	62.93%	189,482	128,011	67.56%	$1.71	$1.09	63.31%
2011	10,098	6,333	62.72%	194,088	131,505	67.76%	$1.99	$1.27	63.89%
2012	10,300	6,093	59.16%	197,081	128,289	65.09%	$2.04	$1.28	62.96%
2013	10,287	6,091	59.21%	193,084	126,253	65.39%	$2.02	$1.27	62.87%
2014	10,299	6,087	59.10%	196,866	128,667	65.36%	$2.08	$1.30	62.30%
2015	10,308	6,096	59.14%	203,138	132,535	65.24%	$1.97	$1.21	61.35%
2016	10,297	6,099	59.23%	204,767	133,760	65.32%	$1.87	$1.15	61.57%

Notes: Staging category data refer to NAFTA U.S. tariff schedule data from NAFTA Annex 302.2. Trade data is 8-digit USITC import data. Panel C aggregates 8-digit exporter-product U.S. imports to the 8-digit level. Imports are measured in trillions of real 2010 USD using the World Development Indicator GDP deflator.

Table B.1: Matching NAFTA tariff schedule to USITC trade data

	Mexico								Canada							
	(1)		(2)		(3)		(4)		(1)		(2)		(3)		(4)	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
1989	-0.041	(0.11)	-0.155	(0.10)	-0.158	(0.10)	-0.151	(0.11)	0.013	(0.05)	-0.113^{a}	(0.05)	-0.127^{a}	(0.05)	-0.137^{a}	(0.05)
1990	-0.035	(0.10)	-0.040	(0.09)	-0.044	(0.09)	-0.066	(0.10)	-0.114^{a}	(0.05)	-0.164^{c}	(0.04)	-0.172^{c}	(0.04)	-0.175^{c}	(0.05)
1991	-0.007	(0.09)	0.054	(0.08)	0.047	(0.08)	0.032	(0.09)	-0.062	(0.04)	-0.102^{b}	(0.04)	-0.111^{b}	(0.04)	-0.115^{b}	(0.04)
1993	0.048	(0.09)	0.110	(0.08)	0.102	(0.08)	0.083	(0.08)	0.108^{b}	(0.04)	-0.000	(0.04)	-0.002	(0.04)	-0.008	(0.04)
1994	-0.064	(0.10)	0.106	(0.09)	0.108	(0.09)	0.092	(0.10)	-0.064	(0.05)	-0.032	(0.04)	-0.032	(0.04)	-0.047	(0.05)
1995	-0.150	(0.10)	0.068	(0.10)	0.079	(0.10)	0.040	(0.10)	-0.070	(0.05)	0.023	(0.05)	0.031	(0.05)	-0.000	(0.05)
1996	-0.086	(0.11)	0.155	(0.10)	0.158	(0.10)	0.117	(0.10)	0.044	(0.06)	0.157^{b}	(0.05)	0.160^{b}	(0.05)	0.138^{b}	(0.05)
1997	0.116	(0.12)	0.306^{b}	(0.10)	0.315^{b}	(0.10)	0.309^{b}	(0.11)	0.117	(0.06)	0.235^{c}	(0.05)	0.241^{c}	(0.05)	0.202^{c}	(0.05)
1998	0.126	(0.12)	0.244^{a}	(0.11)	0.262^{a}	(0.11)	0.232^{a}	(0.11)	0.126^{a}	(0.06)	0.289^{c}	(0.05)	0.303^{c}	(0.05)	0.249^{c}	(0.06)
1999	0.258^{a}	(0.12)	0.375^{c}	(0.11)	0.391^{c}	(0.11)	0.388^{c}	(0.11)	0.209^{c}	(0.06)	0.368^{c}	(0.06)	0.379^{c}	(0.06)	0.321^{c}	(0.06)
2000	0.248^{a}	(0.12)	0.414^{c}	(0.11)	0.428^{c}	(0.11)	0.438^{c}	(0.11)	0.183^{b}	(0.06)	0.327^{c}	(0.06)	0.346^{c}	(0.06)	0.309^{c}	(0.06)
2001	0.376^{b}	(0.12)	0.417^{c}	(0.12)	0.434^{c}	(0.12)	0.429^{c}	(0.11)	0.164^{a}	(0.06)	0.313^{c}	(0.06)	0.336^{c}	(0.06)	0.289^{c}	(0.06)
2002	0.195	(0.13)	0.315^{b}	(0.11)	0.325^{b}	(0.11)	0.380^{c}	(0.11)	0.091	(0.07)	0.270^{c}	(0.06)	0.293^{c}	(0.06)	0.267^{c}	(0.06)
2003	0.359^{b}	(0.12)	0.501^{c}	(0.12)	0.506^{c}	(0.12)	0.511^{c}	(0.11)	0.181^{b}	(0.07)	0.321^{c}	(0.06)	0.340^{c}	(0.06)	0.316^{c}	(0.06)
2004	0.316^{a}	(0.13)	0.501^{c}	(0.12)	0.504^{c}	(0.12)	0.551^{c}	(0.12)	0.225^{b}	(0.07)	0.340^{c}	(0.06)	0.356^{c}	(0.06)	0.278^{c}	(0.06)
2005	0.293^{a}	(0.13)	0.461^{c}	(0.12)	0.459^{c}	(0.12)	0.451^{c}	(0.12)	0.251^{c}	(0.07)	0.312^{c}	(0.06)	0.320^{c}	(0.06)	0.221^{c}	(0.06)
2006	0.332^{a}	(0.13)	0.431^{c}	(0.13)	0.427^{c}	(0.13)	0.401^{b}	(0.12)	0.269^{c}	(0.07)	0.268^{c}	(0.07)	0.271^{c}	(0.07)	0.175^{b}	(0.06)
2007	0.342^{a}	(0.14)	0.420^{b}	(0.13)	0.416^{b}	(0.13)	0.388^{b}	(0.12)	0.238^{b}	(0.08)	0.294^{c}	(0.07)	0.290^{c}	(0.07)	0.162^{a}	(0.07)
2008	0.299^{a}	(0.14)	0.422^{b}	(0.13)	0.413^{b}	(0.13)	0.382^{b}	(0.13)	0.193^{a}	(0.08)	0.230^{c}	(0.07)	0.225^{b}	(0.07)	0.102	(0.07)
2009	0.214	(0.15)	0.374^{b}	(0.13)	0.363^{b}	(0.13)	0.315^{a}	(0.12)	0.131	(0.08)	0.222^{b}	(0.07)	0.225^{b}	(0.07)	0.107	(0.07)
2010	0.240	(0.15)	0.431^{b}	(0.14)	0.424^{b}	(0.14)	0.393^{b}	(0.13)	0.233^{b}	(0.08)	0.302^{c}	(0.07)	0.305^{c}	(0.07)	0.166^{a}	(0.07)
2011	0.418^{b}	(0.15)	0.517^{c}	(0.14)	0.518^{c}	(0.14)	0.448^{c}	(0.13)	0.082	(0.08)	0.216^{b}	(0.07)	0.219^{b}	(0.07)	0.099	(0.07)
2012	0.416^{a}	(0.16)	0.418^{b}	(0.14)	0.425^{b}	(0.14)	0.340^{a}	(0.13)	0.107	(0.08)	0.237^{b}	(0.07)	0.246^{c}	(0.07)	0.113	(0.07)
2013	0.338^{a}	(0.15)	0.387^{b}	(0.14)	0.402^{b}	(0.14)	0.336^{a}	(0.13)	0.205^{a}	(0.08)	0.224^{b}	(0.07)	0.244^{b}	(0.07)	0.064	(0.07)
2014	0.310^{a}	(0.16)	0.426^{b}	(0.14)	0.443^{b}	(0.14)	0.392^{b}	(0.13)	0.117	(0.08)	0.206^{b}	(0.07)	0.225^{b}	(0.07)	0.077	(0.07)
2015	0.403^{b}	(0.15)	0.436^{b}	(0.14)	0.461^{c}	(0.14)	0.426^{b}	(0.13)	0.139	(0.08)	0.255^{c}	(0.07)	0.281^{c}	(0.07)	0.138	(0.07)
2016	0.337^{a}	(0.16)	0.448^{b}	(0.14)	0.477^{c}	(0.14)	0.426^{b}	(0.13)	0.225^{b}	(0.08)	0.257^{c}	(0.07)	0.288^{c}	(0.07)	0.130	(0.07)
Observations	2,914,270		2,825,272		2,825,015		2,816,958		2,769,023		2,686,266		2,686,013		2,678,207	
Adjusted R^2	0.014		0.728		0.740		0.763		0.034		0.732		0.744		0.766	
Country × Product FE	No		Yes		Yes		Yes		No		Yes		Yes		Yes	
Country × Year FE	No		No		Yes		Yes		No		No		Yes		Yes	
Product × Year FE	No		No		No		Yes		No		No		No		Yes	

Notes: After replacing replacing the $Post_t$ dummy with a vector of year dummies $\mathbf{Year}_t = (1989, 1990, 1991, 1993, ..., 2016)$, DDD point estimates in columns (1) and (4) correspond to, respectively, equations (1) and (2). Two-way clustered standard errors are used, clustering on both country-year and product-year. ^{a} p <0.05, ^{b} p <0.01, ^{c} p <0.001.

Table B.2: DDD regression: time varying, homogeneous cumulative treatment effects

	Mexico								Canada					
	Immediate cut		5-year phase-out		10-year phase-out		GSP		Immediate cut		5-year phase-out		10-year phase-out	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
1989	-0.157	(0.18)	-0.492 ^{b}	(0.17)	-0.188	(0.17)	-0.108	(0.11)	0.008	(0.06)	-0.364 ^{c}	(0.07)	-0.548 ^{a}	(0.27)
1990	-0.003	(0.16)	-0.148	(0.15)	-0.074	(0.16)	-0.068	(0.10)	-0.120 ^{a}	(0.05)	-0.252 ^{c}	(0.07)	-0.553 ^{b}	(0.20)
1991	0.080	(0.15)	0.004	(0.14)	-0.011	(0.14)	0.032	(0.09)	-0.076	(0.05)	-0.170 ^{b}	(0.06)	-0.337	(0.18)
1993	0.243	(0.14)	0.057	(0.14)	0.227	(0.14)	0.043	(0.09)	-0.023	(0.05)	0.015	(0.06)	-0.026	(0.19)
1994	0.416 ^{a}	(0.16)	0.256	(0.15)	-0.084	(0.15)	0.057	(0.10)	-0.038	(0.05)	-0.049	(0.07)	-0.301	(0.18)
1995	0.441 ^{a}	(0.17)	0.347 ^{a}	(0.15)	0.064	(0.15)	-0.058	(0.10)	-0.049	(0.06)	0.093	(0.07)	-0.287	(0.18)
1996	0.457 ^{a}	(0.18)	0.683 ^{c}	(0.16)	0.118	(0.16)	-0.023	(0.11)	0.078	(0.06)	0.263 ^{c}	(0.08)	-0.170	(0.19)
1997	0.865 ^{c}	(0.18)	0.844 ^{c}	(0.16)	0.194	(0.16)	0.164	(0.11)	0.116	(0.06)	0.354 ^{c}	(0.08)	0.168	(0.21)
1998	0.915 ^{c}	(0.19)	0.886 ^{c}	(0.16)	0.153	(0.17)	0.034	(0.11)	0.194 ^{b}	(0.06)	0.358 ^{c}	(0.08)	0.006	(0.24)
1999	0.932 ^{c}	(0.18)	1.132 ^{c}	(0.16)	0.302	(0.17)	0.187	(0.11)	0.216 ^{c}	(0.06)	0.523 ^{c}	(0.08)	-0.111	(0.24)
2000	0.900 ^{c}	(0.19)	1.237 ^{c}	(0.17)	0.300	(0.18)	0.248 ^{a}	(0.12)	0.226 ^{c}	(0.06)	0.472 ^{c}	(0.08)	-0.049	(0.23)
2001	1.028 ^{c}	(0.19)	1.194 ^{c}	(0.17)	0.280	(0.18)	0.234 ^{a}	(0.12)	0.250 ^{c}	(0.07)	0.378 ^{c}	(0.08)	-0.084	(0.24)
2002	0.967 ^{c}	(0.19)	1.145 ^{c}	(0.17)	0.237	(0.18)	0.180	(0.12)	0.204 ^{b}	(0.07)	0.399 ^{c}	(0.09)	-0.124	(0.24)
2003	1.026 ^{c}	(0.19)	1.148 ^{c}	(0.17)	0.529 ^{b}	(0.19)	0.321 ^{b}	(0.12)	0.211 ^{b}	(0.07)	0.527 ^{c}	(0.09)	-0.140	(0.24)
2004	1.067 ^{c}	(0.20)	1.146 ^{c}	(0.17)	0.385 ^{a}	(0.19)	0.404 ^{c}	(0.12)	0.220 ^{b}	(0.07)	0.400 ^{c}	(0.09)	-0.098	(0.26)
2005	0.899 ^{c}	(0.20)	0.947 ^{c}	(0.18)	0.335	(0.19)	0.331 ^{b}	(0.12)	0.200 ^{b}	(0.07)	0.261 ^{b}	(0.09)	0.158	(0.26)
2006	0.971 ^{c}	(0.21)	0.724 ^{c}	(0.18)	0.374	(0.19)	0.285 ^{a}	(0.13)	0.156 ^{a}	(0.07)	0.227 ^{a}	(0.09)	-0.119	(0.27)
2007	0.764 ^{c}	(0.21)	0.740 ^{c}	(0.19)	0.328	(0.19)	0.299 ^{a}	(0.13)	0.171 ^{a}	(0.07)	0.169	(0.09)	-0.243	(0.27)
2008	0.662 ^{b}	(0.21)	0.605 ^{b}	(0.19)	0.466 ^{a}	(0.20)	0.307 ^{a}	(0.13)	0.132	(0.07)	0.078	(0.10)	-0.426	(0.28)
2009	0.654 ^{b}	(0.22)	0.517 ^{b}	(0.19)	0.357	(0.20)	0.244	(0.13)	0.140	(0.08)	0.075	(0.10)	-0.362	(0.28)
2010	0.563 ^{b}	(0.21)	0.511 ^{b}	(0.20)	0.427 ^{a}	(0.21)	0.364 ^{b}	(0.13)	0.195 ^{a}	(0.08)	0.131	(0.10)	-0.154	(0.29)
2011	0.791 ^{c}	(0.21)	0.489 ^{a}	(0.20)	0.539 ^{b}	(0.21)	0.401 ^{b}	(0.14)	0.125	(0.08)	0.089	(0.10)	-0.447	(0.29)
2012	0.693 ^{b}	(0.22)	0.311	(0.19)	0.445 ^{a}	(0.21)	0.300 ^{a}	(0.14)	0.145	(0.08)	0.081	(0.10)	-0.334	(0.29)
2013	0.621 ^{b}	(0.21)	0.348	(0.19)	0.587 ^{b}	(0.20)	0.272 ^{a}	(0.13)	0.128	(0.08)	-0.027	(0.10)	-0.424	(0.31)
2014	0.594 ^{b}	(0.22)	0.438 ^{a}	(0.19)	0.589 ^{b}	(0.20)	0.343 ^{a}	(0.14)	0.109	(0.08)	0.058	(0.10)	-0.609 ^{a}	(0.31)
2015	0.732 ^{c}	(0.21)	0.423 ^{a}	(0.20)	0.515 ^{a}	(0.21)	0.389 ^{b}	(0.13)	0.177 ^{a}	(0.08)	0.083	(0.10)	-0.262	(0.30)
2016	0.655 ^{b}	(0.22)	0.545 ^{b}	(0.20)	0.576 ^{b}	(0.20)	0.366 ^{b}	(0.14)	0.175 ^{a}	(0.08)	0.070	(0.10)	-0.344	(0.28)
Observations	2,816,958								2,678,207					
Adjusted R^2	0.763								0.766					
Country × Product FE	Yes								Yes					
Country × Year FE	Yes								Yes					
Product × Year FE	Yes								Yes					

Notes: DDD estimates correspond to equation (3). Two-way clustered standard errors are used, clustering on both country-year and product-year. ^{a} $p<0.05$, ^{b} $p<0.01$, ^{c} $p<0.001$.

Table B.3: DDD regression: time varying, heterogeneous cumulative treatment effects

B.3 Supplemental figures for Canada

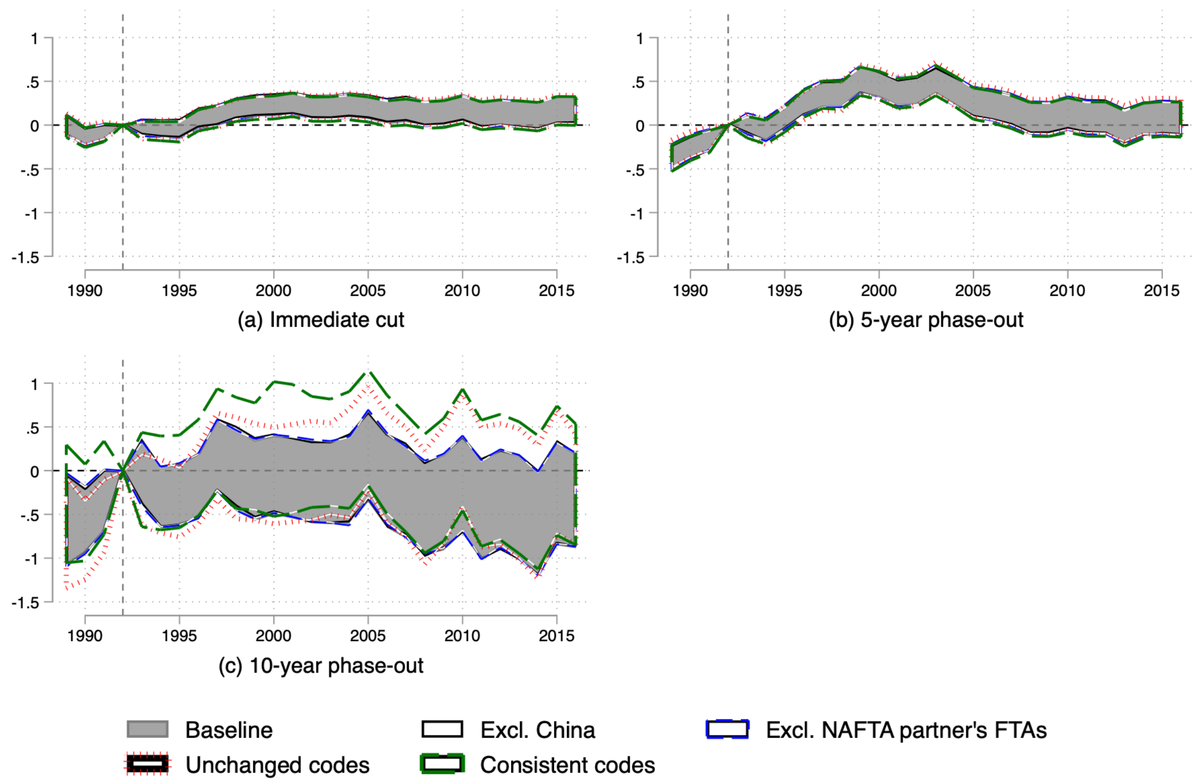


Figure B.1: Time-varying heterogeneous DDD estimates for Canada: Robustness.

Notes: DDD point estimates correspond to equation (3). Plots represent 95% confidence intervals. Two-way clustered standard errors are used, clustering on both country-year and product-year. Excl. China indicates China excluded from ROW definition; Excl. NAFTA partner's FTAs indicates that Canada's FTA partners excluded from ROW for entire sample; Unchanged codes indicates products where product code changes over time excluded from sample; Consistent codes indicates we use extended version, through 2016, of Pierce and Schott (2012) concordance.

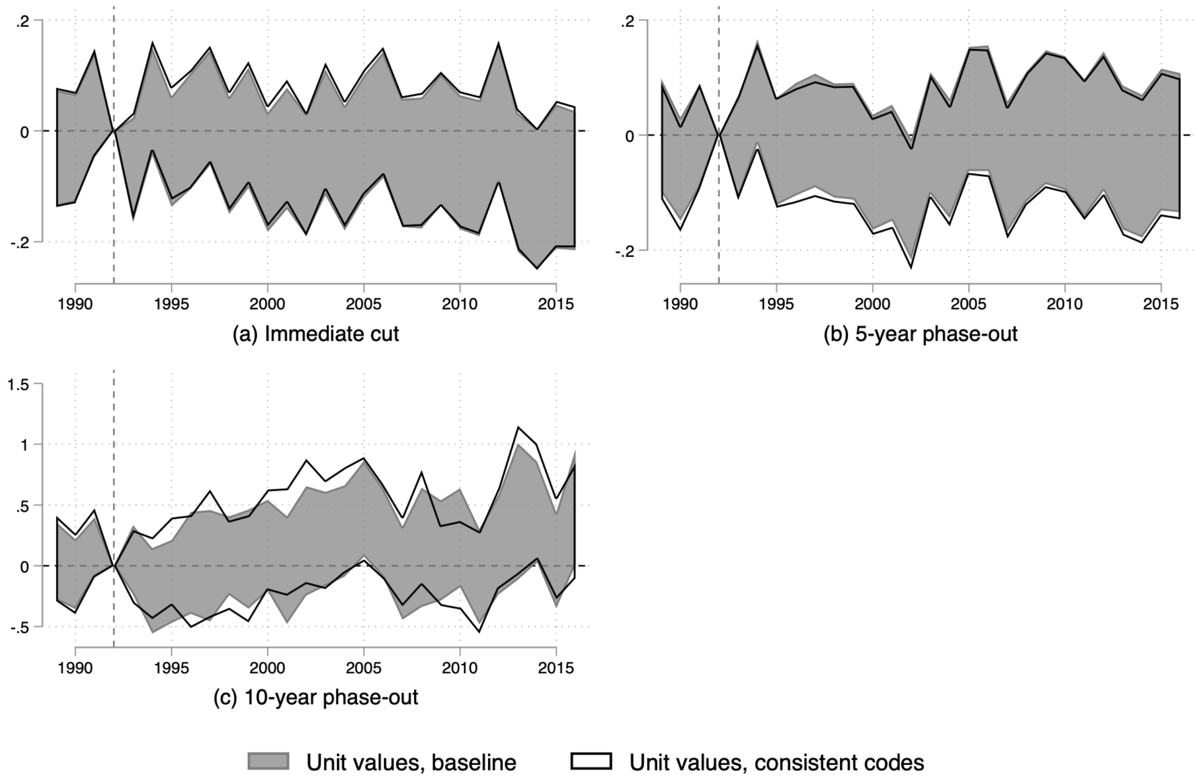


Figure B.2: Time-varying heterogeneous DDD estimates for Canada: Unit values.

Notes: DDD point estimates correspond to equation (3) with log unit values as dependent variable. Plots represent 95% confidence intervals. Two-way clustered standard errors are used, clustering on both country-year and product-year. Consistent codes indicates we use extended version, through 2016, of Pierce and Schott (2012) concordance.