How Does Trade Respond to Anticipated Tariff Changes? Evidence from NAFTA*

Shafaat Y. Khan† and Armen Khederlarian‡

First Draft: September, 2018

This Draft: November 18, 2019

Click here for the latest version

Abstract

We study how anticipation to policy changes affects the estimate of the elasticity of substitution, the most important parameter in international trade. Standard identification of this parameter uses the tariff variation from the Free Trade Agreements (FTA) and assumes that the trade flows equal their consumption. However, FTAs eliminate initial tariffs through announced phaseouts over multiple years. Firms respond to these future policy changes by delaying their purchases until the tariff cut is effective while consuming past purchases held in inventories. These anticipatory dynamics bias the elasticity of substitution as imports diverge from their consumption. We document that around NAFTA's staged tariff reductions, imports experienced sizable *anticipatory slumps* followed by *liberalization bumps*. A trade model with inventories replicates these dynamics and illustrates that consumption of imports provides unbiased estimates of the elasticity of substitution. To overcome the lack of consumption data, we propose an empirical measure of consumed imports, which eliminates the bias in the model simulations. Application to the data shows that using import flows instead of consumed imports overestimates the annual elasticity by 68% and the average elasticity of substitution by 16%. The use of consumed imports increases the ratio of long-run to short-run response from 2 to 3.5.

JEL Classifications: F12, F13, F14

Keywords: Anticipation, Elasticity of Substitution, Inventories, NAFTA, Phaseouts

*We are greatly indebted to our advisor George Alessandria for invaluable discussions, guidance, and support throughout this project. We also are extremely grateful to Mark Bils, Yan Bai and Dan Lu for numerous comments and suggestions. Moreover, we would like to thank Mario Crucini, Alexandre Skiba, B. Ravikumar, George Schaur, Mostafa Beshkar, Ahmad Lashkaripour, Russel Hilberry, Michael Waugh, John Romalis, Lorenzo Caliendo, Kei Mu Yi, Carter Mix, Oscar Avila and seminar participants at the University of Rochester and conference participants at the ETSG 2018 (Warsaw), the Spring 2019 Midwest Macroeconomics Meetings (UGA), the Spring 2019 Midwest International Trade Meetings (IUB), the 2019 Economics Graduate Student's Conference (WashU) and Lisbon Meetings in Game Theory & Applications (LSEM).

†Shafaat Yar Khan, shafaatyar.k@gmail.com, University of Rochester

‡Armen Khederlarian, armen.khederlarian@gmail.com, University of Rochester

1 Introduction

Most trade policy changes are announced before their implementation giving firms the opportunity to shift their purchases to periods with lower costs.[1] In particular, Free Trade Agreements (FTAs) are usually put into effect gradually with scheduled phaseouts of the initial tariffs. In the case of the North American Free Trade Agreement (NAFTA), 96% of the tariff reductions were known at least 1 year in advance. We build on the insight that knowing future tariffs provides firms with incentives to delay their purchases until the reduction is effective. In the meantime, firms satisfy their demand by running down their inventories. Once the tariff reduction sets in, firms replenish their inventories by importing a large amount. This creates a gap between the import flows and the consumption of imported goods. The elasticity that is critical for welfare and policy analyses is informed by the latter. Because standard estimation approaches use variation in import flows and tariffs from the FTAs, these anticipatory dynamics are a potential source of bias.

We first document that in the early years of NAFTA, US imports from Mexico dropped strongly in anticipation and overshot sharply right after the tariff reduction. Secondly, we propose a measure for the consumption of imports that estimates elasticities in the presence of these anticipatory dynamics. Thirdly, we apply our measure to the data and find that the documented anticipation causes substantial upward biases in the estimates of the average and the annual elasticity of substitution.

Most approaches to estimating the elasticity of substitution from changes in tariffs build on static models of trade. In doing so, two key assumptions are made. First, the

[1] Recently there have been numerous examples of anticipated trade policy changes. Imports of solar panels soared in the three months before the tariff safeguards took effect (Wall Street Journal Jan 03, 2018). Similarly, Iran stockpiled oil tankers during the six months before the actual lifting of export sanctions in January 2016 (The Economist Jan 23,2016).

tariff changes from FTAs are not anticipated. In fact, tariff eliminations from FTAs take multiple years after the agreements come into force. This gives firms ample time to anticipate them. Secondly, the goods are not storable and hence they are consumed in the period in which they are imported. However, in response to aggregate shocks, models in which imports crossing the border and their consumption diverge have been very successful in accounting for the observed trade dynamics.[2] In this paper, we relax these assumptions by using the estimation approach from a model in which firms incorporate the knowledge of future tariff cuts and in which goods are stored as inventories. We find that relaxing these assumptions reveals sizeable biases in the elasticity of substitution estimated using the standard approach.

In a trade model with inventories, anticipated tariff reductions lead to import declines before staged tariff reductions — '*anticipatory slumps*' — and a subsequent import rebound — '*liberalization bumps*'. During this period imports diverge from their consumption as importers run down their inventories. The ability of firms to make use of announced tariff reductions hinges on the degree to which the good is held in inventories. Only if firms hold inventories, are they able to gain from delaying input purchases without disrupting their operations. Through the model simulations we show that when this is the case, these anticipatory dynamics lead to biases in the elasticity of substitution when using import flows instead of their consumption. While the import flows display *slumps* and *bumps* around the time of the reduction, consumed imports have a smoother path. We find that consumption of imports identifies the deep preference parameter of the elasticity of substitution regardless of the anticipated nature of shock and the storability of the good.

The implementation of NAFTA with gradual tariff phaseouts provides a clear case of

[2] See Alessandria et al. (2010a,b), Charnavoki (2017)

anticipated trade policy changes. Among the goods that were not already tariff-free before NAFTA, 75% were scheduled to be phased-out gradually over up to 15 years.[3] While there was little time to anticipate the first round of tariff reductions,[4] in the following years, all the phaseout stages were scheduled to take place on January 1$^{\text{st}}$ every year, giving importers sufficient time and certainty to anticipate them. Because elasticities are generally estimated at the annual calendar frequency, the existence of anticipatory *slumps* and liberalization *bumps* around the beginning of the year would cause the trade elasticity (import flows) to exceed the elasticity of substitution (consumption of imports).

Indeed, we find that the data on US imports from Mexico contains strong evidence of anticipation to tariff cuts for the phased-out goods. We estimate anticipatory elasticities by considering within year growth rates of import flows in response to upcoming tariff changes. We use a standard double-difference approach to construct trade flows free of exporter- and importer-specific factors (Romalis (2007), Head and Ries (2001), Caliendo and Parro (2015)). We find that in anticipation of an upcoming tariff reduction of 1 percentage point (pp) imports experienced a sizeable 6% anticipatory trade *slump* in November and December relative to the middle of the year. On the other hand, we find a notable liberalization *bump* of 12% in imports in the first few months after the tariff reduction. More importantly, we find that these results are increasing in the storability of the goods. These results stand in contrast with the standard identification assumptions from static models of trade. The dynamics of trade flows around anticipated tariff change along with the model insights indicate that the identification of the elasticity of substitution requires the usage of import consumption data.

Unfortunately the consumption of imports is not observed, especially at the level of

[3]The original text of the agreement specified the exact schedule of tariff phaseouts at HS-8 product level, including the number of stages, the size and the date of each staged tariff reduction.

[4]NAFTA was ratified by the US Congress only 40 days before coming into force on Jan 1$^{\text{st}}$, 1994.

aggregation used in the estimation of the elasticity of substitution. To overcome this lack of data, we introduce a measure for the consumption of imports. The key ingredients of the consumption measure are high-frequency monthly trade data and inventory-sales ratios. In particular, a fraction of current inventory holdings is assumed to be used for consumption. The inventory-sales ratio is obtained at the product, source and destination level by building on the relationship between the inventory-sales ratio and the lumpiness of imports. Hence the process for consumption of imports only requires monthly trade flows and can be implemented at any aggregation level. We test the validity of the measure by demonstrating that (1) it is as effective as the consumption of imports in the model simulations, and (2) it eliminates the anticipation documented in the data for NAFTA's phased-out goods.

Next, we apply our measure of consumed imports to estimate the elasticity of substitution. We estimate the average (cross-sectional) elasticity as well as the dynamic short-run and long-run elasticities. Using our measure of consumption of imports instead of actual imports, we find a significantly lower average elasticity. The difference is driven by the phased-out goods, for which we documented significant anticipatory *slumps* and liberalization *bumps*. When we use consumption of imports instead of actual imports, the aggregate static elasticity of substitution across different country-varieties drops by 16% from 8.9 to 7.7. For the phased-out goods in particular, we find a bigger drop of 21%. We also explore how these biases are related to the storability of the good. We find that for the more storable goods, the estimates contain a bias of around 28% as their elasticity falls from 10.4 with the import flows to 8.2 with the consumed imports. The stronger biases for the goods which contained significant anticipatory dynamics highlight the role of the mechanism proposed in this paper.

We also correct for the biases in the estimates of elasticities at different time horizons. We apply an Error-Correction Mechanism (ECM) model to estimate the long- and short-run elasticities. The short-run (annual) elasticity based on the variation in import flows is estimated to be 4.2. Whereas when we use the consumption of imports, the short-run elasticity falls to 2.5, eliminating a bias of 68%. This bias is again driven by the goods which experienced gradual phaseouts. We also find that the anticipatory dynamics overestimate the short-run elasticity by more than its long-run counterpart as the biases in the long-run elasticity are negligible. This generates an understatement of the dynamic adjustment of consumed imports. Using the consumption of imports instead of the import flows, we find that the ratio of long- to short-run elasticity rises from 2 to 3.5. While the anticipatory *slump* is always present, the liberalization *bump* after the tariff reduction loses its importance as one considers response at longer horizon. This leads to less contamination of the long-run estimate. Since the average elasticity estimate is based on a pooled-cross-section, it contains *slumps* and *bumps* for a larger number of observations. Nevertheless, the use of consumption of imports removes the bias emanating from the anticipatory behavior present in the trade flows.

The paper is organized as follows. Section 2 illustrates why standard estimation approaches may fail to identify the elasticity of substitution when using imports instead of their consumption. In section 3 we briefly describe why NAFTA represents an ideal setting for this study and document that for the goods that provide most variation in tariffs - phased-out goods - anticipation was strong. In section 4 we define our measure of consumed imports and show that using actual imports instead of their consumption produces substantial upward biases in estimates of the elasticity of substitution. In the final section, we conclude.

Related Literature

The elasticity of substitution is a crucial parameter in computable general equilibrium models of international trade given its implication on the evaluation of policy experiments and welfare gains from trade (Arkolakis et al. (2012), Alessandria et al. (2018)). However as a result of different estimation techniques (Hillbery and Hummels (2012)), samples and response horizons, the range of estimates in the literature is wide[5]. Our paper points to a potential upward bias in estimation methods that exploit cross-sectional and time variation in tariffs. These approaches usually rely on the structure of static models of trade and use annual trade flows[6]. In the upper bound of estimates, using panel data from NAFTA's trade liberalization, Romalis (2007) and Head and Ries (2001) obtain estimates between 7 and 11. In the lower bound, using a single year and cross-sectional data, Simonovska and Waugh (2014) and Caliendo and Parro (2015) situate the estimate between 4 and 5. The bias we document in this paper results in lower estimates from methods used in the former, but might also be a source of bias in the latter.

In a recent paper, Fajgelbaum et al. (2019) examine anticipation to tariff increases by the current US administration, and find significant increases in trade before their implementation. However, most studies of anticipation to policy changes have focused on consumption taxes. Baker et al. (2018) show that households stockpile in anticipation of future sales tax rate rises and propose an inventory mechanism that rationalizes this pattern. Coglianese et al. (2017) document that consumers and distributors of gasoline increase their purchases before the implementation of tax increases. As in our paper, these anticipatory dynamics bias estimates of short-run price elasticities of gasoline de-

[5]Anderson and van Wincoop (2004) situate it between five and ten

[6]An exception to using annual trade flow is Gallaway et al. (2003) who estimate short-run and long-run elasticities using prices at quarterly frequency.

mand when using tax changes as instruments. Coglianese et al. (2017) control the bias by introducing leads and lags of tax rates. In contrast, we propose a measure of the consumption process to overcome the bias.

There is a large literature that explains high-frequency dynamics of macroeconomic and international trade variables using inventories. In Alessandria et al. (2010a), inventory management allows to reproduce the sudden drop in trade during the Great Recession, while in Alessandria et al. (2010b) it captures implosions of trade and pricing dynamics of retail goods following large devaluations. In Bekes et al. (2017) demand volatility raises the motive for precautionary inventory holdings and explains variation in trade lumpiness across French exporter markets. These papers as well as ours build on fixed ordering costs, that have been widely documented (Kropf and Saure (2014), Blum et al. (2019)). Closest to the present paper is Alessandria et al. (2011). Using detailed inventory of the US auto industry, Alessandria et al. (2011) find that accounting for changes in inventory adjustments lowers price elasticity estimates, especially in the short-run.

Finally, our paper is related to the strand of literature that studies the economics of tariff phaseouts. Gradualism in trade liberalizations has a long tradition (see Staiger (1995)). Kowalczyk and Davis (1998) document that US tariffs and little intra-industry trade are associated with long NAFTA phaseout periods for US imports from Mexico. In a recent paper, Besedes et al. (2018) show that imports of NAFTA's phased-out goods experienced similar long-run growth pattern as non-phased-out goods. This contrasts with the popular argument that phaseouts explain the delayed effects of preferential FTAs (Baier and Bergstrand (2007)).

2 Mechanism

In static models of trade, imports and their consumption are identical. Because goods are assumed to be non-storable, imports only respond to changes in tariffs once they are effective. However, in response to aggregate shocks imports and their consumption can diverge significantly in the short-run as AKM (2010, 2011) and others, have documented rigorously. In particular, trade models including inventories have been very successful in accounting for these dynamics. In this section, simulations of the $(\underline{s},\overline{s})$ inventory management model illustrate that, when tariff changes are anticipated and trade is intensive in inventories, imports drop in advance of anticipated tariff reductions and spike thereafter. During these brief periods imports and their consumption deviate, leading to an overestimated elasticity of substitution. We first describe how in static models of trade, difference-in-difference approaches identify the elasticity of substitution. Secondly, we lay out the $(\underline{s},\overline{s})$ model in which goods are storable allowing imports to become lumpy. We then use simulations to show when the standard estimation approaches of the elasticity of substitution yield biases.

2.1 Static Models of Trade

Most approaches to the estimation of the elasticity of substitution are founded on an import demand equation that can be expressed as a gravity equation, where trade is a function of bilateral trade costs and source and destination specific aggregate variables.[7] Under such structure and some assumptions, taking a difference-in-difference approach identifies the elasticity of substitution. One assumption that has been widely overlooked in the literature is that imports are immediately consumed. When using variation in

[7]See for example Romalis (2007), Caliendo and Parro (2015), Head and Ries (2001), Feenstra et al. (2004), Gallaway et al. (2003), etc.

tariffs in the context of FTAs, this assumption is likely to fail if importers internalize upcoming tariff reduction and adjust their ordering accordingly.

To illustrate how static models of trade identify the elasticity of substitution, without loss of generality we focus on an import demand from the Dixit-Stiglitz structure of CES preferences. According to this formulation, importer i's demand in value for good z from country c is expressed as:

$$v_{iczt} = p_{czt} \left(\frac{(1+\tau_{iczt})p_{czt}}{P_{izt}} \right)^{-\sigma} C_{izt} \tag{1}$$

where τ_{iczt} is the applied tariff levied on the imports of good z sourced from country c by country i. The exporter-specific price of the good is denoted by p_{czt}. P_{izt} and C_{izt} are country i's price index and total consumption of good z. Our aim is to estimate the elasticity of substitution across different varieties, given by σ. This preference parameter indicates the degree of differentiation across varieties from different countries. In appendix B, we discuss the underlying preference structure behind this formulation.

This demand equation contains exporter-specific (unit-price) term and a few importer-specific terms (demand) terms. To control for the effects on imports coming from these terms, we will take differences of (1) with control country groups. The following expression of imports is obtained by taking a double-difference of (1) with respect to reference importer (i') and exporter (c'), denoted by m^{DD}_{zt}: [8,9]

$$\underbrace{\ln \left(\frac{v_{iczt}}{v_{ic'zt}} \Big/ \frac{v_{i'czt}}{v_{i'c'zt}} \right)}_{m^{DD}_{zt}} = -\,\sigma \ln \left(\frac{1+\tau_{iczt}}{1+\tau_{ic'zt}} \Big/ \frac{1+\tau_{i'czt}}{1+\tau_{i'c'zt}} \right) \tag{2}$$

[8]This is the dependent variable in Romalis (2007).

[9]See appendix B in the Appendix for the derivation.

In (2), the elasticity of substitution across varieties is identified through variation in tariff levels after eliminating effects from product prices and aggregate terms.

Two key simplifications of the demand structure above might lead to biases when estimating the elasticity of substitution. First, importers don't act on future tariff reductions. Second, all imports are used for consumption when they cross the border because they are non-storable. We argue that these simplifications are empirically consequential in the context of FTAs with scheduled phaseouts. This is because (1) FTAs generally schedule tariff reductions to become effective on the 1st of January and (2) trade flows are commonly aggregated at the annual calendar frequency. If importers delay their purchases until the reductions are effective, then imports might deviate from their consumption, even at the annual frequency. In the model formulated below, the import demand equation is generalized to account for these two simplifications.

2.2 Inventory Trade Model with ($\underline{s}$,$\overline{s}$) Ordering

To illustrate how anticipated tariff reductions might bias elasticity estimates we generalize the model described above, separating imports from their consumption by including a ($\underline{s}$,$\overline{s}$) ordering policy. In particular we adopt AKM's partial-equilibrium model in which a homogeneous storable good is imported by a continuum monopolistically competitive retailers that face a CES demand and decide whether to import or not every period. In advance of an upcoming tariff reduction, importers delay their orders, but continue to satisfy their demand by running down their inventories.

Ordering implies a fixed shipment cost, causing firms to order infrequent but large shipments. On top of the fixed cost, retailers face demand uncertainty and a one period delivery lag, leading to precautionary inventory holdings. Under this setup, retailers run down their stocks to $\underline{s}$ and then replenish it up to $\overline{s}$. Retailers are identical except for their history of demand shocks, that determines their current inventory holdings. Let $p_{j,t}$ denote the importer j specific retail prices and $\nu_{j,t}$ the demand shock in period t. Importer j faces the following demand for its variety:

$$c_{j,t} = e^{\nu_{j,t}} p_{j,t}^{-\sigma} \tag{3}$$

The variable cost of importing is $\omega_t = \omega(1 + \tau_t)$, common across all importers. We assume that exporters are perfectly competitive, so that the pass-through (to retailers) of the tariff reduction is complete. At the beginning of each period retailers observe their inventory holdings, $s_{j,t}$ and their demand shock, $\nu_{j,t} \sim^{iid} N(0, \sigma_\nu^2)$, assumed to be i.i.d. across firms and time[10], and then price their good and decide to import or not. To import, retailers need to pay a fixed cost f. We assume that importing is irreversible, $m_{j,t} \geq 0$. Because of demand uncertainty, importers will never run down their inventories to zero, $\underline{s} > 0$, and because of a one period delivery lag, sales can never exceed current inventory holdings:

$$q_{j,t} = \min[e^{\nu_{j,t}} p_{j,t}^{-\sigma}, s_{j,t}] \tag{4}$$

Assuming the goods in transit ($m_{j,t}$) depreciate at the same rate, δ, as in the warehouse, the law of motion for the inventories is:

$$s_{j,t+1} = (1 - \delta)[s_{j,t} + m_{j,t} - q_{j,t}] \tag{5}$$

Denote the firm's value of adjusting inventories by $V^a(s, \nu)$ and of not adjusting by

[10]The iid demand shock is necessary to obtain variation in the anticipation to a tariff reduction. With perfectly correlated demand shocks all firms would respond equally to the incentives of anticipating the demand shock.

$V^n(s,\nu)$. Every period retailers optimize by choosing $V(s,\nu) = \max[V^a(s,\nu), V^n(s,\nu)]$, where

$$V^a(s,\nu) = \max_{p,m>0} q(p,s,\nu)p - \omega m - f + \beta EV[s',\nu'] \tag{6}$$

$$V^n(s,\nu) = \max_{p} q(p,s,\nu)p + \beta EV[s',\nu']$$

are subject to (5) and (4). Solving for the optimal policies generates an ($\underline{s}$,$\overline{s}$) policy of ordering and the orders are a function of current inventory holdings and the demand shock i.e. $m = m(s,\nu)$. Similarly, the pricing schedule is characterized by a constant markup over the marginal value of an additional unit of inventory, $p = \frac{\theta}{\theta-1}V_s(s,\nu)$. In contrast with static models, demand for the good, $q_{j,t}$, can be satisfied using inventories. Moreover, because firms optimize the timing of their purchases, $m_{j,t}$, responds to the incentives of future price declines.

When facing an upcoming decline of the variable cost of importing, retailers face the following trade-offs. On the one hand, goods are cheaper in the future and the ordering cost is delayed. On the other hand, because inventories are declining as demand continues to be satisfied, the marginal value of inventories increases. Hence, the price increases and some sales are lost. Additionally, because inventories move away from their stationary level, higher inventory holding costs are incurred over a brief period. In the simulations we show that the former outweigh the latter and that for reasonable intensity in inventory holdings, anticipated tariff reductions lead to reversals of trade flows[11].

[11] In section D of the Appendix we show that for an individual retailer that does not face demand uncertainty there is a closed form expression of the anticipatory drop in imports. Indeed, the drop is increasing in the equilibrium inventory-sales ratio and the upcoming tariff cut.

2.3 Simulations

We simulate the model described above to illustrate the deviation between the trade elasticity and the elasticity of substitution. We show that when (1) imports are intensively held as inventories, and (2) tariff changes are anticipated, trade flows *slump* and *bump* around the time of the tariff change. This happens while the consumption of imports remains relatively constant. This creates a bias if one uses the trade elasticity as a meaure of the underlying preference parameter of the elasticity of substitution.

2.3.1 Calibration

We calibrate the model at the monthly frequency by setting $\beta = 0.96^{1/12}$. Thereby the simulated time series is at the same frequency of standard customs data. The simulated dataset is generated by shocking firms with a one time anticipated decrease in τ and firm-specific iid taste shocks. We consider 5 possible tariff reductions of 0pp, 1pp, 2pp, 5pp and 10pp to get sufficient variation in our simulated dataset. To highlight the simplifications that jeopardize the identification using the standard approach we consider two alternative simulations in addition to the benchmark: (1) a calibration that yields a low inventory-sales ratio (or less storability) and (2) tariff changes that are unanticipated.

The parameters used in each of the three simulations are shown in Table 1. In our benchmark simulation the monthly I/S (inventory-sales ratio) is 2.54 and tariff reductions are anticipated. The I/S is jointly determined by a fixed cost that amounts to 3.60% over a firm's mean monthly revenues and a relatively low depreciation rate of 2.5% per month or around 30% per year. In contrast, in the low I/S simulation, the fixed cost is negligible at 0.37% over mean firm revenues and the good perishes faster, at 10% monthly, so that firms order almost every period. The simulation of unanticipated tariff changes is calibrated as the benchmark case, so that they are indistinct at their steady state. All other parameters are the same in the three simulations. The elasticity of substitution, our object of interest, is set to 4. We simulate 2,000 firms for each path of tariff changes.[12] The simulation is over 4 years and the tariff change takes place at the beginning of the third year.

Before reporting the results of estimating the elasticity of substitution with the simulated dataset, we discuss the dynamics around a tariff reduction. Figure 2 shows the aggregate response to a 1pp decrease in tariffs. The top left panel demonstrates that in the benchmark case (blue line) imports show strong anticipation to the upcoming tariffs cut, dropping by more than 20% in the month before the tariff reduction.[13] After the change has materialized, imports spike. In the low I/S and the unanticipated simulation, imports only respond after the change has taken place. The top right panel demonstrates that even though trade flows respond strongly in the benchmark case, consumption of imports only decreases slightly before the reduction. The tamed response of consumption of imports is possible because importers satisfy their input demand making use of their inventory holdings. The bottom right panel illustrates that the small dip in consumption of imports is due to a small increase in prices. This is because at low levels of inventory holdings, importers charge high stock-out prices. Overall, for the storable good facing anticipated tariff changes, we see a magnified response of imports that dwarfs its consumption response.

[12] The model is solved using a shooting algorithm where importers solve their problems backwards starting from a new steady state. We first simulate 2,000 firms over 5,000 periods using their initial policy functions so that we can initiate the distribution of firms at the stationary inventory holdings distribution.

[13] With 5 months until the tariff change, firms start synchronizing their orders with the goal of minimizing per-order fixed cost in the months before the change. These echo effects generate a short-lived spike in imports in months 7 and 8.

2.3.2 Bias in the Elasticity of Substitution

Standard estimation methods use annual trade flows and levels of tariffs. Therefore, we aggregate firm-level monthly flows of imports, m_{jt}, and their consumption, c_{jt}, to the annual frequency and estimate elasticities under the following two equations:

$$\ln(y_{jt}) = \sigma \ln(1 + \tau_t) + u_{jt} \tag{7}$$

$$\Delta \ln(y_{jt}) = \sigma \Delta \ln(1 + \tau_t) + u_{jt} \tag{8}$$

where the trade elasticity is estimated by putting $y_{jt} = m_{jt}$ and the elasticity of substitution is estimated by putting $y_{jt} = q_{jt}$. The average elasticity is estimated using (7) in levels and (8) estimates the annual elasticity by considering annual changes.

In Table 2 we report the results from (7) with imports, m_{jt}, and sales of imported goods, q_{jt}, as dependent variables. Column one and three provide the estimate when the years immediately before and right after the tariff reduction are neglected to exclude the anticipation dynamics.[14] During these years, for all the goods, the trade elasticity and the elasticity of substitution are nearly identical.[15] However, in most empirical implementations, and especially those using FTAs, elasticities are estimated over years in which tariffs are scheduled to change and trade flows around these changes are used. In column two and four we report the elasticity estimates when considering the two years around the tariff reduction. In the benchmark case, using imports in stead of their consumption overestimates the elasticity of substitution by 50%, compared to the true

[14]Note that this approach to side stepping the bias requires knowledge about the existence of anticipation. In contrast, using q_{jt} does not require information about the existence of anticipation.

[15]Notice that the estimated elasticity of substitution differs from σ in all three simulations. This is because under the pricing policy described in section 2, as the variable cost declines mean revenues increase and importers economize more efficiently on the fixed cost of ordering. Therefore, the pass-through of tariff on consumer prices is more than one. We abstract from this and focus on the bias that stems from misspecified dependent variable.

elasticity of substitution (column four). Instead, using consumption of imports closes the gap significantly. Moreover, in the other two simulations in which the source of bias is absent, the estimate is unaffected. In Table 3 we demonstrate that when estimating (8) the bias becomes even more severe when the good is storable and the tariff change is anticipated and this bias is overcome by using consumption of imports as the dependent variable.

These findings indicate that if firms anticipate upcoming tariff reductions then using imports instead of their consumption as the dependent variable leads to an overestimated elasticity of substitution, even when using annual flows. In the next section we study the empirical relevance of the mechanism by considering how US imports from Mexico show anticipatory behavior around NAFTA's scheduled tariff phaseouts.

3 Evidence of Anticipation during NAFTA

In the previous section we illustrated that when tariff reductions are anticipated, standard methods of estimating the elasticity of substitution may be biased because imports and their consumption diverge at the time of the tariff reduction. During NAFTA indeed the entire path of tariff phaseouts was public knowledge, providing importers with strong incentives to delay their purchases in advance of staged phaseouts. This section provides evidence that importers indeed responded to these incentives and that at product level anticipation was increasing in the degree of inventory holding. We first establish that during the NAFTA, phased-out tariffs were the main source of variation in tariff levels. Next, we lay out and implement our empirical approach to estimating anticipation. Finally, we show that anticipation was strongest for goods that are characterized by higher a inventory-sales ratio.

3.1 Background

NAFTA was signed by the three presidents of the US, Canada and Mexico in December 1992 after a lengthy negotiation. It took another year to be approved by legislators in the three countries. US Congress ratified it on November 20, 1993, and the US President Clinton signed it into law on December 8. NAFTA finally came into force on January 1, 1994. The agreement incorporated Mexico to the preferential FTA Canada and the US signed in 1988, and brought a major elimination of tariffs for trade with Mexico. By then, Canadian-US tariffs had mostly been removed. As a result of these tariff reductions, Mexico's share over total US imports almost doubled between 1994 and 1999, as shown in Figure A.1 of the Appendix. Because NAFTA came at a time when there were few preferential FTAs, its tariffs reductions undercut the Most-Favored-Nation (MFN) rates, giving place to variation in the tariff rates. For this reason, NAFTA has been widely studied to evaluate the gains from trade and, hence, to estimate the elasticity of substitution[16]. We argue that neglecting the anticipated nature of NAFTA's tariff reductions biases these elasticity estimates upwards.

Some of the tariff cuts took place immediately on January 1st of 1994 and the rest were scheduled to be phased out gradually over various stages of tariff reductions. Step reductions took place on January 1st for up to 15 years. In fact, 96% of NAFTA's tariff reductions were known at least one year in advance since NAFTA's original text specified the phaseout schedules of all goods at HS-8 level. Broadly, goods were classified into 5 categories, with class A to be immediately zeroed, classes B, C and C+ to be phased out over 5, 10 and 15 years and class D, goods with zero tariffs before NAFTA, to remain at zero[17]. As you can see in Table 4, most of the HS-8 goods were already zeroed by 1993,

[16] See for example Kruegger (1999), Head and Ries (2001), Fukao et al. (2003), Romalis (2007), Kehoe and Ruhl (2013), Caliendo and Parro (2015)

[17] Although most tariffs were eliminated in equal annual stages, some schedules differed in the number

providing little variation to estimate elasticities. Among those that were non-zero by 1993, around three quarters were scheduled to be phased out, while the rest became tariff-free the day NAFTA came into force. Given that there was considerable uncertainty about if NAFTA would become effective[18] it is unclear that the latter reductions could have been anticipated. But once signed, scheduled reductions of class B, C and C+ goods were certainly anticipated to become effective[19]. Before we document that indeed importers internalized this knowledge and anticipated upcoming tariff reductions during the early years of NAFTA, we describe our methodological approach to identifying anticipation in the next subsection.

3.2 Data

Our estimation of anticipation to tariff reductions focuses on US imports from Mexico between 1990 and 1999. To study the response of trade flows to scheduled tariff phaseouts we use trade flows at the monthly frequency and expand the estimation strategy described in the section 2.1. In particular, we take changes over the sub-periods in a year of the (double-differenced) trade flows. If upcoming tariff reductions were anticipated as predicted by the import demand of the subsection 2.2, then imports in the months before the reduction should drop in comparison to the earlier months of the same year. We would also expect to see a sharp increase in the imports subsequent to the tariff reduction, in line with the stocking-up effect.

After taking the double-difference of the standard import demand equation in section

of years with rate reductions and the size of each reduction. Because of this and different amount of goods per phaseout class, the median scheduled tariff rates declined non-monotonically during the 1990s as you can see in Figure A.2 of the Appendix.

[18]Given how the voting went, there was little political agreement in the US on NAFTA

[19]We cannot ascertain at which degree agents internalized these scheduled reduction. However, a correlation of 87.42% between scheduled rates and applied duties at HS-8 level for goods that were phased out indicates that the scheduled reductions indeed materialized.

2.1, we were left with the following expression:

$$\underbrace{\ln\left(\frac{v_{iczt}}{v_{ic'zt}}\Big/\frac{v_{i'czt}}{v_{i'c'zt}}\right)}_{m_{zt}^{DD}} = -\,\sigma\,\underbrace{\ln\left(\frac{1+\tau_{iczt}}{1+\tau_{ic'zt}}\Big/\frac{1+\tau_{i'czt}}{1+\tau_{i'c'zt}}\right)}_{\tau_{z,t}^{DD}} \tag{2}$$

In the implementation of the double-difference approach we follow Romalis (2007). This allows us to map the anticipation demonstrated later in this section to the biased elasticity of substitution that we document in the next section. The reference importer i' is chosen such that exporters c and c' face the same tariffs in destination i' at any period t, i.e. $\tau_{i'czt} = \tau_{i'c'zt}\ \forall z, t$. Given the four directions of trade flows and the choice of reference importers, the only relevant tariff changes are the ones pertaining to importer i. In line with this, the reference importer is an aggregate of 12 countries of the European Union (EU12).[20] These countries did not have a FTA with Mexico, nor with any of the reference exporters during our sample period. The choice of EU12 restricts our sample period up until 1999, since in 2000 the European Union and Mexico signed a FTA.

In addition to this, the choice of reference groups is also dictated by the need of large country groups. This helps in countering the issue of trade lumpiness that becomes pervasive at high frequency and disaggregate HS-6 product categories.[21] For this reason, the reference exporter is an aggregate of 137 countries - rest of the world (RoW) - with which neither the US nor the EU12 established a preferential FTA during 1990-1999.[22]

We obtain imports at HS-6 level and monthly frequency. We consider CIF (Customs, Insurance and Freight) imports for consumption which disregards imports for re-export

[20]These are Austria, Belgium, France, Finland, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Spain, and Portugal.

[21]Because our identification of the anticipatory effects rely on the variation in monthly trade flows, the relevant observations are those HS-6 with non-zero consecutive trade flows between monthly periods of the 4 directions to obtain growth rates. This reduces the sample size while also selecting the sample towards those goods that are more frequently traded. In section 3.5 we dig deeper into this sample selection. We find that anticipatory effects are even stronger for less frequently traded goods.

[22]See the list of 137 countries in the Table A.2.

purpose and includes insurance and freight charges that controls for the changes in shipping costs. The tariff considered are the applied rates, that is, the ratio of duties over Freight-On-Board (FOB) value of imports.[23] US import data and applied tariffs are obtained from the USITC's database. European trade flows are obtained from Eurostat. Finally, because the phaseout classification is at HS-8 level and there is variation within HS-6 levels, we trade-weight each HS-8 good classification using average trade flows over our full sample period. The characteristics of the resulting sample by phaseout classification are very similar to the original at HS-8 level[24].

3.3 Methodology

We estimate the anticipatory effects by taking sub-period differences of (2). In what follows, $\overline{n}$ and $\underline{n}$ will denote sub-periods of the year t.

$$\underbrace{m^{DD}_{z,t,\overline{n}} - m^{DD}_{z,t,\underline{n}}}_{\Delta_{\overline{n}-\underline{n}} m^{DD}_{z,t}} = -\sigma \left[\Delta_{\overline{n}-\underline{n}} \ln \left(\frac{1+\tau_{iczt}}{1+\tau_{ic'zt}} \right) \right] + \Delta_{\overline{n}-\underline{n}} u_{z,t} \tag{9}$$

The right hand side of (9), under the standard assumption of no anticipation, is null. This is due to the fact that scheduled tariffs only change on an annual basis, i.e. $\Delta_{\overline{n}-\underline{n}} \tau_{iczt} = 0 \; \forall c, z$. This illustrates that according to static models of trade, $\Delta_{\overline{n}-\underline{n}} m^{DD}_{z,t}$ does not respond to upcoming tariff changes. To look at how trade flows exhibit anticipatory behavior, we estimate the degree of trade reversal and rebound in months around the tariff change.

We show that the import decline at the end of the year is significantly associated to the

[23]We drop observations at HS10, month, country of origin, district of entry level - the most disaggregate level available in census data - when applied duties are larger than 100%. These are outliers that were not subject MFN nor NAFTA rate provisions. Results are not sensitive to this choice.

[24]We report Table 4 at HS-6 level in Table A.1 of the Appendix

upcoming tariff reduction. Moreover, the subsequent spike in imports at the beginning of the year is significantly related to the just-bygone tariff reduction. Precisely, our baseline anticipatory effect are estimated using the following specifications:

$$\Delta_{\overline{n}-\underline{n}} m_{z,t}^{DD} = \sigma^D \left[\ln\left(\frac{1+\tau_{i,c,z,t+1}}{1+\tau_{i,c',z,t+1}}\right) - \ln\left(\frac{1+\tau_{i,c,z,t}}{1+\tau_{i,c',z,t}}\right)\right] + \delta_z + u_{z,t} \tag{10}$$

$$\Delta_{\overline{n}-\underline{n}} m_{z,t}^{DD} = \sigma^R \underbrace{\left[\ln\left(\frac{1+\tau_{i,c,z,t}}{1+\tau_{i,c',z,t}}\right) - \ln\left(\frac{1+\tau_{i,c,z,t-1}}{1+\tau_{i,c',z,t-1}}\right)\right]}_{\equiv \Delta\tau_{z,t}^{DD}} + \delta_z + \delta_t + u_{z,t} \tag{11}$$

We use (10) to quantify the within-year anticipatory decline i.e. $\overline{n}, \underline{n} \in t$. The estimate of σ^D is the anticipatory elasticity and is identified through the time variation in tariff reductions over the years. A positive value of σ^D would imply that the imports fall in the months before the tariff reduction.

We estimate (11) to capture the rebound in imports in the period immediately after the tariff reduction. In this equation, we compare the import flows in the first few months relative to the sub-period in the previous year (before the tariff fall) i.e. $\overline{n} \in t$ and $\underline{n} \in t-1$. Since we are taking changes over the year, we keep the year fixed effects in this specification. The estimate of σ^R signifies the post anticipation spike in imports as firms replenish their stocks.

In the equations above, δ_z and δ_t are the product and year fixed effects that account for seasonal or growth trend effects of z between period $\overline{n}$ and $\underline{n}$ and that are not eliminated by the double-differences.[25] Monthly averages over periods $\overline{n}$ and $\underline{n}$ are taken to construct the growth rate $\Delta_{\overline{n}-\underline{n}} m_{z,t}^{DD}$. For example $\Delta_{Oct:Dec-Apr:Sept}\ m_{zt}^{DD}$ takes the log difference between the monthly average of m_{zt}^{DD} from October to December ($\overline{n}$) and the monthly

[25] In the baseline estimation equation z will be a HS-6 good. Results are robust to using fixed effects at HS-4, HS-2 or HS-section level.

average of m_{zt}^{DD} from April to September ($\underline{n}$). In years in which exporters from c to i were facing expected tariff reductions on January 1st of the next year, we expect trade to plunge before the reduction ($\sigma^D > 0$) and rise sharply after it ($\sigma^R < 0$), while in years without any significant tariff reduction we expect no movements in trade flows around the same period.

3.4 Baseline Results

In this section we show that indeed imports did strongly respond to the incentives of de-stocking and stocking-up around the tariff reductions as predicted by the dynamic trade model of section 2.2. For goods that were phased out, trade flows a few months before the tariff reduction plunged with respect to a reference period of the same year at rates that are comparable to the trade elasticity estimates in the literature. We also find a rebound effect in the subsequent period to the tariff reduction.

Results are presented in Table 5. Panel A reports the results of the response by phaseout categories. To obtain this, we interact the tariff variable with an indicator variable that is one if the initial tariff on the HS-6 good was phased out (that is if it belongs to class B, C or C+). First three columns report the estimates of σ^D from (10) with different base periods ($\underline{n}$). The first column shows that in anticipation of an upcoming tariff reduction of 1 percentage point (pp), imports fell by a substantial 6% in the last two months relative to the middle of the year. The last three columns report the estimates of σ^R from (11) with different rebound periods ($\overline{n}$). In the period right after the tariff reduction, 1 percentage tariff reduction is associated with a massive 12% increase in imports in the months right after the reduction relative to the months right before it. More importantly, we find no significant effect on the goods which did not experience tariff phaseouts as suggested by the second row of panel A.

Table 5 also shows that the estimated anticipatory effects are robust to considering different time windows. As it was the case in the model of section 2.2, the response increases as we consider the sub-period closer to the tariff reduction (from column 3 to 1 and columns 6 to 4). In Table 6 we report that these results are unchanged when we consider different choices of product/industry fixed effects that control for seasonality. We see that our baseline findings are robust to moving the window of sub-periods and to the choice of seasonality controls.

These results indicate that in contrast with what is assumed by standard estimation strategies of the elasticity of substitution, during 1990 and 1999 US imports from Mexico plunged and then rebounded in the around the tariff reductions. To put the magnitude of the response in perspective, in column 2 of Table 11 we report the estimate of the annual elasticity. The annual elasticity for phased-out goods is around 3 which is a fourth of the immediate rebound effect from anticipated tariff change. We conclude that anticipation for phased-out goods was significant. Because of these anticipatory effects, imports will appear to respond even stronger once the reduction is effective. That is, in years with high levels of tariffs but impending tariff reductions, imports will appear to be small because of the demonstrated anticipation. Likewise, in periods with lower tariffs, annual imports value would be inflated due to the post-anticipation rise.

3.5 Anticipation and Storability

In this subsection we investigate how the anticipatory effects vary with the degree of product storability. The degree of anticipatory effects strongly rely on firms' ability to delay their shipments. This is also evident from the model simulations in Figure 2 where the highly storable good display a bigger anticipatory decline and a sharper rebound around the anticipated tariff change. Hence we expect to find a positive relationship between anticipatory effects and storability.

Measuring storability has proven to be difficult in the literature.[26] We proxy it by considering the lumpiness of its imports. In the model of section 2.2, more storable goods (lower depreciation rates) are characterized higher inventory-sales ratios and more infrequent orders. We build on this insight and relate lumpier trade with high storability of the goods as inventories.[27] We measure trade lumpiness using the Herfindahl–Hirschman index of monthly concentration of the annual US imports, calculated as follows:

$$HH_g = \sum_{m=1}^{12} \left(\frac{v(g,m)}{\sum_m^{12} v(g,m)} \right)^2 \tag{12}$$

where $v(g,m)$ are value of imports of good g in month m.[28] HH_g is defined over the interval $[1/12, 1]$. For $HH_g = 1/12$, g is traded equally every month of the year, and, for $HH_g = 1$, g is traded in only one month of the year. For this measure, we define goods at a very disaggregate level (HS 10-digit, district of entry and source country) and consider HH_g in the second year g enters our sample.[29] Then we take the median HH_g at HS 6-digit level to obtain HH_z.

Table 7 contains results of estimating (10) and (11) with an interaction of the indicator variable that assumes value 1 for above median HH_z. Comparison of results in Table 5

[26]Some papers such as Alessandria et al. (2010) or Chang et al. (2009) propose continuous measures of storability but only for a limited number of goods and at more aggregate levels than HS-6.

[27]If demand for imported goods is continuous but trade flows are lumpy it must be that demand is satisfied by using inventories. The importance of inventories in international trade has been widely documented as we emphasized in the introduction.

[28]Alternatively we could have used number of shipments. This doesn't change the results.

[29]To achieve variation in the HH index it is necessary to consider it at very disaggregate levels. We take the second year to eliminate the bias coming from the month it was first imported. We exclude Canada and Mexico to control for the average distance of sourcing countries. Results are robust to including averages over more than the second year and including Canada and Mexico.

and 7 strongly suggest that the anticipatory decline and the subsequent rebound effects are driven by the goods which are more storable. Similar to the results presented earlier, anticipatory effects are not significantly present for other good categories.

We also use the continuous nature HH index in regression analysis to use more information. To see how the continuous measure of storability interacts with anticipatory decline before the tariff reduction, we estimate the following equation:

$$\Delta_{n-n'} m^{DD}_{z,t} = \sigma^D_0 \; \Delta\tau^{MEX,RoW}_{US,z,t+1} + \sigma^D_1 \; \Delta\tau^{MEX,RoW}_{US,z,t+1} \times HH_z + \sigma^D_2 \; \Delta\tau^{MEX,RoW}_{US,z,t+1} \times HH^2_z + u_{z,t,n-n'} \tag{13}$$

We limit this exercise to the goods that were phased-out, for which we found significant and economically sizeable anticipatory elasticities. This is the same estimation equation as (10), incorporating the interaction of $\Delta\tau^{MEX,RoW}_{US,z,t+1}$ with HH_z and HH^2_z and excluding the product fixed effects since the HH index is product-specific. Including the squared term of the HH index is important in capturing the interaction as results reported in Table A.3 of the Appendix show. The anticipatory elasticity now is the combination of the coefficients $\sigma^A_{i \in \{0,1,2\}}$.

In Figure 1 we report the response of $\Delta_{11:12-3:10} m^{DD}_{z,t}$ to a one percentage point drop in tariffs in the upcoming year for different percentiles of the HH Index. Between the 20th and 70th percentile of the HH index the decline in imports before the tariff change is increasing. The anticipatory elasticity peaks at a value of around 10 at the 70th percentile. Moreover, precision increases over the x-axis, until the 70th percentile. Afterwards, the response is imprecise. The results are similar when we consider $\Delta_{11:12-5:10} m^{DD}_{z,t}$ or $\Delta_{10:12-4:9} m^{DD}_{z,t}$ and are reported in Figure A.3 of the Appendix.

To summarize, the results in Tables 5, Table 7 and Figure 1 document that imports of the goods that experienced phaseouts fell significantly in advance of tariff reductions and rebounded sharply afterwards. More importantly, this effect was strongest for goods that are more intensively held as inventories. We take these results as evidence for the fact that during the window around tariff reductions imports and their consumption diverged significantly. Unfortunately, we can't test this hypothesis because neither inventory nor consumption data are available at the level of disaggregation of the trade data. In the next section, we propose a solution to overcome this lack of data and estimate the elasticity of substitution by introducing a measure of consumed imports.

4 Biases in the Elasticity of Substitution

In the previous section we showed that imports plunge and then rebound strongly in the months around the anticipated tariff reductions. In the model presented in section 2.2, these dynamics cause the trade elasticity to exceed the elasticity of substitution. This leads to an overestimated elasticity of substitution with the standard approach of using import flows to estimate it. The model simulations suggest that regressions based on the consumption of imports rather than the import flows estimates the underlying elasticity of substitution, regardless of the nature of the good and the type of the shock. However, the data on product- and source-specific consumption is not available. To estimate the elasticity of substitution we propose a measure consumption of imports. The proposed process for the consumption uses the high-frequency trade data and the relationship between the inventory-sales ratio and the lumpiness of imports. We find that the bias is strong at annual frequency and diminishes in the long-run estimate. As a result, there is considerable in the average elasticity estimate. We first describe the consumption measure. We then apply it to estimate the elasticity of substitution. Finally, we consider several alternative proxies that accommodate different assumptions on the consumption process.

4.1 Measuring Consumption of Imports

The difference between imports and their consumption can arise due to the presence of inventories. At high frequencies, the two variables can diverge significantly as the literature has demonstrated extensively (Alessandria et al. (2010a, 2011a, 2011b)). To measure consumption of imports we begin with the monthly law of motion of inventory holdings specified in (5) of the model. The end of the month inventory holdings are,

$$\tilde{s}_{icz,n} = \tilde{s}_{icz,n-1} + m_{icz,n} - \tilde{q}_{icz,n} \tag{14}$$

where i indexes the destination and c the source country. We denote the month by n and z is a product at HS6 level. The tilde on top of inventory holdings, $\tilde{s}$, and consumption of imports, $\tilde{q}$, indicate that they are not available at this level of disaggregation.[30] Monthly imports, m, are obtained from the data. We begin by making an innocuous assumption of initial inventory holdings and then roll it forward using (14).[31] The key here is the process of consumption of imports, $\tilde{q}$. In our baseline, the process for consumption in country i of imports from country c of product z is given by,

$$\tilde{q}_{icz,n} = \frac{\tilde{s}_{icz,n-1} + m_{icz,n}}{k_{icz}} \tag{15}$$

The right hand side of (15), establishes that a constant fraction $1/k$ of current inven-

[30] In the model consumption of imports are sales. This is a simplification. We interpret consumption of imports as the fraction of the imported good $m_{icz,t}$ used as sales or in production by the importer.

[31] In particular we assume that in the first month inventory holdings are the monthly average of imports in the first year multiplied by the inventory-sales ratio. Our results do not reply on this assumption.

tory holdings, $\tilde{s} + m$, is used for consumption. The key ingredients of the consumption process are; the monthly import data, and the constant rate of usage, namely the inverse of the inventory-sales ratio, denoted by k.[32] Constant inventory-sales ratio is justified by our inventory model used for simulations just like all the inventory models.[33] The level of inventory-sales ratio essentially depends on the product-specific factors such as the depreciation rate and the fixed cost of ordering. Additionally, from an empirical perspective the frequency of trade has been shown to depend on the delivery lags and the time to trade (Hummels and Schaur (2013), Bekes et al. (2017)). Therefore we allow for different inventory-sales ratio across products and directions of trade.

We obtain the inventory-sales ratio, k_{icz}, by exploiting its relationship with the lumpiness of trade. In particular, we use the fact that a high inventory-sales ratio will be associated with less frequent shipments. therefore, we multiply the Herfindahl-Hirschman (HH) index of monthly concentration of annual imports by 12 to obtain the number of months worth of sales purchased in the average import order. As in section 3.5, we calculate the HH index at HS-6 level by taking the median over HH indexes calculated at HS10 and entry-district level in the second year the good appears in our sample[34]. In section 4.3 we demonstrate that our results are robust to different calculations of k. Finally note that we are assuming no time to market, that is, no lag between the reception of goods and its availability for consumption, since imports received at n are readily available for their consumption. We will later relax this assumption as well in the robustness of this measure.

[32] A convenient feature of the assumed process of consumption of imports is that it generates non-zero values at any time frequency even when actual imports are zero, thereby overcoming issues associated with missing values (See Silva and Tenreyro (2006)).

[33] We later perform robustness checks for this assumptions about constant inventory-sales ratio by adding a shock to it.

[34] Because we don't observe this level of disaggregation of the European trade data, we assume that $k_{US,c,z} = k_{EU,c,z}$.

In Figure 3 we illustrate the behavior of our baseline measure in the case of a specific HS-6 good, namely vehicles used for transport of goods exported from Mexico to the US. In first place, since the numerator of our baseline measure includes contemporaneous imports, consumption of imports tracts the pattern of actual imports. In fact, over the full period of our sample the monthly correlation between US imports from Mexico and our baseline measure of its consumption is 98%. Secondly, by allowing sales from the beginning-of-period inventory holdings, the time series of consumption of imports avoids the lumpiness of actual imports. This can be observed in the abrupt halt in importing in the middle of 1994. During this gap in imports, our measure infers a gradual running down of inventories. As imports resume, both consumption of imports and inventories grow. These two ingredients are crucial to our purpose, since they balance high frequency drops and rises with the overall level of trend in trade volumes. For example, in the three months before January of 1997 imports fell and then spiked after January. However, our measure of consumption of imports does not reflect this reversal since in the previous part of the year imports had been on the rise.

A direct approach to examine whether our baseline measure overcomes the dynamics documented in section 3.4 is by estimating the anticipatory elasticity obtained by (10) using our baseline measure of consumption of imports in the construction of $\Delta_{\overline{n}-\underline{n}} y_{z,t}^{DD}$, the within-year growth rate between $\overline{n}$ and $\underline{n}$ of the double-difference measure of US imports from Mexico. In Table 8 we show that now the anticipatory elasticity is economically negligible and statistically insignificant for phased-out goods, as well as for others.

Finally, we examine how our measure performs in the model simulations.[35] The correlation between actual consumption of imports and our baseline measure is 0.94 at

[35]In the model, contemporaneous orders cross border in the next period and are not available for consumption. Hence, in the model simulations, m_t drops out of (15). This makes it comparable to the measure we apply in the data since import flows in the data are the ones recorded at the border crossing.

the annual frequency and 0.24 at the monthly frequency. In columns three and four of Table 2 and column three of Table 3 we report the results of estimating (7) and (8) using our baseline measure as the dependent variable, respectively. The estimates in all cases are very close to the ones we get using actual consumption of imports in the simulation. In subsection 4.3 we robustness for our measure of consumption of imports. In the next subsection we apply baseline measure of imports consumption to estimate the elasticity of substitution of US imports from Mexico during the early years of NAFTA.

4.2 Elasticity Estimates with Consumption of Imports

We now examine whether the anticipation we documented in section 3.4 affects the estimated elasticity of substitution. To do so, we follow the double-difference approach described in 2.1, but vary the dependent variable, y_{zt}^{DD}. We use import flows to estimate the trade elasticities, and our measure of consumed imports to estimate the elasticity of substitution between domestic and foreign goods. Precisely, We set $y_{zt}^{DD} = \tilde{q}_{zt}^{DD} \equiv \ln\left(\frac{\tilde{q}_{US,MEX,z,t}}{\tilde{q}_{US,RoW,z,t}} / \frac{\tilde{q}_{EU,MEX,z,t}}{\tilde{q}_{EU,RoW,z,t}}\right)$ to estimate the elasticity of substitution. Where $\tilde{q}_{i,j,z,t}$ is the aggregation of monthly flows from (15) to the annual frequency. We first provide estimates of the average, or cross-sectional, elasticity and then consider the distinction between short-run and long-run elasticities.

4.2.1 Static Elasticity Estimates

Estimates of the cross-sectional elasticity of substitution are considered to capture the average elasticity by pooling together all the year and products alike. This is an important parameter for the outcomes of static trade models. Following Romalis (2007), and Head and Ries (2001); we use the following estimation equation derived in (2) to estimate the elasticity of substitution:

$$y_{zt}^{DD} = \sigma \tau_{z,t}^{DD} + \delta_z + \delta_t + u_{zt} \tag{16}$$

Product and time fixed effects, δ_z and δ_t, are added to capture product-level export prices differences and time trends in shipping costs that have differential effect on US imports from Mexico. We report the results in Table 10.[36] In the first and second columns we report the elasticity estimate for all goods. By using consumption of imports the elasticity of substitution drops from 8.9 to 7.7.[37] This implies a considerable upward bias of 16% in the estimates that are based on inport flows. In columns three and four we consider how the elasticity of substitution of phased-out goods is affected. We expect these goods to be the main drivers of the bias, since the anticipatory dynamics were stronger for them. We find that the bias is indeed larger for the at around 21%. Their elasticity of substitution drops from 13.2 to 10.9. For non-phased-out goods the bias is negligible as their elasticity falls from 6.6 to 6. In the robustness section we show that this pattern persists and is sometimes amplified under alternative version of the imports consumption process.

4.2.2 Dynamic Elasticity Estimates

We now make the distinction between short-run and long-run response of imports consumption. We generalize (16) by allowing tariffs to have a long-run impact different from the immediate effect. In particular, we apply an unrestricted Error-Correction

[36] We delete the first year of our sample since our measure does not output sales in the first month of the sample. This does not affect the results.

[37] Our estimate using import flows is below that obtained by Romalis (2007), who obtains 10.9. This is most likely from the fact that we exclude 1989 from our sample period. In fact as we exclude earlier years from our sample the estimate falls, suggesting that those early years with less trade loom large in obtaining large trade elasticities.

Mechanism (ECM) model and estimate the following equation:

$$\Delta y_{zt}^{DD} = \sigma^S \ \Delta\tau_{i,z,t}^{c,c'} + \sigma^L \ \tau_{i,z,t-1}^{c,c'} + \alpha \ y_{z,t-1}^{DD} + \delta_t + \delta_z + u_{zt} \tag{17}$$

where $\Delta y_{z,t}^{DD} = y_{z,t}^{DD} - y_{z,t-1}^{DD}$.[38] Short-run or one-year elasticity is denoted by σ^S and $-\sigma^L/\alpha$ denotes the long-run elasticity.[39] This formulation generalizes (16) by relaxing the assumptions $\sigma^S = \sigma^L$ and $\alpha = -1$. We consider both the imports flows and their consumption as dependent variables to highlight the difference between trade elasticity and the elasticity of substitution.

In Table 11 we report the short-run and long-run elasticities from (17). There is a sizeable bias in the short-run elasticity, of around 68%. The estimate drops from 4.2 when using import flows to 2.5 when using our measure of consumption of imports. Columns three and four show that the bias is driven by a largely overestimated short-run elasticity for phased-out goods. For these goods, the short-run elasticity with import flows is around 2.45 times the one estimated using the consumption of imports, it falls from 7.1 to 2.9. These results are in agreement with the strong anticipatory responses of phased-out goods found in Section 3.4. In contrast, non-phased-out goods yield similar estimates when using imports flows or consumption of imports. It is also noticeable that the estimated short-run elasticity for the phased-out goods becomes almost equal to the one for the non-phased-out goods once we control for the anticipatory effects.

We now discuss the effect of anticipatory dynamics on the long-run elasticity. Anticipation causes firms to delay their imports to right after the tariff reduction. This would affect the immediate response more than the long-run response, which is governed

[38] This estimation equation assumes imports and tariffs have order of integration 1 and are cointegrated.

[39] In a restricted version, we would add a lag of the dependent term to (16). That would impose the restriction $\sigma^S = \sigma^L$.

by longer-run dynamics such as firm entry. Table 11 shows that the use of import flows overstates the aggregate long-run elasticities by only 5%. This finding along with the short-run response discussed above shows how anticipation affects elasticity differently at different horizons.

Anticipatory dynamics affect the measured size and duration of long-run adjustment. The ratio of long-run to short-run elasticity is an important feature of quantitative models. Table 11 shows that long-run elasticity is around twice as big as the short-run elasticity with import flows as the dependent variable. This is in line with the estimates of this ratio found in the literature (Gallaway et al. (2003), Baier and Bergstrand (2007), Jung (2012) and Yilmazkuday (2019)). However, we find a bigger ratio of long- to short-run elasticity of around 3.7. The aggregate short-run elasticity is estimated to be 2.5 and the long-run elasticity is around 9.2.

In Figure 4 we plot the dynamic response of import flows and import consumption to a 1 percentage point permanent tariff reduction. The elasticity grows with the distance between periods. Although the bias becomes relatively less important over time, a small bias persists. However, a notable feature of the plot is the speed with which imports adjust to their long-run value. While import flows show a convergence time of 4 years, import consumption takes around 7 years to converge to its long-run value. This is line with the estimated speed of adjustment parameter reported in Table A.6. In Figure 5 we report the dynamic response for the phased-out goods. Two major differences are evident from the figure. First, the reduction in short-run elasticity is bigger. Second, because of a larger anticipatory dip for phased-out goods, even the long-run elasticity has a bigger bias of around 17%.

4.2.3 Discussion

We have shown that the anticipation documented in section 3 strongly affects the estimated response of consumption of imports to tariff changes. In the short-run, the number of years in which imports fall in anticipation of phaseouts - i.e. the number of years with *base-period effects* - is large. Hence anticipation and the deviation of imports from their consumption looms large in the elasticity estimate. As the number of years since tariff change increases, the importance of the base periods effects falls and the trade elasticity and elasticity of substitution become similar. Moreover, anticipation causes a bigger amplification of import flows over their consumption in the short-run relative to the long-run. This leads to the speed of adjustment being biased downwards when one uses the import flows. In the next subsection we demonstrate that these results are robust to alternative processes of the consumption of imports.

4.3 Robustness

The results described in the previous section are robust to several alternative modelling decisions and its empirical implementation on the processes of consumption of imports. Results of estimating (16) and (17) under the alternative processes described below are presented in Tables 12 and 13.

One Period Time to Market Lag. - Under the baseline measure, once imports are received at the destination, they are immediately available for consumption. However, there can be lags in domestic delivery or lead times to market the imports. This would make current imports unavailable for sale. Hence, the process that assumes a one period lag in the lead time between reception and consumption of imports would have the following representation:

$$\tilde{q}_{icz,t} = \frac{\tilde{s}_{icz,t-1}}{k_{icz}} \tag{18}$$

In Panel A of Table 12 we observe that the use of this measure also causes reduction in the estimate. The bias in aggregate static elasticity is around 22%, falling from 8.9 to 7.3. It is driven by phased-out goods as can be seen in panel B, the estimate for the phased-out goods drops from 13 for m^{DD} to 10 for $\tilde{q}^{DD}$.

We see a similar effect on dynamic elasticities while using the consumption measure that allows for time to market to our baseline measure. The aggregate short-run and long-run elasticities are the same as the ones obtained using baseline consumption measure. The short-run response for the phase-out goods is slightly lower as compared to the baseline.

Different Inventory-Sales Ratios. - In our baseline measure we use a different inventory-sales ratio, k_{icz}, only for different exporter countries i, since we don't consider more disaggregate EU-12 trade flows to calculate HH indexes. To check whether our results are sensitive to this lack of data we impose a common k_z across all four directions of trade in our baseline measure.

Assuming the same inventory-sales ratio does not affect our estimation. The estimates of aggregate static elasticity are very close to the one using our baseline consumption measure. The estimated static elasticity for phased-out goods shows a slightly lower reduction from 13 to 11 rather than 10 with our baseline measure.

For dynamic elasticities, having a common inventory-sales ratio does not affect our main results. The aggregate short-run and long-run elasticities are marginally different from the ones obtained using our baseline measures of consumption of imports. The estimates using this measure points towards the similar differential effect of anticipation on short-run and long-run elasticities.

Demand Shock. - We now extend our baseline version of the process of $\tilde{q}$ by including a demand shock, ν_t. This demand shock is equivalent to an aggregate demand shock in the model presented earlier since it is a HS6-specific shock for all firms. The process for predicted consumption of imports then takes the following form:

$$\hat{q}_{icz,t} = \underbrace{\frac{\tilde{s}_{icz,t} + m_{z,t}}{k_{icz}}}_{\text{expected}} + \underbrace{\frac{m_{icz,t+1}}{k_{icz}} - \mathbb{E}_t(m_{icz,t+1})}_{\text{shock } (\nu_t)} \tag{19}$$

$$\tilde{q}_{icz,t} = \min\left[\hat{q}_{icz,t},\ \tilde{s}_{icz,t-1} + m_{icz,t}\right] \tag{20}$$

The first term on the right hand side of (19) is our baseline measure. The second term on the right hand side is the shock component of $\hat{q}$. The demand shock assumes that next period's import volume reveals information about the contemporaneous demand shock. We model it as the deviation between actual monthly imports and its expected value the period before. In other words, we infer a favourable contemporaneous demand shock if we observe orders higher than the good's average imports in the following period. This assumes the existence of a delivery lag between purchase orders and reception of imports of one month, as it was the case in the model. We divide m_{t+1} by k to account for the fact that purchase orders are intended to satisfy consumption for an average k periods. We calculate $\mathbb{E}_t m_{t+1}$ by taking the average of imports of the contemporaneous and the previous five months, i.e. $1/6 \sum_{i=-5}^{0} m_{t+i}$. Finally, (20) imposes that consumption of imports can't exceed contemporaneous inventory holdings, as we assumed in the model by requiring (4).

As explained above, in the model contemporaneous imports are orders, so that in the model analogous process of our measure with demand shocks, the expected term is $\frac{\tilde{s}_{j,t}}{k}$ and the shock is $\frac{m_{j,t}}{k} - \mathbb{E}_{t-1}(m_{j,t})$. In Table A.5 of the Appendix we report that in the model simulations this measure is similarly effective in reducing the bias from anticipation[40]. In Figure A.4 of the Appendix, we show how the measure with demand shocks behaves under the same import pattern as illustrated in Figure 3. Now imports precede consumption of imports, since demand shocks are obtained from deviations between next period's imports and lagged imports.

In column five of Table 12 we report the results of estimating (16) using actual imports and the measure with demand shocks as the dependent variable. In the aggregate estimate in panel A, the bias of cross-sectional elasticity increases to 22%, falling from 8.9 to 7.3. Again, it is driven mostly by the phased-out goods for which the estimate drops from 13 to 10.4. In column six, we also consider the case in which the expected imports per months of sales, $\mathbb{E}_{t-1}(m_{j,t})$ is calculated using 6 months of lags and leads. In this case we obtain a bigger bias yielding a 39% difference in the estimate for m^{DD} and $\tilde{q}^{DD}$.

In column five of Table 13 we report the results of estimating (17) with the consumption measure with demand shocks. The aggregate short-run elasticity is substantially lower with this measure compared with the baseline measure. However, as anticipated, the effect on the long-run elasticity is not substantial. This is in line with the hypothesis that anticipation largely affects the short-run elasticity. Because of larger reduction in the short-run estimate, the ratio of long- to short-run elasticity comes out to be much bigger with this measure. We also find that the specific effect on the phased-out goods is not robust to adding a demand shock to the consumption process. This can be a result of rigid assumptions about the expectations used to measure the demand shock. The

[40]However, its effectiveness in the model is sensitive to varying the number of leads and lags included in the calculation of the expected imports per months of sales, $\mathbb{E}_{t-1}(m_{j,t})$. Therefore, we consider the measure without demand shock as our baseline.

assumed duration of expectation (of 6 months) varies with the variability of the demand shock at the product level. Nevertheless, the effect on static and aggregate dynamic elasticities is fairly robust to adding a demand shock to our baseline measure of consumption of imports.

5 Conclusion

In this paper we document that anticipation of upcoming tariff changes causes significant trade reversals and subsequent amplification of trade flows. In the NAFTA's case, we find that the import flows from Mexico to the US *slump* and then *bump* significantly around the tariff reductions. Anticipatory effects are strongest for goods which are more storable and whose tariffs were phased out gradually over the years. A standard inventory trade model with a ($\underline{s}$,$\overline{s}$) ordering policy reproduces the observed response of trade flows to anticipated tariff changes. Through the model simulations we show that the imports and their consumption diverge significantly during these periods. Hence using import flows instead of consumption of imports leads to biases in the elasticity of substitution. However, the data on the consumption of imports is not available at the level of disaggregation used in the standard estimation approach. Therefore, this paper introduces a measure of consumed imports. This measure is based on high-frequency disaggregate import data and inventory-sales ratios. We find that not accounting for the deviation between imports and their consumption around anticipated tariff reductions results in strongly overestimated average elasticities of substitution. The bias is driven by the goods which faced tariff phaseouts. The cross-sectional elasticity, which is a parameter for the static trade models, is biased upwards by 16% when import flows are used.

We also address the overestimation of the elasticities used in the dynamic models of trade. We find that anticipatory *slumps* and liberalization *bumps* affect the short-run elasticities more than their long-run counterpart. We find that using import flows instead of consumption of imports biases the short-run elasticity by 68% whereas it only affects long-run estimate by a negligle 5%. This happens because as we move forward from the tariff change, the anticipatory *bumps* are dominated by the long-run response of imports. Nonetheless, the effect of the anticipatory *slump* remains even at the longer horizon. By combining these findings, we see that using consumption of imports reveals that the ratio of long- to short-run elasticities is 75% higher at around 3.5 instead of 2. Using our measure of consumption, we overcome the bias generated from using the trade elasticity as measure of the elasticity of substitution.

There are several insights from this paper that open up interesting future research avenues. On the one hand, in this paper, we focus on the interaction between high-frequency investment decisions, such as inventory holdings, and the anticipated nature of gradual tariff phaseouts. However, more generally, anticipated trade policy changes can generate incentives to shift longer-term investment decisions such as exporting decision or FDI, that can potentially have strong aggregate implications (Alessandria and Mix (2018), Threinen (n.d.)). On the other hand, this paper studies anticipated policy changes where uncertainty was not a factor. Nonetheless, anticipatory effects could be sizeable in the case of uncertain policy changes. In Alessandria et al. (2019) we show that within-year variation in trade around congressional votes during the 1990's on China's MFN status can be used to estimate the probability of its non-renewal. Similarly, the possibility of a disruption of supplies from Europe under a hard Brexit has led British firms to stockpile in anticipation.[41] The mechanism highlighted in this paper elicits information on how to use the observed stockpiling to study expectations about the path of future trade policy.

[41] Wall Street Journal (May 10, 2019).

References

Alessandria, George and Carter Mix, "Trade Policy is Real News: An Analysis of Past, Present, and Future Trade Costs," 2018.

_ , Horag Choi, and Kim J. Ruhl, "Trade Adjustment Dynamics and the Welfare Gains from Trade," 2018.

_ , Joseph Kaboski, and Virgiliu Midrigan, "The Great Trade Collapse of 2008-09: An Inventory Adjustment?," *IMF Economic Review*, 2010, *58*, 254–294.

_ , _ , and _ , "Inventories, lumpy trade, and large devaluations," *American Economic Review*, 2010, *100(5)*, 2304–2339.

_ , _ , and _ , "Trade and Inventory Dynamics," *European Central Bank E-Book*, 2011.

_ , Shafaat Y. Khan, and Armen Khederlarian, "Taking Stock of Trade Policy Uncertainty: Evidence from China's Pre-WTO Accession," *WP*, 2019.

Anderson, James and Eric van Wincoop, "Trade Costs," *Journal of Economic Literature*, 2004, *42(3)*, 691–751.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare, "New Trade Models, Same Old Gains?," *American Economic Review*, 2012, *102(1)*, 94–130.

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

Baker, Scott, Lorenz Keung, and Stephanie Johnson, "Shopping for Lower Sales Tax Rates," 2018.

Bekes, Gabor, Lionel Fontagne, Balazs Murakozy, and Vincent Vicard, "Shipment Frequency of exporters and demand uncertainty: An inventory management approach," *Review of World Economics*, 2017, *153*, 779–807.

Besedes, Tibor, Tristian Kohl, and James Lake, "Phase Out Tariffs, Phase In Trade?," 2018.

Blum, Bernardo, Sebastian Claro, Kunal Dasgupta, and Ignatius Horstmann, "Inventory Management, Product Quality, and Cross-Country Income Differences," *American Economic Journal: Macroeconomics*, 2019, *11(1).*

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

Charnavoki, Valery, "Retail Sales of Durable Goods, Inventories and Imports after Large Devaluations," 2017.

Coglianese, John, Lucas Davis, Lutz Kilian, and James Stock, "Anticipation, Tax Avoidance, and the Price Elasticity of Gasoline Demand," *Journal of Applied Econometrics*, 2017.

Fajgelbaum, Pablo, Pinelopi Goldberg, Patrick Kennedy, and Amit Khandelwal, "The Return to Protectionism," *NBER WP w25638*, 2019.

Feenstra, Robert C., Phillip Luck, Maurice Obstfeld, and Katheryn N. Russ, "In Search of the Armington Elasticity," *Journal of Economic Literature*, 2004, *42(3)*, 691–751.

Fukao, Kyoji, Toshihiro Okubo, and Robert M. Stern, "An econometric analysis of trade diversion under NAFTA," *The North American Journal of Economics and Finance*, 2003, *14(1)*, 3–24.

Gallaway, Michael P., Christine A. McDaniel, and Sandra A. Rivera, "Short-run and Long-run Industry level Estimates of U.S. Armington Elasticities," *The North American Journal of Economics and Finance*, 2003, *14(1)*, 49–68.

Head, Kieth and John Ries, "Increasing Returns versus National Product Differentiation as an Explanation for the Pattern of U.S.-Canada Trade," *American Economic Review*, 2001, *91(4)*, 858–876.

Hillbery, Russel and David Hummels, "Trade Elasticity Parameters for a Computable General Equilibrium Model," *Dixon and Jorgenson (eds) Handbook of Computable General Equilibrium Modeling, Elsevier*, 2012.

Hummels, David and Georg Schaur, "Time as a Trade Barrier," *American Economic Review*, 2013, *103(7)*, 2935–2959.

Jung, Benjamin, "Gradualism and dynamic trade adjustment: Revisiting the pro-trade effect of free trade agreements," *Economic Letters*, 2012, *115*, 63–66.

Kehoe, Timothy and Kim Ruhl, "How Important is the New Goods Margin in International Trade," *Journal of Political Economy*, 2013, *121(2)*, 358–392.

Kowalczyk, Carsten and Donal Davis, "Tariff phase-outs: Theory and evidence from GATT and NAFTA," *The Regionalization of the World Economy. University of Chicago Press*, 1998, pp. 227–258.

Kropf, Andreas and Philip Saure, "Fixed Cost per Shipment," *Journal of International Economics*, 2014, *92*, 166–184.

Kruegger, Anne O., "Trade Diversion and Trade Creation under NAFTA," *Latin American Econometric Society Meeting, Cancun, Mexico*, 1999.

Romalis, John, "NAFTA's and CUSFTA's Impact on International Trade," *Review of Economics and Statistics*, 2007, *89(3)*, 416–435.

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

Simonovska, Ina and Michael Waugh, "The Elasticity of Trade: Estimates and Evidence," *Journal of International Economics*, 2014, *92(1)*, 34–50.

Staiger, Robert, "A Theory of Gradual Trade Liberalization," *Deardorff, A., Levinsohn, J., and Stern, R. (eds.) New Directions in Trade Theory, University of Michigan Press*, 1995.

Threinen, Lucas, "The Dynamic Response to Trade Policy: Evidence from the US Textile and Clothing Industry."

Yilmazkuday, Hakan, "Estimating the Trade Elasticity over Time," 2019.

Table 1: Parameters and Moments of the Simulation

Parameter		High I/S	Low I/S	Unanticipated
f	Fixed Cost Ordering	0.05	0.005	0.05
σ_ν	Variance of Taste Shocks	0.6^2	0.6^2	0.6^2
δ	Monthly Depreciation Rate	2.50%	10%	2.50%
σ	Elasticity of Substitution	4	4	4
β	Monthly Interest Rate	$0.96^{(1/12)}$	$0.96^{(1/12)}$	$0.96^{(1/12)}$
Moments				
	Equilibrium monthly I/S Ratio	2.54	1.35	2.54
	HH Index	0.21	0.09	0.21
	Fixed Cost over Mean Revenues	3.60%	0.37%	3.60%

Table 2: Model Simulation Average Elasticities

Sample Period:	Years 1 & 4	Years 2 & 3	Years 1 & 4	Years 2 & 3
Dep. Var. :	m_t	m_t	q_t	q_t
$1\{\text{Benchmark}\} \times \log(1+\tau_t)$	-4.10***	-6.08***	-4.18***	-4.34***
	(0.04)	(0.04)	(0.04)	(0.04)
$1\{\text{Low I/S}\} \times \log(1+\tau_t)$	-4.00***	-4.14***	-4.05***	-3.97***
	(0.01)	(0.01)	(0.01)	(0.01)
$1\{\text{Unanticipated}\} \times \log(1+\tau_t)$	-4.12***	-4.20***	-4.18***	-4.04***
	(0.04)	(0.04)	(0.04)	(0.04)
N	60000	60000	60000	60000
Adjusted R^2	0.42	0.52	0.48	0.48

Note: Estimates are obtained from estimation equation (7) varying the dependent variable and the sample period included. Columns one and two use annual imports as the dependent variable. Columns three and four use annual consumption of imports or sales. Tariff changes take place in the first month of the third year in the simulations. Hence column one and three using sample years 1 and 4 are not subject to the bias from anticipation. Standard errors, in parentheses, are clustered at firm level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Model Simulation 1-Year Trade Elasticity

Dep. Var.:	Δm_t	Δq_t	$\Delta \tilde{q}_t$ (Imputed)
$1\{\text{ Benchmark }\} \times \Delta \log(1+\tau_t)$	-7.18***	-4.18***	-4.11***
	(0.06)	(0.06)	(0.05)
$1\{\text{ Low I/S }\} \times \Delta \log(1+\tau_t)$	-4.23***	-3.78***	-3.83***
	(0.02)	(0.02)	(0.02)
$1\{\text{ Unanticipated }\} \times \Delta \log(1+\tau_t)$	-4.41***	-3.91***	-3.55***
	(0.07)	(0.06)	(0.05)
N	60000	60000	60000

Note: Estimates are obtained from estimation equation (8) varying the dependent variable. Column one uses imports from the model simulation. Column tow uses consumption of imports or sales from the model simulation. Column three uses our baseline measure described in (15). Sample years are year 2 and 3 of the simulation, the only years with variation in $\Delta \ln(1+\tau)$. Standard errors, in parentheses, are clustered at firm level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Phaseout Categories of HS-8 Goods Imported to US from Mexico

	1993			1999		
Classes	Number Goods	Scheduled Tariff	Import Share	Number Goods	Median Tariff	Import Share
A - 1 stage	438	5.51%	15.36%	367	0%	13.93%
B - 5 stages	760	8.95%	9.08%	695	0%	9.76%
C - 10 stages	641	6.49%	3.31%	498	2.64%	3.26%
C+ - 15 stages	68	19.00%	0.54%	66	10.39%	0.71%
D - Zeroed by 1993	4,814	0.01%	62.02%	4,249	0%	36.77%
Unclassified	264	0.89%	9.67%	1,920	0.02%	35.58%
TOTAL	6,985	2.17%	100%	7,772	0.26%	100%

Table 5: Anticipatory Elasticities

Dep. Var.: $\Delta_{\underline{n}:\overline{n}} m_{zt}^{DD}$	Fall	Fall	Fall	Rise	Rise	Rise
$\overline{n}$	11:12	11:12	10:12	1:4	1:3	1:4
$\underline{n}$	4:8	4:10	4:9	11:12	11:12	10:12
Panel A: Categories						
$\mathbf{1}\{\text{Phased}\}\Delta\tau_z^{DD}$	6.12***	4.63***	4.10***	-11.7**	-8.55**	-7.56
	(2.10)	(1.34)	(1.46)	(4.67)	(4.07)	(3.21)
$\mathbf{1}\{\text{Other}\}\Delta\tau_z^{DD}$	-1.56	-0.97	-0.18	-2.69	-1.50	-3.46
	(2.76)	(2.25)	(2.38)	(2.77)	(2.70)	(2.46)
Panel B: Aggregate						
$\Delta\tau_z^{DD}$	0.72	0.66	1.08	-5.47**	-3.67	-4.70***
	(1.48)	(1.40)	(1.73)	(2.58)	(2.48)	(2.15)
HS6 FE	✓	✓	✓	✓	✓	
Year FE			✓	✓	✓	
N	6023	6285	7076	7014	7014	8323
Adj R2	0.075	0.103	0.087	0.266	0.244	0.263

Note: Estimates are obtained from (10) and (11) using different within-year periods growth rates. Standard errors, in parentheses, are clustered at HS-2 product level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Anticipatory Elasticities - Robustness

Dep. Var.: $\Delta_{\underline{n}:\overline{n}} m_{zt}^{DD}$	Fall	Fall	Fall	Rise	Rise	Rise
$\overline{n}$	11:12	11:12	11:12	1:4	1:4	1:4
$\underline{n}$	4:8	4:10	4:10	4:10	11:12	11:12
$\mathbf{1}\{Phased\}\Delta\tau_z^{DD}$	6.12***	4.01**	4.87***	-11.7**	-8.87**	-13.1***
	(2.10)	(1.93)	(1.77)	(4.67)	(3.43)	(4.58)
$\mathbf{1}\{Others\}\Delta\tau_z^{DD}$	-1.56	-1.69	-1.38	-2.69	1.53	-1.54
	(2.76)	(2.53)	(2.48)	(2.77)	(2.54)	(3.34)
Year FE				✓	✓	✓
HS6 FE	✓			✓		
HS4 FE		✓			✓	
SITC FE			✓			✓
N	6023	6307	6176	7014	7421	7251
adj. R^2	0.075	0.045	0.066	0.266	0.140	0.209

Note: Estimates are obtained from (10) and (11) using different fixed effects. Standard errors, in parentheses, are clustered at HS-2 product level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Anticipatory Elasticities - Robustness

Dep. Var.: $\Delta_{\underline{n}:\overline{n}} m_{zt}^{DD}$	Fall	Fall	Fall	Rise	Rise	Rise
$\overline{n}$	11:12	11:12	10:12	1:4	1:3	1:4
$\underline{n}$	4:8	4:10	4:9	11:12	11:12	10:12
$\mathbf{1}\{\text{HH} > \text{Med}(HH)\}\mathbf{1}\{\text{Phased}\}\Delta\tau_z^{DD}$	7.07**	5.76***	4.94***	-17.0**	-12.7*	-11.5**
	(2.75)	(1.99)	(1.63)	(7.45)	(6.64)	(5.79)
$\mathbf{1}\{\text{Other}\}\Delta\tau_z^{DD}$	-0.83	-0.56	0.13	-2.50	-1.33	-2.94
	(2.22)	(2.00)	(2.06)	(2.55)	(2.59)	(2.43)
HS6 FE	✓	✓	✓	✓	✓	
Year FE			✓	✓	✓	
N	6023	6285	7076	7014	7014	8323
Adj R2	0.075	0.103	0.087	0.266	0.244	0.263

Note: Estimates are obtained from (10) and (11) with interaction with storability. Storability is proxied by the lumpiness of imports given by the HH index of concentration of imports over the year. High HH index is associated with high storablity. Standard errors, in parentheses, are clustered at HS-2 product level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Anticipatory Elasticities - Baseline Consumption Measure

Dep. Var.: $\Delta_{\overline{n}-\underline{n}} \tilde{q}_{zt}^{DD}$	$\overline{n}$	Oct-Dec	Nov-Dec	Nov-Dec	Oct-Dec	Nov-Dec	Nov-Dec
	$\underline{n}$	Apr-Sep	Mar-Oct	Mar-Aug	Apr-Sep	Mar-Oct	Mar-Aug
$\Delta\tau_{US,z,t+1}^{MEX,RoW}$		0.28	0.03	0.22			
		(0.41)	(0.54)	(0.67)			
1 { Phased-Out } $\times$ $\Delta\tau_{US,z,t+1}^{MEX,RoW}$					0.84	0.57	0.90
					(0.76)	(0.83)	(1.36)
1 { Others } $\times$ $\Delta\tau_{US,z,t+1}^{MEX,RoW}$					0.04	-0.19	-0.07
					(0.74)	(0.81)	(1.00)
HS6 FE		✓	✓	✓	✓	✓	✓
Observations	7475	6861	6586	7475	6861	6586	

Note: Estimates are obtained from (10) using different within-year periods to construct the growth rate of trade between $\underline{n}$ and $\overline{n}$. In contrast with estimates reported in Table 5, here we use our baseline measure of consumption of imports from (15) as the dependent variable. We restrict the sample to be the same as in that of estimating with $\Delta_{\overline{n}-\underline{n}} m_{zt}^{DD}$. Standard errors, in parentheses, are clustered at HS-2 product level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: Model Simulation Long-Run Elasticities - Baseline measure

	Observed	Observed	Imputed	Imputed
Sample Period:	Years 1 & 4	Years 2 & 3	Years 1 & 4	Years 2 & 3
Dep. Var. :	q_t	q_t	$\tilde{q}_t$	$\tilde{q}_t$
1 { Benchmark } $\times \log(1 + \tau_t)$	-4.18***	-4.34***	-4.13***	-4.39***
	(0.04)	(0.04)	(0.03)	(0.03)
1 { Low I/S } $\times \log(1 + \tau_t)$	-4.05***	-3.97***	-4.00***	-3.97***
	(0.01)	(0.01)	(0.01)	(0.01)
1 { Unanticipated } $\times \log(1 + \tau_t)$	-4.18***	-4.04***	-4.12***	-3.91***
	(0.04)	(0.04)	(0.03)	(0.03)
N	60000	60000	60000	60000
Adjusted R^2	0.48	0.48	0.52	0.52

Note: Estimates are obtained from estimation equation (7) varying the dependent variable and the sample period included. Columns one and two are the same as columns three and four in Table 2. Columns three and four use our baseline predicted measure of consumption of imports, described in (15). Standard errors, in parentheses, are clustered at firm level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 10: Static Elasticities - Baseline

US Imports Mexico	(1)	(2)	(3)	(4)	
Dep. Var. :	$m^{DD}_{z,t}$	$\tilde{q}^{DD}_{z,t}$	$m^{DD}_{z,t}$	$\tilde{q}^{DD}_{z,t}$	Bias
$\tau^{Mex,RoW}_{US,z,t}$	-8.9***	-7.7***			16%
	(1.10)	(1.05)			
1 { Phased-Out } $\times$ $\tau^{Mex,RoW}_{US,z,t}$			-13.2***	-10.9***	21%
			(1.61)	(1.57)	
1 { Others } $\times$ $\tau^{Mex,RoW}_{US,z,t}$			-6.6***	-6.0***	10%
			(1.35)	(1.33)	
Year FE	✓	✓	✓	✓	
HS6 FE	✓	✓	✓	✓	
Observations	15153	15153	15153	15153	

Note: All estimates are obtained form equation 16 and by varying the dependent variable. Columns one and three use imports, while columns two and four use our baseline predicted measure of consumption of imports, described in (15). Standard errors, in parentheses, are clustered at HS-6 level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 11: Dynamic Elasticities - Baseline Measure

Dep. Var. :	$m^{DD}_{z,t}$	$\tilde{q}^{DD}_{z,t}$	Bias
$m^{DD}_{z,t}$	$\tilde{q}^{DD}_{z,t}$	Bias	
Panel A: All Goods			
Short-run (σ^S)	-4.2***	-2.5***	68%
	(1.25)	(0.89)	
Long-run ($-\sigma^L/\alpha$)	-9.7***	-9.2***	5%
	(1.52)	(1.62)	
Panel B: Phaseout Goods			
Short-run (σ^S)	-7.1***	-2.9**	145%
	(1.98)	(1.37)	
Long-run ($-\sigma^L/\alpha$)	-14.0***	-12.0***	17%
	(2.18)	(2.28)	
Year FE	✓	✓	
HS6 FE	✓	✓	
N	11290	11290	
adj. R^2	0.345	0.314	

Note: All estimates are obtained form (17) and by varying the dependent variable. Columns one and three use imports, while columns two and four use our baseline predicted measure of consumption of imports, described in (15). Standard errors, in parentheses, are clustered at HS-6 level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 12: Static Elasticities - Robustness

Dep. Var. :	$\Delta m_{z,t}^{DD}$	$\Delta \tilde{q}_{z,t}^{DD}$				
		Baseline	Time to Market	Common k_z	Demand Shock with 6 Lags	Demand Shock with 6 Lags & Leads
Panel A: All Goods						
$\tau_{US,z,t}^{Mex,Row}$ (σ)	-8.9***	-7.7***	-7.3***	-7.9***	-7.3***	-6.4***
	(1.10)	(1.04)	(1.06)	(1.04)	(1.08)	(1.11)
Bias		16%	22%	13%	22%	39%
Panel B: Phased Out Goods						
$\tau_{US,z,t}^{Mex,Row}$ (σ)	-13.0***	-10.7**	-10.1***	-11.2***	-10.4***	-9.0***
	(1.90)	(1.35)	(1.38)	(1.43)	(1.26)	(1.57)
Bias		21%	29%	16%	25%	44%
Year FE	✓	✓	✓	✓	✓	✓
HS6 FE	✓	✓	✓	✓	✓	✓
N	14646	14646	14646	14646	14646	14646
adj. R^2	0.613	0.697	0.668	0.682	0.691	0.642

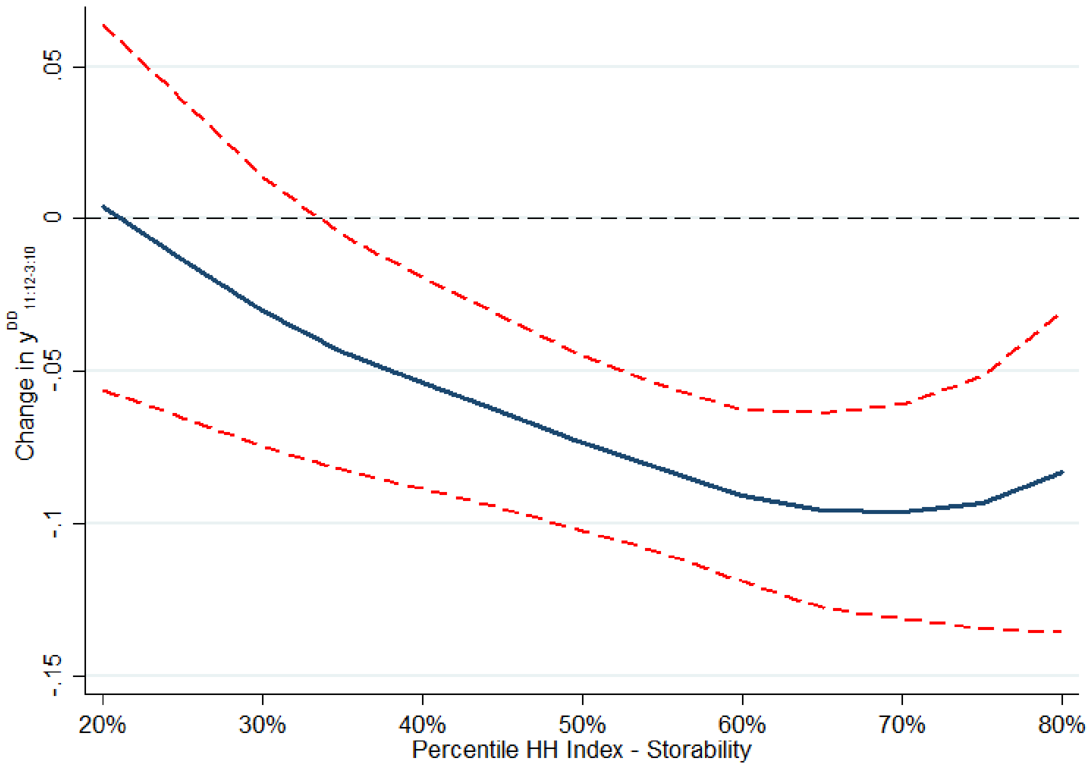
Note: All estimates are obtained form (16) and by varying the dependent variable. Columns one uses actual imports, while columns two onwards uses different measures of consumption of imports, described in section 4.3. Standard errors, in parentheses, are clustered at HS-6 level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 13: Dynamic Elasticities - Robustness

Dep. Var. :	$\Delta m_{z,t}^{DD}$	$\Delta \tilde{q}_{z,t}^{DD}$				
		Baseline	Time to Market	Common k_z	Demand Shock with 6 Lags	Demand Shock with 6 Lags & Leads
Panel A: All Goods						
Short-run (σ^S)	-4.2***	-2.6***	-2.6***	-2.8***	-1.3*	-1.5*
	(1.25)	(0.82)	(0.82)	(0.86)	(0.73)	(0.90)
Long-run ($-\sigma^L/\alpha$)	-9.6***	-9.2***	-9.3***	-9.5***	-9.0***	-9.1***
Panel B: Phased Out Goods						
Short-run (σ^S)	-6.8***	-2.8**	-2.4*	-3.0**	-0.7	-0.7
	(1.90)	(1.35)	(1.38)	(1.43)	(1.26)	(1.57)
Long-run ($-\sigma^L/\alpha$)	-13.7***	-12.0***	-12.0***	-13.0***	-9.8***	-9.7***
Year FE	✓	✓	✓	✓	✓	✓
HS6 FE	✓	✓	✓	✓	✓	✓
N	11019	11019	11019	11019	11019	11019
adj. R^2	0.342	0.293	0.342	0.293	0.342	0.293

Note: All estimates are obtained form (17) and by varying the dependent variable. Columns one and three use actual imports, while columns two and four use different proxies of consumption of imports, described in section 4.3. Robust standard errors are in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 1: Anticipatory Elasticity and Storability

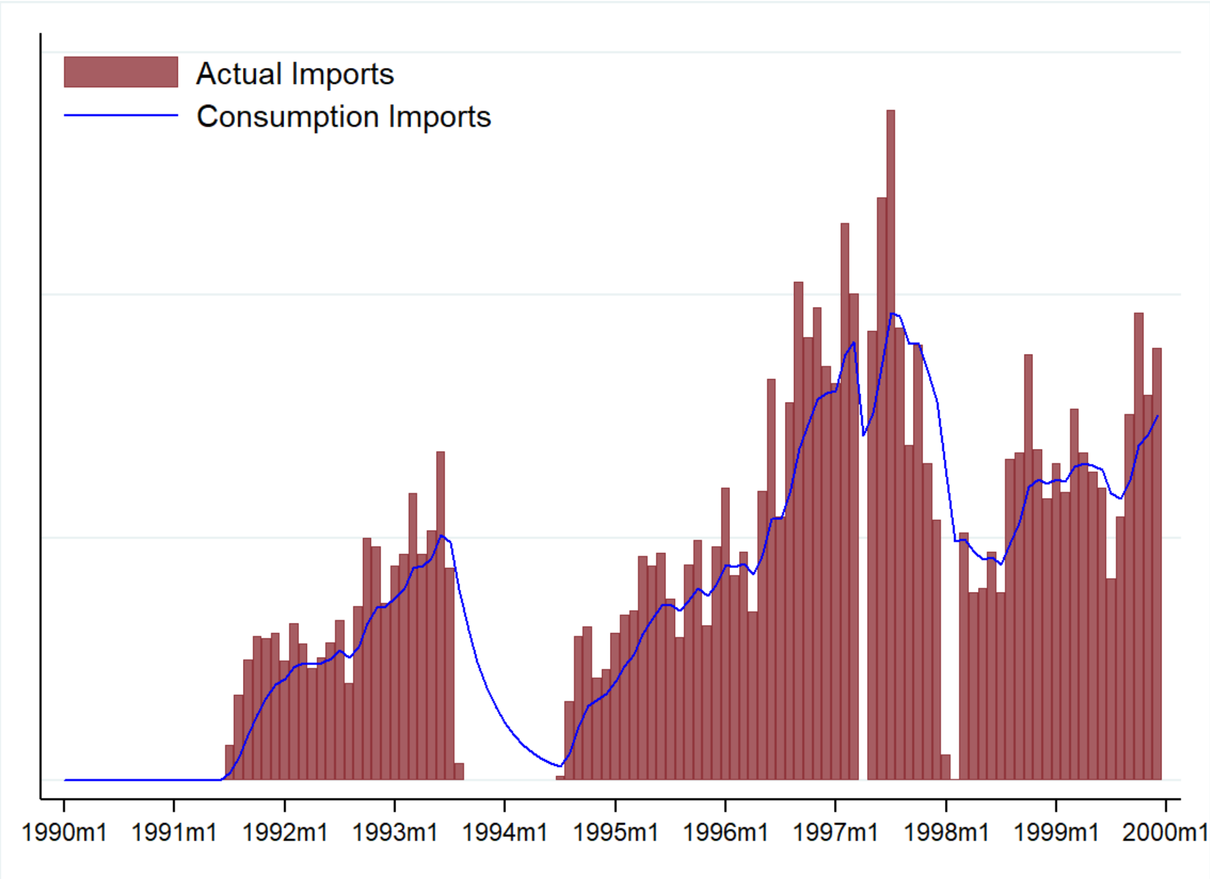


Note: The yaxis corresponds to the predicted $\Delta y^{DD}_{z,t,11:12-3:10}$ from estimating equation (13) given $\Delta\tau^{DD}_{t+1} = -0.01$ and at different percentiles of the HH Index, our measure of storability. The sample includes only HS-6 goods that were phased out. It is calculated as $\hat{\sigma}^A = \hat{\sigma}^A_0 \times \Delta\tau^{DD}_{t+1} + \hat{\sigma}^A_1 \times \Delta\tau^{DD}_{t+1} \times P(HH) + \hat{\sigma}^A_2 \times \Delta\tau^{DD}_{t+1} \times P(HH)^2$. The xaxis is the $P(HH)$. The estimation results with coefficients for $\sigma^A_{i=\{0,1,2\}}$, are reported in Table A.3 of the Appendix. The HH index used here is calculated taking the median HH index of annual imports of goods at HS-10, district of entry and source country level in the second year the good appears in the sample, excluding Canada and Mexico. The red dashed lines are the 95% confidence interval, calculated using the delta method.

Figure 2: Impulse Response Function of Aggregate Variables in the 3 Simulations

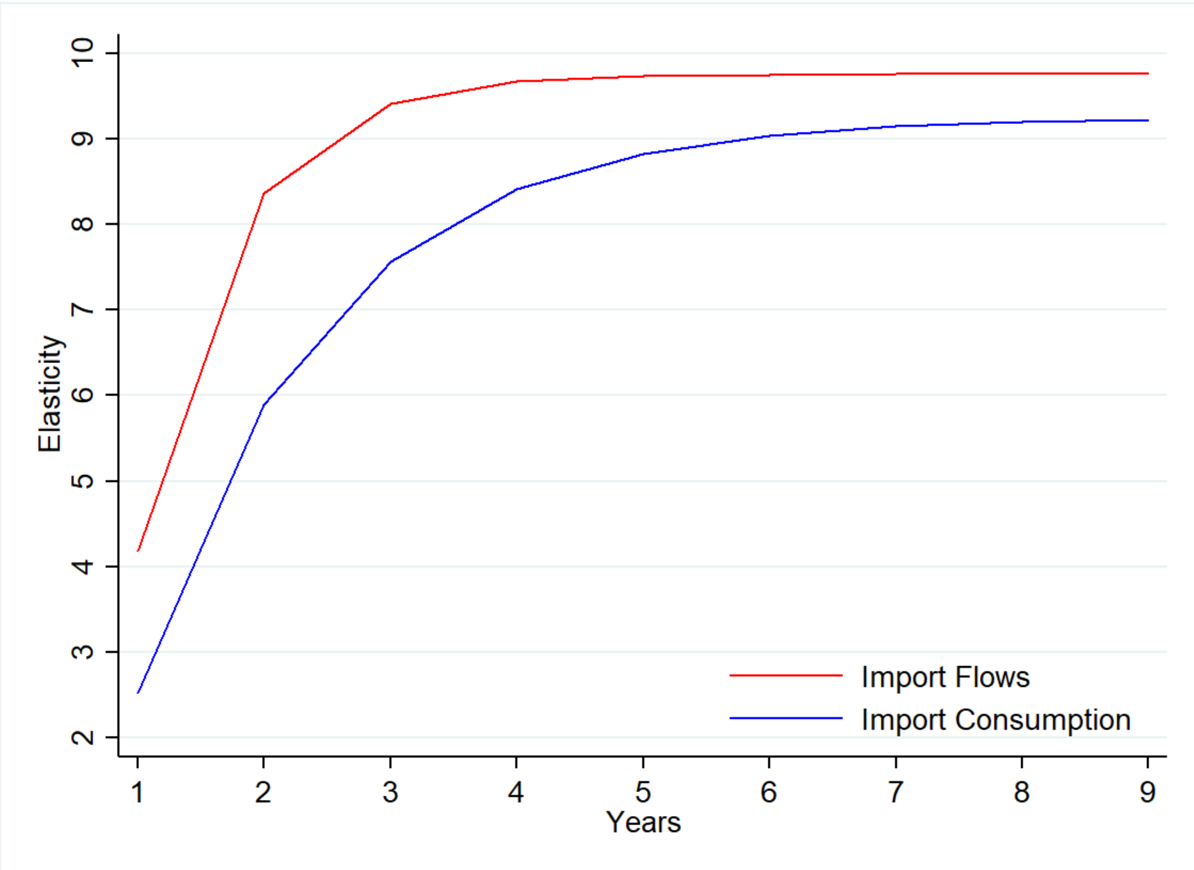
Note: In all four panels the yaxis is log changes with respect to the average level of the corresponding variable in the first 12 months of the simulation. All variables are the summation of firm-level values of purchases, sales, end of period inventory holdings and prices charged to consumers, respectively. The shock corresponds a 1pp decrease in tariffs in month 13. Parameter values used to obtain these responses are specified in Table 1.

Figure 3: Example of our Baseline Measure



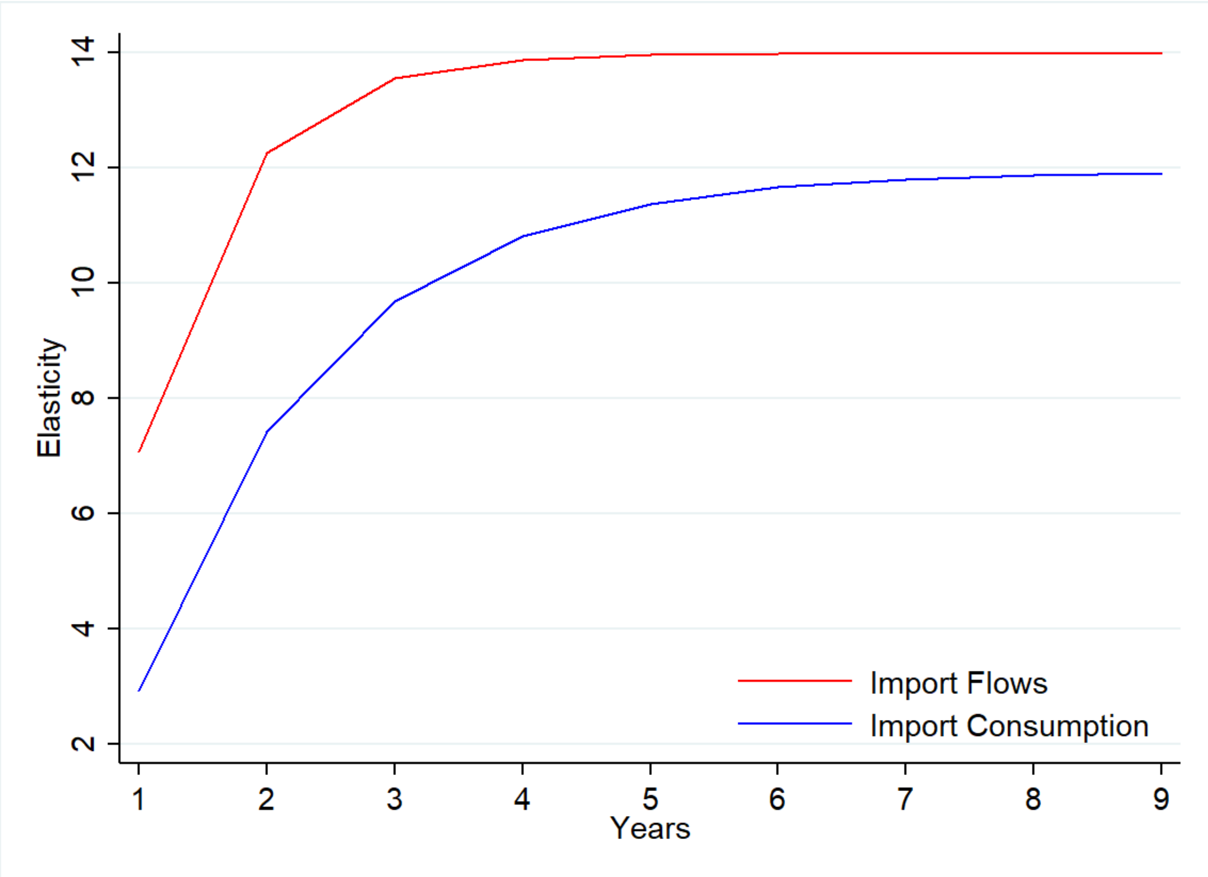
Note: The yaxis is in levels. The example here shown corresponds to HS6 code 870421, described as "Vehicles; compression-ignition internal combustion piston engine (diesel or semi-diesel), for transport of goods, (of a gvw not exceeding 5 tonnes)". Consumption of imports are calculated as in (15).

Figure 4: Dynamic Elasticity - All HS6 Goods



Note: On the y-axis are the trade elasticity estimated obtained from (17) for $N = [1, 5]$. The sample only all HS-6 goods. The red line uses actual imports as the dependent variable, while the blue line uses our baseline measure of consumption of imports, described in (15). The dashed lines are the 68% confidence interval. Standard errors are robust.

Figure 5: Dynamic Elasticity - Phased-Out Goods



Note: On the y-axis are the trade elasticity estimated obtained from (17) for $N = [1, 5]$. The sample only includes HS-6 goods that were phased out. The red line uses actual imports as the dependent variable, while the blue line uses our baseline measure of consumption of imports, described in (15). The dashed lines are the 68% confidence interval. Standard errors are robust

Appendix

A Tables and Figures

Table A.1: Sample Characteristics

Categories	HS6-Goods	HS6-Year Pair	%	Mexico's Import Share 1990-1993	Median Duty 1993
Phased-out	451	5,283	16.51%	12.33%	7,71 %
Class A	129	1,541	4.81%	14.04%	6.50%
Non NAFTA	2,274	15,182	78.68%	73.63%	0.15 %
TOTAL	2,854	32,006	100%	0.81%	

Table A.2: Reference Exporter Countries

Afghanistan	Gabon	Norfolk Is	Angola	Gambia	North Korea
Antigua Barbuda	Ghana	Norway	Argentina	Greenland	Oman
Aruba	Grenada Is	Pakistan	Australia	Guatemala	Palau
Bahamas	Guinea	Panama	Bahrain	Guinea-Bissau	Papua New Guin
Bangladesh	Guyana	Paraguay	Barbados	Haiti	Peru
Belize	Honduras	Philippines	Benin	Hong Kong	Pitcairn Is
Bermuda	India	Qatar	Bhutan	Indonesia	Rwanda
Bolivia	Iran	Samoa	Botswana	Jamaica	Saudi Arabia
Brazil	Japan	Senegal	Brunei	Kenya	Seychelles
Burkina Faso	Kiribati	Sierra Leone	Burundi	Korea	Singapore
Cambodia	Laos	Solomon Is	Cameroon	Lesotho	Somalia
Cape verde	Liberia	Sri Lanka	Cayman Is	Libya	St Kitts-Nevis
Cen African Rep	Macao	St Lucia Is	Chad	Madagascar	St Vinc & Gren
Chile	Malawi	Sudan	China	Malaysia	Suriname
Christmas Is	Maldive Is	Swaziland	Cocos Is	Mali	Switzerland
Colombia	Marshall Is	Taiwan	Comoros	Mauritania	Tanzania
Congo (DROC)	Mauritius	Thailand	Congo (ROC)	Mongolia	Togo
Cook Is	Montserrat Is	Tonga	Costa Rica	Mozambique	Trin & Tobago
Cote d'Ivoire	Namibia	Tuvalu	Cuba	Nauru	Uganda
Djibouti	Nepal	United Arab Em	Dominica Is	Netherlands Ant	Uruguay
Dominican Rep	New Caledonia	Venezuela	Ecuador	New Zealand	Vietnam
El Salvador	Nicaragua	Yemen	Eq Guinea	Niger	Zambia
Ethiopia	Nigeria	Zimbabwe	Fiji	Niue	

Table A.3: Anticipatory Elasticities interacted with Trade Lumpiness

Dep. Var.: $\Delta_{n:n'} m_{zt}^{DD}$	n n'	Oct-Dec Apr-Sep	Nov-Dec Mar-Oct	Nov-Dec May-Oct	Oct-Dec Apr-Sep	Nov-Dec Mar-Oct	Nov-Dec May-Oct
$\Delta\tau_{US,z,t+1}^{MEX,RoW}$		-42.4***	-22.8	-25.7*			
		(14.05)	(13.70)	(13.64)			
$\Delta\tau_{US,z,t+1}^{MEX,RoW} \times HH_z$		280.0***	157.5*	174.1**			
		(92.18)	(84.61)	(85.67)			
$\Delta\tau_{US,z,t+1}^{MEX,RoW} \times HH_z^2$		-415.6***	-244.1*	-266.7**			
		(156.46)	(129.00)	(133.89)			
Phased-Out $\times$ $\Delta\tau_{US,z,t+1}^{MEX,RoW}$					-53.4***	-71.0***	-78.0***
					(18.27)	(22.33)	(25.74)
Phased-Out $\times$ $\Delta\tau_{US,z,t+1}^{MEX,RoW} \times HH_z$					372.6***	478.6***	518.5***
					(99.47)	(138.39)	(173.83)
Phased-Out $\times$ $\Delta\tau_{US,z,t+1}^{MEX,RoW} \times HH_z^2$					-573.8***	-710.3***	-775.6***
					(132.29)	(215.61)	(293.30)
HS6 FE		✓	✓	✓	✓	✓	✓
Observations		7076	6453	6285	7076	6453	6285
Adjusted R^2		0.090	0.121	0.104	0.093	0.122	0.105

Note: Standard errors, in parentheses, are clustered at HS-2 level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.4: Elasticities Consumer vs. Non-Consumer Goods

US Imports Mexico	(1)	(2)	(3)	(4)	
Dep. Var. :	$m_{z,t}^{DD}$	$\tilde{q}_{z,t}^{DD}$	$\Delta m_{z,t}^{DD}$	$\Delta \tilde{q}_{z,t}^{DD}$	Bias
$1\{$ Non-Consumer Goods $\} \times \tau_{US,z,t}^{Mex,RoW}$	-10.4***	-8.15***			28%
	(1.60)	(1.59)			
$1\{$ Consumer Goods $\} \times \tau_{US,z,t}^{Mex,RoW}$	-7.90***	-7.43***			6%
	(1.41)	(1.36)			
$1\{$ Non-Consumer Goods $\} \times \Delta\tau_{US,z,t}^{Mex,RoW}$			-6.82***	-4.31***	58%
			(1.70)	(1.17)	
$1\{$ Consumer Goods $\} \times \Delta\tau_{US,z,t}^{Mex,RoW}$			-2.28*	-1.68*	36%
			(1.33)	(0.96)	
Year FE	✓	✓	✓	✓	
HS6 FE	✓	✓	No	No	
N	15800	15800	11693	11693	
Adjusted R^2	0.02	0.02	0.003	0.006	

Note: Estimates in column 1 and 2 are obtained form equation 16. Estimates in column 1 and 2 are obtained form equation 17. Columns two and four use our baseline predicted measure of consumption of imports, described in (15). HS-6 goods are classified into consumer goods according to the BEC industry classification. Standard errors in columns one and two, in parentheses, are clustered at HS-6 level, and in columns three and four are robust. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.5: Model Simulation - Measure with Demand Shock in Long Run Elasticity

	Simulated	Simulated	Proxy	Proxy
Sample Period:	Years 1 & 4	Years 2 & 3	Years 1 & 4	Years 2 & 3
Dep. Var. :	q_t	q_t	$\tilde{q}_t$	$\tilde{q}_t$
1 { Benchmark } $\times \log(1+\tau_t)$	-4.18***	-4.34***	-4.10***	-4.43***
	(0.04)	(0.04)	(0.03)	(0.03)
1 { Low I/S } $\times \log(1+\tau_t)$	-4.05***	-3.97***	-4.00***	-3.99***
	(0.01)	(0.01)	(0.01)	(0.01)
1 { Unanticipated } $\times \log(1+\tau_t)$	-4.18***	-4.04***	-4.10***	-3.76***
	(0.04)	(0.04)	(0.03)	(0.03)
N	60000	60000	60000	60000
Adjusted R^2	0.48	0.48	0.58	0.58

Note: All estimates are obtained form equation 7 and by varying the dependent variable and the sample period included. Columns one and three are the same as columns three and four in Table 2. Columns three and four use as the measure of consumption of imports that includes a demand shock with 6 lags in the expected value of imports, described in (19) and (20). Standard errors, in parentheses, are clustered at firm level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.6: Dynamic Elasticity Estimates

Dep. Var. :	(1) $\Delta m_{z,t}^{DD}$	(2) $\Delta \tilde{q}_{z,t}^{DD}$	(3) $\Delta m_{z,t}^{DD}$	(4) $\Delta \tilde{q}_{z,t}^{DD}$
$\Delta\tau_{US,z,t}^{Mex,RoW}$	-4.19***	-2.53***		
	(1.25)	(0.89)		
$\tau_{US,z,t-1}^{Mex,RoW}$	-7.31***	-4.62***		
	(1.15)	(0.81)		
$1\{Phased\} \times \Delta\tau_{US,z,t}^{Mex,RoW}$			-7.07***	-2.92**
			(1.98)	(1.37)
$1\{Phased\} \times \tau_{US,z,t-1}^{Mex,RoW}$			-10.5***	-5.97***
			(1.64)	(1.14)
$1\{Others\} \times \Delta\tau_{US,z,t}^{Mex,RoW}$			-2.70*	-2.18*
			(1.48)	(1.11)
$1\{Others\} \times \tau_{US,z,t-1}^{Mex,RoW}$			-5.57***	-3.89***
			(1.43)	(1.04)
$y_{z,t-1}^{DD}$	-0.75***	-0.50***	-0.75***	-0.50***
	(0.01)	(0.01)	(0.01)	(0.01)
Year FE	✓	✓	✓	✓
HS6 FE	✓	✓	✓	✓
N	11290	11290	11290	11290
adj. R^2	0.344	0.314	0.345	0.314

Note: All estimates are obtained from (17). Columns one and three use the actual import flows as dependent variable. Columns two and four use the consumption of imports as the dependent variable. The last row shows the speed of adjustment towards the long-run value. Standard errors, in parentheses, are clustered at HS6 level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A.1: Share of US Imports by Mexican Imports 1990-1999

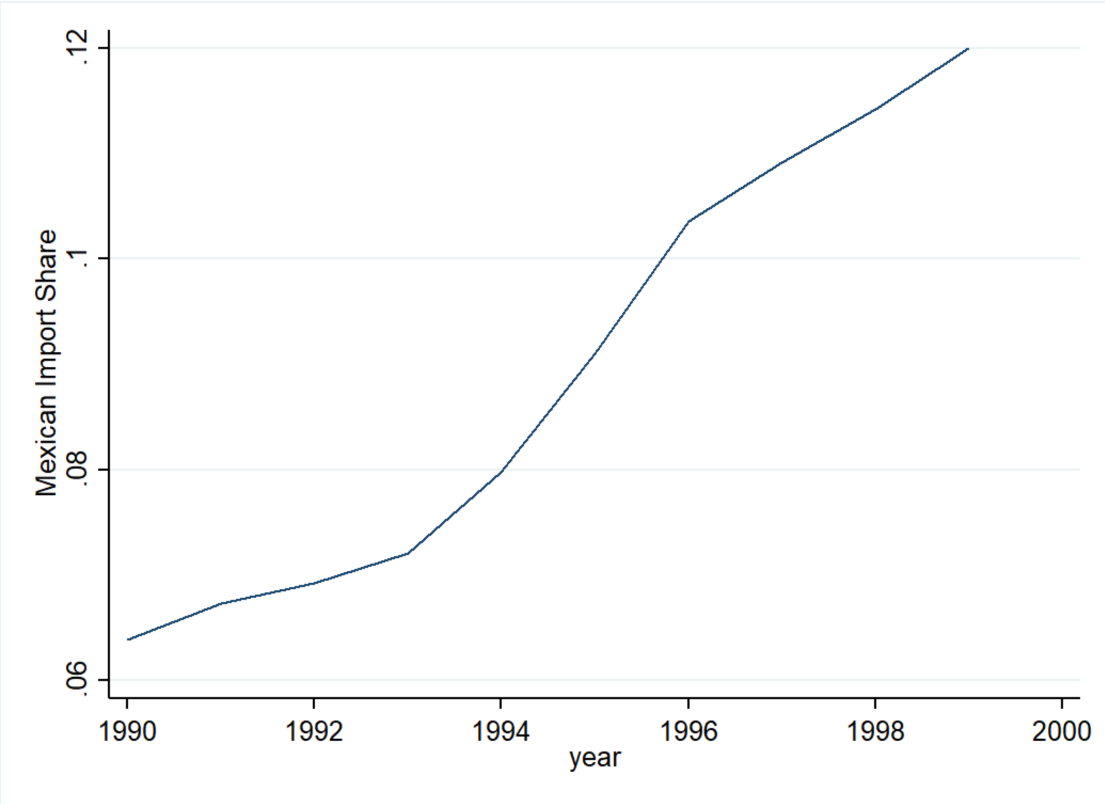


Figure A.2: Median HS-8 Scheduled US Tariff Rates on Mexican Imports by phaseout Classification

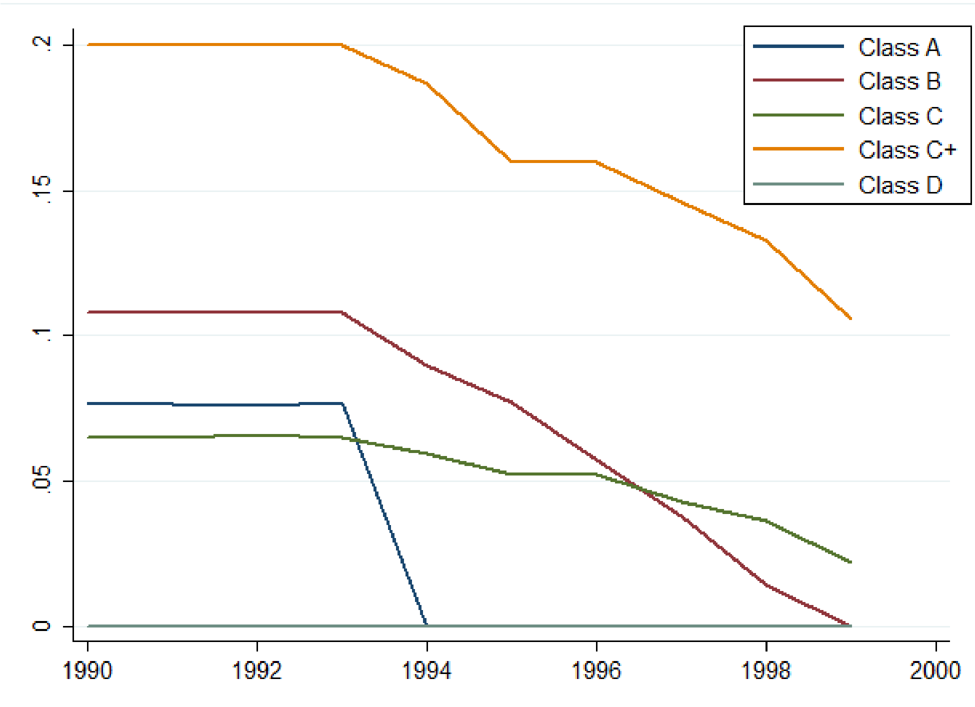
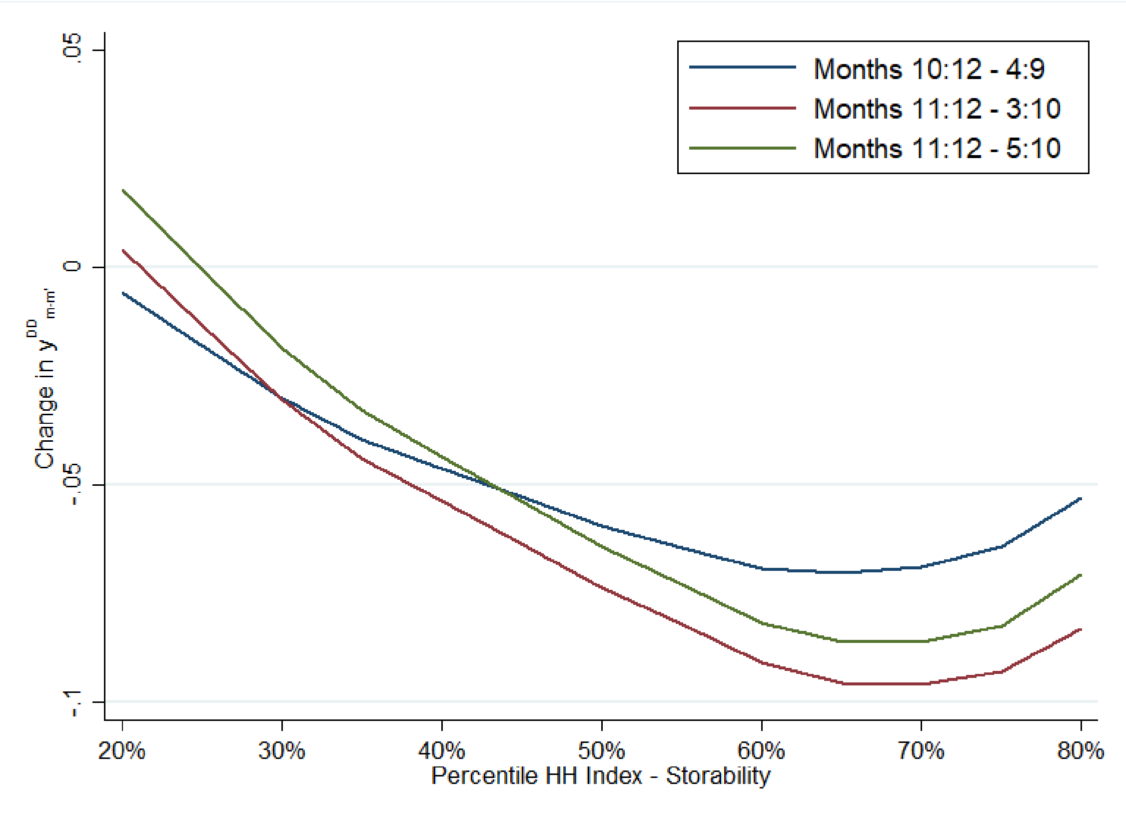
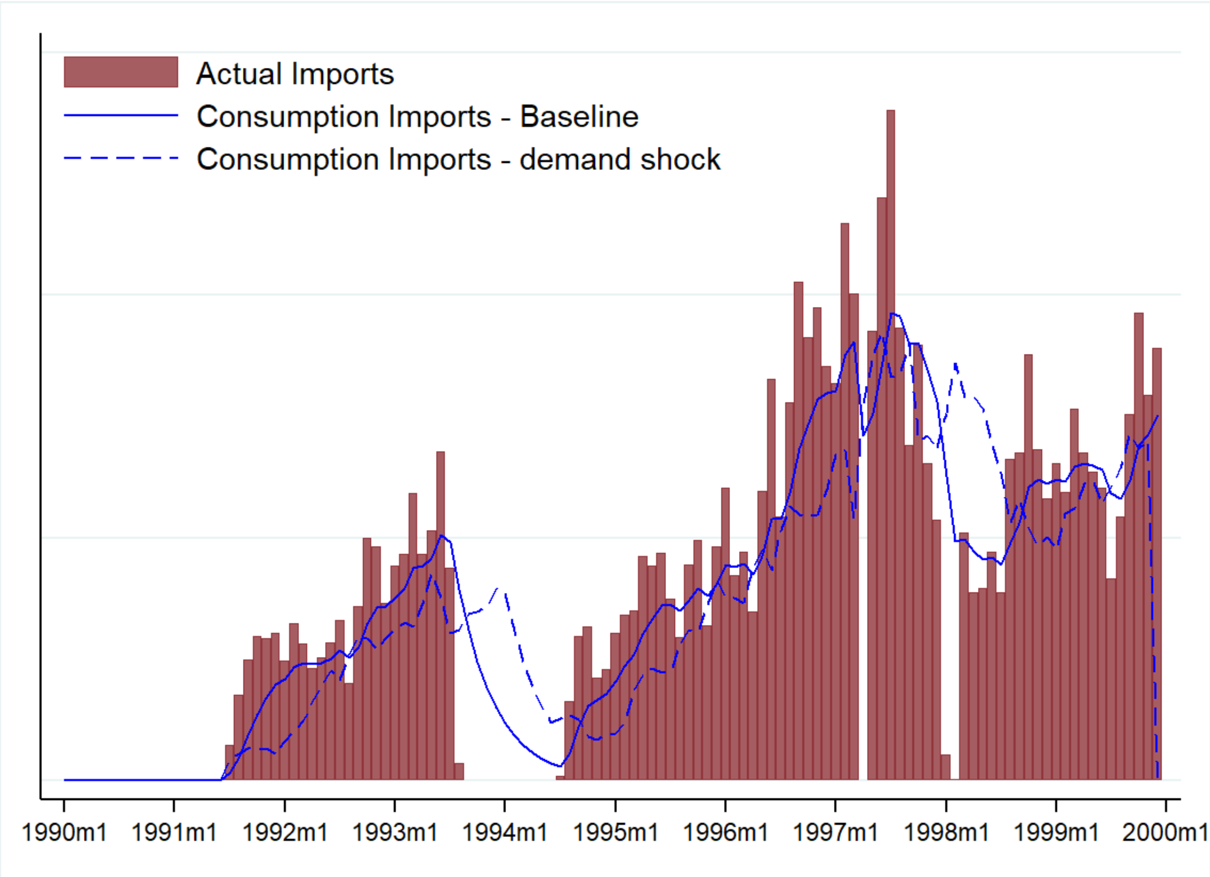


Figure A.3: Anticipatory Elasticity and Storability for different $\Delta y^{DD}_{t,m-m'}$



Note: The yaxis corresponds to the predicted $\Delta y^{DD}_{z,t,m-m'}$ from estimating equation 13 given $\Delta\tau^{DD}_{t+1} = -0.01$ and at different percentiles of the HH Index, our proxy of storability. It is calculated as $\hat{\sigma}^A = \hat{\sigma}^A_0 \times \Delta\tau^{DD}_{t+1} + \hat{\sigma}^A_1 \times \Delta\tau^{DD}_{t+1} \times P(HH) + \hat{\sigma}^A_2 \times \Delta\tau^{DD}_{t+1} \times P(HH)^2$. The xaxis is the $P(HH)$. The estimation results with coefficients for $\sigma^A_{i=\{0,1,2\}}$, are reported in Table A.3 of the Appendix. The HH index used here is calculated taking the median HH index of annual imports of goods at HS-10, district of entry and source country level in the second year the good appears in the sample, excluding Canada and Mexico.

Figure A.4: Example of our Measure with Demand Shocks



Note: The yaxis is in levels. The example here shown corresponds to HS6 code 870421, described as “Vehicles; compression-ignition internal combustion piston engine (diesel or semi-diesel), for transport of goods, (of a gvw not exceeding 5 tonnes)”. Consumption of imports are calculated as in (19).

B Difference-in-Difference Approach

In the background of the demand equation (1) is a representative consumer whose period-t utility is given by a Cobb-Douglas aggregator over products:

$$C_{it} = \prod_z C_{izt}^{\alpha_{iz}}$$

where α_{iz} is the expenditure share of product z in country i and C_{izt} denotes the total consumption of good z in country i in period t. Differentiated varieties of product z can be sourced from different countries. These varieties are combined in a Dixit-Stiglitz form,

$$C_{izt} = \left(\sum_c m_{iczt}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$$

where m_{icz} is the quantity of product z consumed in country i sourced from exporter c in period t. This leads to an import demand equation given by,

$$m_{iczt} = \left(\frac{p_{czt}(1 + \tau_{iczt})}{P_{izt}} \right)^{-\sigma} C_{izt}$$

where P_{izt} is the importer-specific price index given by $P_{izt} = \left(\sum_c (p_{czt}(1 + \tau_{iczt}))^{1-\sigma} \right)^{1/(1-\sigma)}$. We transform import demand equation into Free-On-Board (FOB) value to overcome measurement problems in quantities.[42] We multiplying both sides of the import demand equation by the unit price i.e. $v_{iczt} = m_{iczt} p_{czt}$. In what comes next, we aim to remove the exporter-specific unit-price term and the importer-specific price indices and aggregate demand term.

Next we take the ratio of country i's import demand from country c and c', thereby

[42] See Hillbery and Hummels (2012).

eliminating importer i-specific effects:

$$\frac{v_{iczt}}{v_{ic'zt}} = \left(\frac{(1+\tau_{iczt})}{(1+\tau_{ic'zt})}\right)^{-\sigma} \left(\frac{p_{czt}}{p_{c'zt}}\right)^{1-\sigma}$$

Secondly, considering a reference exporter i' and taking logs on both sides elimiates the exporter-specific price terms:

$$\ln\left(\frac{v_{iczt}}{v_{ic'zt}} \Big/ \frac{v_{i'czt}}{v_{i'c'zt}}\right) = -\sigma \ln\left(\frac{1+\tau_{iczt}}{1+\tau_{ic'zt}} \Big/ \frac{1+\tau_{i'czt}}{1+\tau_{i'c'zt}}\right) \tag{21}$$

C General Equilibrium Considerations

C.1 Anticipatory Effects in Exporter Prices

In the model in section 2.2 we assume that the price exporters charge, ω is exogenous and constant, which is the same as assuming perfectly competitive suppliers[43]. In reality, however, it could be the case that when suppliers observe a temporary drop in their sales they offered discounts countervailing the incentives of anticipation to tariff reductions. Similarly, it could be that the drop in utilization of shipment infrastructure results in a price drop from lower transportation costs. We test this by considering changes in the HS-6 unit values of Mexican exports. To control for Mexican fixed effects we take the ratio of the unit values of imports to the US and EU12[44]. As in our approach to estimating anticipatory elasticities, we consider within-year growth rates between sub-periods m and

[43] In the model, introducing an ad-hoc supply elasticity only changes the results reported in Table **??** quantitatively.

[44] Additionally, we considered unit values of Mexican alone and the overall result does not change.

m'. The estimation equation is:

$$\Delta UV^{D}_{z,t,m-m'} \equiv \log\left(\frac{UV^{Mex,US}_{z,t,m}}{UV^{Mex,EU12}_{z,t,m}} \Big/ \frac{UV^{Mex,US}_{z,t,m'}}{UV^{Mex,EU12}_{z,t,m'}}\right) = \sigma^{A,UV}\Delta\tau^{Mex,US}_{z,t+1} + \delta_{z,m-m'} + u_{z,t}$$

$\sigma^{A,UV}$ is the anticipatory elasticity of unit values to an upcoming tariff change. This estimation equation is very similar to 10. We restrict the sample to the sample considered in Table 5, in which we found anticipatory effects in phased-out goods using the triple difference approach. Calculating unit values further restricts the sample since quantities are generally less available than values. Results are reported in Table C.1. Neither in the aggregate, nor for phased-out goods, we observe any significant movement in unit values before an upcoming tariff change. When we don't restrict the sample results are the same.

Table C.1: Anticipatory Effects in Mexican Exporter Prices

Dep. Var.: $\Delta_{n:n'}m^{DD}_{zt}$	n n'	Oct-Dec Apr-Sep	Nov-Dec Mar-Oct	Nov-Dec Mar-Aug	Oct-Dec Apr-Sep	Nov-Dec Mar-Oct	Nov-Dec Mar-Aug
$\Delta\tau^{MEX,RoW}_{US,z,t+1}$		-0.03	-0.10	-0.24			
		(0.96)	(0.77)	(0.83)			
1 { Phased-Out } $\times$ $\Delta\tau^{MEX,RoW}_{US,z,t+1}$					0.40	1.11	1.21
					(1.03)	(1.27)	(1.14)
Observations		4977	4371	4290	4977	4371	4290
Adjusted R^2		0.04	0.05	0.06	0.04	0.05	0.06

Note: Standard errors, in parentheses, are clustered at HS-2 level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.2 Trade Diversion: Spillovers to Canadian Exports to the US?

Another interesting question that arises given the drop in exports from Mexico to the US in anticipation of upcoming tariff reductions is whether these spill over to exports from other trading partner. If the goods in question are complementary, e.g. they are intermediates combined in the importers' production process, the imports of the non-Mexican good would also drop. Alternatively, if the goods are substitutes, consumers will increase their purchases of the non-Mexican goods if the Mexican good becomes unavailable before the tariff drop. To gain some insights on these general equilibrium effects, we study how imports to the US from Canada respond to the anticipatory drop in imports from Mexico of phased-out goods. We estimate the following regression:

$$\Delta y^{DD,CAN}_{z,t,m-m'} = \sigma^{A}\ \Delta x^{DD,CAN}_{z,t+1} + \sigma^{A,Diversion}\ \Delta x^{DD,MEX}_{z,t+1} \delta_{z,m-m'} + u_{z,t}$$

This is very similar to **??**, except that we now include the change in τ^{DD} for another trading partner and we are interested not in σ^{A}, but in $\sigma^{A,Diversion}$. Naturally we distinguish those goods for which we found significant drops in imports, that is the goods for which tariffs on Mexico were phased-out. Results in Table C.2 show that imports for those goods on which Mexican imports declined, Canadian exports increased, that is, we find evidence of trade diversion.

Table C.2: Response of Canadian Exports to the US to Mexican Tariff phaseouts

Dep. Var.: $\Delta_{n:n'} m_{zt}^{DD}$ n / n'	Oct-Dec Apr-Sep	Nov-Dec Mar-Oct	Nov-Dec May-Oct	Oct-Dec Apr-Sep	Nov-Dec Mar-Oct	Nov-Dec May-Oct
$\Delta\tau_{US,z,t+1}^{MEX,RoW}$	-0.65	-0.34	-0.85			
	(1.10)	(0.89)	(1.03)			
$\Delta\tau_{US,z,t+1}^{CAN,RoW}$	-1.24	0.30	0.14			
	(1.65)	(1.46)	(1.67)			
Phased-Out $\times$ $\Delta\tau_{US,z,t+1}^{MEX,RoW}$				-2.18	-2.89**	-3.38***
				(1.54)	(1.14)	(1.01)
Class A $\times$ $\Delta\tau_{US,z,t+1}^{MEX,RoW}$				0.16	3.36	2.52
				(2.19)	(2.53)	(2.82)
Non NAFTA $\times$ $\Delta\tau_{US,z,t+1}^{MEX,RoW}$				0.03	0.30	-0.16
				(1.50)	(1.10)	(1.34)
Phased-Out $\times$ $\Delta\tau_{US,z,t+1}^{CAN,RoW}$				-0.65	3.16	3.67
				(1.90)	(3.21)	(3.31)
Class A $\times$ $\Delta\tau_{US,z,t+1}^{CAN,RoW}$				1.14	-3.87	-4.31
				(5.84)	(7.76)	(8.48)
Non NAFTA $\times$ $\Delta\tau_{US,z,t+1}^{CAN,RoW}$				-2.03	-0.70	-1.18
				(2.13)	(1.68)	(1.95)
HS6 FE	✓	✓	✓	✓	✓	✓
Observations	13270	12379	12213	13270	12379	12213
Adjusted R^2	0.099	0.098	0.101	0.099	0.099	0.101

Note: Standard errors, in parentheses, are clustered at HS-2 level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D Model: Closed Form of Anticipation

D.1 Solving the Model

This version has no uncertainty so the primary reason of holding inventory is to avoid fixed cost of importing and the import demand equation is given by $q = p^{-\sigma}$.

$$
\begin{aligned}
V(s) &= \max[V^a(s), V^n(s)] \\
&\text{where} \\
V^a(s) &= \max_{p,i} = pq - \omega i - f + \beta V(s') \\
V^n(s) &= \max_{p} = pq + \beta V(s') \\
s.t. \quad s' &= (1-\delta)(s + i - q)
\end{aligned}
$$

Solving this gives the ($\underline{s}$,$\overline{s}$) ordering policy and the policy functions for price and orders,

$$
\begin{aligned}
[p:] \quad p &= \frac{\sigma}{\sigma - 1} V_1(s) && (22) \\
q &= \left[\frac{\sigma}{\sigma - 1} V_1(s)\right]^{-\sigma} \\
[i:] \quad \omega &= \beta(1-\delta) V_1((1-\delta)(s + i - q)) && (23)
\end{aligned}
$$

We start solving analytically by first characterizing the value and policy functions in the no-ordering period. Using the guess, $V^n(s) = \alpha s^{(\sigma-1)/\sigma}$, in the no-order period we have;

$$V^n(s) = \max_{p} p^{1-\sigma} + \beta\alpha[(1-\delta)(s - p^{-\sigma})]^{(\sigma-1)/\sigma} \tag{24}$$

This gives a pricing and sales function,

$$q = p^{-\sigma} = \frac{(\alpha\beta)^{-\sigma}(1-\delta)^{(1-\sigma)}s}{1+(\alpha\beta)^{-\sigma}(1-\delta)^{1-\sigma}}$$

Which when put into (24), gives us the value of α implicitly as $\alpha^{-\sigma} = 1-\beta^{\sigma}(1-\delta)^{\sigma-1}$. We get the sales function by plugging the value function in (22),

$$\begin{aligned} q = p^{-\sigma} &= \alpha^{-\sigma}s \qquad (25) \\ &= (1-\beta^{\sigma}(1-\delta)^{\sigma-1})s \end{aligned}$$

Here, (25) denotes the analytical form for the inventory sales ratio in this model. Next, we use law of motion for stock to get,

$$\begin{aligned} s' &= (1-\delta)(s-q) \qquad (26) \\ &= (1-\delta)s(1-\alpha^{-\sigma}) \\ &= \beta^{\sigma}(1-\delta)^{\sigma}s \end{aligned}$$

We can also derive the number of periods it takes for a firm to run down its inventory to a given level or the time interval between orders using (26).

Now given the $V^n(s)$, we characterize the $V^a(s)$ which involves a fixed cost. Here, we assume that firms do not order in the consecutive periods.

$$\begin{aligned} V^a(s) &= \max_{p,i} \quad p^{1-\sigma} - \omega i - f + \beta V^n((1-\delta)(s+i-p^{-\sigma})) \qquad (27) \\ &= \max_{p,i} \quad p^{1-\sigma} - \omega i - f + \beta\alpha\left[(1-\delta)(s+i-p^{-\sigma})\right]^{(\sigma-1)/\sigma} \end{aligned}$$

We obtain the price function from this $p = \omega\sigma/(\sigma-1)$ which implies that $V_1^a(s) = \omega$ i.e. the marginal value of stock in the period the firm orders is equal to the purchase price. In contrast, when the firm is not ordering the marginal value of stock decreases with stock holding so the firm sells stock in order to have the same discounted marginal value every period. Building on this intuition, the firm sells everything it started with in the ordering period and orders the amount equal to $\overline{s}$ i.e. $\underline{s}$ = q in the ordering period. This implies $\underline{s} = (\omega\sigma/(\sigma-1))^{-\sigma}$. We get the optimal i from (27),

$$\begin{aligned} i(s) &= \left(\frac{\omega\sigma}{\sigma-1}\right)^{-\sigma}\alpha^{\sigma} - s \\ &= \left(\frac{\omega\sigma}{\sigma-1}\right)^{-\sigma}\alpha^{\sigma} - \underline{s} \\ &= \left(\frac{\omega\sigma}{\sigma-1}\right)^{-\sigma}(\alpha^{\sigma}-1) \end{aligned} \tag{28}$$

We plug the optimal price and order in (27) to obtain the functional form of $V^a(s)$,

$$\begin{aligned} V^a(s) &= \left(\frac{\omega\sigma}{\sigma-1}\right)^{1-\sigma}\frac{\alpha^{\sigma}}{\sigma} - \omega s - f \\ &= \left(\frac{\omega\sigma}{\sigma-1}\right)^{1-\sigma}\left(\frac{\alpha^{\sigma}-(\sigma-1)}{\sigma}\right) - f \end{aligned}$$

D.2 Anticipation Effects

Suppose that the firm begins the period t with inventory $\underline{s}$ which means it sells $\underline{s}$ and must order i amount this period given by (28). At the beginning of the period, the firm learns that price of the good is falling next period i.e. $\omega_{t+1} < \omega_t$. We present here a simple case of how anticipation would effect purchases in this period by assuming a 1-period transition. The firm would try to be at the new lower-limit of the inventory holding policy, denoted by $\underline{s}$ in period t+1 so that it can order according to the new optimal policy i.e. the firm plans to have $s_{t+1} = \left(\frac{\sigma\omega_{t+1}}{\sigma-1}\right)^{-\sigma}$. However, since the problem becomes non-stationary in the transition, the firm will not sell all the beginning period t inventories. We now solve for the optimal sales and orders in this transition period[45]. For what follows, we denote by $\tilde{i}$ the purchases after incorporating the knowledge of the future change. After the shock, the firms solve

$$\begin{aligned} V^a(s_t) &= \max_{p_t,\tilde{i}_t>0} p_t^{1-\sigma} - \omega_t\tilde{i}_t - f + \beta V^a_{t+1}[(1-\delta)(s_t + \tilde{i}_t - p_t^{-\sigma})] \qquad (29) \\ &= \max_{p_t,\tilde{i}_t>0} p_t^{1-\sigma} - \omega_t\tilde{i}_t - f + \beta\left[\left(\frac{\sigma\omega_{t+1}}{\sigma-1}\right)^{-\sigma}\left(\frac{\alpha^\sigma}{\sigma}\right) - \omega_{t+1}(1-\delta)(s_t + \tilde{i}_t - p_t^{-\sigma}) - f\right] \end{aligned}$$

This outputs the policy functions,

$$p_t = \beta(1-\delta)\frac{\omega_{t+1}\sigma}{\sigma-1} \qquad (30)$$

$$\tilde{i}_i = \left(\frac{\omega_t\sigma}{\sigma-1}\right)^{-\sigma}\left[\left(\frac{\omega_{t+1}}{\omega_t}\right)^{-\sigma}\frac{(1+\beta^{-\sigma}(1-\delta)^{1-\sigma})}{(1-\delta)} - 1\right] \qquad (31)$$

Considering that without anticipation $i_t = (\omega_t\sigma/\sigma-1)^{-\sigma}(\alpha^\sigma - 1)$ and taking change in imports with and without the anticipation, we get

$$\frac{\tilde{i}_t}{i_t} = (\alpha^\sigma - 1)^{-1}\left[\left(\frac{\omega_{t+1}}{\omega_t}\right)^{-\sigma}\frac{(1+\beta^{-\sigma}(1-\delta)^{1-\sigma})}{(1-\delta)} - 1\right] \qquad (32)$$

[45] Although not ordering in period t might dominate ordering any amount, here we are imposing that the firm orders. This provides an analogue of the richer model with the demand shocks where firms may be forced to order in period t if they have high enough demand shock.

We can see in (32) how the term in bracket from (31) affects the purchases.

$$\frac{\tilde{i}_t}{i_t} < 1 \iff \left(\frac{\omega_{t+1}}{\omega_t}\right)^{-\sigma} < \frac{\alpha^\sigma(1-\delta)(\alpha^\sigma - 1)}{2\alpha^\sigma - 1} \tag{33}$$

The condition above is satisfied within a long range of plausible calibrations of β and δ. Hence, (32) displays analytically the short-run trade reversals in anticipation of policy change.

Moreover, we can also look at how trade elasticity is affected by this anticipation. In period $t+1$ as the firm jumps up to the new steady state, it orders according to the new policy rule $\tilde{i}_{t+1} = \left(\frac{\sigma\omega_{t+1}}{\sigma-1}\right)^{-\sigma}(\alpha^\sigma - 1)$. Comparing $\tilde{i}_{t+1}$ and $\tilde{i}_t$, we get

$$\frac{\tilde{i}_{t+1}}{\tilde{i}_t} = \left(\frac{\omega_{t+1}}{\omega_t}\right)^{-\sigma} \underbrace{\left[\left(\frac{\omega_{t+1}}{\omega_t}\right)^{-\sigma} \frac{(1+\beta^{-\sigma}(1-\delta)^{1-\sigma})}{(1-\delta)} - 1\right]^{-1} (\alpha^\sigma - 1)}_{*} \tag{34}$$

The part (*) in (34) is the amplification in response to the tariff change coming from the anticipation (notice that it is the inverse of anticipation effect in (32)). We can account for this amplification using inventory holdings around the tariff change

$$\frac{\tilde{m}_t}{m_t} = (\alpha^\sigma - 1)^{-1} \left[\left(\frac{\omega_{t+1}}{\omega_t}\right)^{-\sigma} \frac{(1+\beta^{-\sigma}(1-\delta)^{1-\sigma})}{(1-\delta)} - 1\right] < 1 \tag{35}$$

Where $\tilde{m}$ are imports after knowing that tariff will drop next period and m are counterfactual imports without a tariff change. Counterfactual imports can be thought of as the imports earlier in the year before the short-run dynamics set in. (35) expresses the anticipatory effect to a future tariff reduction. The magnitude of anticipation is increasing in the equilibrium inventory-sales ratio, α^σ, and the reduction in tariffs, $\frac{\omega_{t+1}}{\omega_t}$. This result explains the findings of the model simulations in section 2.