Phase Out Tariffs, Phase In Trade?*

Tibor Besedes†
Georgia Institute of Technology

Tristan Kohl‡
University of Groningen

James Lake§
Southern Methodist University

May 30, 2020

Abstract

An important stylized fact in the empirical Free Trade Agreement (FTA) literature is that member trade flows gradually increase over time following an FTA. Baier & Bergstrand (2007) suggest two explanations: tariff phase-out and delayed pass-through of tariffs into import prices. We examine these hypotheses using 1989-2016 U.S. import growth and product-level data on the tariff phase-out negotiated under NAFTA and the earlier Canada-U.S. FTA. We find evidence supporting a weak form of the tariff phase-out hypothesis. But, we find little evidence supporting a strong form of this hypothesis because the bulk of delayed import growth stems from products granted tariff-free access upon NAFTA rather than products with phased-out tariffs. Additionally, we do not find evidence supporting the delayed tariff pass-through hypothesis. Instead, we find evidence for an important role played by NAFTA tariff cuts reducing the impact of frictions that, in turn, allow a spatial expansion of imports across the U.S.

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

JEL classification: F1

*We 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, and participants at the ASSA, ETSG, LETC, MITC, SETC, SEA, Aarhus University, Bank of Mexico, Clemson University, Radboud University Nijmegen, Southern Methodist University, University of Bayreuth, University of North Dakota as well as two anonymous referees and the editor for helpful comments and discussion. Kohl gratefully acknowledges Georgia Institute of Technology and Southern Methodist University for hospitality during visits that have contributed to this research. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

†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 two-and-a-half years of politically controversial negotiations, the U.S., Canada, and Mexico recently concluded 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 incorporated 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 & Parro 2014).

An interesting phenomenon in the FTA literature is that FTAs have delayed effects on trade flows. The classic stylized fact in this regard goes back to Baier & Bergstrand (2007, p. 7) who use lagged FTA dummy variables in a gravity model of international trade and find that "... an FTA approximately doubles two members' bilateral trade after 10 years." A simple glance at the evolution of CUSFTA/NAFTA's trade flows illustrates this stylized fact. 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.

Baier & Bergstrand (2007, p. 89-90) suggest two hypotheses to explain these prolonged

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

differential growth rates of real trade flows. Their first hypothesis is a "tariff phase-out" hypothesis which, citing the original European Economic Community (EEC) agreement and NAFTA as examples, revolves around the observation that "... virtually every [FTA] is 'phased-in,' typically over 10 years." Thus, one could naturally expect FTA trade flows to increase over time as the FTA gradually phases out bilateral tariffs. To help conceptualize this hypothesis, we distinguish between what we call "weak" and "strong" forms of the tariff phase-out hypothesis. For the weak form, we say that the gradual phase-out of tariffs over time, e.g. over 10 years, should steadily increase trade flows over time. For the strong form, we say that longer tariff phase-out periods should lead to a longer period of gradual growth in trade flows. The second hypothesis of Baier & Bergstrand (2007) is a "delayed tariff pass-through" hypothesis whereby tariff changes filter through to prices gradually over time. In this case, one could again naturally expect FTA trade flows to increase over time as tariff cuts gradually filter through to import prices.

By now, inclusion of the Baier & Bergstrand (2007) lagged FTA terms has become standard in applied work (see, for example, Baier et al. 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.

Our paper, to the best of our knowledge, is the first to explore the relevance of either Baier & Bergstrand (2007) hypothesis — "tariff phase-out" or "delayed tariff 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 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 approach or, in other words, a 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.[2]

The DDD approach applies naturally in our tariff phase-out context. Intuitively, our main empirical strategy looks at import growth between the pre- and post-NAFTA period (i.e. the first difference) from NAFTA partners vis-à-vis ROW (i.e. the second difference). But, to control for broader non-tariff related NAFTA effects, the DDD approach looks at this measure of import growth for products where tariffs are phased out *relative* to products that were already duty free before, and hence continue duty free after, NAFTA (i.e. the third difference).

In terms of the tariff phase-out hypothesis, our main finding is twofold. First, imports of products receiving tariff cuts show delayed import growth and grow more, both in the short- and long-run, than the continue-duty-free products where tariffs were already zero before NAFTA. This supports the weak form of the tariff phase-out hypothesis that tariff phase-out should lead to delayed import growth. 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, as our second main result, we find that the bulk of this delayed impact comes from products that *immediately* had their tariff cut to zero and, in the case of Mexico, had their tariff-free access via the Generalized System of Preferences (GSP) program converted into *permanent* tariff-free access. In other words, products where the tariff was phased out over 5-10 years

[2] See Frazer & Biesebroeck (2010), Coelli (2018), and Friederich & Zator (2018).

do *not* show a more pronounced pattern of delayed import growth relative to products where tariff-free access was granted upon implementation of NAFTA. This is inconsistent with the strong form of the tariff phase-out hypothesis.

In terms of the delayed tariff pass-through hypothesis, we do not find any evidence that tariff cuts gradually filter through to import prices over time and, in turn, lead to gradual import growth. Specifically, when looking at tariff-exclusive prices paid to foreign exporters, as proxied by unit values, we find that the dynamic path of prices does not differ between products receiving NAFTA tariff cuts and products that were duty free before, and continue duty free after, NAFTA. This suggests the tariff cut is fully and immediately passed through to U.S. importers. Thus, for the delayed tariff pass-through hypothesis to be working in practice it would have to be that importers are only passing through the tariff cut to U.S. consumers gradually over time which, in turn, gradually increases imports over time. Of course, further exploration of this idea would require transaction-level U.S. sales data.

Our findings on import prices relate to recent empirical evidence on tariff pass-through during the recent wave of U.S. protection under the Trump administration. Using quite different methodologies, Amiti et al. (2019) and Fajgelbaum et al. (2019) find almost complete pass-through of these new U.S. tariffs to prices, proxied by unit values, paid by U.S. importers. This contrasts with standard terms-of-trade theory whereby prices paid by U.S. importers should fall for goods where the U.S. has market power. A possible explanation for this puzzle is that import prices will eventually adjust to tariffs in the long-run. However, Amiti et al. (2020) find their earlier results in Amiti et al. (2019) hold with an extra year of data. And, our results suggest that import prices may not respond to tariffs even in the long-run. In turn, our results push back against resorting to the long run to reconcile the puzzle created by the results of Amiti et al. (2019) and Fajgelbaum et al. (2019) that contrast with standard terms-of-trade theory.

As an alternative explanation of the delayed import growth in the wake of NAFTA, we present indirect evidence that NAFTA tariff cuts reduce the impact of frictions related to the spatial expansion of imports across the U.S. Our conclusion relies on U.S. import data that report the customs district where a product enters consumption channels. Regardless of the specific phase-out category, we see a given imported product from a NAFTA partner gradually spreading out geographically across the U.S. over time when NAFTA cuts its tariff. A key takeaway from our analysis is that a product's "spatial" margin is crucial for understanding post-NAFTA delayed import growth driven by NAFTA tariff cuts.[3] We show that our results for the spatial margin are especially pronounced for homogeneous goods (per the Rauch classification) and final goods (per the BEC classification).

The broader trade literature has also addressed the interplay between the spatial margin of trade and the timing of trade flows. One strand of the literature focuses on firms learning about their demand (e.g. Albornoz et al., 2012; Fernandes & Tang, 2014; Timoshenko, 2015; Arkolakis et al., 2018; Berman et al., 2019). Very much like our "spatial" margin results, Albornoz et al. (2012) show empirically how this generates a firm-level "sequential exporting" phenomenon where successful firms expand their reach across more destination markets over time. Alternative strands of the literature focus on explaining the dynamics of demand-driven firm expansion through the necessity of firm-level marketing expenditures in new markets (Arkolakis, 2010, 2016; Fitzgerald & Priolo, 2018; Fitzgerald et al., 2019) or dynamic pricing strategies due to informational and reputational frictions (Gourio & Rudanko, 2014; Foster et al., 2016). Our contribution to this broad literature is to say that spatial margin effects triggered by FTA tariff cuts, rather than the original Baier & Bergstrand (2007) hypotheses of tariff phase-outs or delayed tariff pass-through to import prices, appear to play an important role in driving the delayed increase in member trade flows observed after FTAs.

[3] Viewing different geographic markets as different destination markets, this spatial margin could be viewed as an extensive margin. For post-NAFTA trade growth, the importance of the extensive margin dates back to at least Hillberry & McDaniel (2002). More generally, Baier et al. (2014, p. 339) find "intensive-margin effects occurring sooner than extensive-margin effects." And, 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.

2 Data

2.1 U.S. import data

We use annual product-level U.S. import data from the United States International Trade Commission (USITC).[4] Our data contain U.S. imports at the HS 10-digit level from every foreign country over the period 1989-2016 and total $7,823,777$ observations. We aggregate these data to the exporter-year 8-digit 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]

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

[5]Cleaning the data entails dropping three sets of observations at the 10-digit level: (i) the 1.17% of observations with import program "Unknown country", (ii) respectively, the 3.65% and then the 0.87% 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

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 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 U.S. imports from the NAFTA partners 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 almost 20% of U.S. imports from

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).

[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.

NAFTA partners but the former for only 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 January 1, 1994.[13] Since Mexican GSP-eligible products entered 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, around 90% 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

[11] The B6 products account for around 80% of the "5-year" products and about 40% 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.

[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 et al. (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).

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 that the total tariff cut implied by eventual tariff elimination delivers is, on average, highest for 5-year products.[16] Nevertheless, panel (b) of Figure 3 shows, as one would expect, that the implied annual tariff cut is, on average, highest for immediate-cut products and successively smaller for 5-year and 10-year products.[17]

2.3 Matching tariff schedules to trade data

Matching issues arise when merging the NAFTA staging categories and 8-digit USITC import data. On the 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.[18] 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. One might initially think that these represent politically sensitive sectors excluded from U.S. tariff elimination. But, this is not the case. Rather, two of these products are Chapter 99 "Temporary Protection " products and the other 13 are all new products introduced by the USITC in 1993.[19] Thus, despite being a common perception of FTAs, the U.S. did not exclude any sectors from eventual tariff elimination under NAFTA.

Failure to match the above 15 products from the 1993 U.S. import data to the U.S. NAFTA staging schedule leaves a match rate of 99.83% for imported products and 98.85% for import value, while these match rates are only slightly lower in the 1989-1992 pre-NAFTA

[16]See Section A.2 in the Appendix for details on constructing pre-NAFTA tariffs.

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

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

[19]These 15 products are 0814.00.80, 2921.42.21, 2921.42.22 2921.42.26, 2921.42.28, 2922.50.11, 4418.20.40, 4418.20.80, 7326.90.35, 8521.10.30, 8521.10.60, 8521.10.90, 9021.19.85, 9999.00.15, and 9999.95.00.

years. However, as Table 2 illustrates, these 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).

The time-varying set of product codes in the U.S. import data creates problems when linking a product code in the import data to a product code in the NAFTA tariff schedule. Thus, we extend the concordance of Pierce & Schott (2012) that creates time-consistent HS8 codes to cover our sample period and use these "consistent codes" throughout our analysis. This forces us to drop about 1% of observations because a new concorded product code can inherit various HS8 codes that have different NAFTA staging categories.[20]

3 Empirical strategy

Our empirical approach can be theoretically motivated by a multi-country, multi-product Ricardian model of trade such as Costinot et al. (2011). With this motivation, Section 3.2 uses empirical intuition to motivate the triple-differences specifications that form the basis of our empirical analysis. Section 3.3 discusses identification.

3.1 Theoretical motivation

In Costinot et al. (2011), each product p has infinite varieties indexed by ω, with exporter-product-variety productivity for an exporting country c given by $z_{cp}(\omega)$ and drawn from a Frechet distribution. Because buyers shop around the world for the cheapest source country of a given variety, only a subset of countries produce this variety. In turn, an exporting coun-

[20] In this case, we exclude the entire country-product family, using the Pierce & Schott (2012) terminology, of affected observations.

try's *observed* product-level average productivity, $\tilde{z}_{cp}$, differs from its average productivity draw across all varieties of a product, z_{cp}. Costinot et al. (2011) show that this theoretical framework delivers the following estimating equation:

$$\ln \tilde{x}_{cjp} = \gamma_{cj} + \gamma_{jp} + \theta \ln \tilde{z}_{cp} + \varepsilon_{cjp}. \tag{1}$$

where the γ terms are fixed effects. Moreover, $\ln \tilde{x}_{cjp} \equiv \ln x_{cjp} - \ln \pi_{ccp}$ is "corrected exports" of product p from exporting country c to importing country j that takes observed exports x_{cjp} and corrects for a measure of "openness" given by the share of country c's production of product p consumed domestically. The residual term ε_{cjp} represents importer-exporter-product specific trade costs.

Unlike Costinot et al. (2011), we are not interested in estimating θ. Thus, taking π_{ccp} to the right-hand side, we can rewrite (1) as

$$\ln x_{cjp} = \gamma_{cj} + \gamma_{jp} + \gamma_{cp} + \varepsilon_{cjp}. \tag{2}$$

where $\gamma_{cp} = \theta \ln \tilde{z}_{cp} + \ln \pi_{ccp}$. Adding time subscripts and restricting attention to U.S. imports, we re-write x_{cjp} as U.S. imports M_{cpt} of product p from exporting country c in year t:

$$\ln M_{cpt} = \gamma_{ct} + \gamma_{pt} + \gamma_{cp} + \varepsilon_{cpt} \tag{3}$$

where ε_{cpt} represents exporter-product-year specific trade costs when exporting to the U.S. and, given the fixed effect γ_{cp}, the time-varying components of productivity and openness at the exporter-product-level. This closely resembles a sector-level gravity equation where one would naturally interpret γ_{pt} and γ_{ct} as inward and outward multilateral resistances and γ_{cp} as exporter-product-specific trade costs and productivity. To equation (3), our triple-differences specification will add a vector of treatment effect variables, defined in the next section as $NAFTA_c \times \mathbf{Phase}_p \times \mathbf{Year}_t$, that allow the treatment effect of NAFTA phase-out to vary heterogeneously across both a vector of staging categories ($\mathbf{Phase}_p$) and a vector of years ($\mathbf{Year}_t$).

3.2 Triple-differences 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.[21] However, by ignoring continue-duty-free products, the NAFTA versus ROW approach ignores the possibility that a product's NAFTA imports grow relative to its ROW imports *regardless* of whether the product's tariff is being phased out. Such a pattern of import growth could be driven by positive supply shocks in the NAFTA partners or broad effects of NAFTA that go beyond tariff reduction. Conversely, by ignoring ROW imports, the phase-out versus continue-duty-free approach ignores the possibility that a phase-out product's imports grow relative to a continue-duty-free product *regardless* of the exporting country. Such a pattern of import growth could be driven by global product-specific supply or demand shocks.

[21]In any type of DD estimation (or triple-difference estimation) like this, ROW is not a pure control group in the sense that NAFTA likely reallocates demand from ROW to NAFTA partners. In other words, this methodology reveals a combined trade diversion and trade creation effect.

To avoid these problems, we use triple-differences (DDD) specifications. The simplest DDD specification is

$$\text{(4)} \quad \ln M_{cpt} = \alpha + \beta_1 NAFTA_c + \beta_2 Phase_p + \beta_3 Post_t$$
$$+\gamma_1 Phase_p \times Post_t + \gamma_2 NAFTA_c \times Post_t + \gamma_3 NAFTA_c \times Phase_p$$
$$+\delta NAFTA_c \times Phase_p \times Post_t + \varepsilon_{cpt}.$$

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.[22] 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 exclude countries that are U.S. FTA partners at any point in time.[23]

Although improving on the intuitive NAFTA versus ROW and phase-out versus continue-duty-free approaches, equation (4) is a "no controls" DDD specification. As such, it still omits many potentially relevant variables. Indeed, the theoretical motivation in Section 3.1 suggests the use of product-year, exporter-year, and country-product fixed effects that can be thought of as inward and outward multilateral resistances as well as time-invariant exporter-product-specific trade costs and productivity. Of course, these fixed effects flexibly control for these and other variables that vary at the product-year, exporter-year, or country-product levels. We now have the following fixed-effects, or "with controls", DDD specification:

$$\text{(5)} \quad \ln M_{cpt} = \alpha + \delta NAFTA_c \times Phase_p \times Post_t + \gamma_{pt} + \gamma_{ct} + \gamma_{cp} + \varepsilon_{cpt}.$$

[22]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.

[23]These FTA partners are the only exporters excluded from our baseline analysis. They are Australia, Bahrain, Chile, Colombia, Costa Rica, Dominican Republic, El Salvador, Guatemala, Hondruas, Israel, Jordan, Korea, Morocco, Nicaragua, Oman, Panama, Peru, and Singapore.

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 (5) 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:

$$\text{(6)} \qquad \ln M_{cpt} = \alpha + \boldsymbol{\delta} NAFTA_c \times \mathbf{Phase}_p \times \mathbf{Year}_t + \gamma_{pt} + \gamma_{ct} + \gamma_{cp} + \varepsilon_{cpt}.$$

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.

3.3 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.

The theoretical motivation from Section 3.1 and equation (3) in particular provide some formal guidance regarding what shocks could violate this identification assumption. Specifically, the error term in equation (3) consists of time-varying exporter-product trade costs and the time-varying components of productivity and openness at the exporter-product level. To see the implications for the identification assumption, consider exporter-product-year-specific trade costs such as an export subsidy implemented by ROW exporters after NAFTA. All else equal, these export subsidies would not violate the identification assumption if they applied to a subset of products where ROW exports to the U.S. are equally distributed between phase-out and continue-duty-free products. In this case, there would be no impact on U.S. import growth from ROW of phase-out products *relative to* continue-duty-free products. However, the identification assumption would be violated if these export subsidies were systematically targeted at either phase-out or continue-duty-free products so that they affected U.S. import growth from ROW of phase-out products *relative to* continue-duty-free products. Naturally, this same distinction applies if one replaces the export subsidy shock with time-varying productivity or openness shocks at the exporter- product level.

To investigate the reasonableness of the identification 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 was 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 to 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.3 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 mitigates endogeneity concerns surrounding the possibility that U.S.-Mexico-specific factors, e.g. Mexican supply shocks, drive the tariff phase-out faced by Mexican exports to the U.S. 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 (4) is just a comparison of mean import growth between phase-out and continue-duty-free products and between NAFTA partners and ROW. Table 3 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 0.434 log points in the post-NAFTA period, mean log imports of phase-out products from ROW *decreased* by 0.207 log points in the post-NAFTA period. Thus, import growth of phase-out products from Mexico vis-à-vis ROW was 0.641 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 0.647 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 0.382 log points for Mexico vis-à-vis ROW and 0.455 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 0.259 log points for Mexico and 0.192 log points for Canada. On the 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 3 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 −0.051 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 −0.095 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 -0.310 (-0.287) 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 0.259 log points for Mexico and 0.192 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 3.3 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 3.2. 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 4 shows the impacts of moving from the "no controls" specification in equation (4) to the "with controls" specification in equation (5). Column (1) shows the "no controls" DDD estimates from equation (4) and, by construction, the DDD point estimate matches that from the means-based DDD in Table 3. 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 (5).

Column (2) shows that the DDD point estimate increases by about 45% 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 U.S., 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. These are based on a modified version of equation (5) that replaces the $Post_t$ dummy with the vector of year dummies $\mathbf{Year}_t$. As we discussed in Section 3.3, statistical insignificance of pre-NAFTA DDD point estimates is important in terms of assuaging potential endogeneity concerns as this says that relative import growth from Mexico (i.e. phase-out products relative to continue-duty free products) is not diverging from ROW relative import growth before NAFTA. Indeed, Figure 4 shows that the Mexican DDD point estimate is statistically insignificant before NAFTA.

Moving on to the time-varying post-NAFTA DDD point estimates, recall that column (4) from Table 4 said the time-*invariant* DDD point estimate was 0.388 log points. In contrast, Panel (a) of Figure 4 plots the time-varying DDD estimates for Mexico with 95% confidence intervals. Remember that, given our DDD methodology, these point estimates reflect, conditional on the fixed effects, import growth from Mexico in phase-out products relative to continue-duty-free products vis-à-vis this relative import growth from ROW. Unsurprisingly given the tariff phase-out argument from Baier & Bergstrand (2007), these results show considerable time heterogeneity.

Indeed, the delayed emergence of statistical significance after 1992 and the steady but gradual growth of the point estimates is strong evidence for what we call the "weak " form of the tariff phase-out hypothesis: gradual phase out of tariffs over time should gradually increase trade flows over time. More specifically, the DDD point estimate only becomes statistically significant in 1997 and keeps growing from 0.129 log points in 1997 to a peak, right around the 10-year mark of NAFTA, of 0.570 log points in 2005. While the effects taper off somewhat post-2005, the point estimates largely hover in the 0.4-0.5 log points in the post-2000 period and are always statistically significant.[24]

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. Such an investigation not only provides further insight into the weak

[24]Table A.5 in the Appendix presents the time-varying regression results for the specifications that parallel those in Table 4.

form of the tariff phase-out hypothesis but also allows investigation of the "strong " form of the tariff phase-out hypothesis whereby notably longer tariff phase-out periods should lead to a notably longer period of gradual growth in trade flows. More specifically, we expect to see a rapid increase in import growth of immediate-cut products that stabilizes quickly. Assuming that NAFTA removes uncertainty over future GSP eligibility, we expect similar dynamics for GSP-eligible products.[25] In contrast, we expect a steady and gradual increase in import growth of 5-year and 10-year phase-out products that stabilizes after 5-10 years. The dashed lines in Figure 5 represent these qualitative hypotheses.

Figure 5 presents time-varying DDD estimates from equation (6) when allowing the heterogeneity in the DDD treatment effect across staging categories: immediate cut (A), 5-year phase-out (B and B6), 10-year phase-out (C, C10, and C+), and GSP.[26] Importantly, note the statistically insignificant DDD point estimates in the pre-NAFTA period across panels (a)-(d). These DDD point estimates say that relative import growth (i.e. phase-out products relative to continue-duty-free products) from Mexico was not diverging from ROW relative import growth before NAFTA.[27]

For each staging category, Figure 5 shows statistically significant, although somewhat delayed, post-NAFTA import growth and provides further evidence for the weak form of the tariff phase-out hypothesis. Panel (a) shows this begins in 1993 for immediate-cut products, reaching around 0.9 log points by the early 2000s and stabilizing shortly thereafter. Panel (b) shows 5-year phase-out products experience up to 40% larger import growth that eventually peaks around 1.226 log points in 2000 and stabilizes shortly thereafter. This 40% larger import growth is consistent with their 25-69% larger tariff cuts from Table A.2. Although

[25] As part of the broader and growing literature on trade policy uncertainty (e.g. Handley 2014, Pierce & Schott 2016, and Handley & Limão 2017), Hakobyan (2020) documents the inherent legislative uncertainty surrounding GSP renewal and the adverse impact of uncertainty on import growth from beneficiary countries.

[26] Table A.6 in the Appendix presents the regression results. These results, as well as the analogous results for Canada are virtually identical, qualitatively and quantitatively, when weighting the regressions by the first observed country-product log imports or mean log country-product imports for 1989-2016.

[27] 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.

notably smaller, panels (c) and (d) show import growth of 10-year phase-out and GSP products stabilizing around the mid 2000s with respective peak import growth around 0.4-0.5 and 0.3-0.4 log points. Collectively, these results support the weak form of the tariff phase-out hypothesis that phased-in NAFTA tariff cuts stimulated gradual import growth from NAFTA partners.

Figure 5 also provides a preliminary assessment about the strong form of the tariff phase-out hypothesis that says longer tariff phase-out periods should lead to a notably longer period of gradual growth in trade flows. First, inconsistent with this hypothesis, panels (a) and (b) suggest that import growth of the immediate-cut products and the 5-year phase-out products seem to stabilize at similar points in time in the late 1990s or early 2000s. Second, again inconsistent with this hypothesis, panels (c) and (d) suggest that import growth of the 10-year phase out products and GSP products both stabilize in the mid 2000s.

Indeed, Table 5 shows these features of panels (a)-(b) and (c)-(d) in Figure 5 are robust statistical properties. Specifically, Table 5 illustrates the extent that we can detect *changes* in cumulative import growth over time and hence whether import growth takes notably longer to stabilize for notably longer phase-out periods. Naturally, it is difficult to detect statistically significant changes in cumulative import growth at an annual frequency. Thus, Table 5 presents estimates of import growth over 4-year rolling windows: the year t point estimate is the difference between the DDD point estimates for cumulative import growth in year t and year $t-3$ from Figure 5.[28] 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 5 shows that import growth of immediate-cut products in these rolling 4-year windows is around 0.3-0.4 log points in all years between 1994 and 1998 with the caveat of not seeing import growth in 1996-1997 relative to 1992-1993. Statistically speaking, the post-NAFTA import growth of immediate-cut products only stabilizes from the year 2000.

[28]That is, the year t point estimate in Table 5 represents the test of equality between the year t and year $t-3$ point estimates in Table A.6 of the Appendix that come from estimating equation (6).

However, 5-year phase-out products grow about 0.3-0.5 log points in these rolling 4-year windows beginning in 1996 and continuing every year until 2001. Statistically speaking, the post-NAFTA import growth for 5-year phase-out products stabilizes from 2002. Ultimately, import growth of immediate-cut products stabilizes 8 years after NAFTA was signed which is only 2 years ahead of when 5-year phase-out products stabilize. From this perspective, the import growth dynamics are remarkably similar for immediate-cut products and 5-year phase-out products which is not what we expect based on the strong form of the tariff phase-out hypothesis.

Panel (b) of Figure 2 shows the importance of 10-year and GSP-eligible products. The former account for about 90% of imports where tariffs on Mexico were actually phased out over time and the latter account for around 30% of all imports from Mexico. Thus, if the strong form of the tariff phase-out hypothesis helps explain delayed import growth from Mexico then it should help explain these products. However, Table 5 shows that, by and large, we cannot detect any increase in import growth for 10-year phase-out products over time in rolling 4-year windows.[29] Further, import growth of GSP-eligible products stabilizes around 2002 after statistically significant increases in cumulative import growth in 1999. That is, GSP-eligible products look more like what we expect from 5-year or 10-year phase-out products than immediate-cut products. Overall, there is little evidence to support the strong form of the tariff phase-out hypothesis.

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, panel (b) of Figure 4 shows such effects. And, Figure 6 illustrates this quite strongly for 5-year phase-out products

[29] This inability to detect robust import growth for the 10-year phase-out products also holds when using 3-year or 5-year rolling windows.

which, per panel (b) of Figure 2, constitute over 90% of Canadian imports that were actually phased out over time.[30] 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 to the Appendix.

Nevertheless, the key points emerging from our Mexican analysis also emerge for Canadian imports: support for the weak form of the tariff phase-out hypothesis but little evidence for the strong form of the tariff phase-out hypothesis. Figure 6 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.[31] Indeed, Table 5 shows we can detect 4-year rolling window import growth of 0.1-0.2 log points for immediate-cut products each year during 1996-2000. 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. This supports the weak form of the tariff phase-out hypothesis.

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 5 and Panel (b) of Figure 6 show import growth of Canada's 5- year phase-out products stabilize in the late 1990s at around 0.4 log points.[32] Statistically speaking, immediate-cut and 5-year phase-out

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

[31] Table A.6 in the Appendix contains regression results.

[32] 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.

products both stop growing in the 1999-2000 period. This is inconsistent with the strong form of the tariff phase-out hypothesis.

4.2.3 Robustness

We now describe various robustness checks to our results presented above.

Alternative samples Figure 7 for Mexico (and Figure A.1 in the Appendix for Canada) show that various restrictions on the sample of countries representing ROW and the sample of products do not not affect our results. These restrictions are motivated by three concerns.

First, one may be concerned about the well-documented surge of U.S. imports from China (e.g. Autor et al. (2013), Pierce & Schott (2016)). Specifically in our context, a concern may be that Chinese exports to the U.S. are surging in products that are systematically related to the classification of either phase-out versus continue-duty-free products or NAFTA staging categories. First, consider the weak form of the tariff phase-out hypothesis. If Chinese exports are systematically surging for continue-duty-free (phase-out) products rather than phase-out (continue-duty-free) products then our DDD approach would underestimate (overestimate) the impact of NAFTA tariff phase-out on import growth. Second, consider the strong form of the tariff phase-out hypothesis. If Chinese exports are systematically surging for 5-year and 10-year phase-out (immediate-cut and GSP) products rather than immediate-cut and GSP (5-year and 10-year phase-out) products then our DDD approach could underestimate (overestimate) the extent that delayed import growth is related to the length of tariff phase-out. Nevertheless, Figure 7 (and Figure A.1 in the Appendix) shows that excluding China from the ROW country sample does not alter our results.

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.[33] But, Figure 7 (and Figure A.1 in the Appendix) show that removing

[33]Canada's FTA partners are Chile, Colombia, Costa Rica, Honduras, Iceland, Israel, Jordan, Korea, Liechtenstein, Norway, Panama, Peru, and Switzerland. Mexico's FTA partners are Belize, Chile, Colombia,

Mexico and Canada's FTA partners from the ROW country sample does not affect our results.

Third, while our analysis so far has used time-consistent HS product codes using our extended version of the Pierce & Schott (2012) concordance, one may wonder about how these results compare to non-concorded product codes. Figure 7 (and Figure A.1 in the Appendix) show our baseline results are unaffected when we use the sample of product codes that never change over time.

This sub-sample of 7,079 product codes also provides a clean opportunity to investigate the extent that the U.S. actually implemented the tariff cuts embodied in the staging schedule. Focusing on imports from Mexico, the U.S. follows the NAFTA staging schedule remarkably closely. In particular, Table 6 shows that the annual tariff phase-out implemented by the U.S. on imports from Mexico matches that specified in the NAFTA staging schedule for 96.2% of products. For a further 0.5% of products, the length of the phase-out period matches the NAFTA staging schedule even though the time path of tariff cuts does not. A further 2.5% of products have their tariff phase-out accelerated by one of four rounds of NAFTA tariff acceleration that took place during the 1994-2003 period; for 5-year phase-out products, this merely meant the tariff was cut to zero one year before specified in the NAFTA staging schedule.[34] These three baskets of products account for 99.3% of all products and at least 94.8% of products within each staging category.

In terms of the remaining discrepancies between actual tariff phase-out and that specified in the NAFTA staging schedule, 0.5% of products, concentrated nearly entirely within 10-year phase-out products, see U.S. MFN tariff cuts leave the NAFTA-specified Mexican preferential tariff above the U.S. MFN applied tariff. In turn, the NAFTA preferential tariff applied to Mexican imports of these products falls to the U.S. MFN applied tariff. For six products, there are one or two years where the NAFTA preferential tariff applied differs from

Costa Rica, El Salvador, EU28, Guatemala, Honduras, Iceland, Israel, Japan, Liechtenstein, Nicaragua, Norway, Panama, Peru, Switzerland, and Uruguay.

[34] For further details, see bit.ly/3etvIRI (1st Round), bit.ly/2VJ6Nkq (2nd Round), bit.ly/3bjrsSJ (3rd Round), and bit.ly/2VJ6Nkq (4th Round).

the NAFTA staging schedule (including one product receiving temporary tariff protection per Chapter 99 of the USHTS). And, finally, nine products appear to completely ignore the NAFTA staging schedule; the NAFTA staging schedule specified eight of these products will receive permanent post-NAFTA tariff-free access via their pre-NAFTA GSP eligibility.

Figure 7 shows our baseline results remain unchanged for the sub-sample of 6,810 products in Table 6 that always follow the staging schedule.

Heterogeneity of tariff cuts within staging categories While we have allowed heterogeneity in our time-varying DDD treatment effects according to NAFTA staging categories, we have ignored treatment effect heterogeneity within a staging category. But, one may expect such heterogeneous treatment effects according to the product-specific magnitude of the U.S. MFN tariff cut associated with eventual tariff elimination.

To this end, we create quartiles of the 1989 U.S. MFN tariff distribution for each staging category. Letting Q_{np} denote the indicator variable for whether product p lies in the n-th quartile of this tariff distribution for its staging category, we augment estimating equation (6) by interacting the treatment effects $NAFTA_c \times Phase_p \times \mathbf{Year}_t$ for a given staging category $Phase_p \in \mathbf{Phase}_p$ with the vector of quartile dummy variables $\mathbf{Q}_p = (Q_{1p}, Q_{2p}, Q_{3p}, Q_{4p})$. Figure 8 illustrates the results for imports from Mexico and shows little heterogeneity across the quartiles of each staging category. As one may expect, to the extent that there is heterogeneity, the treatment effects are stronger for products in the top quartile of the 1989 U.S. MFN tariff distribution for immediate cut products and 5-year phase-out products. Additionally, Figure A.2 in the Appendix shows there is little noticeable heterogeneity across quartiles for imports of immediate-cut, 5-year, and 10-year phase-out products from Canada.

5 Results: Delayed tariff pass-through

To the extent that tariffs pass through to prices paid by U.S. importers, changes in the value of trade can come from changes in quantities or prices. Thus, we now modify equation (6) by using log real unit values as the dependent variable to proxy for import prices.[35] To be clear, unit values proxy for the (tariff exclusive) price of imports received by the foreign exporter rather than the tariff inclusive price of imports paid by the importer. This is because unit values are measured based on tariff exclusive import values. Thus, one would expect that tariff cuts should increase the price received by NAFTA partner countries exporting to the U.S. Investigating the dynamics of these prices addresses the second hypothesis of Baier & Bergstrand (2007) that delayed pass-through of tariff cuts to import prices can explain the delayed trade flow effects of FTAs.

Figure 9 (and Figure A.3 in the Appendix) present the results where we modify equation (6) to use log real unit values as the dependent variable. 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 lower prices 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.

Recently, Amiti et al. (2019) and Fajgelbaum et al. (2019) have found that the spate of U.S. tariffs in 2018 and 2019 (predominately on imports from China) were nearly entirely passed through to U.S. importers in the short run. This goes against conventional terms-of-trade theory which suggests that this should not happen for goods where the U.S. holds market power. One resolution of this puzzle would be that unit values adjust to tariffs in the longer run. However, our results suggest that unit values can be invariant to tariffs both in the short- and long-run and hence push back against resorting to the long-run to reconcile the puzzle.

[35] 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 margins of delayed import growth

At its most intuitive level, the idea that tariff phase-out or delayed pass-through of tariff cuts to prices could drive delayed import growth is implicitly a story about the intensive margin. That is, imports of products that were imported before NAFTA gradually grow over time as NAFTA phases out their tariffs. Thus, we now look to the extensive margin of trade to explore explanations beyond tariff phase-out and delayed pass-through of tariff that can help explain the delayed growth of trade after NAFTA. Of course, the importance of the extensive margin has pervaded a large body of theoretical (Arkolakis et al. 2012) and empirical (Bernard et al. 2009) research in the broader trade literature. In practice, the role of the extensive margin emerges from frictions created by fixed cost barriers that firms face when entering foreign markets or adding new products. Moreover, these barriers are likely to become less burdensome for Mexican firms as NAFTA reduces tariff and non-tariff barriers.

6.1 New products extensive margin

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 from Mexico in *every* year of our sample.[36]

Panel (a) of Figure 10 shows how the value of U.S. imports from Mexico decomposes into these three margins. The intensive margin increases from around 60% of imports to

[36]Note, our new products extensive margin is actually about products that are newly traded. This is different than the new goods margin of Kehoe & Ruhl (2013). After rank ordering products according 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.

around 80% during our sample period and essentially mirrors the reduction of the infrequent extensive margin from around 40% of imports to around 20%. The new products extensive margin never accounts for a significant amount, usually hovering around 1% of imports.

In terms of the number of products imported from Mexico, panel (b) of Figure 10 shows the intensive margin accounts for roughly 38% of products. While this share dips as NAFTA ramps up, it returns to its original level by the end of the sample. The infrequent extensive margin accounts for slightly more than 60% of products before NAFTA and steadily decreases to about 42%. The new products extensive margin quickly increases as NAFTA starts to about 17% of products in 1995 but then only increases slowly to about 20% by the end of our sample. While newly imported products from Mexico increase their presence in terms of how many are imported, they never increase to a meaningful share of imported value. Thus, the extensive margin, as traditionally defined, cannot account for either the weak or strong forms of the tariff phase-out hypothesis and the resulting delay in growth of imports.

We should note that two reasons imply the data described in Figure 10 could understate the role of the new products extensive margin. First, around 12.5% of products have a "Mixed" staging category. But, as discussed in Section 2.2, these products are split into "sub-HS8 codes" with different staging categories attached to each sub-code and, hence, we cannot include them in our analysis. Second, for every newly introduced HS8 code by the Census, the Pierce & Schott (2012) concordance maps this new product to at least one 1989 HS8 product code. Since we use this concordance, a product can only appear in our new product extensive margin if (i) it existed and was not imported from Mexico before NAFTA or (ii) it did not exist before NAFTA but its concorded 1989 HS8 code(s) was not not imported from Mexico before NAFTA.[37] Indeed, these two reasons are quantitatively important.

Modifying Figure 10 to not concord products using the Pierce & Schott (2012) concor-

[37] For example, one likely considers the development of a cell phone as a new product during our post-NAFTA sample period. But, the Pierce & Schott (2012) concordance maps any new code for cell phones back to a 1989 HS8 code for telephones. So, cell phones imported from Mexico would only appear in our new product extensive margin if telephones were not imported from Mexico before NAFTA.

dance but include products in the Mixed staging category reveals that the new products extensive margin increases to about 35% of imported value by 2016. This is substantially higher than the 1% share in panel (a) of Figure 10.[38] Despite the qualitative importance of these issues for quantifying the new products extensive margin, our results are based on the data underlying Figure 10 and, per Section 4.2.3, are robust to only using product codes that remain unchanged over our sample period. Thus, the new products extensive margin cannot account for our results regarding the weak and strong versions of the tariff phase-out hypothesis and the delayed tariff pass-through hypothesis.

6.2 Spatial margin

Rather than thinking of the extensive margin in terms of new products being imported from Mexico after NAFTA, one could also think of a "spatial" extensive margin where a given imported product from Mexico starts spreading out geographically across the U.S. after NAFTA. We hereafter refer to this as simply the "spatial margin." 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.[39] 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. five in Texas and three 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 $\ln D_{cpt}$ 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 im-

[38]The 35% share reduces to around 15% after concording the data using the Pierce & Schott (2012) concordance. This 15% share then reduces to the 1% share in Figure 10 when removing products with a Mixed staging category.

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

pacts the spread of imports from Mexico across the U.S. over time by replacing the imports variable $\ln M_{cpt}$ in equation (6) 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 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. For each staging category in the post-NAFTA period, the temporal pattern of spatial expansion for Mexican exports across the U.S. looks remarkably similar to the temporal patterns in our baseline analysis regarding the value of those exports. Each staging category shows a steady, but delayed, increase in the number of customs districts reached by Mexican exports.

What potential mechanisms could underlie this finding of a gradual spatial expansion of Mexican exports across the U.S. that is largely independent of the staging category? To begin, remember that the treatment effects illustrated in Figure 11, like all our treatment effects, are measured relative to continue-duty-free products. This is important for two reasons. First, Figure 11 says tariffs matter for spatial expansion: products receiving NAFTA tariff cuts experience stronger spatial expansion than continue-duty-free products. Second, many NAFTA provisions that reduced non-tariff barriers apply both to products receiving tariff cuts as well as continue-duty free products including reduced uncertainty over future tariffs, dispute settlement systems (both investor-state and state-state dispute settlement), moves towards harmonization and mutual recognition of product standards (for both agricultural and non-agricultural goods), temporary entry of business persons, national treatment and non-discrimination for government procurement contracts, and intellectual property rights protection. Thus, while these reductions in non-tariff barriers may stimulate imports from Mexico after NAFTA, they cannot explain our results regarding differential growth rates of phase-out versus continue-duty-free products.

Rather than think of NAFTA as directly reducing frictions that restrict export growth along the spatial margin, NAFTA tariff cuts can increase the profitability of Mexican exporters and thereby reduce the burden of incurring costs associated with spatial expansion across the U.S. In doing so, NAFTA tariff cuts could encourage Mexican exporters to export to more geographic parts of the U.S over time. A growing literature in international trade and macroeconomics not only emphasizes the importance of demand expansion dynamics for understanding how firms grow but also provides theoretical models and empirical evidence for underlying mechanisms.

As explained by Foster et al. (2016), new businesses and extensions of existing businesses that expand into new markets are smaller but grow faster than their more established competitors. Often, this is despite these smaller firms being as efficient as their more established competitors. In turn, the literature has emphasized initially low demand caused by, e.g., informational and/or reputational frictions that diminish over time. Foster et al. (2016) empirically show how firms use dynamic pricing strategies to build future demand at the expense of current profits. Arkolakis (2010) models firm-level marketing expenditures while presenting evidence that such expenditures total around 5% of 2001-2004 U.S. GDP. Adding idiosyncratic firm-level productivity growth, Arkolakis (2016) shows the model helps explain important firm dynamics in the U.S. for U.S. firms and Brazilian exporters. Building on Foster et al. (2016) and Arkolakis (2010), Gourio & Rudanko (2014) provide a microfoundation for the role of marketing expenditures through a search and matching model of firms and consumers where firms use dynamic pricing strategies. Moreover, Compustat data supports the model's key predictions. Using Irish data and U.S. scanner-level retail food data respectively, Fitzgerald et al. (2019) and Fitzgerald & Priolo (2018) argue that the marketing mechanism but not the dynamic pricing mechanism can explain post-entry firm behavior of quantities and markups.

A related but alternative mechanism in the literature that explains the dynamics of demand-driven firm expansion revolves around firms learning about demand. Berman et al. (2019) find supporting empirical evidence for a model where a firm forms beliefs about their demand over time in a product-destination market. In contrast, Fernandes & Tang

(2014) find empirical evidence of belief spillovers across exporters in a destination market while Timoshenko (2015) provides evidence of belief spillovers for a multi-product exporter across their range of products in a destination market. Finally, perhaps closest in spirit to our spatial margin results, Albornoz et al. (2012) present evidence of belief spillovers for an exporter across their destination markets whereby an exporter sequentially expands into new destination markets as they learn about their demand.

Our spatial margin results suggest the NAFTA setting is ripe for investigating the kind of mechanisms just discussed. While a substantive investigation of these mechanisms lies beyond the scope of this paper, we now document two dimensions of heterogeneity in our spatial margin results that could motivate future research.

First, using the Rauch classification, Figure 12 allows our treatment effects from Figure 11 to vary for differentiated and homogeneous products.[40] Perhaps surprisingly, homogeneous goods show notably stronger delayed spatial expansion than differentiated goods with the latter generally statistically insignificant. An interpretation of this result may be that consumers largely care about the price of homogeneous goods but they also care about other non-price product attributes of differentiated goods.[41] Thus, at the margin, the impact of NAFTA tariff cuts may be larger for homogeneous than differentiated goods and allow homogeneous goods to increase their geographic footprint relative to differentiated goods.

Second, Figure 13 allows our treatment effects from Figure 11 to vary for final, intermediate and primary goods. To the extent that heterogeneity exists, it appears that final goods show stronger delayed spatial expansion than intermediate or primary goods. A potential explanation is that, unlike consumption of final goods, the assembly of imported intermediate goods into final goods and the processing of imported primary goods may often take place in geographically concentrated areas. This could happen because of external economies of scale, proximity of shipping routes for re-export, or proximity to final goods producers with

[40] We include reference priced goods in our homogeneous goods category.

[41] Also highlighting the additional sensitivity of homeogenous goods to tariffs, Cavallo et al. (2019) find that export prices received by U.S. firms facing tariff retaliation in China fall more for homogeneous than differentiated products.

just-in-time inventory management systems.

7 Conclusion

Since the seminal work of Baier & 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 & 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 & Bergstrand (2007) for the delayed import growth following FTA formation.

On the one hand, we find support for what we call the "weak" form of the tariff phase-out hypothesis. Specifically, NAFTA's phased-out tariff cuts lead to a delayed and gradual growth of imports when either lumping all products together that faced these phased-out tariff cuts or separating these products out according to the length of their phase-out period.

On the other hand, we do not find supporting evidence for what we call the "strong" form of the tariff phase-out hypothesis that says products with notably longer phase-out periods should experience a notably longer period of gradual import growth. In fact, the import dynamics of 5-year phase-out products look remarkably similar to that of products where tariffs were immediately cut to zero. Similarly, the import dynamics of 10-year phase-out products look remarkably similar to products that had their prior GSP tariff- free access converted into immediate and permanent tariff-free access. Indeed, in terms of import value, the bulk of delayed import growth actually comes from products that received immediate and permanent tariff-free access upon implementation of NAFTA rather than products experiencing 5-year or 10-year phase-out periods.

Further, we do not find evidence of delayed pass through of tariff cuts to import prices. Because we proxy import prices with unit values and these are based on tariff exclusive import values, our measure of import prices is the (tariff exclusive) price received by the foreign exporter. Indeed, we cannot detect any effect of NAFTA tariff cuts on these import prices which suggests that NAFTA tariff cuts were passed immediately and fully to U.S. importers. The flavor of this result is consistent with the apparent puzzle in recent research that finds nearly complete pass through of 2018–2019 U.S. tariffs to U.S. importers (Amiti et al., 2019; Fajgelbaum et al., 2019). However, our results cover the short-run and the long-run and hence push back against the idea that the puzzle can be resolved merely by resorting to what might happen in the long-run.

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 "spatial" margin of trade. Largely independent of an imported product's length of tariff phase-out, NAFTA tariff cuts lead to 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 is crucial to understanding how trade flows respond after FTA formation. We show that this spatial expansion is driven particularly by homogeneous goods as well as final goods.

Our findings do not invalidate the use of lagged FTA dummies in the standard gravity approach when looking at the aggregate effects of FTAs. Indeed, the evidence we find for the weak form of the tariff phase-out hypothesis confirms FTAs do indeed have lagged effects on trade. But, the lack of evidence for the strong form of the tariff phase-out hypothesis implies these lagged effects go beyond tariff phase-out. Better understanding these mechanisms is an interesting area for future research.

References

Albornoz, F., Pardo, H. F. C., Corcos, G., & Ornelas, E. (2012). Sequential exporting. *Journal of International Economics*, *88*, 17 – 31. URL: http://www.sciencedirect.com/science/article/pii/S0022199612000220.

Amiti, M., Redding, S. J., & Weinstein, D. E. (2019). The impact of the 2018 tariffs on prices and welfare. *Journal of Economic Perspectives*, *33*, 187–210. URL: http://www.aeaweb.org/articles?id=10.1257/jep.33.4.187.

Amiti, M., Redding, S. J., & Weinstein, D. E. (2020). Who's paying for the US tariffs? A longer-term perspective. *National Bureau of Economic Research Working Paper No.26610*, . URL: https://www.nber.org/papers/w26610.

Arkolakis, C. (2010). Market penetration costs and the new consumers margin in international trade. *Journal of Political Economy*, *118*, 1151–1199.

Arkolakis, C. (2016). A unified theory of firm selection and growth. *The Quarterly Journal of Economics*, *131*, 89–155. URL: https://academic.oup.com/qje/article-pdf/131/1/89/30636098/qjv039.pdf.

Arkolakis, C., Costinot, A., & Rodríguez-Clare, A. (2012). New trade models, same old gains? *American Economic Review*, *102*, 94–130. URL: http://www.aeaweb.org/articles?id=10.1257/aer.102.1.94.

Arkolakis, C., Papageorgiou, T., & Timoshenko, O. A. (2018). Firm learning and growth. *Review of Economic Dynamics*, *27*, 146 – 168. URL: http://www.sciencedirect.com/science/article/pii/S1094202517300534.

Autor, D. H., Dorn, D., & Hanson, G. H. (2013). The China syndrome: Local labor market effects of import competition in the United States. *American Economic Review*, *103*, 2121–2168. URL: https://www.aeaweb.org/articles?id=10.1257/aer.103.6.2121.

Baier, S. L., & Bergstrand, J. H. (2007). Do free trade agreements actually increase members' international trade? *Journal of International Economics*, *71*, 72 – 95. URL: http://www.sciencedirect.com/science/article/pii/S0022199606000596.

Baier, S. L., Bergstrand, J. H., & Feng, M. (2014). Economic integration agreements and the margins of international trade. *Journal of International Economics*, *93*, 339 – 350. URL: http://www.sciencedirect.com/science/article/pii/S0022199614000506.

Berman, N., Rebeyrol, V., & Vicard, V. (2019). Demand learning and firm dynamics: Evidence from exporters. *The Review of Economics and Statistics*, *101*, 91–106.

Bernard, A. B., Jensen, J. B., Redding, S. J., & Schott, P. K. (2009). The margins of US trade. *American Economic Review*, *99*, 487–93. URL: http://www.aeaweb.org/articles?id=10.1257/aer.99.2.487.

Broda, C., & Weinstein, D. E. (2006). Globalization and the gains from variety. *The Quarterly Journal of Economics*, *121*, 541–585. URL: http://www.jstor.org/stable/25098800.

Caliendo, L., & Parro, F. (2014). Estimates of the trade and welfare effects of NAFTA. *The Review of Economic Studies*, *82*, 1–44. URL: https://doi.org/10.1093/restud/rdu035.

Cavallo, A., Gopinath, G., Neiman, B., & Tang, J. (2019). Tariff passthrough at the border and at the store: evidence from US trade policy. *National Bureau of Economic Research Working Paper No.26396*, .

Coelli, F. (2018). Trade policy uncertainty and innovation: Evidence from China. *Mimeo*, . URL: https://www.dropbox.com/s/t2i3evje7nbhb94/TPU_INNO_FC_THESIS.pdf?dl=0.

Costinot, A., Donaldson, D., & Komunjer, I. (2011). What goods do countries trade? A quantitative exploration of Ricardo's ideas. *The Review of Economic Studies*, *79*, 581–608. URL: https://doi.org/10.1093/restud/rdr033.

Fajgelbaum, P. D., Goldberg, P. K., Kennedy, P. J., & Khandelwal, A. K. (2019). The Return to protectionism. *The Quarterly Journal of Economics*, *135*, 1–55. URL: https://doi.org/10.1093/qje/qjz036.

Feenstra, R. C., Romalis, J., & Schott, P. K. (2002). US imports, exports, and tariff data, 1989-2001. *National Bureau of Economic Research*, . URL: https://www.nber.org/papers/w9387.

Fernandes, A. P., & Tang, H. (2014). Learning to export from neighbors. *Journal of International Economics*, *94*, 67 – 84. doi:https://doi.org/10.1016/j.jinteco.2014.06.003.

Fitzgerald, D., Haller, S., & Yedid-Levi, Y. (2019). How exporters grow. *National Bureau of Economic Research Working Paper No. 21935*, . URL: https://www.nber.org/papers/w21935.

Fitzgerald, D., & Priolo, A. (2018). How do firms build market share? *National Bureau of Economic Research Working Paper No. 24794*, . URL: https://www.nber.org/papers/w24794.

Foster, L., Haltiwanger, J., & Syverson, C. (2016). The slow growth of new plants: Learning about demand? *Economica*, *83*, 91–129. URL: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ecca.12172.

Frazer, G., & Biesebroeck, J. V. (2010). Trade growth under the African Growth and Opportunity Act. *The Review of Economics and Statistics*, *92*, 128–144. URL: http://www.jstor.org/stable/25651394.

Friederich, B. U., & Zator, M. (2018). Adaptation to shocks and the role of capital structure: Danish firms during the cartoon crisis. *Mimeo*, . URL: https://ideas.repec.org/p/aah/aarhec/2018-12.html.

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

Gourio, F., & Rudanko, L. (2014). Customer capital. *The Review of Economic Studies*, *81*, 1102–1136. URL: https://academic.oup.com/restud/article-pdf/81/3/1102/18397883/rdu007.pdf.

Gruber, J. (1994). The incidence of mandated maternity benefits. *The American Economic Review*, *84*, 622–641. URL: http://www.jstor.org/stable/2118071.

Hakobyan, S. (2020). GSP Expiration and declining exports from developing countries. *Canadian Journal of Economics*, *53*. URL: https://shakobyan.weebly.com/uploads/3/6/1/4/3614012/hakobyan_gsp_expiration_rr_clean.pdf.

Handley, K. (2014). Exporting under trade policy uncertainty: Theory and evidence. *Journal of International Economics*, *94*, 50 – 66. URL: http://www.sciencedirect.com/science/article/pii/S002219961400083X.

Handley, K., & Limão, N. (2017). Policy uncertainty, trade, and welfare: Theory and evidence for china and the united states. *American Economic Review*, *107*, 2731–83. URL: http://www.aeaweb.org/articles?id=10.1257/aer.20141419.

Hillberry, R. H., & McDaniel, C. A. (2002). A decomposition of North American trade growth since NAFTA. *US International Trade Commission International Economic Review*, . URL: https://www.usitc.gov/publications/ier/pub3527.pdf.

Kehoe, T. J., & Ruhl, K. J. (2013). How important is the new goods margin in international trade? *Journal of Political Economy*, *121*, 358–392. URL: http://www.jstor.org/stable/10.1086/670272.

Kohl, T. (2014). Do we really know that trade agreements increase trade? *Review of World Economics / Weltwirtschaftliches Archiv*, *150*, 443–469. URL: http://www.jstor.org/stable/44211813.

Pierce, J. R., & Schott, P. K. (2012). Concording US harmonized system codes over time. *Journal of Official Statistics*, *28*, 53–68. URL: http://www.justinrpierce.com/index_files/Pierce_Schott_JOS_2012.pdf.

Pierce, J. R., & Schott, P. K. (2016). The surprisingly swift decline of US manufacturing employment. *American Economic Review*, *106*, 1632–62. URL: http://www.aeaweb.org/articles?id=10.1257/aer.20131578.

Rauch, J. E. (1999). Networks versus markets in international trade. *Journal of International Economics*, *48*, 7 – 35. URL: http://www.sciencedirect.com/science/article/pii/S0022199698000099.

Romalis, J. (2007). NAFTA's and CUSFTA's impact on international trade. *The Review of Economics and Statistics*, *89*, 416–435. URL: http://www.jstor.org/stable/40043039.

Ruhl, K. J. (2008). The international elasticity puzzle. *Mimeo*, . URL: http://kimjruhl.com/s/ElasticityMarch08.pdf.

Timoshenko, O. A. (2015). Product switching in a model of learning. *Journal of International Economics*, *95*, 233 – 249. URL: http://www.sciencedirect.com/science/article/pii/S0022199614001548.

Trefler, D. (2004). The long and short of the Canada-U.S. free trade agreement. *American Economic Review*, *94*, 870–895. URL: http://www.aeaweb.org/articles?id=10.1257/0002828042002633.

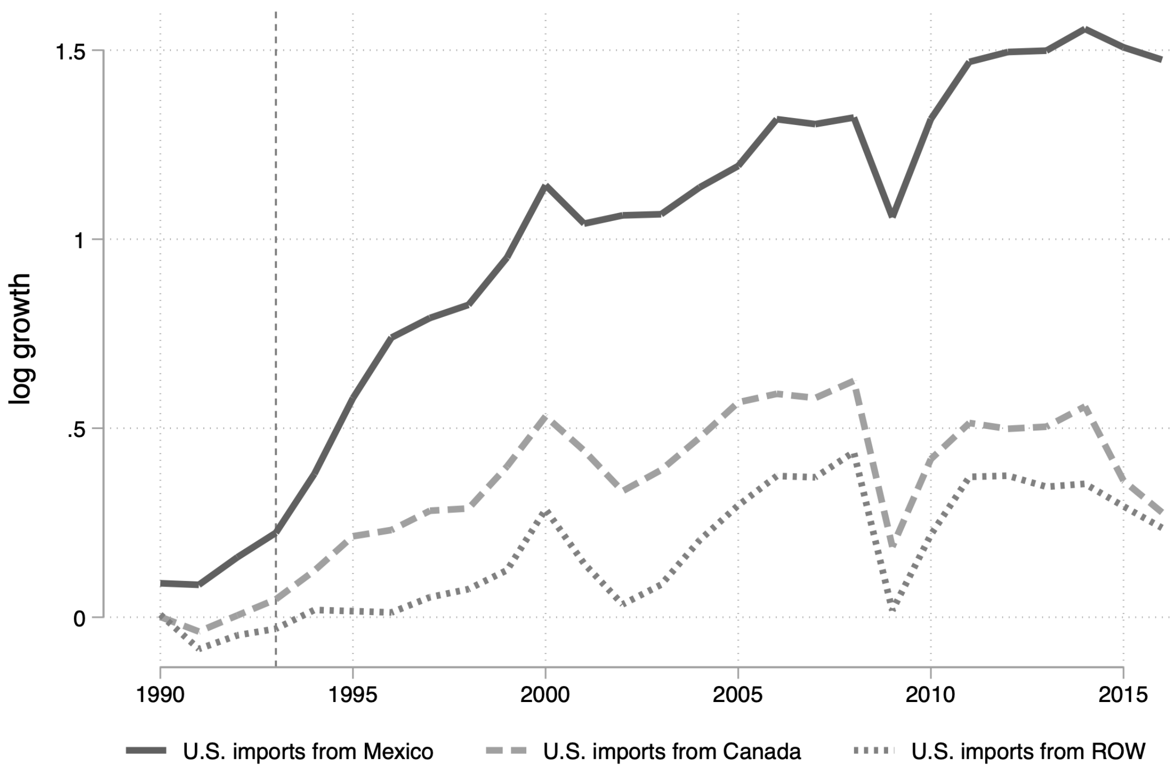

Figure 1: Cumulative real growth of U.S. imports from 1989.

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.

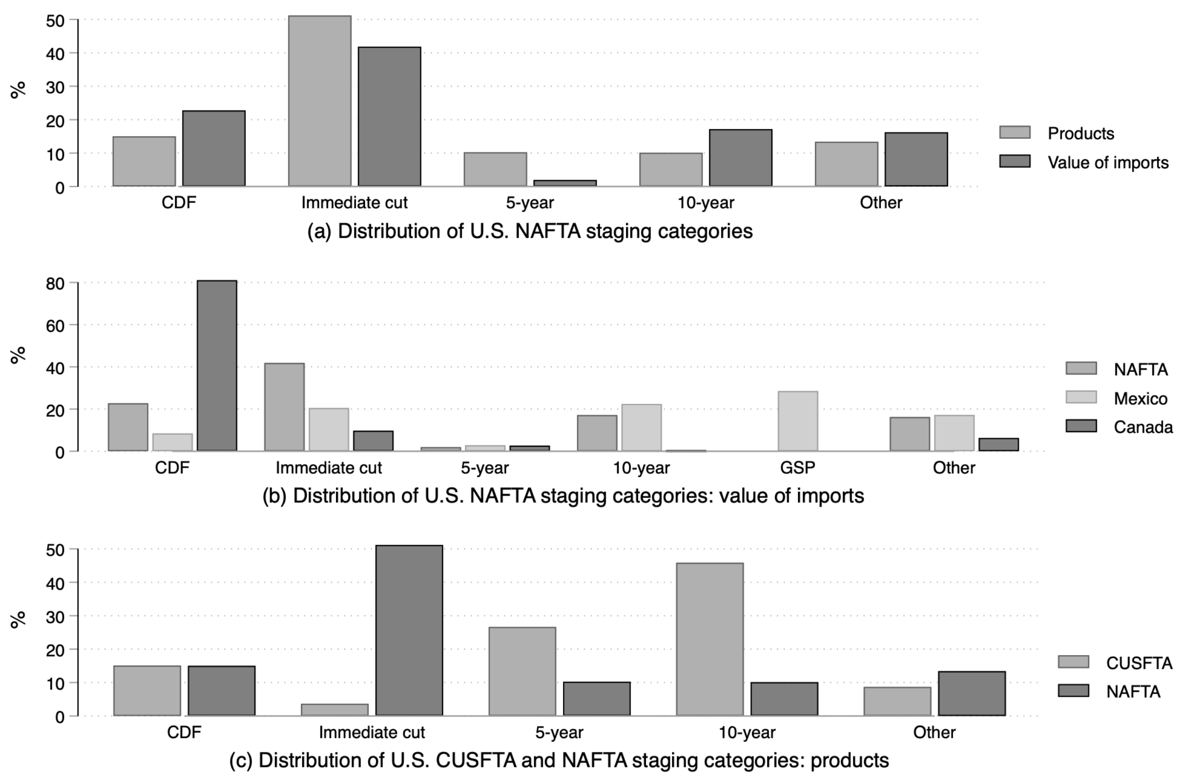

Figure 2: Distributions of U.S. NAFTA staging categories.

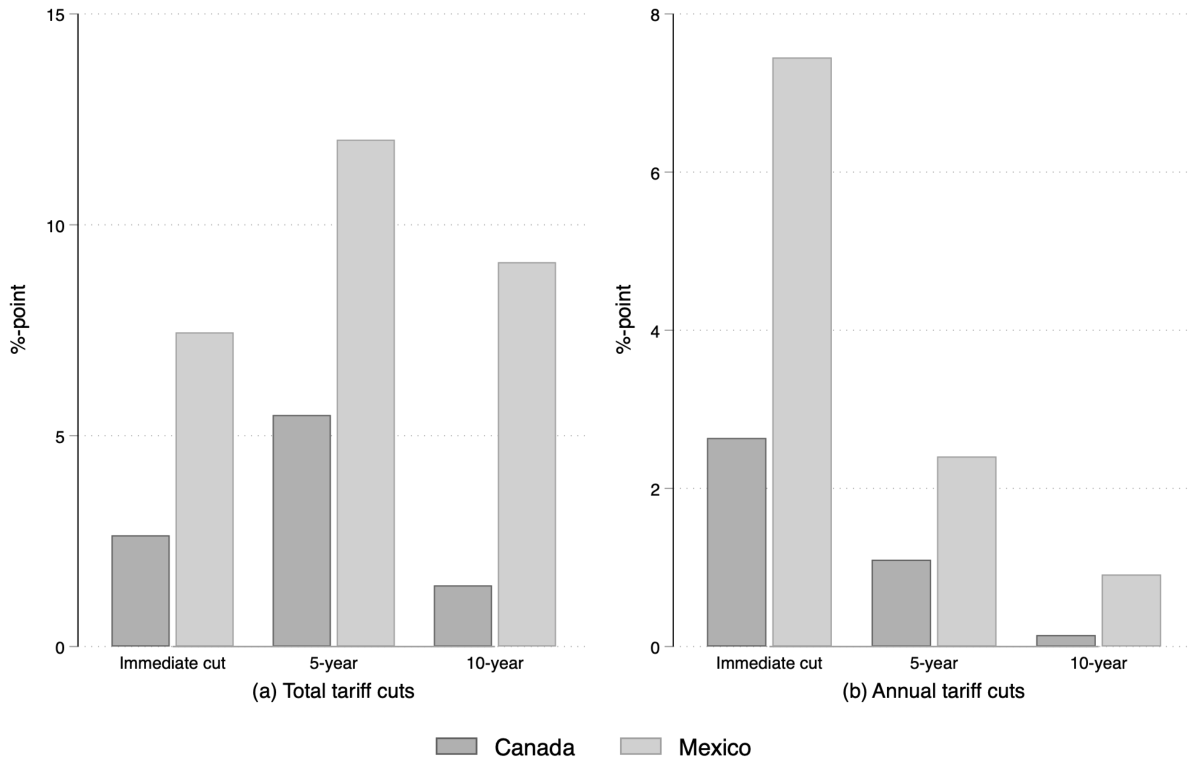

Figure 3: Tariff cuts implied by U.S. NAFTA staging categories.

Notes: Total tariff cut is the mean percentage-point reduction in U.S. tariffs per staging category implied by tariff elimination from the pre-NAFTA tariffs imposed on Canada and Mexico; see Appendix A.2 for more details on pre-NAFTA tariffs. The annual tariff cut is the total tariff cut divided by the staging category's length of phase-out in years.

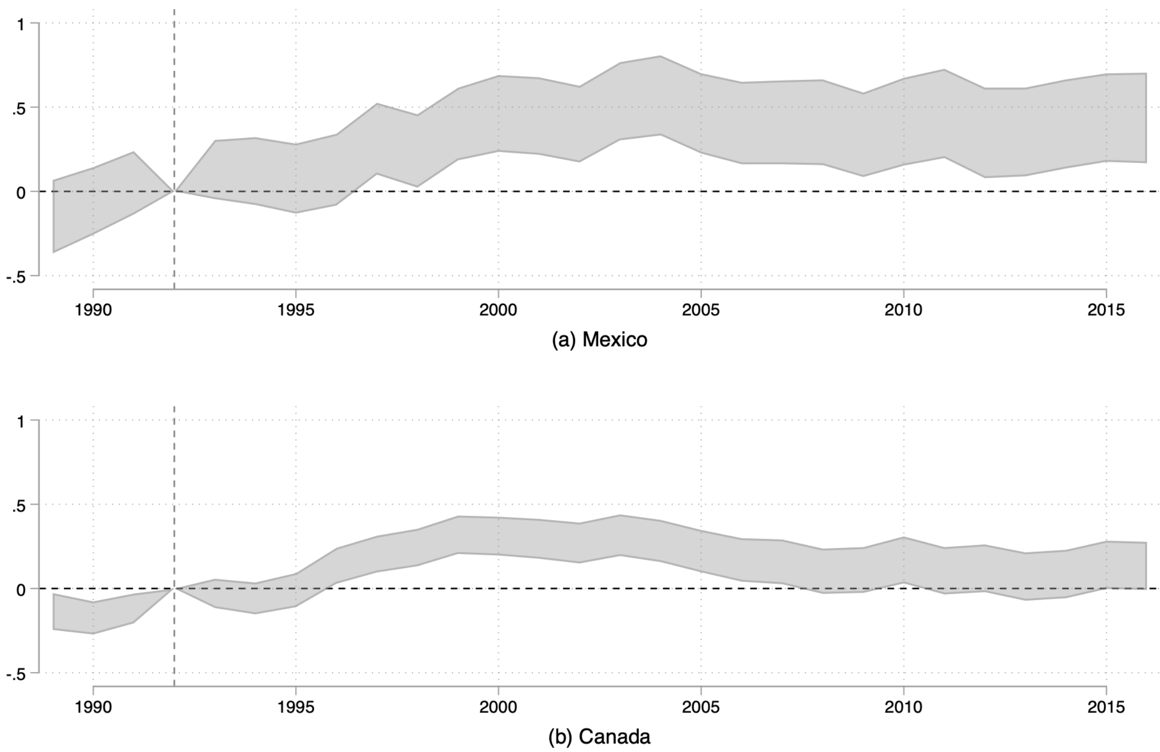

Figure 4: Time-varying homogeneous DDD estimates for Mexico and Canada.

Notes: DDD point estimates correspond to modified version of equation (5) with $Post_t$ replaced by $\mathbf{Year}_t$. Plots represent 95% confidence intervals. Two-way clustered standard errors are used, clustering on both country-year and product-year.

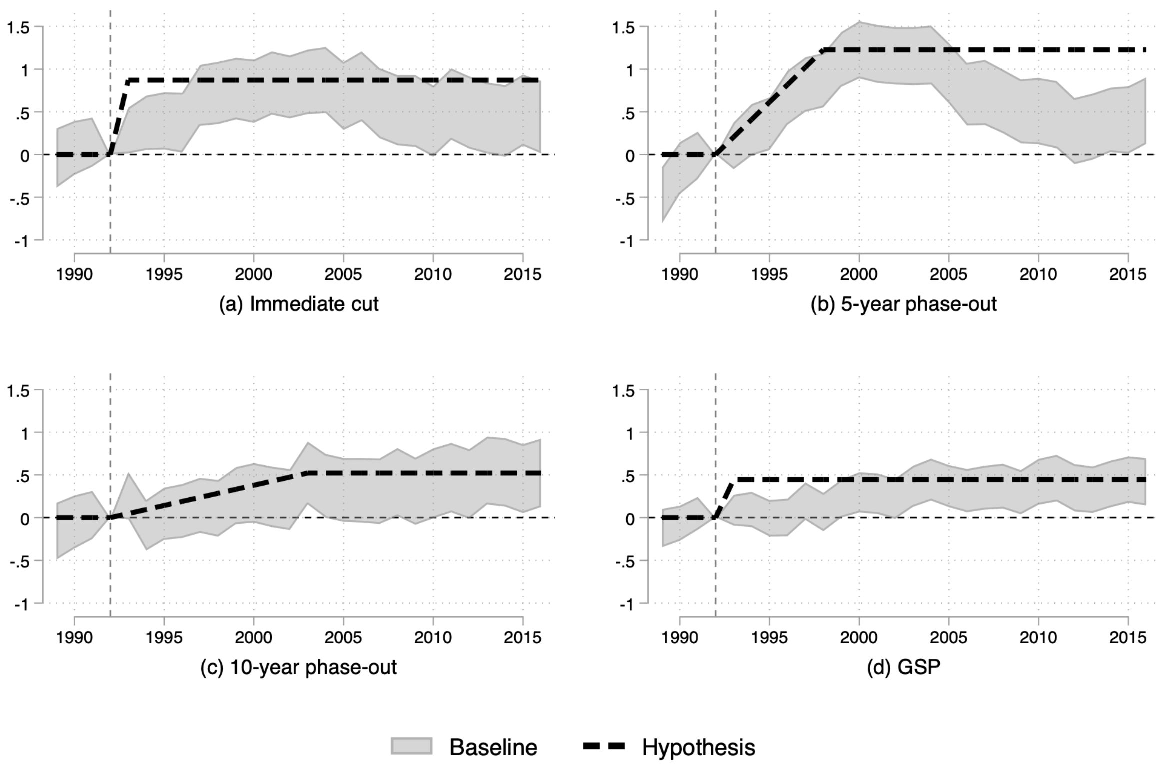

Figure 5: Time-varying heterogeneous DDD estimates for Mexico.

Notes: DDD point estimates correspond to equation (6). Plots represent 95% confidence intervals. Two-way clustered standard errors are used, clustering on both country-year and product-year. Height of Hypothesis line corresponds to largest DDD point estimate from equation (6) in the window 1993-2004; upward-sloping segment depicts gradual, linear increase of trade anticipated under specified staging category.

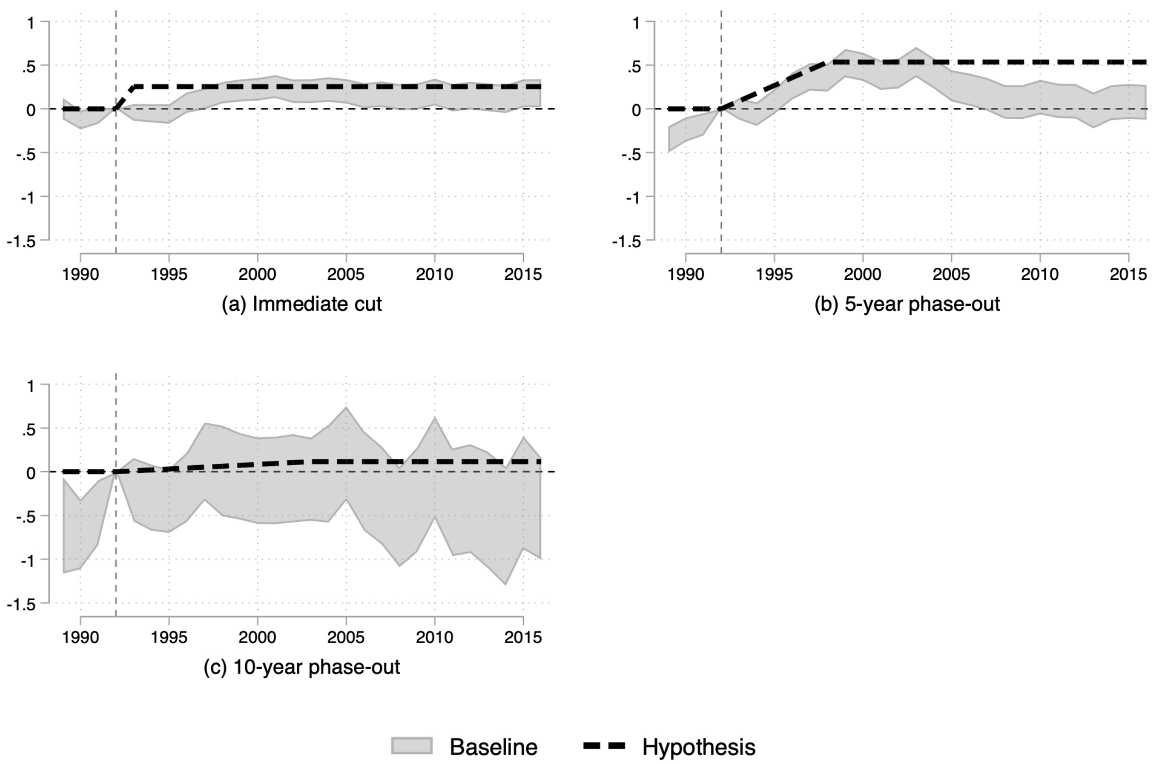

Figure 6: Time-varying heterogeneous DDD estimates for Canada.

Notes: DDD point estimates correspond to equation (6). Plots represent 95% confidence intervals. Two-way clustered standard errors are used, clustering on both country-year and product-year. Height of Hypothesis line corresponds to largest DDD point estimate from equation (6) in the window 1993-2004; upward-sloping segment depicts gradual, linear increase of trade anticipated under specified staging category.

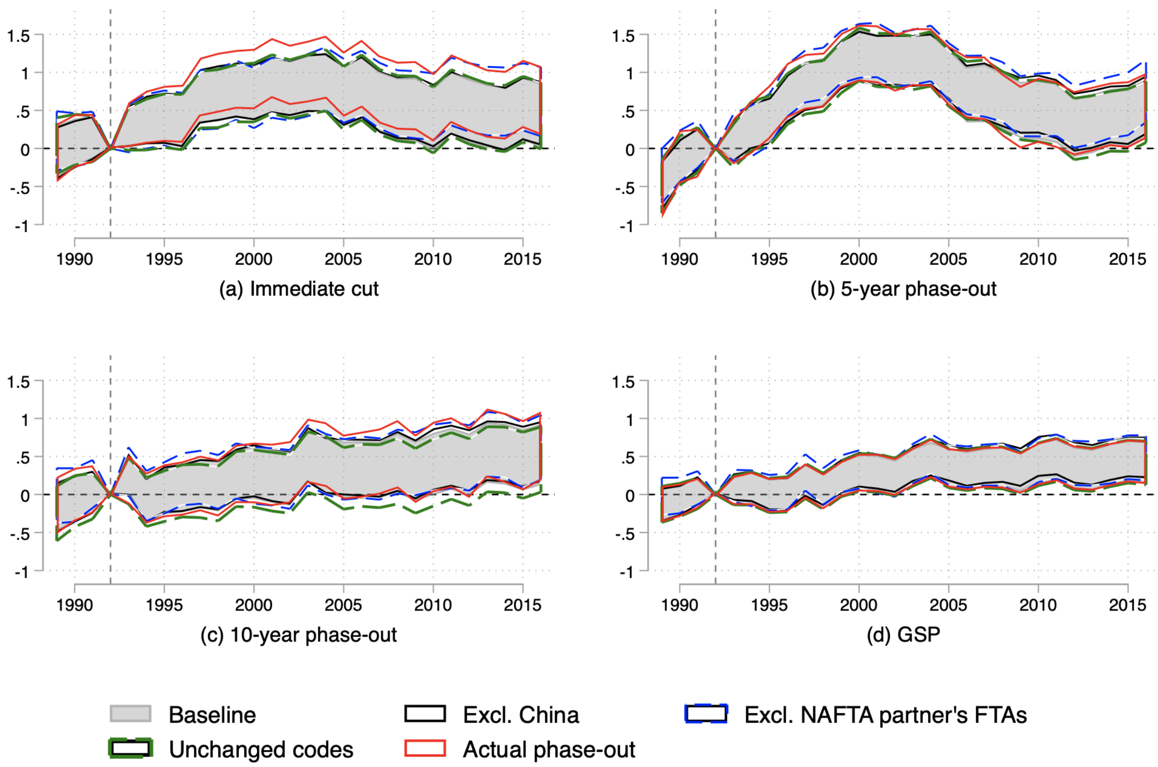

Figure 7: Time-varying heterogeneous DDD estimates for Mexico: Robustness.

Notes: DDD point estimates correspond to equation (6). 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 sub-sample where product code remains unchanged over sample period; Actual phase-out indicates sub-sample where product phase-out always follows staging schedule.

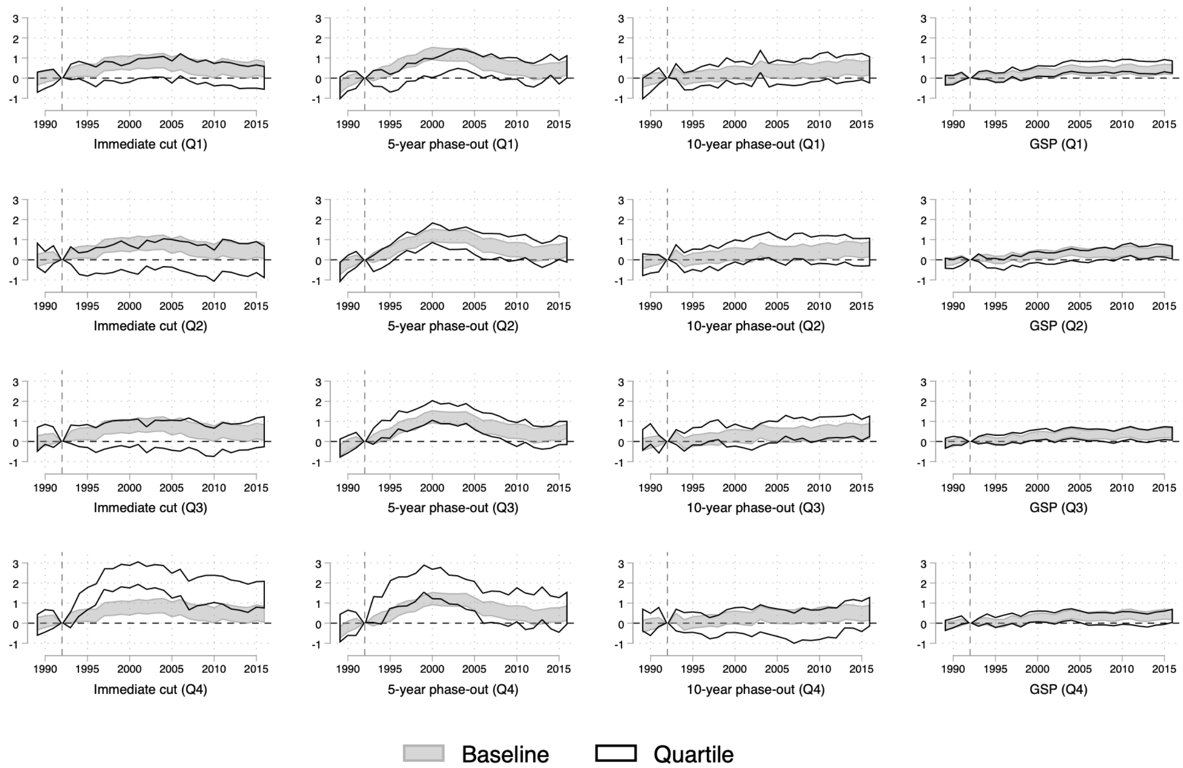

Figure 8: Time-varying heterogeneous DDD estimates for Mexico: Tariff cuts.

Notes: DDD point estimates correspond to equation (6). Plots represent 95% confidence intervals. Two-way clustered standard errors are used, clustering on both country-year and product-year. Plots by staging-category-specific quartiles of 1989 MFN tariff distribution.

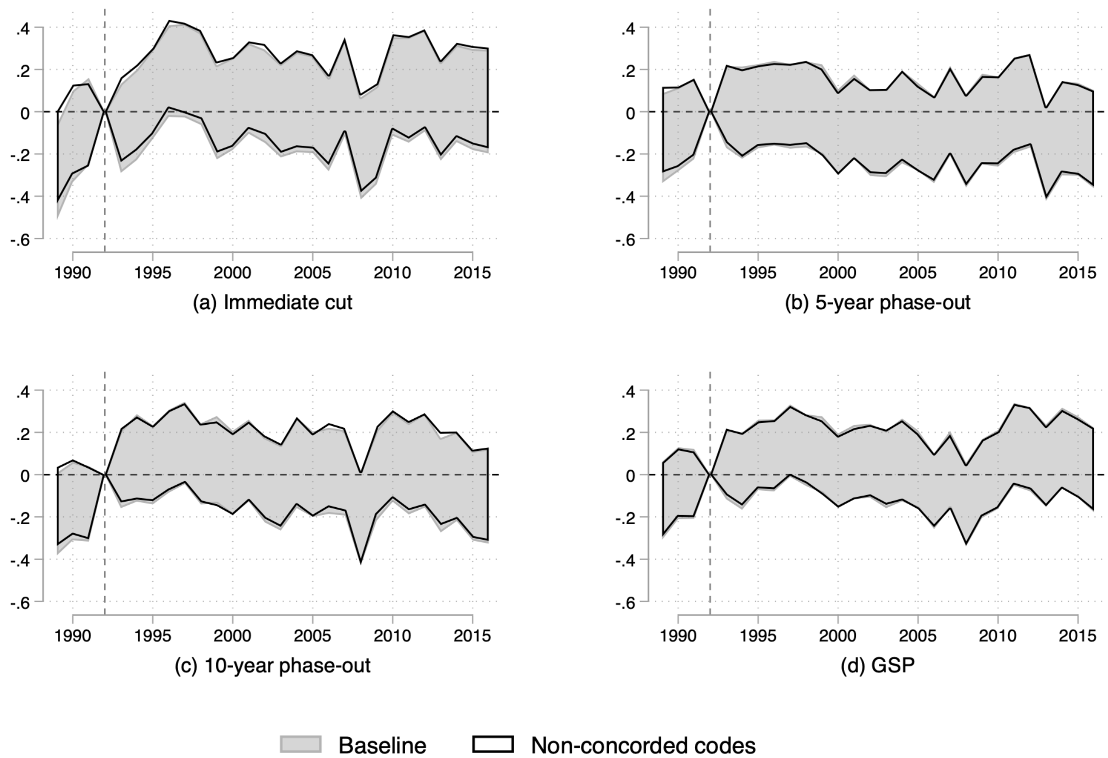

Figure 9: Time-varying heterogeneous DDD estimates for Mexico: Unit values.

Notes: DDD point estimates correspond to equation (6) with log real unit values as dependent variable. Plots represent 95% confidence intervals.

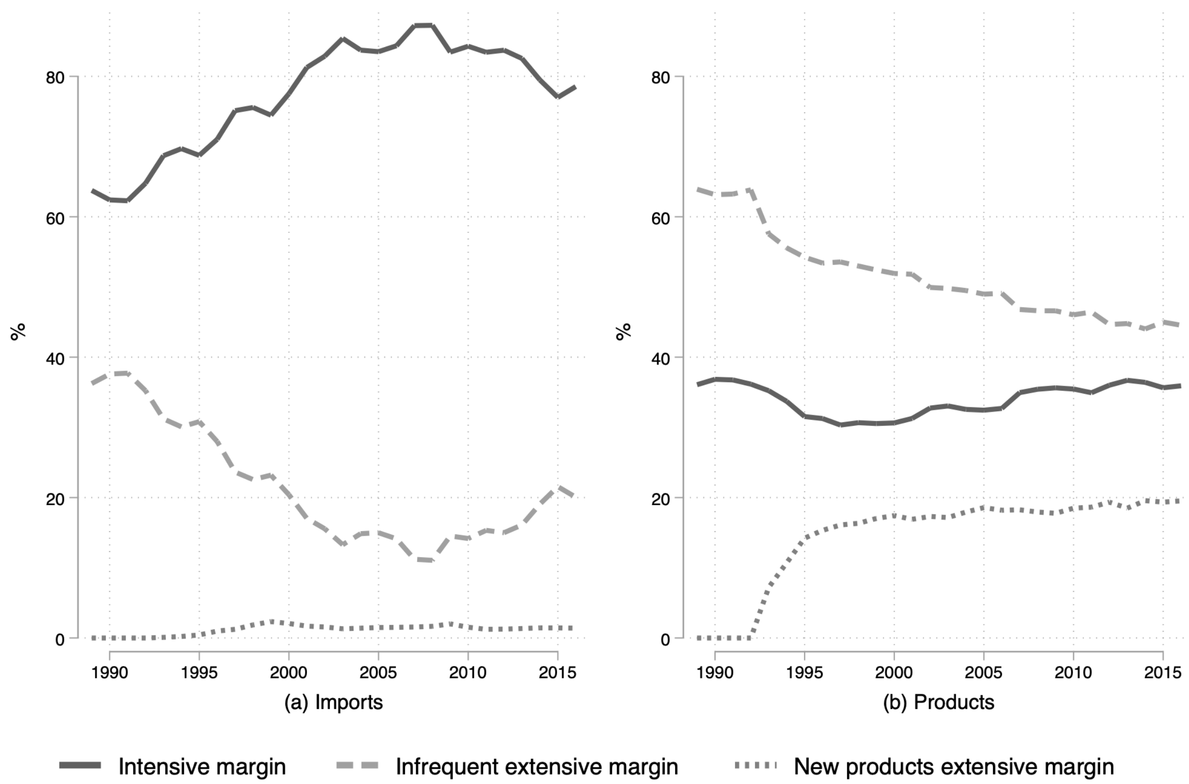

Figure 10: Imports from Mexico: Intensive margin, infrequent extensive margin, and new products extensive margin.

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.

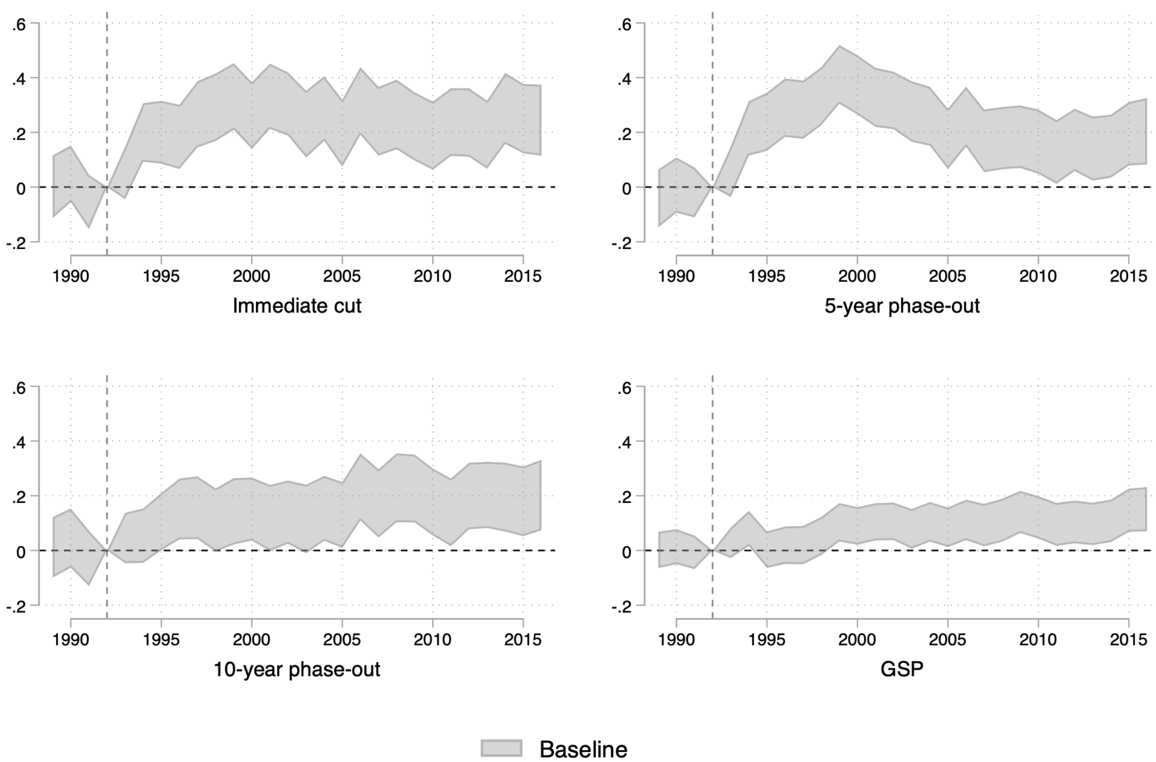

Figure 11: Time-varying heterogeneous DDD estimates for Mexico: Districts entered.

Notes: DDD point estimates correspond to equation (6) with log number of districts entered, $\ln D_{cpt}$, by exports of product p from exporter c in year t as dependent variable. Plots represent 95% confidence intervals. Two-way clustered standard errors are used, clustering on both country-year and product-year.

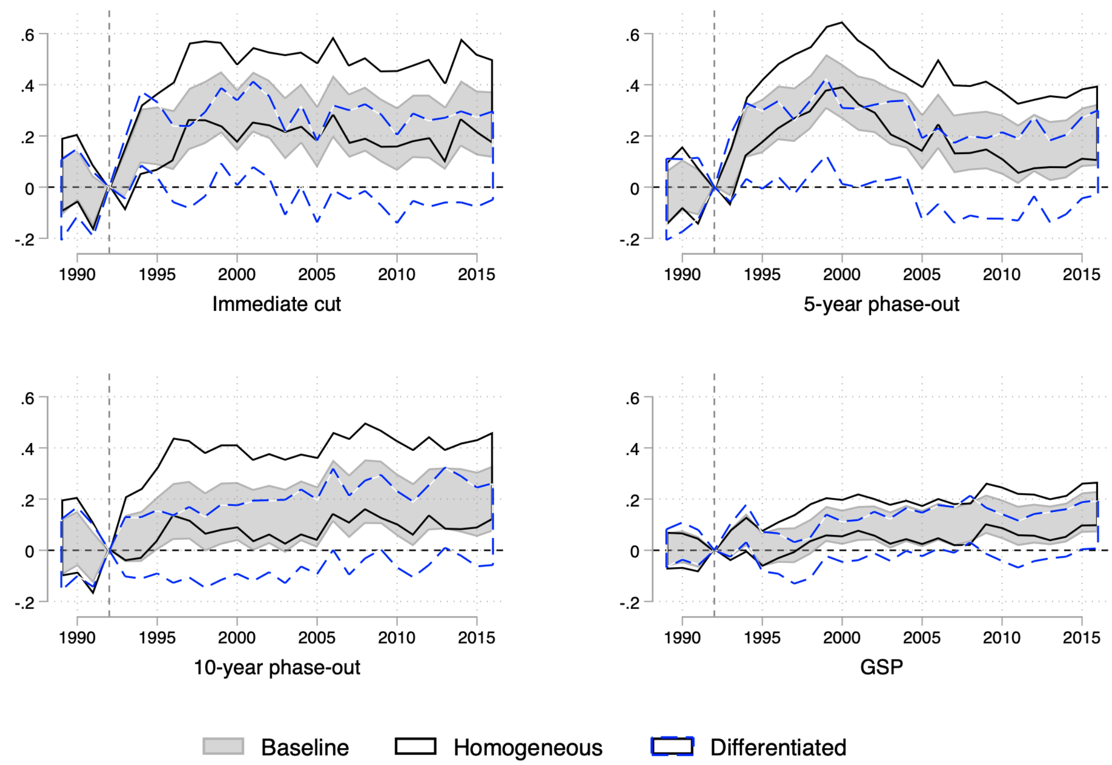

Figure 12: Time-varying heterogeneous DDD estimates for Mexico: Districts entered for Rauch product types.

Notes: DDD point estimates correspond to equation (6) with log number of districts entered, $\ln D_{cpt}$, by exports of product p from exporter c in year t as dependent variable. Plots represent 95% confidence intervals. Two-way clustered standard errors are used, clustering on both country-year and product-year.

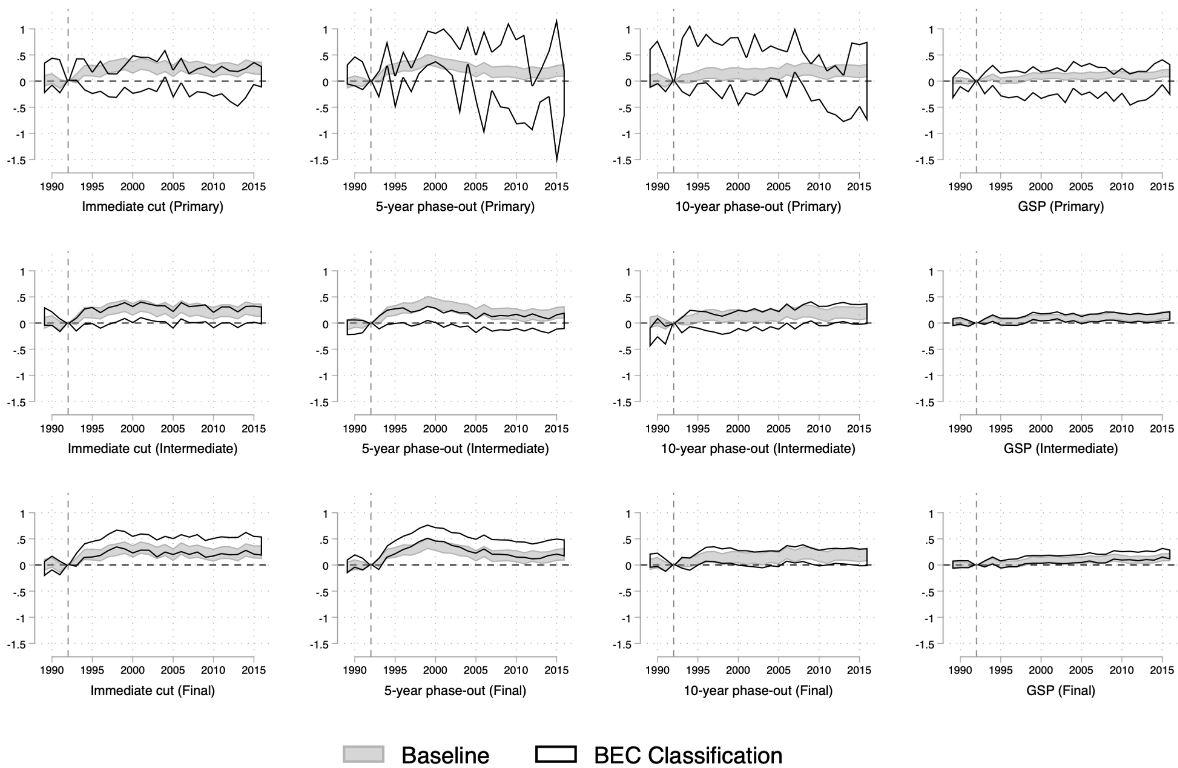

Figure 13: Time-varying heterogeneous DDD estimates for Mexico: Districts entered for BEC categories.

Notes: DDD point estimates correspond to equation (6) with log number of districts entered, $\ln D_{cpt}$, by exports of product p from exporter c in year t 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	Tariff 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

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%	$0.70	$0.68	97.11%
1990	8,677	8,456	97.45%	126,447	122,960	97.24%	$0.71	$0.69	97.22%
1991	8,659	8,523	98.43%	125,963	123,708	98.21%	$0.67	$0.66	97.79%
1992	8,745	8,642	98.82%	129,326	127,600	98.67%	$0.71	$0.70	98.13%
1993	8,690	8,675	99.83%	134,926	134,541	99.71%	$0.75	$0.74	98.85%
1994	8,994	8,492	94.42%	145,319	136,326	93.81%	$0.85	$0.73	85.69%
1995	9,568	7,911	82.68%	151,752	129,641	85.43%	$0.93	$0.73	78.68%
1996	9,770	7,449	76.24%	158,050	125,800	79.60%	$0.98	$0.71	72.37%
1997	9,997	7,461	74.63%	168,033	130,389	77.60%	$1.04	$0.74	71.34%
1998	9,896	7,392	74.70%	168,495	130,903	77.69%	$1.09	$0.76	70.18%
1999	9,876	7,406	74.99%	170,030	132,860	78.14%	$1.18	$0.82	69.72%
2000	9,908	7,412	74.81%	178,080	138,807	77.95%	$1.37	$0.97	70.41%
2001	9,917	7,406	74.68%	178,476	138,543	77.63%	$1.24	$0.91	73.47%
2002	10,163	6,955	68.43%	185,114	134,846	72.84%	$1.24	$0.84	67.79%
2003	10,179	6,953	68.31%	188,279	136,934	72.73%	$1.32	$0.90	67.91%
2004	10,155	6,950	68.44%	191,986	139,445	72.63%	$1.51	$1.02	67.44%
2005	10,172	6,944	68.27%	195,741	141,474	72.28%	$1.67	$1.13	67.60%
2006	10,188	6,951	68.23%	198,368	142,945	72.06%	$1.80	$1.22	67.83%
2007	10,116	6,343	62.70%	197,675	133,373	67.47%	$1.85	$1.19	64.51%
2008	10,095	6,339	62.79%	192,709	130,455	67.70%	$1.96	$1.29	65.64%
2009	10,043	6,326	62.99%	183,535	124,129	67.63%	$1.41	$0.88	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%	$1.95	$1.25	63.89%
2012	10,300	6,093	59.16%	197,081	128,289	65.09%	$1.96	$1.23	62.96%
2013	10,287	6,091	59.21%	193,084	126,253	65.39%	$1.91	$1.20	62.87%
2014	10,299	6,087	59.10%	196,866	128,667	65.36%	$1.94	$1.21	62.30%
2015	10,308	6,096	59.14%	203,138	132,535	65.24%	$1.81	$1.11	61.35%
2016	10,297	6,099	59.23%	204,767	133,760	65.32%	$1.70	$1.04	61.57%

Notes: Staging category data refer to NAFTA U.S. tariff schedule data from NAFTA Annex 302.2. All trade data is non-concorded 8-digit HS data. Import data from USITC and measured in trillions of real 2010 USD using the World Development Indicators GDP deflator.

Table 2: Matching NAFTA tariff schedule to USITC trade data

Panel A: NAFTA vs ROW approach

A1. Phase-out products

	Mexico				Canada		
	Pre-NAFTA	Post-NAFTA	Growth		Pre-NAFTA	Post-NAFTA	Growth
NAFTA partner	12.375	12.809	0.434	NAFTA partner	12.264	12.635	0.370
	(0.026)	(0.011)	(0.028)		(0.026)	(0.013)	(0.029)
	[11,665]	[77,159]			[11,077]	[59,874]	
ROW	11.544	11.337	-0.207	ROW	11.354	11.077	-0.277
	(0.005)	(0.002)	(0.005)		(0.006)	(0.002)	(0.007)
	[295,869]	[2,024,041]			[186,686]	[1,340,522]	
DD			0.641	DD			0.647
			(0.037)				(0.034)

A2. CDF products

	Mexico				Canada		
	Pre-NAFTA	Post-NAFTA	Growth		Pre-NAFTA	Post-NAFTA	Growth
NAFTA partner	12.354	12.838	0.485	NAFTA partner	13.042	13.507	0.465
	(0.060)	(0.026)	(0.065)		(0.029)	(0.014)	(0.032)
	[2,137]	[13,120]			[10,176]	[51,392]	
ROW	11.825	11.928	0.103	ROW	11.847	11.856	0.010
	(0.012)	(0.005)	(0.013)		(0.007)	(0.003)	(0.007)
	[49,344]	[328,656]			[139,683]	[877,854]	
DD			0.382	DD			0.455
			(0.082)				(0.037)
DDD			0.259	DDD			0.192
			(0.090)				(0.050)

Panel B: Phase-out vs CDF-products approach

B1. NAFTA partner

	Mexico				Canada		
	Pre-NAFTA	Post-NAFTA	Growth		Pre-NAFTA	Post-NAFTA	Growth
Phase-out products	12.375	12.809	0.434	Phase-out products	12.264	12.635	0.370
	(0.026)	(0.011)	(0.028)		(0.026)	(0.013)	(0.029)
	[11,665]	[77,159]			[11,077]	[59,874]	
CDF products	12.354	12.838	0.485	CDF products	13.042	13.507	0.465
	(0.060)	(0.026)	(0.065)		(0.029)	(0.014)	(0.032)
	[2,137]	[13,120]			[10,176]	[51,392]	
DD			-0.051	DD			-0.095
			(0.091)				(0.051)

B2. ROW

	Mexico				Canada		
	Pre-NAFTA	Post-NAFTA	Growth		Pre-NAFTA	Post-NAFTA	Growth
Phase-out products	11.544	11.337	-0.207	Phase-out products	11.354	11.077	-0.277
	(0.005)	(0.002)	(0.005)		(0.006)	(0.002)	(0.007)
	[295,869]	[2,024,041]			[186,686]	[1,340,522]	
CDF products	11.825	11.928	0.103	CDF products	11.846	11.856	0.010
	(0.012)	(0.005)	(0.013)		(0.007)	(0.003)	(0.007)
	[49,344]	[328,656]			[139,683]	[877,854]	
DD			-0.310	DD			-0.287
			(0.032)				(0.021)
DDD			0.259	DDD			0.192
			(0.090)				(0.050)

Notes: Cells contain mean log imports for the relevant group of countries, products and years. 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).

Table 3: Time-invariant DDD estimates of NAFTA

	Mexico				Canada			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post	0.103 [c]	0.255 [c]			0.010	0.240 [c]		
	(0.03)	(0.02)			(0.02)	(0.01)		
NAFTA	0.529 [c]				1.196 [c]			
	(0.10)				(0.05)			
Phase	-0.281 [c]				-0.492 [c]			
	(0.03)				(0.02)			
Post × NAFTA	0.382 [c]	0.349 [c]			0.455 [c]	0.129 [c]		
	(0.08)	(0.08)			(0.04)	(0.03)		
Post × Phase	-0.310 [c]	-0.071 [c]	-0.083 [c]		-0.287 [c]	-0.139 [c]	-0.166 [c]	
	(0.03)	(0.02)	(0.02)		(0.02)	(0.01)	(0.01)	
NAFTA × Phase	0.303 [b]				-0.285 [c]			
	(0.11)				(0.07)			
Post × NAFTA × Phase	0.259 [b]	0.371 [c]	0.391 [c]	0.388 [c]	0.192 [c]	0.296 [c]	0.323 [c]	0.262 [c]
	(0.09)	(0.08)	(0.08)	(0.08)	(0.05)	(0.04)	(0.04)	(0.04)
Observations	2,801,991	2,717,217	2,716,953	2,708,951	2,677,264	2,597,551	2,597,297	2,589,435
Adjusted R^2	0.013	0.722	0.737	0.760	0.032	0.728	0.741	0.763
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: Columns (1) and (5) based on equation (4) and columns (4) and (8) based on equation (5). 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 4: 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	0.041	0.479 [b]	0.159	0.122	-0.002	0.353 [c]	0.613 [a]
1993	0.203	0.262	0.308	0.153	0.085	0.236 [b]	0.507 [a]
1994	0.226	0.299	-0.120	0.046	0.024	0.116	0.172
1995	0.394 [a]	0.360 [a]	0.045	-0.008	-0.059	0.085	-0.335
1996	0.092	0.563 [c]	-0.180	-0.085	0.113 [a]	0.261 [c]	0.030
1997	0.321	0.531 [c]	0.232	0.097	0.165 [b]	0.426 [c]	0.412 [a]
1998	0.327 [a]	0.510 [c]	0.063	0.073	0.243 [c]	0.272 [c]	0.343
1999	0.399 [b]	0.449 [c]	0.180	0.221 [a]	0.138 [b]	0.262 [c]	0.124
2000	0.049	0.407 [b]	0.146	0.103	0.109 [a]	0.115	-0.218
2001	0.116	0.308 [a]	0.134	0.215 [a]	0.070	0.028	-0.107
2002	0.020	0.043	-0.047	-0.003	-0.008	-0.119	-0.023
2003	0.110	-0.074	0.233	0.071	-0.022	0.055	0.018
2004	0.034	-0.013	0.129	0.165	-0.033	0.020	0.076
2005	-0.105	-0.205	0.114	0.148	-0.002	-0.141 [a]	0.285
2006	-0.053	-0.446 [c]	-0.202	-0.051	-0.052	-0.315 [c]	-0.021
2007	-0.269	-0.438 [b]	-0.063	-0.095	-0.054	-0.239 [b]	-0.249
2008	-0.167	-0.328 [a]	0.091	0.001	-0.066	-0.184 [a]	-0.730 [b]
2009	-0.290 [a]	-0.201	-0.011	-0.019	-0.009	-0.143	-0.215
2010	-0.211	-0.219	0.093	0.069	0.026	-0.033	0.320
2011	0.071	-0.157	0.053	0.093	-0.007	0.013	0.169
2012	-0.017	-0.232	0.084	0.053	0.011	0.010	0.013
2013	0.037	-0.182	0.149	-0.094	-0.061	-0.152	-0.486 [a]
2014	-0.197	-0.060	0.062	-0.068	-0.013	-0.021	-0.275
2015	0.028	0.131	0.064	0.095	0.027	-0.002	0.067
2016	0.011	0.188	-0.026	0.093	0.048	0.093	0.016
Observations	2,708,951				2,589,435		
Country × Product FE	Yes				Yes		
Country × Year FE	Yes				Yes		
Product × Year FE	Yes				Yes		

Notes: Point estimates for year t coefficient is the difference in year t and year $t-3$ point estimates from Figure 5, 6 and Table A.6; 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 5: DDD regression: time-varying, heterogeneous cumulative treatment effects for 4-year rolling windows

Description	Continue duty free		Immediate cut		5-year phase-out		10-year phase-out		GSP		Total	
	Products	%	Products	%	Products	%	Products	%	Products	%	Products	%
Always follows staging schedule	1,138	100%	850	99.8%	614	76.8%	659	89.9%	3,549	99.8%	6,810	96.2%
NAFTA tariff acceleration	-		-		158	19.8%	21	2.9%	-		179	2.5%
MFN tariff reductions	-		-		3	0.4%	35	4.8%	-		38	0.5%
Only follows phase-out period length	-		-		22	2.8%	15	2.0%	-		37	0.5%
Phase-out path temporarily not followed	-		1	0.1%	2	0.3%	3	0.4%	-		6	0.1%
Different staging category imposed	-		1	0.1%	-		-		8	0.2%	9	0.1%
Total	1,138	100%	852	100%	799	100%	733	100%	3,557	100%	7,079	100%

Notes: Comparison of actual tariffs imposed by U.S. over post-NAFTA sample period versus tariff path specified by NAFTA phase-out schedule. Sample is the 7,079 HS8 products whose product codes are unchanged over 1989-2016 sample period. Actual U.S. tariff data from USITC tariff database and USITC HTS Archive. See main text for further details.

Table 6: Tariff phase-out: schedule vs actual

A Appendix

A.1 Supplemental tables for staging categories

		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%	1,015	11.4%	$2.67	41.9%	$0.40	9.8%	$0.47	20.5%
B	5 equal annual cuts to 0	2,285	26.7%	179	2.0%	792	9.0%	176	2.0%	$0.08	1.2%	$0.09	2.3%	$0.04	1.7%
B6	1 immediate cut + 5 equal annual cuts to 0			728	8.2%	728	8.2%	726	8.2%	$0.05	0.8%	$0.02	0.4%	$0.03	1.2%
C	10 equal annual cuts to 0	3,932	45.9%	750	8.5%	94	1.1%	748	8.4%	$1.09	17.0%	$0.01	0.2%	$0.51	22.1%
C10	Non-linear cuts to zero over 10 years			71	0.8%	0	0.0%	71	0.8%	$0.00	0.1%	$0.00	0.0%	$0.00	0.2%
C+	15 equal annual cuts to 0			74	0.8%	3	0.0%	74	0.8%	$0.00	0.1%	$0.00	0.0%	$0.00	0.1%
D	Continue duty free	1,295	15.1%	1,329	15.0%	3,871	43.8%	1,301	14.7%	$1.46	22.8%	$3.32	81.1%	$0.19	8.5%
GSP								4,111	46.3%					$0.66	28.5%
Mixed		745	8.7%	1,120	12.7%	755	8.5%	586	6.6%	$0.99	15.6%	$0.25	6.1%	$0.36	15.8%
Missing		2	0.0%	66	0.7%	65	0.7%	66	0.7%	$0.04	0.7%	$0.01	0.2%	$0.03	1.5%
Total		8,574	100%	8,843	100%	8,843	100%	8,874	100%	$6.39	100%	$4.09	100%	$2.30	100%

Notes: Staging category data come 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 U.S. 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 these 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	315	3.6%	3.6%	2,535	2.6%	2.6%	1,105	7.4%	7.4%
B	5 equal annual cuts to 0	2,285	5.8%	1.2%	792	4.7%	0.9%	176	9.3%	1.9%
B6	1 immediate cut + 5 equal annual cuts to 0				728	6.3%	1.1%	726	12.7%	2.1%
C	10 equal annual cuts 0	3,932	8.5%	0.9%	94	1.5%	0.1%	748	7.6%	0.8%
C10	Non-linear cuts to 0 over 10 years							71	14.1%	1.4%
C+	15 equal annual cuts to 0				3	0.0%	0.0%	74	19.7%	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								4,111	N/A	N/A
Mixed		745	N/A	N/A	586	N/A	N/A	586	N/A	N/A
Missing		2	N/A	N/A	66	N/A	N/A	66	N/A	N/A
Total		8,574			8,843			8,874		

Notes: Appendix A.2 explains the construction of the pre-CUSFTA and pre-NAFTA tariffs.

Table A.2: Tariff cuts by staging category

A.2 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 in line with the following procedure.

As a starting point for pre-CUSFTA tariffs faced by Canada, we take the 1989 U.S. MFN tariffs per Romalis' data described in Feenstra et al. (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.

We match 8,574 products from the CUSFTA staging schedule to (non-concorded) 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, five 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 begin in 1989. The remaining 37 products have "complex" base rates that cannot be transformed into an ad valorem equivalent tariff with USITC import data.[42]

For Canada's pre-NAFTA tariff, we follow a two-step procedure. First, a product's pre-

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

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 an ad valorem tariff, 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, five have complex tariff structures and two 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 five 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.[43] For remaining products, we 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 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.

[43] According to Romalis' data, the U.S. does not change any ad valorem MFN tariffs between 1991-1993.

A.3 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 shows how the U.S. phased out tariffs over time 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 notably 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).[44] Given USITC HS data start in 1989, our CUSFTA analysis only uses 1989 data even though this is not ideal because CUSFTA was implemented on January 1, 1989. Since NAFTA negotiations concluded in mid-1992, our NAFTA analysis uses 1991 data.

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

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 U.S. imports to U.S. 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 Appendix A.2), (v) a dummy variable for whether the product is homogeneous (as opposed to differentiated, per Rauch 1999), (vi) dummy variables for whether the product is an intermediate good or a primary good (as opposed to a final good, per the BEC classification), (vii) for NAFTA, the average percentage change in unit values from 1989 to 1991, and (viii) the product-level elasticity of substitution (Broda & Weinstein, 2006).[45] 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 the probability of a product's phase-out period being at least as long as it is relative to the probability of being shorter. For convenience of terminology, we hereafter refer to the impact on the odds ratio from a one standard deviation increase in an explanatory variable as the *odds ratio factor.*

Columns (1) and (2) describe the CUSFTA results. First, higher GLIs and, especially, higher CUSFTA base rates deliver longer phase out periods with odds factor ratios of 1.09-1.15 and 5.34-6.14 respectively.[46] Second, homogeneous goods (relative to differentiated goods) as well as intermediate and primary goods (relative to final goods) have substantially

[45] Homogeneous goods are reference-priced goods or goods traded on an organized exchange. Additionally, we measure import-export ratios and GLIs at the 6-digit level because HS codes for imports and exports are only comparable up to the 6-digit level. Given limitations in creating comparable and consistently-coded product codes for both imports and exports simultaneously, all estimates rely on the subset of product codes that remain unchanged at the 6-digit level throughout our sample.

[46] Note, $1.283^{6.72} = 5.34$ and $1.31^{6.72} = 6.14$.

shorter phase-out periods with respective odds factor ratios of 0.71-0.76, 0.68-0.77, and 0.52-0.60. Similarly, products with higher import shares from Canada and higher import-to-export ratios have shorter phase-outs with respective odds factor ratios of 0.75- 0.81 and 0.71-0.76. In contrast, the elasticity of substitution and the share of U.S. exports to Canada are not statistically significant determinants of CUSFTA phase-out length. Nevertheless, overall, the U.S. CUSFTA staging schedule appears closely related to key economic variables.

The NAFTA results show an overall dampening of the CUSFTA results. The odds ratio factor for the NAFTA partner base rate falls to 2.51-2.28 for Canada and 1.37-1.95 for Mexico. Additionally, the import-to-export ratio and the GLI are no longer robust determinants of phase-out. In terms of robust determinants, homogeneous, intermediate and primary goods still have robust and shorter phase-out periods and, unlike the CUSFTA results, products with higher elasticities of substitution have longer phase-out with a modest odds ratio factor of 1.04-1.07. When controlling for the elasticity of substitution, products with higher import shares from Canada and Mexico still have shorter phase-out periods with similar odds ratio factors to the CUSFTA results. In terms of non-robust determinants, changes in import and exports shares remain non-robust determinants and changes in unit values are also non-robust determinants.

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 an overly strong concern, especially for Mexican imports. Among other factors, 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, the elasticity of substitution, and whether goods are homogeneous (versus differentiated) and intermediate or primary (versus final).

Variable	Mean	Std. dev	Min.	Max.	Obs.
Canada-CUSFTA					
Staging category	3.15	1.07	1.00	4.00	7,118
Import share	13.28	23.53	0.00	100.00	7,118
Export share	18.95	19.04	0.00	100.00	7,118
Log import/export ratio	0.02	5.35	-32.77	34.26	7,118
GLI	0.33	0.31	0.00	1.00	7,118
Base rate	6.33	6.72	0.00	79.01	7,118
Homogeneous	0.44	0.50	0.00	1.00	7,118
Intermediate	0.52	0.50	0.00	1.00	7,118
Primary	0.06	0.24	0.00	1.00	7,118
Elasticity of substitution	9.88	33.19	1.06	1110.16	5,368
Canada-NAFTA					
Staging category	1.73	0.80	1.00	4.00	7,302
Import share	13.87	24.10	0.00	100.00	7,302
Export share	24.38	22.09	0.00	100.00	7,302
Log import/export ratio	-0.30	5.45	-33.60	34.91	7,302
GLI	0.31	0.32	0.00	1.00	7,302
Base rate	1.98	3.38	0.00	40.44	7,302
Homogeneous	0.42	0.49	0.00	1.00	7,302
Intermediate	0.53	0.50	0.00	1.00	7,302
Primary	0.06	0.23	0.00	1.00	7,302
Change import share	0.47	14.02	-100.00	100.00	7,302
Change export share	5.08	18.59	-100.00	100.00	7,302
Change unit values	0.33	28.27	-99.97	99.57	3,740
Elasticity of substitution	9.42	32.37	1.06	1110.16	5,555
Mexico-NAFTA					
Staging category	2.35	1.15	1.00	5.00	3,674
Import share	5.98	16.96	0.00	100.00	3,674
Export share	12.67	17.87	0.00	100.00	3,674
Log import/export ratio	-0.05	6.14	-33.60	34.91	3,674
GLI	0.19	0.28	0.00	1.00	3,674
Base rate	6.89	8.18	0.00	80.87	3,674
Homogeneous	0.45	0.50	0.00	1.00	3,674
Intermediate	0.49	0.50	0.00	1.00	3,674
Primary	0.09	0.29	0.00	1.00	3,674
Change import share	0.33	9.39	-100.00	100.00	3,674
Change export share	-0.70	15.26	-100.00	100.00	3,674
Change unit values	-0.08	32.29	-99.39	99.45	1,022
Elasticity of substitution	9.91	35.92	1.10	1110.16	2,921

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

Table A.3: Descriptive statistics

	Canada CUSFTA (1)	Canada CUSFTA (2)	Canada NAFTA (3)	Canada NAFTA (4)	Mexico NAFTA (5)	Mexico NAFTA (6)
Import share	0.991[c]	0.988[c]	0.996[a]	0.992[b]	1.001	0.998
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
Export share	1.001	1.001	1.004[a]	1.004	0.991[c]	0.994
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
Log import/export ratio	0.949[c]	0.938[c]	0.972[c]	0.968	0.960[c]	0.971
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.03)
GLI	1.339[a]	1.589[b]	1.102	0.950	1.158	0.733
	(0.16)	(0.23)	(0.13)	(0.18)	(0.19)	(0.24)
Base rate	1.283[c]	1.310[c]	1.313[c]	1.367[c]	1.085[c]	1.039[a]
	(0.01)	(0.02)	(0.03)	(0.06)	(0.01)	(0.02)
Homogeneous	0.577[c]	0.506[c]	0.764[a]	0.626[a]	0.754[a]	0.391[a]
	(0.08)	(0.08)	(0.09)	(0.11)	(0.10)	(0.15)
Intermediate	0.600[c]	0.466[c]	0.925	0.665	0.570[b]	0.203[b]
	(0.08)	(0.09)	(0.09)	(0.14)	(0.10)	(0.11)
Primary	0.121[c]	0.067[c]	0.271[c]	0.084[c]	0.188[c]	0.008[b]
	(0.03)	(0.03)	(0.07)	(0.04)	(0.05)	(0.01)
Change import share			1.006[a]	1.008	1.001	1.002
			(0.00)	(0.00)	(0.00)	(0.01)
Change export share			1.000	0.999	1.005[a]	0.997
			(0.00)	(0.00)	(0.00)	(0.01)
Change unit values				1.001		1.000
				(0.00)		(0.00)
Elasticity of substitution		1.001		1.002[a]		1.001[a]
		(0.00)		(0.00)		(0.00)
Observations	7,118	5,368	7,302	3,522	3,674	976
Possible Total	7,827	7,827	8,023	8,023	4,111	4,111
SIC industry FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Ordered logistic regression results (odds ratios). Outcomes ranked from least to most restrictive staging category. Robust standard errors in parentheses, clustered at the HS8 level. Possible Total is number of products in staging schedule data after excluding GSP, Mixed and Missing categories (see Appendix Table A.1). [a] $p < 0.05$, [b] $p < 0.01$, [c] $p < 0.001$.

Table A.4: Determinants of staging outcomes

A.4 Supplemental tables for regression results

	Mexico								Canada							
	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
1989	-0.054	(0.11)	-0.158	(0.11)	-0.161	(0.11)	-0.150	(0.11)	0.014	(0.05)	-0.115 [a]	(0.05)	-0.130 [a]	(0.05)	-0.137 [a]	(0.06)
1990	-0.029	(0.10)	-0.030	(0.10)	-0.035	(0.10)	-0.056	(0.10)	-0.121 [a]	(0.05)	-0.168 [c]	(0.05)	-0.178 [c]	(0.05)	-0.175 [c]	(0.05)
1991	0.009	(0.09)	0.073	(0.09)	0.064	(0.09)	0.051	(0.10)	-0.067	(0.04)	-0.102 [b]	(0.04)	-0.113 [b]	(0.04)	-0.119 [b]	(0.04)
1992	0.000	(0.00)	0.000	(0.00)	0.000	(0.00)	0.000	(0.00)	0.000	(0.00)	0.000	(0.00)	0.000	(0.00)	0.000	(0.00)
1993	0.064	(0.09)	0.150	(0.08)	0.140	(0.08)	0.130	(0.09)	0.084	(0.04)	-0.029	(0.04)	-0.031	(0.04)	-0.029	(0.04)
1994	-0.039	(0.10)	0.120	(0.10)	0.121	(0.10)	0.120	(0.10)	-0.101 [a]	(0.05)	-0.056	(0.04)	-0.057	(0.04)	-0.059	(0.05)
1995	-0.122	(0.11)	0.090	(0.10)	0.101	(0.10)	0.076	(0.11)	-0.095	(0.05)	0.003	(0.05)	0.010	(0.05)	-0.010	(0.05)
1996	-0.099	(0.11)	0.155	(0.11)	0.158	(0.11)	0.129	(0.11)	0.051	(0.06)	0.146 [b]	(0.05)	0.148 [b]	(0.05)	0.135 [a]	(0.05)
1997	0.102	(0.12)	0.297 [b]	(0.11)	0.305 [b]	(0.11)	0.313 [b]	(0.11)	0.117	(0.06)	0.231 [c]	(0.05)	0.236 [c]	(0.05)	0.204 [c]	(0.06)
1998	0.126	(0.12)	0.239 [a]	(0.11)	0.259 [a]	(0.11)	0.240 [a]	(0.11)	0.118	(0.06)	0.275 [c]	(0.06)	0.288 [c]	(0.06)	0.243 [c]	(0.06)
1999	0.246 [a]	(0.12)	0.379 [c]	(0.11)	0.396 [c]	(0.11)	0.400 [c]	(0.11)	0.208 [c]	(0.06)	0.359 [c]	(0.06)	0.368 [c]	(0.06)	0.319 [c]	(0.06)
2000	0.222	(0.13)	0.427 [c]	(0.12)	0.441 [c]	(0.12)	0.463 [c]	(0.12)	0.185 [b]	(0.06)	0.318 [c]	(0.06)	0.336 [c]	(0.06)	0.311 [c]	(0.06)
2001	0.357 [b]	(0.13)	0.423 [c]	(0.12)	0.440 [c]	(0.12)	0.447 [c]	(0.12)	0.168 [a]	(0.07)	0.306 [c]	(0.06)	0.328 [c]	(0.06)	0.295 [c]	(0.06)
2002	0.192	(0.13)	0.325 [b]	(0.12)	0.334 [b]	(0.12)	0.399 [c]	(0.12)	0.072	(0.07)	0.264 [c]	(0.06)	0.285 [c]	(0.06)	0.270 [c]	(0.06)
2003	0.360 [b]	(0.13)	0.517 [c]	(0.12)	0.521 [c]	(0.12)	0.535 [c]	(0.12)	0.157 [a]	(0.07)	0.314 [c]	(0.06)	0.331 [c]	(0.06)	0.316 [c]	(0.06)
2004	0.326 [a]	(0.13)	0.515 [c]	(0.12)	0.518 [c]	(0.13)	0.570 [c]	(0.12)	0.205 [b]	(0.07)	0.336 [c]	(0.06)	0.352 [c]	(0.06)	0.282 [c]	(0.06)
2005	0.295 [a]	(0.13)	0.467 [c]	(0.13)	0.465 [c]	(0.13)	0.463 [c]	(0.12)	0.232 [b]	(0.07)	0.305 [c]	(0.07)	0.312 [c]	(0.07)	0.222 [c]	(0.06)
2006	0.327 [a]	(0.14)	0.430 [c]	(0.13)	0.426 [b]	(0.13)	0.406 [b]	(0.12)	0.250 [c]	(0.07)	0.259 [c]	(0.07)	0.261 [c]	(0.07)	0.169 [b]	(0.07)
2007	0.342 [a]	(0.14)	0.437 [b]	(0.13)	0.431 [b]	(0.13)	0.410 [b]	(0.13)	0.224 [b]	(0.08)	0.285 [c]	(0.07)	0.279 [c]	(0.07)	0.158 [a]	(0.07)
2008	0.312 [a]	(0.14)	0.443 [b]	(0.14)	0.433 [b]	(0.14)	0.410 [b]	(0.13)	0.184 [a]	(0.08)	0.226 [b]	(0.07)	0.219 [b]	(0.07)	0.102	(0.07)
2009	0.211	(0.15)	0.385 [b]	(0.14)	0.372 [b]	(0.14)	0.336 [b]	(0.13)	0.131	(0.08)	0.221 [b]	(0.07)	0.223 [b]	(0.07)	0.110	(0.07)
2010	0.234	(0.15)	0.443 [b]	(0.14)	0.436 [b]	(0.14)	0.414 [b]	(0.13)	0.222 [b]	(0.08)	0.299 [c]	(0.07)	0.301 [c]	(0.07)	0.169 [a]	(0.07)
2011	0.407 [b]	(0.15)	0.522 [c]	(0.14)	0.521 [c]	(0.14)	0.463 [c]	(0.14)	0.072	(0.08)	0.216 [b]	(0.07)	0.217 [b]	(0.07)	0.105	(0.07)
2012	0.403 [a]	(0.16)	0.415 [b]	(0.15)	0.421 [b]	(0.15)	0.347 [a]	(0.14)	0.109	(0.08)	0.241 [b]	(0.07)	0.248 [c]	(0.07)	0.120	(0.07)
2013	0.334 [a]	(0.15)	0.388 [b]	(0.14)	0.402 [b]	(0.14)	0.353 [b]	(0.13)	0.198 [a]	(0.08)	0.224 [b]	(0.08)	0.243 [b]	(0.08)	0.071	(0.07)
2014	0.303	(0.16)	0.425 [b]	(0.14)	0.440 [b]	(0.14)	0.400 [b]	(0.13)	0.121	(0.08)	0.209 [b]	(0.08)	0.226 [b]	(0.08)	0.086	(0.07)
2015	0.398 [b]	(0.15)	0.437 [b]	(0.14)	0.460 [b]	(0.14)	0.438 [b]	(0.13)	0.126	(0.08)	0.248 [b]	(0.08)	0.272 [c]	(0.08)	0.141	(0.07)
2016	0.327 [a]	(0.16)	0.446 [b]	(0.14)	0.473 [b]	(0.14)	0.436 [b]	(0.14)	0.225 [b]	(0.08)	0.254 [c]	(0.08)	0.283 [c]	(0.08)	0.134	(0.07)
Observations	2,801,991		2,717,217		2,716,953		2,708,951		2,677,264		2,597,551		2,597,297		2,589,435	
Adjusted R^2	0.014		0.725		0.737		0.760		0.033		0.730		0.742		0.763	
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 the $Post_t$ dummy with a vector of year dummies $\mathbf{Year}_t = (1989, 1990, 1991, 1993, \ldots, 2016)$, DDD point estimates in columns (1) and (5) and columns (4) and (8) correspond to, respectively, equations (4) and (5). 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 A.5: 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	-0.041	(0.18)	-0.479 ^{b}	(0.17)	-0.159	(0.17)	-0.122	(0.11)	0.002	(0.06)	-0.353 ^{c}	(0.08)	-0.613 ^{a}	(0.28)
1990	0.077	(0.16)	-0.161	(0.16)	-0.051	(0.16)	-0.066	(0.10)	-0.127 ^{a}	(0.06)	-0.236 ^{c}	(0.07)	-0.714 ^{c}	(0.20)
1991	0.145	(0.15)	-0.011	(0.14)	0.031	(0.14)	0.050	(0.10)	-0.074	(0.05)	-0.177 ^{b}	(0.06)	-0.468 ^{a}	(0.19)
1993	0.280 ^{a}	(0.14)	0.101	(0.14)	0.258	(0.14)	0.087	(0.09)	-0.042	(0.05)	-0.001	(0.06)	-0.207	(0.19)
1994	0.371 ^{a}	(0.16)	0.288	(0.15)	-0.088	(0.15)	0.095	(0.11)	-0.051	(0.05)	-0.061	(0.07)	-0.297	(0.19)
1995	0.394 ^{a}	(0.17)	0.360 ^{a}	(0.16)	0.045	(0.16)	-0.008	(0.11)	-0.059	(0.06)	0.085	(0.07)	-0.335	(0.19)
1996	0.372 ^{a}	(0.18)	0.664 ^{c}	(0.16)	0.078	(0.16)	0.002	(0.11)	0.072	(0.06)	0.261 ^{c}	(0.08)	-0.177	(0.20)
1997	0.692 ^{c}	(0.18)	0.819 ^{c}	(0.16)	0.144	(0.16)	0.192	(0.11)	0.114	(0.06)	0.365 ^{c}	(0.08)	0.116	(0.23)
1998	0.721 ^{c}	(0.19)	0.870 ^{c}	(0.16)	0.109	(0.17)	0.065	(0.11)	0.184 ^{b}	(0.06)	0.357 ^{c}	(0.08)	0.008	(0.26)
1999	0.771 ^{c}	(0.18)	1.113 ^{c}	(0.16)	0.258	(0.17)	0.223 ^{a}	(0.11)	0.210 ^{b}	(0.06)	0.522 ^{c}	(0.08)	-0.053	(0.25)
2000	0.741 ^{c}	(0.19)	1.226 ^{c}	(0.17)	0.290	(0.18)	0.296 ^{a}	(0.12)	0.223 ^{c}	(0.07)	0.480 ^{c}	(0.08)	-0.103	(0.25)
2001	0.837 ^{c}	(0.19)	1.178 ^{c}	(0.17)	0.243	(0.18)	0.280 ^{a}	(0.12)	0.254 ^{c}	(0.07)	0.385 ^{c}	(0.09)	-0.099	(0.26)
2002	0.791 ^{c}	(0.19)	1.156 ^{c}	(0.17)	0.211	(0.18)	0.220	(0.12)	0.202 ^{b}	(0.07)	0.404 ^{c}	(0.09)	-0.075	(0.26)
2003	0.851 ^{c}	(0.19)	1.152 ^{c}	(0.17)	0.522 ^{b}	(0.19)	0.367 ^{b}	(0.12)	0.201 ^{b}	(0.07)	0.535 ^{c}	(0.09)	-0.085	(0.24)
2004	0.871 ^{c}	(0.20)	1.165 ^{c}	(0.18)	0.372	(0.19)	0.445 ^{c}	(0.13)	0.221 ^{b}	(0.07)	0.405 ^{c}	(0.09)	-0.022	(0.29)
2005	0.686 ^{c}	(0.20)	0.950 ^{c}	(0.18)	0.325	(0.19)	0.369 ^{b}	(0.13)	0.200 ^{b}	(0.07)	0.263 ^{b}	(0.09)	0.211	(0.28)
2006	0.798 ^{c}	(0.21)	0.706 ^{c}	(0.19)	0.321	(0.19)	0.316 ^{a}	(0.13)	0.149 ^{a}	(0.07)	0.220 ^{a}	(0.09)	-0.105	(0.29)
2007	0.602 ^{b}	(0.21)	0.726 ^{c}	(0.19)	0.309	(0.20)	0.350 ^{b}	(0.13)	0.166 ^{a}	(0.07)	0.166	(0.10)	-0.272	(0.29)
2008	0.519 ^{a}	(0.21)	0.622 ^{b}	(0.19)	0.415 ^{a}	(0.20)	0.369 ^{b}	(0.13)	0.133	(0.08)	0.079	(0.10)	-0.519	(0.29)
2009	0.509 ^{a}	(0.21)	0.506 ^{b}	(0.19)	0.309	(0.20)	0.297 ^{a}	(0.13)	0.140	(0.08)	0.076	(0.10)	-0.320	(0.31)
2010	0.390	(0.21)	0.508 ^{a}	(0.20)	0.402	(0.21)	0.420 ^{b}	(0.14)	0.192 ^{a}	(0.08)	0.133	(0.10)	0.048	(0.30)
2011	0.590 ^{b}	(0.21)	0.465 ^{a}	(0.20)	0.468 ^{a}	(0.21)	0.462 ^{c}	(0.14)	0.126	(0.08)	0.091	(0.10)	-0.350	(0.31)
2012	0.492 ^{a}	(0.22)	0.274	(0.20)	0.393	(0.21)	0.350 ^{a}	(0.14)	0.151	(0.08)	0.086	(0.10)	-0.307	(0.32)
2013	0.427 ^{a}	(0.21)	0.326	(0.20)	0.551 ^{b}	(0.20)	0.326 ^{a}	(0.14)	0.131	(0.08)	-0.019	(0.10)	-0.438	(0.34)
2014	0.393	(0.21)	0.405 ^{a}	(0.19)	0.530 ^{b}	(0.20)	0.394 ^{b}	(0.14)	0.113	(0.08)	0.070	(0.10)	-0.625	(0.35)
2015	0.520 ^{a}	(0.21)	0.405 ^{a}	(0.20)	0.457 ^{a}	(0.21)	0.445 ^{b}	(0.14)	0.178 ^{a}	(0.08)	0.084	(0.10)	-0.241	(0.33)
2016	0.438 ^{a}	(0.22)	0.514 ^{b}	(0.20)	0.524 ^{a}	(0.20)	0.420 ^{b}	(0.14)	0.179 ^{a}	(0.08)	0.074	(0.10)	-0.422	(0.30)
Observations	2,708,951								2,589,435					
Adjusted R^2	0.760								0.763					
Country × Product FE	Yes								Yes					
Country × Year FE	Yes								Yes					
Product × Year FE	Yes								Yes					

Notes: DDD estimates correspond to equation (6). 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 A.6: DDD regression: time-varying, heterogeneous cumulative treatment effects

A.5 Supplemental figures for Canada

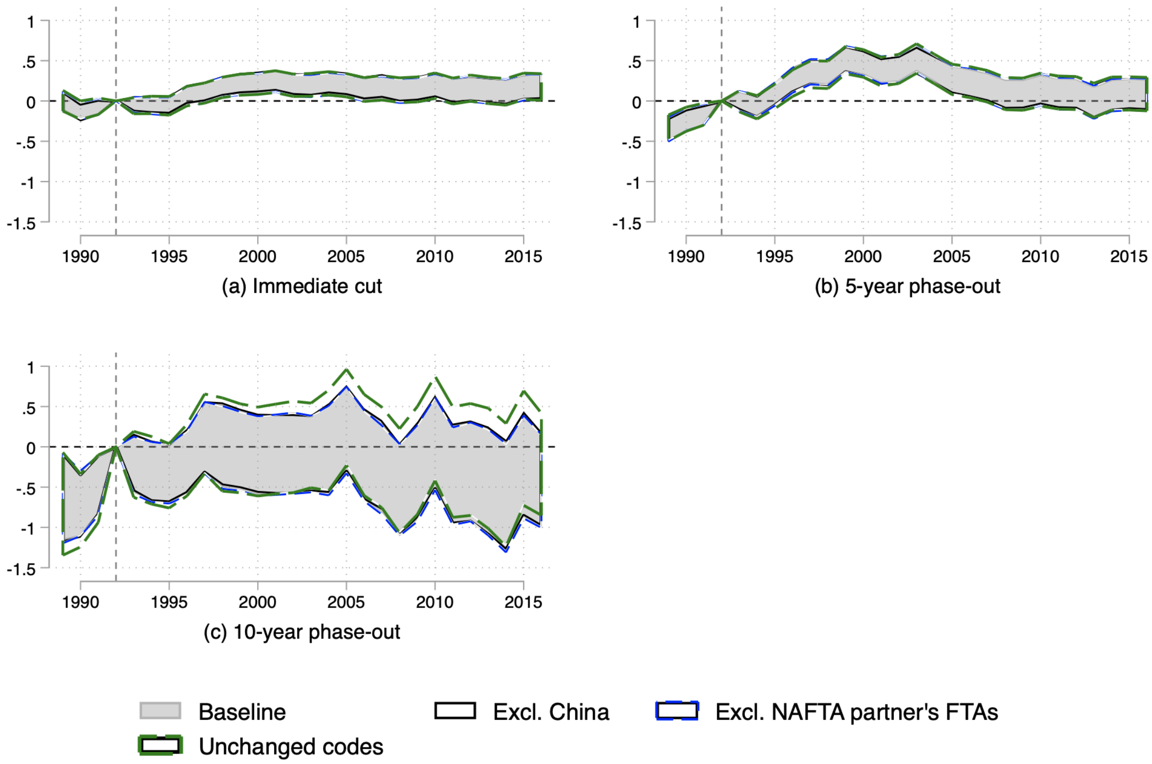

Figure A.1: Time-varying heterogeneous DDD estimates for Canada: Robustness.

Notes: DDD point estimates correspond to equation (6). 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 sub-sample where product code remains unchanged over sample period.

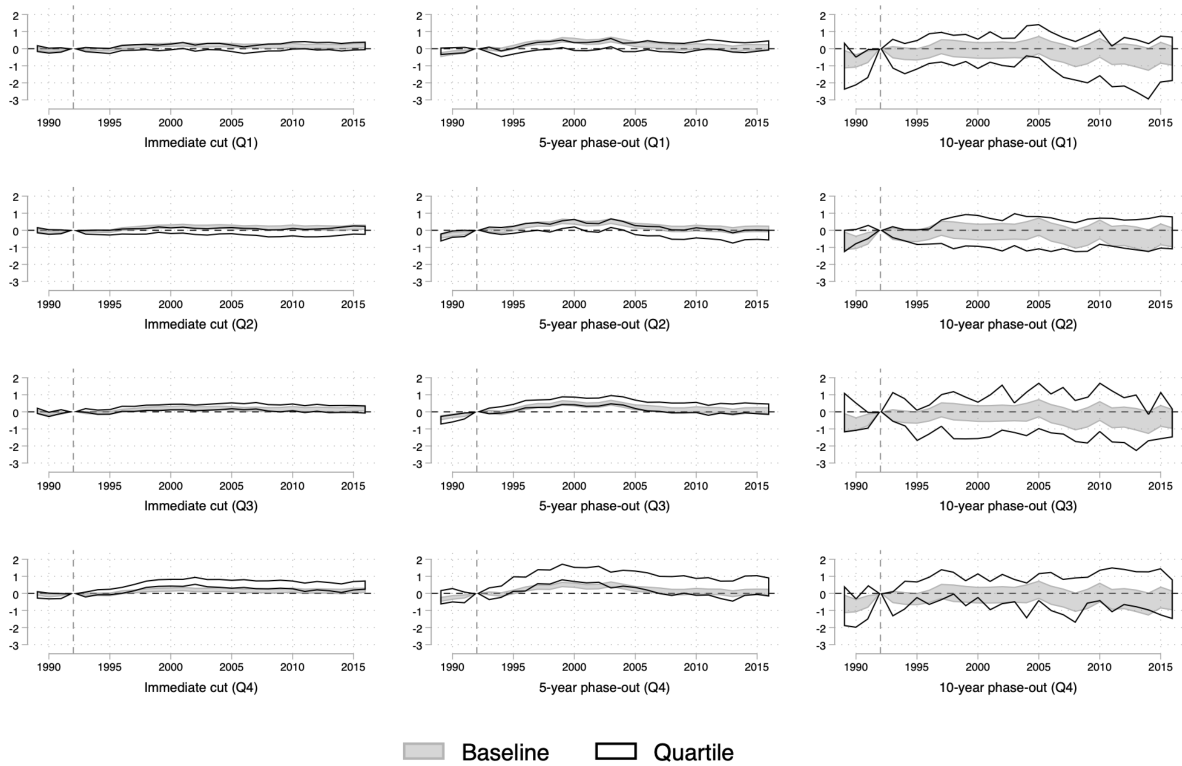

Figure A.2: Time-varying heterogeneous DDD estimates for Canada: Tariff cuts.

Notes: DDD point estimates correspond to equation (6). Plots represent 95% confidence intervals. Two-way clustered standard errors are used, clustering on both country-year and product-year. Plots by staging-category-specific quartiles of 1989 MFN tariff distribution.

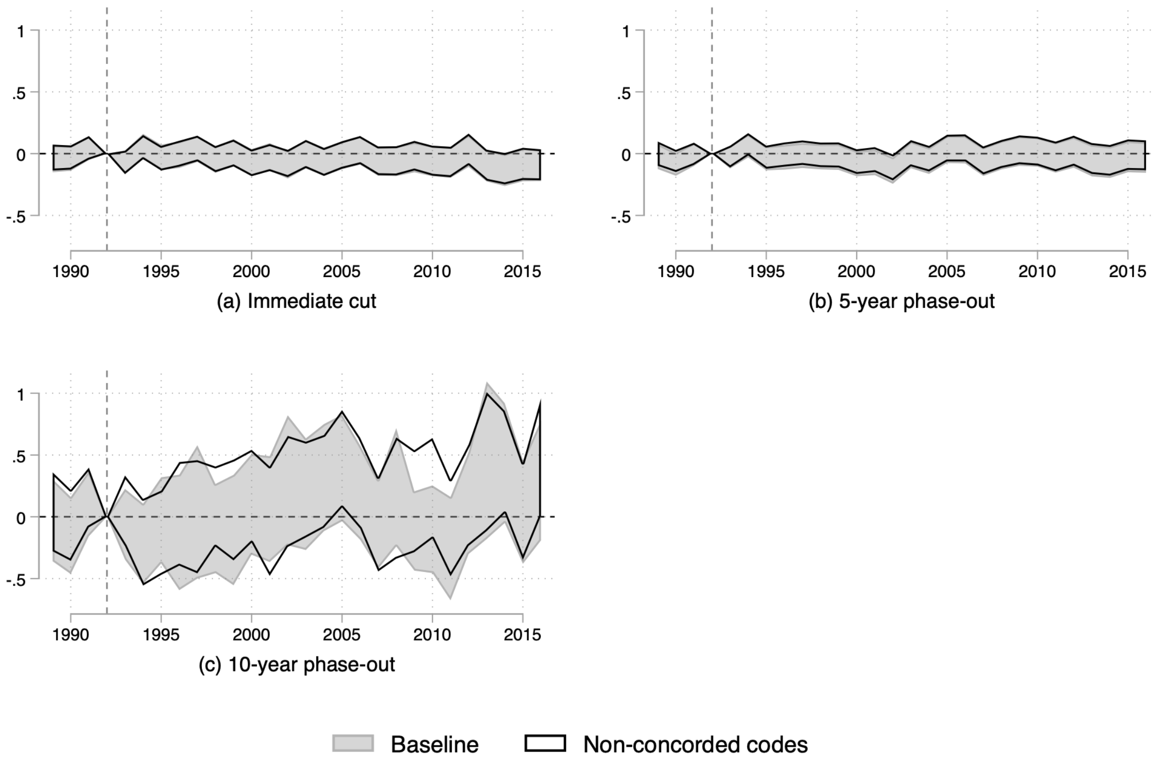

Figure A.3: Time-varying heterogeneous DDD estimates for Canada: Unit values.

Notes: DDD point estimates correspond to equation (6) with log real 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.