5-2024

Essays on Trade Policy

Ward Reesman
wreesma@clemson.edu

Essays on Trade Policy

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Economics

by
Ward M. Reesman
May 2024

Accepted by:
Dr. Scott L. Baier, Committee Chair
Dr. Cheng Chen
Dr. Aspen Gorry
Dr. Matthew Lewis

Abstract

This dissertation consists of two essays on international trade and trade policy, focusing on a quantitative evaluation of the causes and consequences of trade barriers and economic integration. In the first essay, I examine the role of a widely-used discriminatory trade barrier in shaping dynamic trade patterns of exporters across a range of affected and unaffected export markets. In the next chapter, I explore the heterogeneous economic determinants impacting the formation of economic integration agreements among countries, paying specific attention to the role of bilateral migration flows.

The first chapter estimates the dynamic trade effects of *temporary trade barriers* (TTBs), a form of targeted trade barrier widely used by members of the World Trade Organization in response to alleged unfair trade practices. TTBs like antidumping (AD) have been shown to have large and persistent effects on trade flows between countries, but there is mixed evidence on the effect of these tariffs on trade to unrelated markets in part driven by unique institutional features that complicate identification. In this essay, we revisit these classic TTB questions with a focus on AD policy using publicly available product-level trade data and a dynamic difference-in-differences framework, paying specific attention to trends in export growth. We find *qualitatively different* trade effects when accounting for growth effects that suggest AD investigations are associated with *global* reductions in within-product trade. We provide evidence that these reductions are not primarily driven by policy-related chilling effects, and argue scale economies in exporting may play a large part. We also identify a significant amount of heterogeneity in the third market trade effects of AD investigation across export markets. Our findings suggest the aggregate impact of AD policy on global trade flows of exporters is potentially large and complicated due to interdependence of export markets and long run dynamics.

The second essay investigates the formation of economic integration agreements (EIAs).

EIAs are widespread multilateral agreements between countries with a variety of characteristics based on the level of integration, ranging from preferential trade agreements to economic unions. However, little research documents heterogeneity in the economic determinants of EIA formation beyond the classical forces of trade creation and diversion. Further, little previous work has explored the role of migration flows in shaping international integration. In this paper, we investigate the economic determinants of EIA formation using a panel of trade, migration, gravity, and EIA data over the period 1990-2015. We first build a simple spatial model of trade and migration that delivers a structural gravity equation for the movement of people across borders, which we estimate by using changes in EIA membership as changes in migration costs. We find novel evidence on the relationship between EIA participation and migration that suggests deeper agreements generate more migration flows, and these flows exhibit a non-linear relationship with respect to bilateral distance. The model also suggests a set of economic variables that impact the welfare gains from certain types of EIAs, and thus form potential determinants to agreement formation. To this end, we build and estimate a random forest to verify the most important factors that influence the formation of certain levels of EIAs. The random forest suggests the importance of certain country or country-pair economic characteristics in determining EIA membership is heterogeneous across agreement types, and a prediction exercise using distance, contiguity, and country-pair income and income differentials can predict out-of-sample EIA formation with a high degree of accuracy.

Dedication

To my parents Ward and Cheryl, to my incredible partner Becca, and to our cats Rudy and Pesto.

Acknowledgments

I am incredibly grateful to all those that assisted in my development as a researcher, and in particular Scott Baier, Cheng Chen, and Aspen Gorry. Scott, your wisdom and advice was instrumental to me throughout my time at Clemson, as was your provision of resources to make me a better trade economist (don't worry, I'll return all the books). Cheng, you always pushed me to learn and grow – your insights were invaluable to the development of my research agenda. Aspen, your advice on presenting research, interviewing, the job market process helped me immensely throughout the past year and I can't thank you enough.

I would also like to thank Matthew Lewis and Chuck Thomas for their advice, knowledge, and support. Your stewardship of the Industrial Organization workshop and field coursework was crucial to spearheading my initial interests and setting me on the right path to become a productive researcher.

Finally, and most importantly, I want to thank my family and friends. To my parents, Ward and Cheryl, who have always believed in me and offered their unconditional support. To all my peers here at Clemson – Mustafa, Jacob, Illia, Ben, Micah, Aspen, Srijan, Titir, and others – I am fortunate to call you all friends, and I cannot thank you enough. And of course, Becca, I could not have done this without you.

Table of Contents

List of Tables

List of Figures

Chapter 1

Temporary Trade Barriers and Trade Growth

1.1 Introduction

Temporary trade barriers (TTBs) like antidumping (AD) duties and countervailing duties are the last remaining source of trade barriers that exist between member nations of the World Trade Organization (WTO). These barriers have proliferated over the course of the past several decades, with over 7,500 TTB investigations reported to the WTO between 1995 and 2022.[1] The resulting tariff rates can often be large – in the case of China, the average AD ad valorem tariff rate imposed by the United States over the period 2000-2009 was 153% (Felbermayr & Sandkamp, 2020).[2] An extensive literature has investigated the effects of these policies on trade flows into both the investigating destinations, as well as unrelated, non-investigating destinations.[3] However, despite high tariff rates and widespread use of these policies, clear evidence on the direction of trade effects to unrelated destinations is mixed – whether TTBs *deflect* or *dampen* trade in sanctioned products is still an open question.[4] This ambiguity is partly due to the complex institutional nature

[1] These data were retrieved from the WTO website, accessed August 25, 2023. AD cases are reported here, countervailing cases are reported here, and safeguard cases are reported here.

[2] Tariff rates against non-market economies are often higher. Additionally, AD duties can come in many other forms such as price undertakings and specific tariffs.

[3] For example, Prusa (2001) and Bown and Crowley (2007, 2010, 2013a). For a broader discussion of this literature see Bown and Crowley (2016) and Blonigen and Prusa (2016).

[4] Trade deflection occurs when a TTB leads exporters to reallocate excess capacity to unrelated, or third markets. Trade dampening is when a TTB results in a reduction of trade to these third markets.

of TTB policy and how it hinders identification. In particular, it has been shown that there is a clear bias in which products are selected by importing countries to target with duties – either in the form of targeting products with downward price trends, or surges in the volume and share of imports.[5] This selection may result in strong pre-treatment trends in the level of trade flows that complicate the identification of policy effects (Steinbach & Khederlarian, 2022).

In this paper we revisit the effect of TTB investigation on trade, focusing on AD policy, using a dynamic difference-in-differences (DiD) event study framework to answer two main questions. First, what are the effects of AD investigation on the level and growth of trade within the investigation destination? Second, what are the effects of AD investigation on the level and growth of trade to non-investigating destinations? We are particularly interested in whether viewing classical AD questions through the lens of export growth addresses identification concerns and delivers qualitatively different results. We use UN Comtrade product-level data over the period 2000-2016 to study the effect of AD investigations on trade, first focusing on China and then extending to other frequently-targeted economies, leveraging both export data and import data of the top destinations responsible for filing the most AD petitions.[6] Our focus on China is spurred primarily by the fact that China is the largest target of TTB actions across the spectrum of petitioning destinations, with four times the number of petitions as the second-largest target destination (Bown, 2011). This trend was exacerbated by China's accession to the WTO in 2001 that rendered other forms of trade barriers a violation of the "most-favored nation" status enshrined in WTO membership.[7]

The presence of pre-treatment trends has significant implications for the approach to estimating the trade effects of AD policy. If there are indeed strong trends in the level of trade flows prior to initiation, it is likely that the parallel trends assumption is violated and DiD estimation will be biased. AD tariffs are levied disproportionately on developing countries that presumably have time-varying trade flows that exhibit strong positive growth trends. This is especially salient for China due to the explosion of exports following its 2001 accession to the WTO (Bown, 2011). Therefore, it is natural to conduct the DiD estimation in growth rates. If the growth rates of trade flows evolve over time in a similar fashion for treated and control products prior to investigation (despite treated products exhibiting higher growth rates), the parallel trends assumption will hold

[5] The former follows from the stated goal of AD policy (USITC, 2015). Bown and Crowley (2013b), Hillberry and McCalman (2016) document the latter.

[6] Beyond China, we examine the effect of AD investigation on exports from other export-oriented developing economies (India, Indonesia, Thailand, and Malaysia), and developed economies (the United States, Japan, and South Korea).

[7] Our focus on China is also in part due to our ability to access both firm-level data and transaction-level trade data for Chinese exporters, which we intend to leverage.

and the DiD estimation will produce more valid estimates. This matters both quantitatively and qualitatively, as the size and *sign* of the estimated effects may be influenced by accounting for differential growth trends.

We contribute to the literature by highlighting the existence of strong pre-trends in trade flows prior to an AD investigation across all exporting destinations, not just within focal markets. Across all exporting destinations, there exists an upward growing trend in the difference in the levels of log trade volume but a relatively stable difference in the growth rates of trade volume between target products and non-target products within the same industry over time. Within the focal market, We find a persistent negative effect on the *growth rate* of targeted products up to 9 years after investigation (even after controlling for unobserved industry-destination heterogeneity), which leads to a negative effect on the levels of those Chinese exports to the focal markets in the long run, as compared to similar non-target products. The growth effects are quantitatively significant, with treated products exhibiting growth rates up to 36 percentage points lower than control products in post-treatment periods. These findings are important and likely non-specific to the setting that we are studying (i.e., AD tariffs), as almost all TTBs (AD tariffs, quantity restrictions, SG tariffs, CVDs etc.) target high-growth exports prior to the sanctions, which calls for further investigation into the long-run effects of all types of TTBs (Bown & Crowley, 2013b; Steinbach & Khederlarian, 2022).[8]

Concerning the third market, we find a strong trade *dampening* effect instead of a trade deflection effect following an AD investigation. This finding is established *only when* we carefully take into account the difference in the growth rates between target products and non-target products within the same industry prior to the AD tariff shock.[9] Specifically, we find a persistent negative effect on the *growth rate* of the AD products in the third market up to 9 years post-investigation, which leads to a *substantial downward deviation* of the export level from its pre-treatment trend after the AD tariff. These growth effects are also quantitatively significant, though smaller in magnitude than the focal market effects – growth rates of export volumes to third markets are consistently 3 to 5 percentage points lower among treated products than control products following AD investigation, and are larger for products with a larger share of exports in the sanctioned market. This finding

[8]Caveat: the identified growth effect is the average treatment effect on the treated (ATT), not the average treatment effect (ATE). Target and non-target products are not randomly selected by the authorities (e.g., the difference in the export growth rate prior to the AD action). In particular, the ATE should be smaller (bigger) than the ATT if the AD tariff has a bigger (smaller) negative impact on products that feature higher export growth.

[9]A simple DID regression (without controlling for the difference in the pre-trends) would lead to the opposite finding.

shows that target products displayed a *global* pattern of high export growth before the AD tariff and substantially reduced export growth afterwards. Robustness exercises suggest this finding is not China-specific and likely applies broadly to exports subject to TTB activity.[10]

To arrive at the above results, we consider a number of specifications. Beyond our concern about the validity of the parallel trends assumption, AD imposition is a staggered event where treatment timing varies by product and a number of recent papers in the DiD literature discuss how differential timing introduces bias into a standard two-way fixed effects (TWFE) DiD estimation.[11] We start with the simple static model and extend to a dynamic event study setting to investigate long-run effects. Both the static and the dynamic DiD models are estimated via ordinary least squares (OLS) and weighted least squares (WLS). We then consider the alternative dynamic estimator proposed by Sun and Abraham (2021) to address the econometric issues discussed in Goodman-Bacon (2021). Under each of these specifications, we also consider a variety of fixed effects to control for varying degrees of unobserved heterogeneity. Our results are robust to these alternatives as we discuss in Section 1.5; we focus on the static and dynamic DiD results estimated via OLS in the main text and relegate additional estimations to the appendix.

This paper contributes to several strands of literature. First, it complements the considerable discussion on the effects of TTBs on trade flows, specifically AD tariffs. Prusa (2001), Lu et al. (2013), Besedeš and Prusa (2017), Sandkamp (2020), and Steinbach and Khederlarian (2022) examine the effects of TTBs within the investigating market, documenting sharp reductions and even total elimination in trade of targeted products. Egger and Nelson (2011) find negative effects of much smaller magnitudes, while Staiger et al. (1994) show that investigation alone can induce trade destruction. Other papers discuss outcomes such as exchange rates and prices (Blonigen & Haynes, 2002; Blonigen & Park, 2004), productivity (Jabbour et al., 2019; Pierce, 2011), and aggregate bilateral trade (Vandenbussche & Zanardi, 2010). Like Steinbach and Khederlarian (2022), we document strong trends in trade flows prior to investigation, but we differ in that we also provide novel evidence of growth rate effects that both support previous findings of the size and duration of the trade effects of AD policy, as well as suggest a (potentially) *permanent* impact of AD policy on exports.

With respect to third market effects of TTBs, Bown and Crowley (2007) and Baylis and

[10]The same caveat mentioned in footnote 8 applies: we identify an ATT, not ATE.

[11]See Goodman-Bacon (2021), Callaway and Sant'Anna (2021), and Sun and Abraham (2021) for more.

Perloff (2010) find trade deflection effects for US AD actions, and Hoai et al. (2017) find deflection for EU actions. For Chinese exports specifically, Chandra (2016), Felbermayr and Sandkamp (2020), and Bao et al. (2021) find evidence of trade deflection, while Bown and Crowley (2010) and Lu et al. (2013) do *not* find evidence of trade deflection. Prusa (1997), Lasagni (2000), and Durling and Prusa (2006) discuss trade diversion effects, or whether other trading partners "fill the void" left by target importers. Our findings challenge the existence of broad trade deflection effects by identifying reductions in the growth rate of exports to non-investigating destinations, which we document over a longer time horizon and across a variety of AD targets and AD petitioners. Importantly, our results show that AD tariffs' growth effect and level effect can be *qualitatively* different for some classical questions studied in the AD literature, like trade deflection, and suggest the importance of accounting for growth trends in the analysis of third market trade effects.

A second strand of literature our paper connects to concerns the endogeneity of TTB policy. A number of papers document links between likelihood of TTB petition and macroeconomic conditions, industry-specific factors, political motivations, strategic retaliation, and the presence of preferential trade agreements.[12] We contribute to a subsection of this literature focused on the role of sudden surges in import growth. Bown and Crowley (2013b) and Hillberry and McCalman (2016) find AD tariffs and safeguard actions are precipitated by rapid growth in imports from the target economies, in line with terms-of-trade motives (Bagwell & Staiger, 1990, 2011; Broda et al., 2008). Our paper is closely related, echoing the link between import growth and TTB investigation and illustrating the significant implications this has for the identification of the effects of TTB policy.

Finally, the findings we present have implications for how firms respond to trade policy, which is the subject of a growing literature. Morales et al. (2019) and Alfaro-Urena et al. (2023) provide firm-level evidence that suggests exporting to one destination lowers the cost of exporting to similar destinations. Albornoz et al. (2021) examine both focal market and third market effects following a tariff shock to Argentine exporters and Fajgelbaum et al. (2023) examine exports from non-target countries in response to the U.S.-China trade war, both of which suggest the existence of within-product interdependence between export destinations. Breinlich et al. (2022) offer scale economies as a possible explanation for export destruction and link this channel to the discussion on industrial policy. Our third market results suggest that the loss of exports to one market due

[12]For work on these topics, see Aggarwal (2004), Blonigen and Bown (2003), Bown and Crowley (2013a), Bown and Tovar (2011), Bown et al. (2023), Crowley (2011), Feinberg (1989), Feinberg and Hirsch (1989), Furceri et al. (2021), Knetter and Prusa (2003), Prusa and Skeath (2002), Prusa et al. (2022), and Reynolds (2006).

to AD duties induces further reductions in exports of the same product to unaffected destinations, which is consistent with exporting complementarity and scale economy channels highlighted in the recent literature.

In the next section, we briefly discuss some institutional features of TTB and AD policy, before moving into data construction in Section 1.3 and empirical methodology in Section 1.4. We then present the results and conclude.

1.2 Institutional Background

Over 7,500 TTB investigations were reported to the WTO between the years 1995 and 2022, of which over 6,500 are AD cases.[13] Increasingly used by lower-income developing countries, these tariffs are discriminatory in nature and specifically target products and firms accused of engaging in unfair trade practices such as dumping or export subsidization. Among TTBs, AD duties are by far the most commonly used policy instrument, comprising a majority of total TTB investigation and imposition across petitioning countries – particularly for newer users of TTBs (Bown & Crowley, 2016). Due to the overwhelming popularity of AD policy as a vehicle for obtaining temporary tariff protection, we focus on AD imposition in this paper.

In order to obtain temporary tariff protection via AD law, domestic firms, industry associations, or labor unions organize and file petitions with key government agencies. In the United States, these agencies are the Department of Commerce and the United States International Trade Commission. In the European Union, this is the European Commission. Agencies review the petitions and make affirmative or negative decisions on two major questions: (1) were the target exporters dumping, and (2) did this activity cause or threaten to cause material injury to domestic competitors?[14] Proving dumping or material injury is likely easier for products that exhibit declining relative prices and rapidly growing import shares, thus it is natural to think strong trends are likely to precede investigation.

If the respective agencies make affirmative determinations to both questions, then duties are imposed. AD proceedings typically last around one year, from investigation initiation to the levying of final duties.Duties remain in place until revoked by the imposing country, though per WTO rules

[13] These data were retrieved from the WTO website, accessed August 25, 2023. AD cases are reported here, countervailing cases are reported here, and safeguard cases are reported here.

[14] Dumping is defined as selling an exported product below "normal value," which typically is calculated as the price that firm charges in a home market, a 3rd market, or an estimated production cost.

are subject to sunset review every 5 years by the presiding agencies. Sunset review requires re-evaluation of the AD case, and result in duty revocation unless it is found that termination would result in the "continuation or recurrence of dumping [...] and of material injury" (USITC, 2015). The size of AD duties are determined by the difference between the observable export price and the calculated "normal value," a differential that is referred to as the dumping margin. All firms in the target destination engaged in the export of the target product are subject to the duties. In market economies, dumping margin is calculated on a firm-specific basis using firm-specific prices. However, for exporting countries with non-market economy status like China, the final dumping margin is the difference between the average export price of all firms and the "normal value" of the product.[15] While there is a possibility for exporting firms to obtain individual or market economy treatment, and thus receive firm-specific tariffs, for many exporting firms the tariff size is invariant to that firm's individual export price. For the purposes of this paper, we choose not to focus on the tariff rates as treatment and instead on whether the product was investigated in a binary sense. The delayed nature of AD proceedings (and thus likelihood of anticipatory behavior), the vast heterogeneity in tariff rates, and the continuous nature of the tariff rate as a treatment all present additional challenges for identification (particularly within a staggered adoption setting) and would require even stronger parallel trends assumptions (Callaway et al., 2021).

1.3 Data

To examine the effects of AD activity on growth, we use annual product-level trade data retrieved from UN Comtrade. These data document import and export flows at the Harmonized System (HS) 6-digit product level for all countries over the period 2000-2016. Data on AD activity comes from the Global Antidumping Database (GAD), which is part of the Temporary Trade Barrier Database maintained by the World Bank (Bown et al., 2020). The GAD contains information on the universe of AD activity across all filing countries. These data contain detailed records on the timing of various stages of AD investigations, targeted countries and HS product codes, as well as tariff rates for cases that advance to either preliminary or final tariffs.

One drawback of these AD data is that products can be targeted at the HS6, HS8, or HS10 level as determined by the importing destination's agencies and customs office. However, HS codes

[15] Non-market economy domestic prices are assumed to be distorted, e.g. due to state subsidies.

are only comparable across countries at the 6-digit level: 8-digit and 10-digit codes are destination-specific, so a given product may be assigned a different code when it leaves as an export than from what it is assigned as it arrives as an import in the destination. As such, we aggregate target product data from the GAD to the HS6 level. While this aggregation removes some detail, it is not too problematic for two reasons: first, many 6-digit categories do not have a large number of component 8-digit and 10-digit codes, and second, many AD petitioners recognize the possibility of relabeling as a loophole and intentionally include related 6-, 8-, and 10-digit codes in the AD petition (that may even have zero trade flows prior to treatment) to anticipate this relabeling behavior.

Before merging the GAD into relevant trade data, we collapse the AD data to the product-petitioning destination-initiation year level. This involves several steps. First, we extract all AD activity from all petitioners against China over the period 2000-2016, with target products aggregated to the HS6 level.[16] This provides us with a data set of all target HS6 products by all petitioners that file against China. Then, we eliminate duplicates – for the focal market analysis, we first focus on the top 10 petitioners, then we eliminate duplicates within petitioning destination-product such that a given HS6 product is only treated at the date of the first case filed against the product by the given destination chronologically. For the third market analysis, we eliminate duplicates within product (but across all filing destinations), such that a given HS6 product is only treated at the date of the first case filed against the product by any destination chronologically.

We include both unsuccessful and successful cases, as previous work has suggested that investigation by itself can have an effect on trade flows, and repeat filing against products following an unsuccessful case is a common practice (Staiger et al., 1994).[17] Prior to removal of duplicate filing the full sample of AD activity against China over the period 2000-2016 consists of 933 cases and 2,359 targeted HS6 products. For the focal market analysis, we focus on the top 10 petitioners and remove duplicates within destination-product, which delivers a sample of 656 cases and 1,516 targeted HS6 products. For the third market analysis, we include the universe of AD filing activity against China and remove duplicates within product, which delivers a sample that consists of 505 cases and 1,036 targeted HS6 products.

Table 1.1 documents summary statistics on the distribution of cases and target products

[16] The top 10 petitioners against China are the U.S., the E.U., India, Argentina, Australia, Brazil, Colombia, Canada, Turkey, and Mexico, which we focus on in the focal sample. We use the full list of petitioners for the third market analysis.

[17] We also estimate just using successful cases; these results are reported in Tables A8 and A15 the appendix.

across petitioning destination for the full sample, as well as the share of total AD cases that advance to final duties and the mean and standard deviation of case duration in years.[18] On average the success rate of AD cases is high, and on average the duration exceeds the 5 year period when cases must be reviewed per WTO rules – suggesting that most of these temporary barriers are extended at least once. The United States and Turkey have the longest average duration in the sample and India, the United States, and the European Union file the largest number of cases against the largest number of products in the sample.

We construct two data sets to explore our two major questions of interest. For the analysis within the focal markets, we extract data from UN Comtrade documenting all HS6 imports into the 10 destinations responsible for the most AD petitions against China. We aggregate trade volume across origin countries to calculate import shares by destination-product category in each year of our sample. Then, we calculate year-to-year growth rates for trade volume and the import share. A common issue impeding the consistent computation of volume and share growth rates is the presence of zero trade flows – trade is lumpy and can sometimes exhibit intermittent zero flows within a product-destination pair over time. To account for these zeroes over the sample period, we rely on a modified growth calculation from Davis et al. (1998), which is formally

$$\Delta Y_{ijt} = \frac{x_{ijt} - x_{ijt-1}}{(x_{ijt} + x_{ijt-1})/2} \tag{1.1}$$

where i indexes product, j indexes destination, t indexes period, and the denominator is the average of variables in periods t and $t-1$. In this sense, when a trade flow x_{ijt} switches from zero to a positive number between years, the growth rate will equal 200%. Likewise, when a trade flow x_{ijt} switches from a positive number to a zero, the growth rate will equal -200%. These rates can be thought of as capturing entry and exit, and allow us to account for frequent zeroes in the data without dropping too many observations, while also serving as a normalization.[19]

To examine third market effects, we extract export data from UN Comtrade that documents all Chinese exports by product and destination. As we are focused on the response of trade flows to non-investigating destinations, we omit exports of the investigated products to any of the petitioning

[18] Many AD actions are still "in force" as of 2019, the last year of the GAD sample used. As such, duration is censored at 19 years.

[19] We also calculate growth using a more standard approach by log-differencing trade volume and import share levels, which omits the extensive margin. This robustness exercise is included in Tables A9 and A16 in the appendix.

destinations that ever file against the product and retain all other product-destination pairs.[20] We merge the AD data in at the product level, applying the same treatment date to all exports of the target product to non-investigating destinations, determined by the initiation year in the first filing destination chronologically. One drawback of using Chinese export data is an inability to compute import shares. Otherwise, we proceed as outlined above using the growth calculation in (1) to compute the per-period growth in trade volume.

As a final step, we compute unit values for both the focal market and third market to decompose both level and growth effects into price-driven and quantity-driven. This sheds more insight on the possible underlying mechanism, and allows us to test for the existence of a chilling effect in prices as hypothesized in previous work (Bao et al., 2021). Unit value is calculated by dividing the trade value by the trade volume, which gives a rough approximation of the average price of the product. To compute the growth rate of unit value, we log difference in all specifications.

Table 1.2 contains summary statistics on the quantity variables and their growth rates, with the first panel displaying statistics within the focal destinations and the second panel displaying statistics across all export destinations. Average export growth is 15% within both focal markets as well as across third markets, though growth to third markets exhibits a larger variance, and log export volume is on average higher within the focal destinations. The third panel documents average HS6-year export shares across various categories of destinations.[21] Here we can see on average exports to linked third markets have smaller shares than exports to the related focal markets, highlighting on average that third markets tend to be peripheral, but are large in number. However, the insight these summary statistics offer is limited so we report more detailed shares over time in Table 3.

Table 1.3 further breaks down product export shares across destinations over time. In panel (a), we calculate the number of new products investigated in that year, the average share of those HS6 exports to the investigating, or focal, destination in that year, and the average share of those HS6 exports to all third markets, or linked destinations in that year. We can see that the share of newly-investigated products exported to the investigating destination fluctuates between 8 and 16%, while the share of said products exported to all other destinations is between 82 and 94%. Because the third market sample eliminates duplicate filing against the same HS6 product across destinations, the focal and linked shares do not always sum to 100%.

[20] There are 1,963 HS6-destination pairs (or markets) excluded from the export data.

[21] Focal markets are HS6-ISOs where AD investigations were initiated. Linked are HS6-ISOs where the HS6 is investigated, but by another destination. Unlinked are HS6-ISOs where the HS6 was not investigated.

Panel (b) of Table 1.3 looks at export shares over a longer horizon. Here, we calculate the share of total exports in each year by whether or not the product was ever investigated over the sample period. This panel illustrates that over the course of 2000-2016, products investigated by at least one of the top 10 focal destinations account for 30% to 35% of total Chinese exports. Products that never see an investigation account for 65 to 70% of total Chinese exports. Over time, the share of Chinese trade in products that ever see investigations is stable, despite underlying heterogeneity in the shares of products that see AD cases in different years over the sample period. Now, we proceed to outline the empirical methodology, first discussing ex-ante trends before outlining the identification strategy for estimating the effect of AD investigation.

1.4 Empirical Methodology

1.4.1 Ex-ante growth rates and probability of investigation

Before examining the trade effects of AD activity, it is worth verifying previous findings in the AD and TTB literature documenting the relationship between pre-treatment import growth and AD initiation within the context of our sample. Table 1.4 presents summary statistics comparing the average export growth in pre-treatment periods for initiated products to average import growth over the full sample for two definitions of control products: first, for all non-target HS6 products, and second, for non-target HS6 products within the same HS4 subcategory as target products. Table A1 in the appendix displays the same, but breaks up comparisons by petitioning destination. Simple comparisons of means consistently suggest that import growth is higher among target products in pre-treatment periods than non-target products, with some significant heterogeneity in the size of these growth differentials by importing destination.

We also consider a simple regression framework to further validate this relationship. Similar to Bown and Crowley (2013b), we estimate

$$AD_{ijt} = \beta_0 + \beta_1 g_{ijt} + \alpha_{st} + \gamma_j + \varepsilon_{ijt} \tag{1.2}$$

where AD_{ijt} is a binary variable equal to 1 if there was a trade policy change (i.e. a new AD investigation) for product i by destination j in period t and g_{ijt} is the mean growth rate of product i

imports into destination j from $t-1$ to $t-3$.[22] We include sector-time and destination fixed effects to attempt to capture unobserved sector- and destination-specific heterogeneity in the probability of AD petition. We estimate (2) via OLS, probit, and logit; the results are reported in Table 1.5 where g_{ijt} is standardized with mean zero and standard deviation one. Estimated coefficient of the linear probability model and the marginal effects of the probit and logit models are in line with reported results of Bown and Crowley (2013b). Our estimates imply a one standard deviation increase in the growth of import volume increases the likelihood of an AD investigation by approximately 20 percent relative to the mean likelihood.

Finally, as an additional step, we model the process as a survival problem and estimate a proportional hazard model, where the hazard rate is the probability of investigation. Formally, we estimate

$$\lambda_{ij}(t) = \lambda_0(t) \times \exp(\beta_1 g_{ij}(t)) \tag{1.3}$$

where $\lambda_{ij}(t)$ is the hazard rate at time t, or the probability of investigation of product i by destination j, $\lambda_0(t)$ is the baseline hazard rate (unobserved heterogeneity in probability of investigation), and $g_{ij}(t)$ is the mean growth rate over the past 3 years. The results of the proportional hazard model are displayed in Table A2 in the appendix. The statistically significant positive coefficient implies the hazard ratio $\exp(\beta_1)$ is greater than 1, meaning an increase in the mean growth rate in the past 3 years contributes to an increase in the hazard rate, or the probability of investigation. This further confirms our above findings that there is a positive relationship between ex-ante import growth and probability of AD investigation. We now turn toward the main question of interest, estimating the effect of AD on post-treatment import growth.

1.4.2 Effect of AD investigation on growth

AD imposition varies by product and importing destination, and units (product-destination pairs) are treated for varying degrees of length depending on case timing and the outcome of regular sunset reviews. In this sense, we want to estimate a dynamic DiD specification that allows us to capture the effect of the treatment at different lengths of exposure, while accounting for the staggered adoption of the treatment across units. Identification of trade effects relies on variation in

[22]We choose the average of the 3 preceding years due to common institutional features of AD policy. For example, the USITC uses data from the past 3 full calendar years plus up to 3 additional quarters in the material injury determination. See USITC (2015) for more information.

AD imposition across products, both within and across destinations – the first difference compares trade flows before and after initiation, and the second difference compares these differences between target and non-target varieties (within the same broader sector). We consider estimating

$$Y_{ijt} = \alpha_{ij} + \lambda_{jt} + \gamma_{st} + \sum_{k=-10}^{-2} \delta_k D_{ij,t-k} + \sum_{k=0}^{10} \beta_k D_{ij,t-k} + \varepsilon_{ijt} \tag{1.4}$$

where Y_{ijt} is log import volume, the import share, or log unit value of product i shipped from China into destination j in year t, α_{ij} represent product-destination fixed effects, λ_{jt} represent destination-year fixed effects, γ_{st} represent sector-year fixed effects (defined at either the HS2 or HS4 level) and ε_{ijt} is an error term. $D_{ij,t-k} = \mathbb{1}\{D_{ij} = 1\}\mathbb{1}\{t - k = t^*_{ij}\}$ is relative time indicator equal to 1 if (1) product i is ever treated by destination j, and (2) product i receives treatment by destination j in period $t - k$, where t^*_{ij} denotes the treatment year. This term indicates the treatment year for $k = 0$, it indicates treatment beginning k periods ago for $k > 0$, and it indicates start of treatment $|k|$ periods in the future for $k < 0$. Thus, δ_k and β_k represent leads and lags of the treatment, so estimates of β_k capture the average treatment effect on the treated (ATT) for varying lengths of exposure, while estimates of δ_k will capture pre-treatment trends and can be used for pre-testing. We define the control group to be HS6 products within the same HS4 category as treated HS6 products in an attempt to reduce some concern regarding selection based on broader industry category.[23] We use product-destination fixed effects as this is the level of treatment, and destination-year and sector-year fixed effects (where sector is denoted by HS2 or HS4 category) to control for sector- and destination-specific trends and macroeconomic conditions.

However, as suggested by the discussion in section 1.4.1, it is likely that the log level of trade volume and the import share exhibit strong trends in the pre-treatment period – higher growth rates among treated products suggest that the level of trade of treated and control products will not evolve in a parallel fashion. Further, if AD cases truly target the phenomenon of dumping, we should also observe pre-treatment trends in average prices. In both cases, estimates of $\delta_k, k \leq -2$ will likely not equal zero, and the parallel trends assumption will be violated. While this concern compromises the previous estimation and implies treatment is not randomly assigned, we still believe from the perspective of the exporting firms, AD investigation is plausibly exogenous as it originates from foreign firms and governments and can arrive unexpectedly. We propose an alternative estimation

[23] Previous work has highlighted trends in AD imposition that suggest certain industries, like metals and chemicals, are more likely to be investigated.

strategy by first-differencing (4). In this manner, we are estimating the effect of AD investigation on the growth rate differential between target products and non-target products within the same industry or sector. This can be thought of as a triple difference, with the first difference comparing differences in growth rates before and after initiation. Our estimation equation becomes

$$\Delta Y_{ijt} = \lambda_{jt} + \gamma_{st} + \sum_{k=-10}^{-2} \delta_k \Delta D_{ij,t-k} + \sum_{k=0}^{10} \beta_k \Delta D_{ij,t-k} + \Delta\varepsilon_{ijt} \tag{1.5}$$

where ΔY_{ijt} is the growth rate of import volume or import share as defined in (1), or unit value of product i from China into destination j from $t-1$ to t, λ_{jt} is a destination-year fixed effect, γ_{st} is a sector-year fixed effect, and $\Delta\varepsilon_{ijt}$ an error term. $\Delta D_{ij,t-k} = D_{ij,t-k} - D_{ij,t-1-k} = \mathbb{1}\{D_{ij} = 1\}(\mathbb{1}\{t-k = t^*_{ij}\} - \mathbb{1}\{(t-1)-k = t^*_{ij}\})$ will remain a relative time indicator for product i being k periods away from initial treatment within destination j at year t, as the last indicator function will always evaluate to zero. As before, β_k capture the ATT for varying lengths of exposure to the treatment, using products within the same industry (HS4 category) as treated HS6 products as the control. By differencing, the time-invariant unit fixed effects drop out. The destination-time fixed effects control for destination-specific macroeconomic conditions like exchange rates that have been shown to influence growth in trade, and the sector-time fixed effects control for unobserved sectoral level growth trends (where as before, we define sector as either HS2 or HS4 category). Given our level specification, sector-destination or sector-destination-year fixed effects are also reasonable if we think that there are sector-destination specific growth trends. We consider these for robustness.

Our first question centers around the effect of AD investigation and imposition on the growth of trade flows within the focal destination. We estimate (4) and (5) using OLS, WLS, and an estimator proposed by Sun and Abraham (2021) designed to address issues that arise in dynamic DiD settings with staggered adoption when treatment effects evolve over time.[24] In this context, t^*_{ij} corresponds to the year an investigation was initiated by the destination j against Chinese imports of product i, and the dependent variable is the level or growth of Chinese import volume, the import share, and the unit value within destination j.

The second question concerns the response of investigated exports to non-investigating destinations. The general estimation strategy is the same as above, with a few important modifications.

[24]Related issues discussed in Goodman-Bacon (2021), Callaway and Sant'Anna (2021).

We now estimate

$$Y_{ijt} = \alpha_{ij} + \lambda_{jt} + \gamma_{st} + \sum_{k=-10}^{-2} \delta_k D_{i,t-k} + \sum_{k=0}^{10} \beta_k D_{i,t-k} + \varepsilon_{ijt} \tag{1.6}$$

$$\Delta Y_{ijt} = \lambda_{jt} + \gamma_{st} + \sum_{e=-10}^{-2} \delta_k \Delta D_{i,t-k} + \sum_{k=0}^{10} \beta_k \Delta D_{i,t-k} + \eta s_i^{AD} + \Delta\varepsilon_{ijt} \tag{1.7}$$

where the first major difference is our treatment variables $D_{i,t-k}$ and $\Delta D_{i,t-k}$ are now indicators for product i being k periods away from initial treatment by *any* of the petitioning destinations, using the earliest chronological case as the treatment date for each product i. Treated product-destination pairs are dropped from the sample, so we estimate the effect of an AD investigation initiated by any petitioning country against China on export growth to all other destinations (including other petitioners, if they did not ever investigate product i). Note that we drop treated product-destination pairs even if they were not the first chronological case against a product i – while we want to eliminate these within-product duplicate cases for the purposes of defining the treatment date, we still consider these product-destination pairs as focal markets and exclude them. Including these markets would bias results downward as they are eventually subject to investigation at later dates.

The second major modification is the inclusion of the variable $s_i^{AD} = x_{if}^{AD} / \sum_k x_{il}^{AD}$ where x_{il} is export volume of product i shipped to destination l in the year that product i was treated via AD investigation, and f indexes the focal destination. This variable represents the share of product i exports sent to the focal destination f in the year that focal destination enacts an AD investigation against the product, and can be thought of as a measure of within-product exposure to the focal destination. We believe it is important to control for significant heterogeneity across products in the growth rate beyond sector-level trends. A higher export share within the focal market implies a lower export share in third markets, which is often associated with faster growth and more entry in the third markets. Further, we *do not* control for product-specific linear growth trends via (HS6) product or product-destination fixed effects in the growth regressions for two reasons. First, we are worried that linear growth trends probably do not hold at such a disaggregated product level. Second, controlling for product-specific linear growth trends would sweep up most of the variation in the growth effect we attempt to identify, which is across HS6 products. Since we still want to control for product-level heterogeneity, we include the focal export share s_i^{AD} as a "weaker" control variable that leaves us room for identification. Note that we do not include this variable in the level specification (6), as it is absorbed by the product-destination fixed effects α_{ij}.

A final difference to note is we do not estimate this model for import share as a response variable since we use Chinese export data, using log level and growth in import volume and unit value as the main dependents of interest. As before, our baseline specifications incorporate product-destination (or market), destination-year, and sector-year fixed effects in the level regression and sector-year and destination-year fixed effects in the growth regression. As in the focal analysis, we consider both HS2 and HS4 levels for the sector-year fixed effects. We also consider including sector-destination-year fixed effects instead of sector-year and destination-year separately. We cluster the standard errors of coefficients at the level of the treatment assignment. Specifically, we cluster the standard error at the HS6 product-destination (ISO) level in the focal market regressions and at the HS6 product level in the third market regressions.[25] In the following section we discuss the results from the estimation strategy outlined above.

1.5 Results

We now present the results of our empirical strategy. We first discuss the effect of AD investigation within the investigating destination on both volume and share of imports, as well as prices by estimating (4) and (5) via OLS. We focus on the average treatment effect on the treated (ATT), aggregated across cohorts and lengths of exposure, and then break down the effect dynamics using event study plots aggregated across cohorts and plotted over length of exposure to AD orders. We also estimate the model using WLS and the Sun and Abraham (2021) estimator for robustness, and consider a wider sample of target exporters – both developing and developed – which suggests our finding is not China-specific.

After summarizing our findings within the investigating destination, we turn to a discussion of the effect of AD investigation on export volume and average export prices of target products to unaffected, non-investigating destinations. As with the previous set of results, we present OLS estimates of (6) and (7), focusing on both aggregate ATTs and dynamic coefficient plots across lengths of exposure to the treatment. We also consider alternative estimators and a wider range of target exporters for robustness. Finally, we consider some extensions to our model to investigate the role of market share, correlation across AD imposition, and regional proximity to the investigating

[25] Note that all third markets of the same product are treated at the same time, when the product is investigated in the focal market.

destination as they relate to the trade effects of AD actions.[26]

1.5.1 The effect of investigation in investigating destination

1.5.1.1 Main results

We first present OLS estimation results of (4) and (5), which focus on the effect of AD investigation within the investigating destination, or focal market, for the three main dependent variables of interest. Table 1.6 documents these results, with three panels for our three dependent variables. In all three panels, columns 1–3 contain estimates of the model run in levels as outlined in (4), and columns 4–6 contain estimates of the model run in differences as outlined in (5). In both cases, we include alternative fixed effects specifications in addition to the "baseline" model outlined in the previous section – the baseline specifications are in columns 1 and 4, respectively. Table 6 presents ATTs of AD investigation aggregated over both cohort and length of exposure.

Focusing first on panel (a), we see that AD investigation is associated with statistically and quantitatively significant reductions in both level and growth in import volume. The first three columns suggest that, following an AD investigation, import quantity falls by 30 to 50 percent. Columns 4–6 suggest that the growth rate of targeted imports into the focal destination falls by 11 to 12 percentage points. However this aggregate ATT fails to capture some of the interesting dynamics in both the level and growth effects, so we plot the coefficients for each length of exposure, aggregated across cohorts (i.e., product-destination pairs with different investigation dates), in Figure 1 for quantity level (panel (a)) and quantity growth (panel (b)).

There are several things to note about Figure 1.1. First, in panel (a) there is a persistent and clear trend in the difference between log quantity in the treated products and non-treated products within the same industry during the pre-investigation periods. This difference narrows over time, as the treated products exhibit higher growth rates. This is suggestive of a violation of the parallel trends assumption that compromises our DiD estimates. After investigation, export volume of treated products falls by over 50 percent relative to non-target products in the same industries, and this difference persists with the gap widening significantly over time. In panel (b), we see a much tighter relationship in the growth differential prior to investigation for treated and control

[26]For variables such as the export shares and the growth rates, we use their numerical values without the percent sign in the regressions (i.e., 10 means a 10% share or growth rate). Import shares in the focal market regressions are reported in decimals.

products, and a sharp decline in growth rates for treated products relative to control products following an investigation. In the investigation year, import growth of treated products falls by 10 percent relative to control products, and in the two years immediately following investigation growth of treated products relative to control products falls by 35 and 29 percent respectively. The growth differential narrows five years post-treatment, and is present but weaker up to ten years post-treatment.[27]

Turning to the import share effects in Table 1.6, panel (b), we find much of the same – AD investigation is associated with both statistically and quantitatively significant reductions in the level of import share and the growth of the import share, though the magnitudes are smaller. Investigation leads to a fall in the import share of 1 to 2 percentage points, while the growth rate of the import share falls by 9 to 10 percentage points. The dynamics of these effects provide further insight. Figure 2, panel (a) plots the coefficients aggregated across cohorts for each length of exposure depicting the import share differential between treated and non-treated products within the same industry, and panel (b) plots the same but for the growth rate in the import share.

As with import volume, strong pre-trends in the level of the import share are present in Figure 1.2(a), reinforcing the selection issue that impacts our estimates. Post-investigation, treated products exhibit import shares that are consistently 7 to 10 percentage points lower than non-treated products within the same industry, until at least 10 years post-investigation. The growth effects depicted in Figure 2(b) are similar to the quantity growth effects, with a sharp immediate reduction in the growth rate of treated products relative to control products of 22 to 31 percentage points. Up to 9 years post-investigation, growth of import share among treated products is lower by 3 to 9 percentage points relative to control products, with most statistically different from zero.

Finally, we turn to prices. Table 1.6, panel (c) contains estimation results for (4) and (5) using log unit value and log-differenced unit value as the dependent variables. We find a positive effect of AD investigation on both log unit value and the growth in unit value. Columns 1–3 suggest that average prices are 6 to 11 percent higher for treated products relative to control products following an AD investigation, and columns 4–6 suggest that growth in unit value is 2 to 3 percent higher for treated products relative to control products following an AD investigation. We decompose these estimates into dynamic effects in Figure 3. Figure 3 shows that the growth effect comes

[27] The weaker effect after 5 years post-treatment is consistent with institutional details: WTO mandates TTB orders reviewed every 5 years after duty imposition. While many cases are extended upon review (see Table 2), some are not, in which case duties are removed and we would expect dissipation of the negative growth effect.

primarily from an immediate shock in the first three periods following investigation, with treated products exhibiting growth rates 5 to 9 percent higher than control products before returning to no significant differences beyond five years post-treatment. Log prices rise quickly in the first three periods, and then level off to a degree 20 to 30 percent higher among treated products relative to control products.

However, unit values in Table 1.6 and Figure 1.3 are constructed from import values. A point of recent discussion in the trade literature concerns the degree to which tariffs are passed through to consumers; for example Amiti et al. (2019, 2020) document complete passthrough of the 2018 US tariffs against China onto US consumers and firms. Our price results using import data lead us to suspect a similar phenomenon occurring more broadly across AD tariffs. To further investigate the validity of this claim, we estimate the same model as before, but using data on Chinese exports to the focal destinations. Results are reported in Table A3, and coefficients for the price regressions are plotted by length of exposure in Figure A1, panels (a) and (b). From here, we cannot identify a clear price effect. We believe this suggests that AD tariffs do not impact firms' export prices, but *do* impact the eventual price of the exports within the destination market – indicative of full tariff passthrough. Our finding that Chinese export prices do not respond to AD activity is consistent with Felbermayr and Sandkamp (2020), while our finding that import prices rise post-tariff is consistent with Amiti et al. (2019, 2020).

1.5.1.2 Robustness

In addition to the OLS results, we also estimate (4) and (5) via WLS and the estimator proposed by Sun and Abraham (2021). Tables A4 and A5 in the appendix display the aggregated ATTs, and Figures A2–A7 in the appendix display the coefficient plots aggregated across cohorts by length of exposure for our three dependent variables import volume, import share, and average prices for the two estimation procedures. All together, the alternative estimators returns qualitatively and quantitatively similar results as the OLS estimator.

For additional robustness, we also use a wider sample of target exporters. Given our setting, one might wonder whether the identified trade effects are China-specific. Therefore, we extract AD activity against other frequent targets. For less developed targets, we focus on India, Indonesia, Malaysia, and Thailand; for developed targets, we focus on the United States, Japan, and South Korea – all of which are among the most frequent targets of AD action. For less developed targets, we create similar data sets for imports into the same focal destinations used for the earlier analysis, as these economies share similar AD petitioners as China.[28] For developed targets, we use all AD activity from all petitioners and export data from all relevant focal petitioners.[29] from Table A6 and A7 report the results for developing and developed targets respectively, which echo the results for Chinese imports into focal destinations. AD investigation is followed by sharp drops in trade volume and trade growth even among a wider set of exporters.

Taken together, our focal market results have a similar flavor to the main findings of Steinbach and Khederlarian (2022), but with some distinct differences. While we also identify persistent trends in the pre-treatment periods that confound the estimation of treatment effects in levels, we are able to identify a growth effect of AD policy beyond extrapolating the pre-treatment trend line. This growth effect is statistically and quantitatively significant, with persistently lower growth among target products for several years. This novel finding suggests the AD policy has a permanent impact within the focal market. We also find that while it appears average prices increase at the product level, export prices do not respond and it seems AD duties are passed through to the importing market (Amiti et al., 2019, 2020). These results are robust to alternative estimators, and seem to apply to a wider range of AD-affected exporting economies beyond China. With an idea of how AD actions impact the growth and level of trade flows within the investigating destination, we now move to address the impact of AD actions on trade to the rest of the world.

1.5.2 The effect of investigation on exports to other destinations

Our second question of interest concerns the effects of AD investigation on exports to destinations where the products are not being investigated. We present OLS estimation results of (6) and (7) for our two key dependent variables in Table 1.7. As in section 1.5.1, the baseline estimations of (6) and (7) are reported in columns 1 and 4, respectively, with alternative sets of fixed effects reported in adjacent columns. We then consider three different specifications with interaction terms to investigate heterogeneity of the third market ATTs along three dimensions: (1) the export share of the related focal market, (2) the pairwise correlation of AD incidence between the third destination and the related focal destination, and (3) geographic proximity via "extended gravity" (Morales

[28]Since India is now considered as a targeted exporter, we do not use India as a focal destination. The other 9 focal destinations remain in the sample.

[29]Our departure from the focal sample used previously is spurred by the considerably different makeup of petitioners targeting developed economies.

et al., 2019).

1.5.2.1 Main results

From Table 1.7(a), we see two contrasting results. In columns 1–3, we find that export volume of investigated products exported to non-investigating destinations increases by 10 to 13 percent relative to non-investigated products within the same HS4 category. Taken alone, this would be suggestive of trade deflection – target exports increase to non-investigating destinations in response to the AD investigation. However, this result ignores differential growth trends. We report estimates of (7) in columns 4–6, where the focal export share s_i^{AD} is standardized and coefficients correspond to a change of one standard deviation. We find lower growth in import volume of investigated products exported to non-investigating destinations – at the average focal export share, growth rates of investigated products in third markets are 1.3 percentage points lower post-investigation than non-investigated products within the same HS4 category. While the level of trade increases, it increases at a *lower* rate than before investigation. This deviation from the pre-investigation growth trend suggests that rather than deflection, we see dampening of trade in investigated products to non-investigating destinations.

Figure 1.4 reinforces the story told by panel (a) in Table 1.7. Figure 1.4(a) depicts coefficients by length of exposure, aggregated across cohorts, for the baseline level regression and Figure 4(b) depicts coefficients by length of exposure, aggregated across cohorts, for the baseline growth regression. As in section 1.5.1, there are strong trends in the level of import volume preceding investigation of the product in the focal market (which are excluded from the sample here). The point estimates illustrate a continuation of the trend post-treatment, but at a slower rate and with larger variance. The 95 percent confidence bands suggest that the quantity differential between treated and control products post-treatment may be zero, which is indicative of a possibly larger reduction in the growth rate of treated products. A plausible linear trend of the coefficients in the pre-treatment period is drawn onto the figure, which further illustrates the deviation following AD investigation.

Figure 1.4(b) substantiates this reduction in the growth trend. Following investigation of the product, the growth rate of exports to non-investigating destinations falls by 3 to 4 percentage points. This growth effect is delayed but persistent, though marginally insignificant in most periods. This evidence suggests that AD investigations against China have a dampening effect on growth of trade to alternative destinations. We include lines denoting the average across pre-treatment coefficients and the average across post-treatment coefficients, which more clearly outline the reduction in growth rates following investigation. However, this dampening effect is smaller in magnitude than the dampening effect within the target destination. Further, the point estimates have large confidence bands that suggest heterogeneity we are not picking up. We investigate this heterogeneity later, after discussing the baseline price effects.

Next, we examine what happens to prices in the non-investigating destinations in response to an investigation from a focal destination. In particular, we are looking for evidence (or lack thereof) of a "chilling effect" in prices that has been discussed in previous work – if exporters believe AD cases are correlated across destinations, one investigation raises the probability of investigation in other markets. To reduce this probability of investigation in the third markets, exporters may wish to raise prices of exports to destinations that have not (yet) initiated AD action against them. Table 1.7, panel (b) and Figure 5 document our results on price level and growth.

Table 1.7(b), columns 1–3 show that we cannot identify any change to the pricing behavior of exporting firms in non-investigating destinations following an investigation of the product, relative to non-target products within the same industry. Columns 4–6 display estimates from the growth rate regressions, which similarly suggest we cannot identify changes to the pricing behavior in third markets. Figure 1.5(a) displays the dynamic coefficients by length of exposure to an AD action in another market for the log unit value dependent variable, and Figure 1.5(b) shows the same for the unit value growth dependent variable. In both panels, we cannot identify significant deviations in average prices for treated products relative to control products.

In addition to the OLS results reported above, we also estimate (6) and (7) using both WLS and the estimator proposed by Sun and Abraham (2021). The overall ATTs are reported in Tables A11 and A12 for the two estimators, and the dynamic coefficients by length of exposure, aggregated across cohorts, are plotted in Figures A8–A11. The results are much the same as the OLS results reported above.

1.5.2.2 Exposure to the focal market

While we find consistently negative growth effects within the third market above, the large standard errors in Table 1.7 and large confidence bands of Figure 1.4(b) suggest further heterogeneity impacting our estimates. One source of this heterogeneity concerns the "importance" of the focal market. If the exporting destination targeted by AD action is responsible for a large share of Chinese exports, would the third market ATT be larger? To investigate the link between the size of Chinese exports within the treated market and the associated third market effect, we consider interacting the treatment variable with the export share variable s_i^{AD} included as a control that measures the share of the target product i exported to the investigating destination in the investigation year.

With this interaction, our question of interest concerns the effect of a higher or lower share of exports within the target destination on Chinese trade in that targeted product to other destinations, and how that modifies the effect of AD investigation – do products with higher exposure to the focal market see larger, or smaller reductions in growth across third markets? Table 1.8, panels (a) and (b) outline the results of this alternative specification across the same range of fixed effects we considered for the baseline results for quantity and unit value, respectively. Columns 1–3 display the estimates for the log level dependent, and 4–6 display the same for the growth dependent.[30] As above, we standardize the export share variable and report coefficients associated with a one standard deviation increase in the variable to facilitate interpretation.

First, columns 1–3 of panel (a) illustrate no significant modifying effect of export share to the ATT in levels. At the average focal market share, the ATTs are quantitatively similar to the baseline third market estimation, with AD investigation in a focal market associated with a 9 to 12 percent increase in the log level of trade volume.[31] However, columns 4–6 show the focal exposure does modify the effect of AD investigation on the growth rate. With a focal market share in the treatment year at the average, an AD investigation within the focal destination reduces growth rates in third markets by 1.3 to 1.4 percentage points. This estimate is of a similar magnitude to our baseline result in Table 7. An increase in the focal share by one standard deviation reduces growth rates by 2.4 percentage points among treated products. For a product with a focal share one standard deviation higher than average, the total growth effect of an AD investigation within the focal market is -3.5 to -3.8 percentage points, compared to non-treated products.

Importantly, if AD investigation induced *trade deflection*, we would expect the coefficient on $AD \times s_i^{AD}$ to be *positive* – the larger the share of exports in the focal market, the more excess capacity exporting firms should try to offload in, or deflect to, third markets. However, we find the exact opposite – the larger the share of exports in the focal market, the *more* exporting firms *reduce*

[30] Note that the HS6-ISO fixed effects of the level specifications absorb the first-order export share term, as in Table 7.

[31] The average focal market share across all products is 9.6%. Table 4 shows average shares by year of initiation.

their growth to third markets. This points to the likelihood of supply-side factors as a fundamental driver of firm export responses to AD policy – firms with larger shares of lost exports due to focal market AD activity should make larger adjustments to their scale or investment, which propagate through the rest of their exporting networks. Our findings here are consistent with this story. Before moving on, we note that panel (b) shows we cannot identify any price effects, even accounting for focal market heterogeneity.

1.5.2.3 Correlated AD imposition and the chilling effect

Next, while we do not identify a clear chilling effect in prices in either the main results or with the export share interaction, we are still concerned about the possibility of correlated AD cases across destinations having an impact on our estimates. In order to investigate this, we compute correlation coefficients from the AD data and re-estimate the model with these coefficients included as an interaction. First, from the AD data we obtain the list of all petitioning destinations that file AD investigations against China over the period 2000-2016 – there are 27 such destinations. Then, for each of these destinations we generate a binary indicator over all HS6 products in the sample that denote whether the destination filed against the HS6 product at least once over the period. We compute a correlation coefficient between each pair of these destination-specific list of binaries. This yields a 27-by-27 correlation matrix of filing behavior between all petitioning country pairs, across all products over the sample period.

We report the correlation matrix in Table A10 of the appendix. Each element signifies the degree of overlap between destinations, in the sense that a high correlation coefficient indicates the two destinations file against similar products. We believe that, if chilling effects are present, they can be captured via this variable – if an AD investigation is filed by some destination j against product i, then exports of i should fall more (or prices should rise more) in third markets where the correlation coefficient between that destination and filing destination j is higher. To integrate this measure into the data set, we reshape the correlation matrix into a pairwise form and merge into the trade data at the HS6 level.[32] As a final note, destinations that do not file have no AD activity to generate correlations, so the sample consists of the 27 destinations that file against China. We also drop non-treated products, as they do not have a linked focal market to generate a pairwise

[32] The first order term is absorbed by HS6-ISO fixed effects (focal market matches are made at the HS6 level, so destination-pairs are equivalent to HS6-ISO pairs), and thus is omitted from the level regressions.

correlation.

Table 1.9 displays the results of interacting the pairwise correlation measure with the treatment variable over the same range of fixed effects previously considered. Panels (a) and (b) report the results for quantity and unit value, respectively, with the log level dependent in columns 1–3 and the growth dependent in columns 4–6. We cannot identify an effect of AD case correlation on log level of export quantity, or growth in export quantity. We also cannot identify an effect of AD case correlation on prices. Prices and quantity do not seem to respond in any systematic way among investigated products when AD imposition is highly correlated between a third market and its linked focal market. This suggests that a supposed chilling effect is unlikely to be a key driver of the main results we document above – as a chilling effect arises through the (perceived) increased probability of AD imposition in a non-investigating destination following an investigation initiated by another destination.

1.5.2.4 Extended gravity

Perhaps rather than correlation in AD filing, broader similarities between destinations impact firm exporting behavior and induce export reallocation. These "extended gravity" considerations may imply an easier effort to reallocate exports from an investigating destination to ones nearby and with similar characteristics (Morales et al., 2019). To determine if this source of heterogeneity may be impacting our results, we construct an indicator variable denoting whether a third market exists within the same broader geographic region as the investigating destination for a given product.[33] Since non-treated products have no destination pair with which to calculate the indicator, we focus only on heterogeneity within treated products.

We report the results to this estimation in Table 1.10. As before, we present the level results in columns 1–3 and the growth results in columns 4–6, where the comparison is between treated products and not-yet-treated products. AD investigation by a focal market within the same region as the third market results in a growth rate of trade volume 1.5 to 1.8 percentage points lower that target products exported to a third market further away, while having no effect on the level of trade or prices. Proximity to the sanctioning destination leads to a *larger* reduction in export growth, rather than inducing more reallocation. Within a region, export markets are complementary – a

[33]We use the following delineations of region: America (North/Central/South), Europe, Africa (Northern/Sub-Saharan), Asia (Central/Eastern/South-eastern/Southern/Western), Caribbean, and Oceania.

trade shock in one market leads to lower growth in exports to other markets within the same region. This also suggests that fixed costs of exporting are not country-specific as argued in Morales et al. (2019) – just as they find easier entry in countries close to an existing trading partner, we find easier exit in countries close to where a trade shock occurs.

1.5.2.5 Target exporters beyond China

As a final robustness step, we also extend the analysis beyond Chinese exports as in section 1.5.1.2. For similar developing economies, we consider AD activity targeting India, Indonesia, Malaysia, and Thailand – four large developing exporters that are also among the most frequent targets of AD action. For developed targets of AD action, we consider the United States, the European Union, Japan, and South Korea which are also among the most frequent targets of AD action. We merge in AD activity initiated by the top 10 filers against these destinations (excluding cases initiated by any of the four), using the first case chronologically, within product and across petitioners, and exclude focal product-destination pairs from the sample. The results for developing and developed targets are reported in Table A13 and A14 in the appendix, respectively. From this, we can confirm that the findings we document here are not China-specific (and further, not even specific to export-oriented developing economies) but rather apply broadly across many common targets of AD action – most importantly the distinction between the level effects and the growth effects of AD activity on trade to non-target markets.

To summarize, in non-investigating destination markets we find several key results. First, while level regressions imply trade deflection, the growth of future trade volume falls among target products and this deviation from the trend implies trade destruction. This dampening effect is larger for products where the investigating destination is an important market. We cannot identify an effect of AD filing on prices in third markets, nor an effect of correlation among AD filing behavior on either prices or export growth for investigated exports to non-investigating destinations. This suggests that a chilling effect on prices or export growth is not the primary source of the reduction in growth we see across third markets. Finally, we show that “extended gravity” considerations (such as geographic proximity) influence reallocation within treated products following AD investigation in a manner consistent with exporting scale economies. These key findings seem to extend beyond our setting of Chinese trade, and apply more broadly to export-oriented developing economies.

1.6 Conclusion

In this paper, we use a dynamic DiD methodology to examine the effect of AD investigation on trade flows to both the investigating destination, as well as non-investigating destinations. We first establish a relationship between pre-treatment export growth and AD imposition, which suggests a selection issue that potentially compromises the parallel trends assumption and thus canonical DiD estimation of the trade effects of AD policy. We find significant trends in the level of import volume and import share prior to treatment within the focal market, as well as significant trends in the level of import volume prior to treatment across third markets. With these trends in mind, we revisit classic questions in the AD literature through the lens of growth, using differenced specifications to estimate the effect of AD investigation on the growth of trade volume, import share, and average prices.

Within the focal markets, we find that AD investigations lead to significant and persistent reductions in the growth rate of import volume and the import share for target products relative to non-target products within the same industry. Within non-investigating markets, we also find significant and persistent reductions in the growth rate of trade volume for target products relative to non-target products within the same industry when the product is faced with an investigation in some other export destination. These growth reductions are larger in magnitude the larger the share of that HS6 product is exported to the relevant investigating destination, and larger in magnitude in export markets within the same region as the sanctioning country. Ignoring these growth effects in the third market leads to the finding that AD investigations induce trade deflection, as the level of trade volume increases post-treatment – but at a much slower rate. However, accounting for the growth trends suggests that in response to AD investigations, firms reduce export growth *globally*. We also find no robust evidence of a chilling effect in prices or export growth within non-investigating destinations, whereby firms raise price or slow growth to reduce the probability of investigation.

Our findings imply that the AD tariff shock in the focal market is likely to generate negative effects on supply-side factors of exporting that are product- but not market-specific. In particular, if exporting exhibits scale economies and export markets are complementary, the loss of one market via AD imposition may lead to the loss of exports to unrelated markets due to scale effects or reductions in investment and innovation. While our results are consistent with scale- and investment-driven channels, we cannot disentangle these mechanisms.

Table 1.1: AD cases, products, success, and duration by petitioning country

				Duration (Years)	
Destination	Cases	HS6 Products	Success	Mean	σ
India	154	207	0.86	8.92	3.96
United States	116	310	0.77	10.98	4.45
European Union	91	215	0.77	9.04	3.50
Brazil	83	130	0.69	7.79	3.10
Turkey	81	156	0.95	12.31	4.58
Argentina	70	144	0.79	9.25	4.17
Mexico	47	87	0.68	8.23	4.54
Colombia	41	105	0.59	7.05	3.26
Australia	39	60	0.59	6.61	3.09
Canada	36	102	0.75	8.68	3.43
South Africa	26	35	0.31	9.67	4.50
South Korea	20	28	0.85	8.35	3.44
Indonesia	19	48	0.58	7.80	2.62
Thailand	19	63	0.84	8.12	3.36
Pakistan	17	45	0.71	6.08	2.35
Peru	15	58	0.80	9.00	5.04
Russia	12	41	0.92	6.27	2.57
Malaysia	11	37	0.82	4.67	1.58
Taiwan	10	40	0.60	8.83	3.25
Ukraine	7	16	1.00	6.86	2.27
Israel	6	12	0.67	5.50	2.89
New Zealand	5	5	0.80	7.00	4.16
Japan	3	3	1.00	7.33	4.16
Trinidad & Tobago	2	12	1.00	6.00	1.41
Jamaica	1	2	1.00	5.00	–
Philippines	1	1	0.00	–	–
Uruguay	1	1	1.00	9.00	–

Table 1.2: Summary statistics

Variable	N	Mean	Std. Dev.	Min	Pctl. 25	Pctl. 75	Max
Sample: focal market							
log(quantity)	277,612	12	3.7	-4.4	9.6	14	31
quantity growth	282,332	15	107	-200	-35	76	200
import share	387,481	0.18	0.25	0	0.000043	0.28	1
import share growth	293,838	14	94	-200	-20	52	200
Sample: third market							
log(quantity)	2,700,595	9.7	3.7	-6.2	7.5	12	26
quantity growth	2,986,767	15	134	-200	-81	129	200
Export share statistics							
focal market shares	796	0.083	0.1	0.00071	0.013	0.12	0.55
linked market shares	796	0.0054	0.004	0.00054	0.003	0.0065	0.066
unlinked market shares	5095	0.02	0.028	0.001	0.0066	0.022	0.53

Table 1.3: Summary statistics: export shares

Panel (a): export shares by year, newly investigated products

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
focal	0.083	0.064	0.16	0.12	0.1	0.1	0.06	0.12	0.096	0.086	0.1	0.12	0.13	0.074	0.08	0.1
linked	0.91	0.94	0.82	0.88	0.9	0.9	0.94	0.88	0.87	0.88	0.9	0.85	0.87	0.93	0.92	0.9
N	101	69	60	45	76	59	93	42	94	83	59	34	27	34	14	31

Panel (b): total export share by ever-treated status

	2003	2007	2011	2015
treated	0.3	0.34	0.35	0.35
never-treated	0.7	0.66	0.65	0.65

Note: Panel (a) reports HS6-year export shares averaged across products; focal ISO is the first destination chronologically investigated the product. Focal and linked shares may not sum to 100% in the case an HS6 product is investigated by another focal destination. N denotes number of products investigated that year. Panel (b) reports the share of exports each year in products ever-treated over the sample period compared to never-treated.

Table 1.4: Average growth rates by treatment status

	All products		Same HS4	
Treatment	volume	share	volume	share
0	13.582	12.451	15.351	14.097
1	31.125	24.594	31.125	24.594

Table 1.5: Ex-ante growth and AD initation

Dependent Variable:		AD_{ijt}	
Model:	(1)	(2)	(3)
	OLS	Probit	Logit
Variables			
g_{ijt}	0.01288***	0.10117***	0.16415***
	(0.00371)	(0.02457)	(0.04494)
Fixed-effects			
HS2-year	Yes	Yes	Yes
dest_iso	Yes	Yes	Yes
Fit statistics			
Observations	7,450	5,524	5,524
Squared Correlation	0.35885	0.17949	0.18076
Pseudo R^2	0.54208	0.18164	0.18129
BIC	6,970.0	5,954.9	5,956.6

Note: Robust standard errors in parentheses, g_{ijt} standardized;
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.6: Effect of AD investigation in the focal market

Dependent:	Level			Growth		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): Quantity						
AD	-0.4423***	-0.5166***	-0.3570***	-12.09***	-11.27***	-10.50***
	(0.0453)	(0.0415)	(0.0474)	(0.6149)	(0.6475)	(0.6572)
Observations	277,612	277,612	277,612	282,332	282,332	282,332
R^2	0.84688	0.85795	0.89699	0.14278	0.14446	0.19185
Within R^2	0.00124	0.00173	0.00068	0.00054	0.00041	0.00036
Panel (b): Import share						
AD	-0.0120**	-0.0182***	-0.0044	-10.26***	-9.257***	-9.142***
	(0.0047)	(0.0047)	(0.0056)	(0.5059)	(0.5391)	(0.5493)
Observations	387,481	387,481	387,481	293,838	293,838	293,838
R^2	0.68358	0.69784	0.74746	0.10913	0.11039	0.14488
Within R^2	0.00008	0.00019	0.00001	0.00047	0.00034	0.00033
Panel (c): Unit Value						
AD	0.1038***	0.1128***	0.0683***	0.0279***	0.0218***	0.0184**
	(0.0165)	(0.0165)	(0.0203)	(0.0039)	(0.0040)	(0.0041)
Observations	277,612	277,612	277,612	243,043	243,043	243,043
R^2	0.95604	0.95760	0.97004	0.10974	0.11128	0.16863
Within R^2	0.00025	0.00029	0.00009	0.00004	0.00002	0.00002
Fixed effects						
ISO-year	✓	✓	–	✓	✓	–
HS-ISO	HS-6	HS-6	HS-6	–	HS-2	–
HS-year	HS-2	HS-4	–	HS-4	HS-4	HS-4
HS-year-ISO	–	–	HS-4	–	–	HS-2

Note: Standard errors in parentheses are clustered at the HS6-ISO level; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.7: Effect of AD investigation within third markets

Dependent:	Level			Growth		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): Quantity						
AD	0.1163***	0.1431***	0.1306***	-1.169**	-1.188**	-1.178**
	(0.0348)	(0.0349)	(0.0401)	(0.5747)	(0.5779)	(0.5876)
s_i^{AD}				0.3288	0.3153	0.3207
				(0.2465)	(0.2479)	(0.2526)
Observations	2,352,405	2,352,405	2,352,405	2,613,188	2,613,188	2,613,188
R^2	0.84548	0.85355	0.90499	0.08407	0.08594	0.15308
Within R^2	0.00028	0.00031	0.00035	0.00001	0.00001	0.00001
Panel (b): Unit Value						
AD	0.0008	-0.0005	0.0032	-0.0016	-0.0017	-0.0020
	(0.0163)	(0.0155)	(0.0180)	(0.0037)	(0.0037)	(0.0038)
s_i^{AD}				0.0025*	0.0026*	0.0027*
				(0.0014)	(0.0014)	(0.0015)
Observations	2,352,405	2,352,405	2,352,405	1,854,017	1,854,017	1,854,017
R^2	0.90416	0.91084	0.93897	0.09378	0.09579	0.14474
Within R^2	0.00000	0.00000	0.00000	0.00000	0.00001	0.00001
Fixed effects						
ISO-year	✓	✓	–	✓	✓	–
HS-ISO	HS-6	HS-6	HS-6	–	HS-2	–
HS-year	HS-2	HS-4	–	HS-4	HS-4	HS-4
HS-year-ISO	–	–	HS-4	–	–	HS-2

Note: s_i^{AD} is standardized. Standard errors in parentheses are clustered at the HS6 level; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.8: Effect of AD investigation within third markets: export share heterogeneity

Dependent:	Level			Growth		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): Quantity						
AD	0.1100^{***}	0.1404^{***}	0.1280^{***}	-1.294^{**}	-1.309^{**}	-1.299^{**}
	(0.0364)	(0.0356)	(0.0406)	(0.5748)	(0.5781)	(0.5877)
$AD \times s_i^{AD}$	0.0135	0.0055	0.0054	-2.428^{***}	-2.428^{***}	-2.358^{***}
	(0.0240)	(0.0259)	(0.0294)	(0.5174)	(0.5209)	(0.5304)
s_i^{AD}				2.078^{***}	2.064^{***}	2.021^{***}
				(0.4032)	(0.4064)	(0.4122)
Observations	2,352,405	2,352,405	2,352,405	2,613,188	2,613,188	2,613,188
R^2	0.84548	0.85355	0.90499	0.08411	0.08598	0.15312
Within R^2	0.00029	0.00032	0.00035	0.00006	0.00006	0.00006
Panel (b): Unit Value						
AD	-0.0086	-0.0074	-0.0051	-0.0013	-0.0014	-0.0017
	(0.0168)	(0.0152)	(0.0176)	(0.0036)	(0.0036)	(0.0038)
$AD \times s_i^{AD}$	0.0201^{*}	0.0139	0.0170	0.0042	0.0041	0.0038
	(0.0107)	(0.0123)	(0.0139)	(0.0030)	(0.0030)	(0.0031)
s_i^{AD}				-0.0007	-0.0005	-0.0001
				(0.0023)	(0.0023)	(0.0025)
Observations	2,352,405	2,352,405	2,352,405	1,854,017	1,854,017	1,854,017
R^2	0.90416	0.91084	0.93898	0.09378	0.09579	0.14474
Within R^2	0.00006	0.00002	0.00005	0.00001	0.00001	0.00001
Fixed effects						
ISO-year	✓	✓	–	✓	✓	–
HS-ISO	HS-6	HS-6	HS-6	–	HS-2	–
HS-year	HS-2	HS-4	–	HS-4	HS-4	HS-4
HS-year-ISO	–	–	HS-4	–	–	HS-2

Note: s_i^{AD} is standardized. Standard errors in parentheses are clustered at the HS6 level; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.9: AD investigation within third markets: AD case correlation, treated products

Dependent:	Level			Growth		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): Quantity						
AD	0.0888*	0.0523	0.0789	-7.491***	-7.526***	-7.500***
	(0.0458)	(0.0540)	(0.0487)	(1.475)	(1.486)	(1.529)
AD × *corr*	0.0008	-0.0029	0.0020	0.0275	0.0145	-0.0004
	(0.0025)	(0.0019)	(0.0030)	(0.0416)	(0.0440)	(0.0482)
corr				-0.0725**	-0.0328	-0.0245
				(0.0367)	(0.0399)	(0.0431)
Observations	308,482	308,482	308,482	312,672	312,672	312,672
R^2	0.83879	0.86008	0.85384	0.17397	0.17621	0.24014
Within R^2	0.00032	0.00001	0.00036	0.00031	0.00031	0.00034
Panel (b): Unit Value						
AD	-0.0094	0.0077	-0.0132	0.0002	0.0003	-0.0014
	(0.0162)	(0.0165)	(0.0171)	(0.0059)	(0.0060)	(0.0062)
AD × *corr*	0.0007	0.0010	0.0010	0.0002	0.0002	0.0004
	(0.0008)	(0.0006)	(0.0010)	(0.0002)	(0.0003)	(0.0003)
corr				0.0002	0.0003	0.0002
				(0.0002)	(0.0002)	(0.0002)
Observations	308,482	308,482	308,482	271,971	271,971	271,971
R^2	0.91331	0.92581	0.91933	0.20484	0.20626	0.26484
Within R^2	0.00001	0.00001	0.00003	0.00002	0.00002	0.00003
Fixed effects						
ISO-year	✓	✓	–	✓	✓	–
HS-ISO	HS-6	HS-6	HS-6	–	HS-2	–
HS-year	HS-2	HS-4	–	HS-4	HS-4	HS-4
HS-year-ISO	–	–	HS-2	–	–	HS-2

Note: estimates only using ever-treated products and destinations that file AD against China. Standard errors in parentheses are clustered at the HS6 level; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

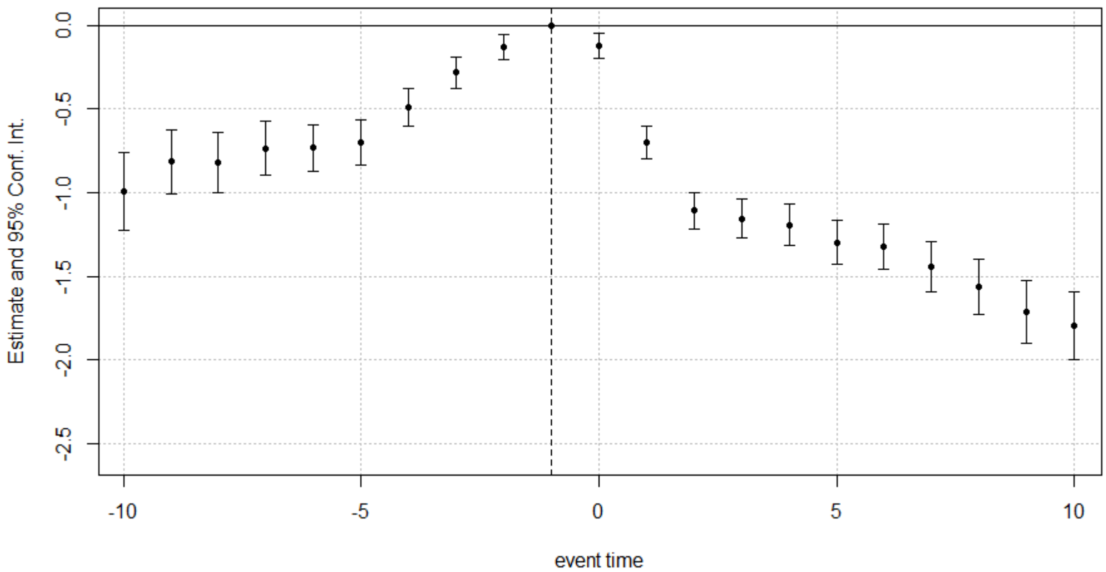
Table 1.10: Effect of AD investigation in third markets, region indicator

Dependent:	Level			Growth		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): Quantity						
AD	0.0463	0.0142	0.0269	-6.942***	-7.039***	-6.991***
	(0.0387)	(0.0506)	(0.0714)	(1.539)	(1.545)	(1.598)
AD × same_region	0.0000	0.0268	0.0449	-1.590**	-1.461*	-1.872**
	(0.0394)	(0.0346)	(0.0788)	(0.8009)	(0.8217)	(0.9204)
same_region				0.7525	1.144	1.424*
				(0.6922)	(0.7183)	(0.8106)
Observations	1,063,019	1,063,019	1,063,019	1,161,452	1,161,452	1,161,452
R^2	0.82555	0.83756	0.92530	0.10672	0.10978	0.22371
Within R^2	0.00006	0.00001	0.00002	0.00014	0.00014	0.00015
Panel (b): Unit Value						
AD	0.0039	0.0009	0.0043	-0.0059	-0.0067	-0.0069
	(0.0186)	(0.0180)	(0.0248)	(0.0056)	(0.0057)	(0.0060)
AD × same_region	-0.0087	-0.0030	-0.0220	-0.0041	-0.0056	-0.0060
	(0.0122)	(0.0104)	(0.0315)	(0.0055)	(0.0058)	(0.0066)
same_region				0.0045	0.0030	0.0032
				(0.0048)	(0.0051)	(0.0056)
Observations	1,063,019	1,063,019	1,063,019	853,680	853,680	853,680
R^2	0.88615	0.89720	0.95341	0.14841	0.15257	0.25488
Within R^2	0.00000	0.00000	0.00001	0.00000	0.00000	0.00000
Fixed effects						
ISO-year	✓	✓	–	✓	✓	–
HS-ISO	HS-6	HS-6	HS-6	–	HS-2	–
HS-year	HS-2	HS-4	–	HS-4	HS-4	HS-4
HS-year-ISO	–	–	HS-4	–	–	HS-2

Note: estimates only using ever-treated products. Standard errors in parentheses are clustered at the HS6 level; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure 1.1: Effect of AD on import volume in focal markets

(a) Log import volume



(b) Growth in import volume

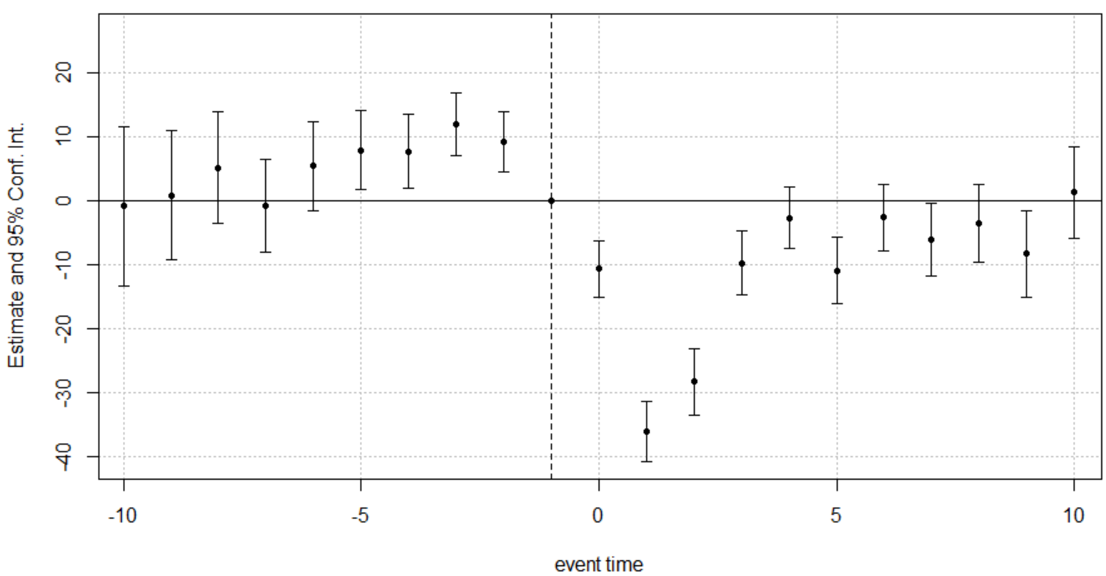
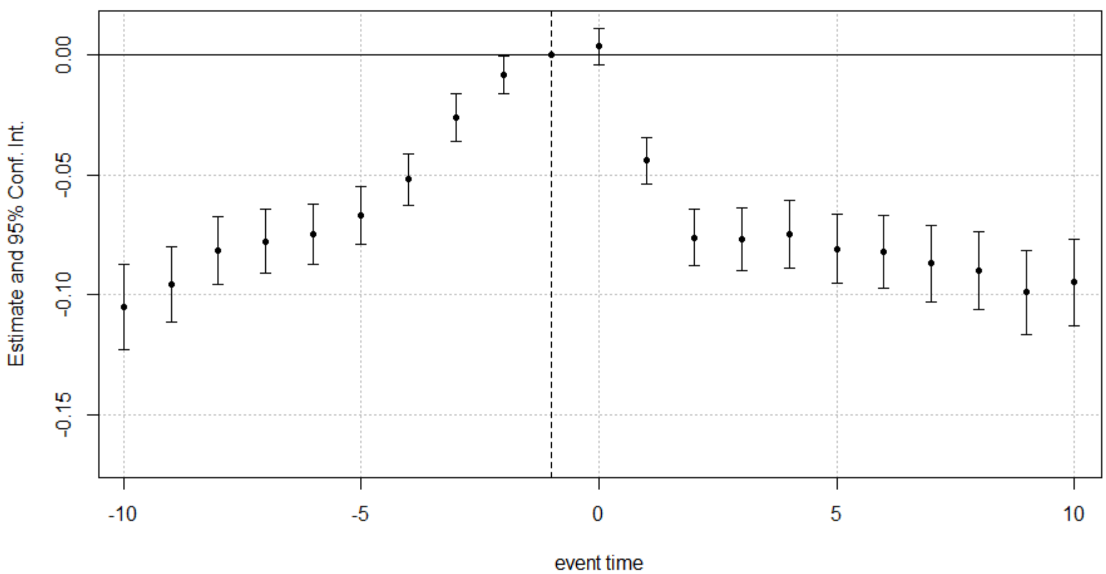


Figure 1.2: Effect of AD on import share in focal markets

(a) Import share



(b) Growth in import share

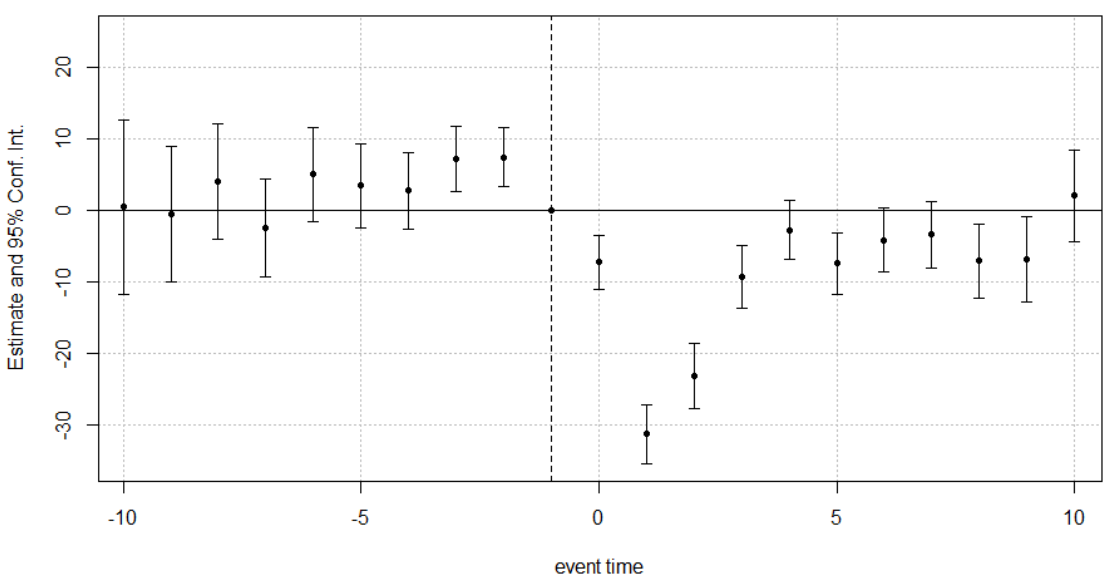
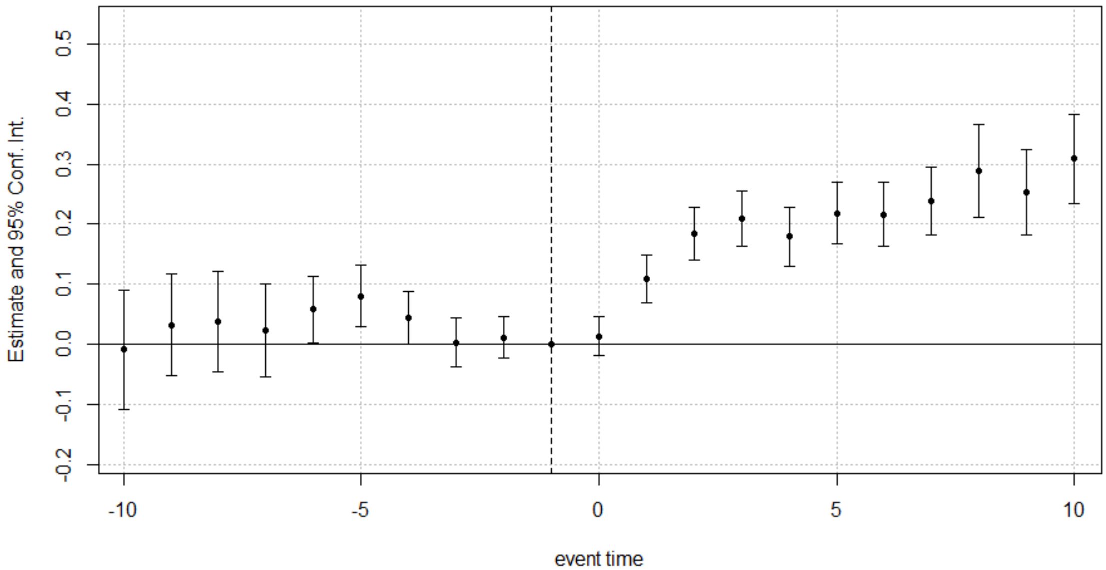


Figure 1.3: Effect of AD on unit value in focal markets

(a) Log unit value



(b) Growth in unit value

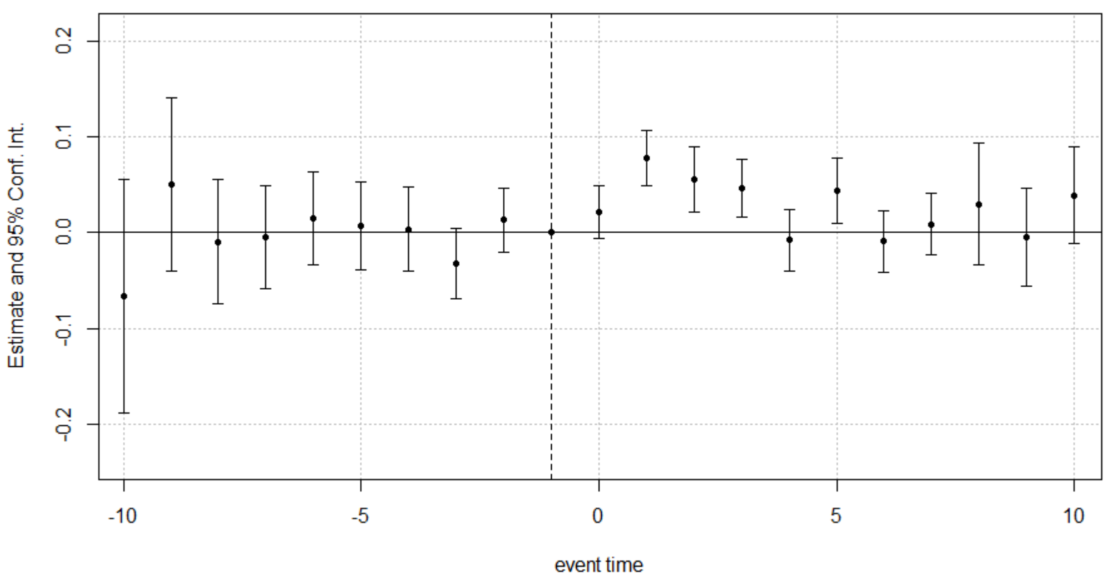
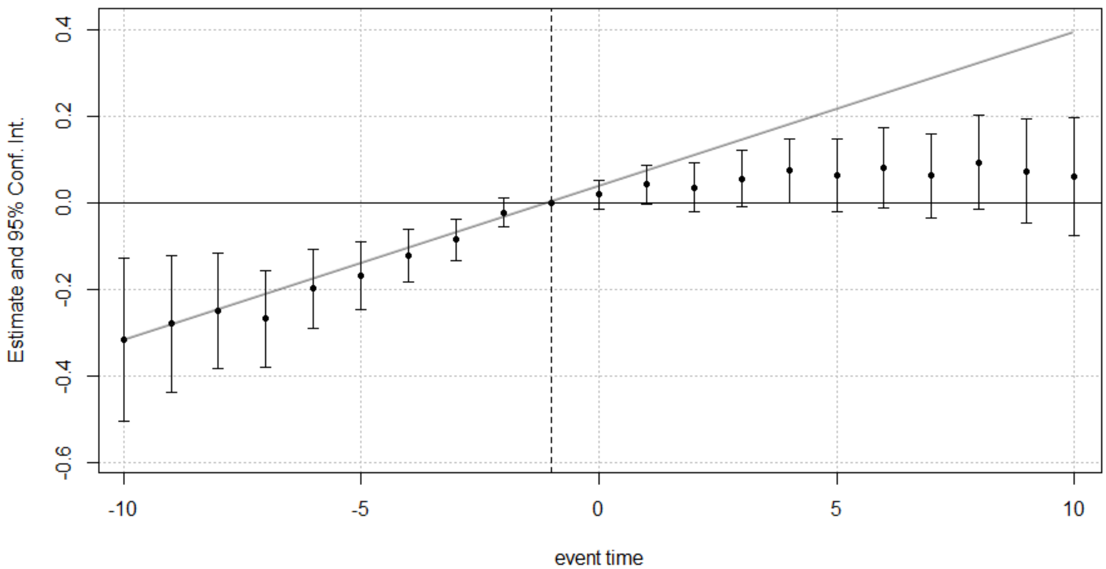


Figure 1.4: Effect of AD on import volume in non-target markets

(a) Log import volume



(b) Growth in import volume

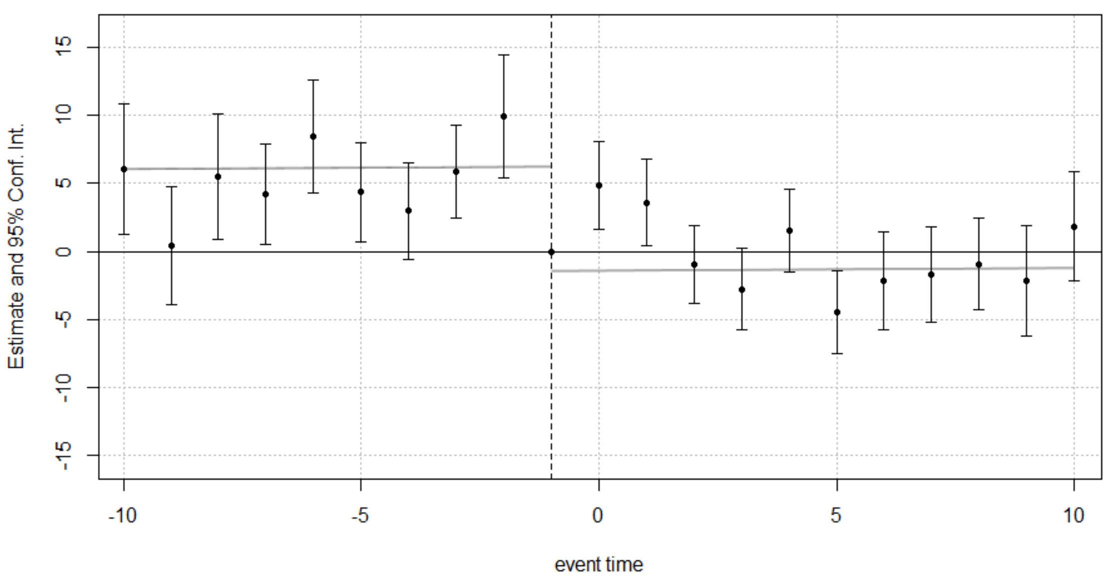
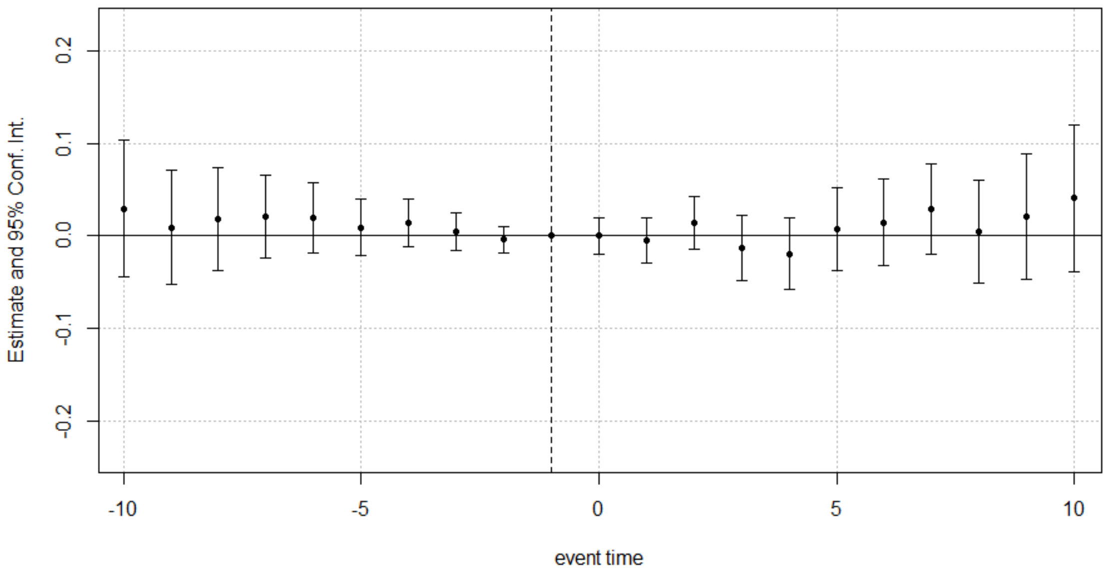
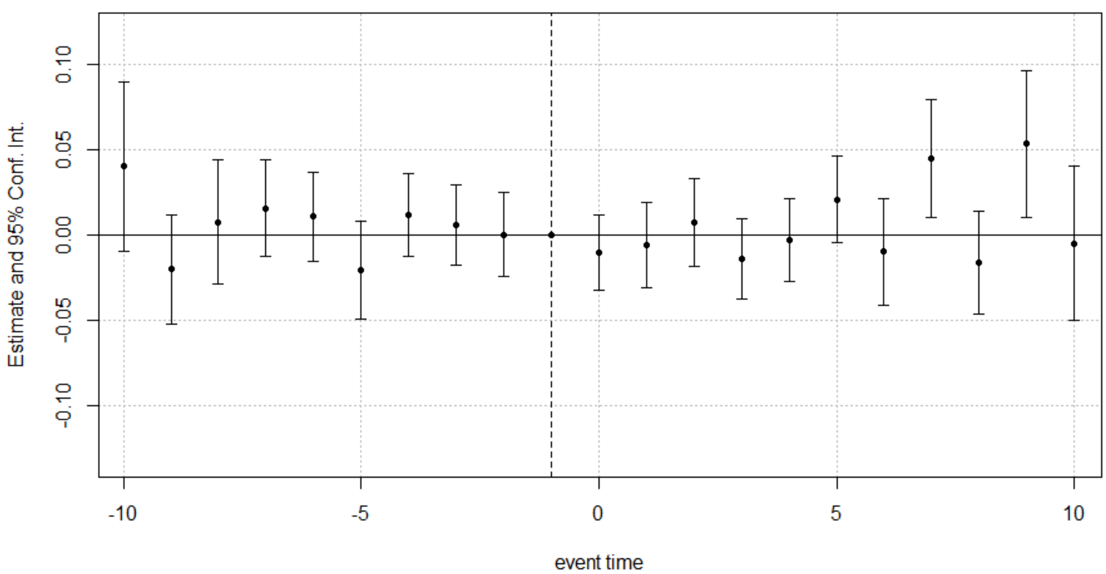


Figure 1.5: Effect of AD on unit value in non-target markets

(a) Log unit value



(b) Growth in unit value



Chapter 2

Economic Determinants of EIA Formation

2.1 Introduction

Economic integration agreements (EIAs) are widespread multilateral agreements between countries with a wide range of possible characteristics based on the level of integration and the provisions adopted by partnered countries. These agreements can range from narrow preferential trade agreements (PTAs) that seek to reduce tariffs in a specific industry to expansive economic unions that liberalize the movement of factors of production and harmonize domestic and macroeconomic policy. Since the end of the Cold War, membership in EIAs dramatically expanded, with particular growth in deeper agreements that go beyond simple tariff reductions. During this time a vast literature developed, concerned with estimating the effect of EIA or trade agreement partnership on a large variety of economic outcomes (Anderson & Yotov, 2016; Baier & Bergstrand, 2007; Baier et al., 2014, 2018, 2019; Egger et al., 2011). In particular, recent work highlighted the *heterogeneous effects* of EIAs depending on the depth of the agreement (Baier et al., 2018). While there is also work examining the economic (and non-economic) determinants of agreement formation, relatively little studies depth of agreements, nor investigates heterogeneity across agreement types.[1] As EIAs become increasingly deep and multifaceted, the role of economic forces and characteristics in shap-

[1] For political economy determinants, see Grossman (2016), Hinz (2023), Maggi and Ossa (2021), and McLaren (2016). For economic determinants, see Baier and Bergstrand (2004), Egger and Larch (2008), and Limão (2016).

ing agreements is vital to a comprehensive understanding of the trade and welfare effects of EIA partnership and formation.

In this paper, we study the economic determinants of EIA formation with a focus on heterogeneity across broad agreement types, ranging from non-reciprocal PTAs to economic unions. First, we build a spatial model of trade with labor mobility that delivers gravity in both goods trade and migration as functions of country characteristics and bilateral trade and migration costs respectively. Then using a bilateral panel of EIA, trade, migration, and gravity data from 1990-2015, we conduct several empirical exercises. First, we estimate the gravity of migration to investigate the relationship between migration flows and formation of different types of EIAs, providing novel evidence of the migration effects of EIAs. Then, we select a few economic variables implied by the model and gravity estimation and estimate their contribution to the probability of EIA formation, before building a random forest model to validate the importance of these variables and predict out-of-sample EIA formation. We conclude by outlining a simple model parameterization and counterfactual exercise as the next steps in this research, which will deliver general equilibrium effects of EIAs on trade and migration and allow us to characterize the welfare effects across different counterfactual agreement types, as well as examine how these welfare effects change when fundamental economic characteristics of trading partners change.

Our first contribution to the literature is a simple model delivering structural gravity of both goods trade and migration. Trade and migration flows depend on explicit trade and migration costs, as well as prices, wages, income, population, and location amenities in both locations. The model allows us to add heterogeneous EIAs as bundles of distinct trade and migration cost reductions, and lets migration flows respond to deeper EIAs directly through a reduction in migration costs as well as indirectly through larger trade effects that induce migration through changes in the indirect utility of moving or staying. The model's structure permits us to estimate both partial equilibrium and general equilibrium effects of EIA formation on bilateral migration flows, which along with trade effects can characterize the welfare changes from differential EIA membership.

In the partial equilibrium exercise, we find novel and heterogeneous effects of EIAs on migration by agreement type, and also document a nonlinear relationship between deeper EIA formation and migration flows depending on the distance of the member countries. Deeper EIAs result in larger migration flows than more shallow EIAs, but shallow EIAs like preferential trade agreements and free trade agreements still result in some statistically significant migration flows.

This channel is supported by the model, which permits migration to arise indirectly through trade-creating effects of EIAs that do not explicitly reduce migration costs. Among deeper EIAs, we find a "U-shaped" relationship between EIA-induced bilateral migration and bilateral distance – countries that share borders see either no change in or *less* migration post-EIA, while non-adjacent country-pairs see an increase in migration flows that declines in magnitude as the country-pair becomes more distant.

The model implies a set of bilateral economic variables associated with changes in trade and migration flows, and thus welfare changes from EIA membership – in particular, bilateral variables that shape trade and migration costs such as distance and contiguity, as well as the economic size and income differentials of country-pairs, which is consistent with the framework of Baier and Bergstrand (2004). We investigate the relationship between these variables first in a standard reduced-form manner, estimating the effect of variation in these underlying characteristics on the probability of EIA membership across agreement types over the period 1990-2015. However, this approach is restrictive and the fit is poor, so we leverage an alternative strategy using a random forest, a supervised machine learning technique popularized by Breiman (2001). EIA formation is modeled as a classification problem and we estimate classification trees that sort observations into various sub-partitions based on a series of decision rules related to our economic variables, which are considered "predictors" of classification.

The random forest method consists of estimating a number of these trees using a randomly selected subset of the predictor space, and yields (1) information about the importance of certain variables in contributing to accurate classification, and (2) a framework to predict out-of-sample EIA formation. The results of the random forest indeed suggest the relationship between economic characteristics and agreement formation are heterogeneous across agreement type – for example, contiguity is much more important for reciprocal PTAs than non-reciprocal PTAs; and income differentials are generally more important in shallower agreements while income sums are more important for deeper agreements. Distance is an important factor across the board, but contributes to accurate classification more so among deeper agreements. A prediction exercise suggests these few economic variables successfully predict new out-of-sample EIA formation across agreement types with over 90% accuracy.

This paper contributes to a relatively new and still-underdeveloped literature examining the determinants of EIA formation. While a deep literature exists studying political economy determinants (Conconi et al., 2014; Grossman, 2016; Hinz, 2023; Maggi & Ossa, 2021; McLaren, 2016), less examines strictly economic determinants of EIAs. Most work studying economic determinants focuses on the classical trade-related determinants of trade creation, trade diversion, and prices popularized by Viner (1950). Limão (2016) provides a survey of the research on agreement determinants, highlighting that evidence for other economic determinants beyond trade-related forces has been limited, though some recent work has examined terms-of-trade externalities (Blanchard et al., 2016) and the role of foreign direct investment (Antràs & Staiger, 2012; Osnago et al., 2019). Baier and Bergstrand (2004) and Egger and Larch (2008) are most similar to the current work, and seek to answer similar questions about strictly economic determinants of agreement formation. However, these papers do not consider agreement depth, and do not address the role of factor (particularly labor) mobility which we believe integral to an understanding of deeper agreements. We contribute to this literature by highlighting the role of migration flows as a major factor in determining the welfare effects of EIAs, and by revisiting the question of agreement formation with a new model and new empirical approaches. While Orefice (2015) also documents a relationship between PTA formation and migration flows, we broaden the analysis, show novel non-linear relationships between these migration effects and distance of trading partners, and tie them into a theory that can quantitatively evaluate agreement formation.

Our findings also have implications for the wider literature estimating the effects of EIAs on economic outcomes (Anderson & Yotov, 2016; Baier & Bergstrand, 2007; Baier et al., 2014, 2018, 2019; Egger et al., 2011). While the issue of trade policy endogeneity is well studied in the trade literature, the proliferation of deeper agreements and expansion in coverage may introduce further endogeneity bias into estimation. Recent work highlights provision-level heterogeneity in the effect on the forces of trade diversion and trade creation (Baier & Regmi, 2023; Mattoo et al., 2022). If deep agreement provisions have heterogeneous trade (and perhaps migration) effects, they will generate different welfare gains for a country-pair and thus have implications for the specific provision-level shape of an agreement.

In the next section, we outline the framework of our theoretical model and derive structural gravity of goods trade and migration in a general equilibrium setting. In section 2.3, we briefly discuss our data before turning to the empirical methodologies in section 2.4. Section 2.5 presents the results, section 2.6 outlines a simulation exercise using the model, and section 2.7 concludes.

2.2 A Model of Trade and Migration

In this section, we present a spatial model of trade and migration inspired by the framework of Monte et al. (2018). Workers live and produce varieties of differentiated goods across a number of regions, and can freely move between regions. The model delivers gravity equations for both goods trade and migration, and provides a setting to introduce EIA formation as a bundle of trade and migration cost shocks.

2.2.1 Preferences

We assume there is a mass of workers L_j that live within a region j who are geographically mobile. The preferences of a worker who begins in region j are defined over final goods consumption (C_j), land use (H_j), and a location amenity (b_j), according to the Cobb-Douglas form

$$U_j = b_j \left(\frac{C_j}{\alpha}\right)^{\alpha} \left(\frac{H_j}{1-\alpha}\right)^{1-\alpha}. \tag{2.1}$$

Following the new economic geography literature, we model goods consumption as a constant elasticity of substitution (CES) function of consumption of a number of tradable varieties m_i sourced from each location i,

$$C_j = \left[\sum_{i=1}^{N} \sum_{m_i=1}^{M_i} c_{ijm_i}^{\rho}\right]^{1/\rho}, \quad \rho = \frac{\sigma-1}{\sigma}. \tag{2.2}$$

Utility maximization implies equilibrium consumption of workers in location j of each variety m_i sourced from location i is given by

$$c_{ijm_i}^{W} = \alpha \frac{E_j}{P_j} \left(\frac{p_{ijm_i}}{P_j}\right)^{-\sigma}$$

where $E_j = w_j L_j$ is aggregate expenditure in location j, p_{ijm_i} is the price of a variety m_i produced in i and consumed in j, and P_j is the price index dual to (2.2), given by

$$P_j = \left[\sum_{i=1}^{N} \sum_{i=m_i}^{M_i} p_{ijm_i}^{1-\sigma}\right]^{\frac{1}{1-\sigma}}.$$

Utility maximization also implies that a fraction $(1-\alpha)$ of worker income is spent on housing. We assume there exists a group of immobile landlords who own the housing stock, and rent out the housing stock to the L_j households in each region j. Landlords spend this received housing expenditure on the composite consumption good C_j, with equilibrium consumption equal to

$$c_{ijm_i}^{LL} = (1-\alpha)\frac{E_j}{P_j}\left(\frac{p_{ijm_i}}{P_j}\right)^{-\sigma}.$$

Combining workers and landlords then delivers total expenditure on consumption goods of variety m_i produced in i, which is given by

$$p_{ijm_i}c_{ijm_i} = \left(\frac{p_{ijm_i}}{P_j}\right)^{1-\sigma} w_j L_j. \tag{2.3}$$

To complete the model on this side, we have total housing expenditures $r_j H_j = (1-\alpha)w_j L_j$, and the supply of housing Z_j^{ϵ}. For simplicity, we will assume housing supply is equivalent to land area supply and is therefore perfectly inelastic.

2.2.2 Production

Production of tradable final goods is done by workers in an industry characterized by monopolistic competition and increasing returns to scale. A variety m_i is produced with fixed cost F_i and constant marginal cost $1/A_i$, where A_i is the productivity available to all firms in i. This implies the total amount of labor ℓ_{im_i} required to produce x_{im_i} units of a variety m_i in location i is given by

$$\ell_{im_i} = F_i + \frac{x_{im_i}}{A_i}$$

Profit maximization implies equilibrium prices are a constant markup over marginal cost, given by

$$p_{ijm_i} = \left(\frac{\sigma}{1-\sigma}\right)\frac{\tau_{ij}w_i}{A_i}$$

where $\tau_{ij} > 1$ represent iceberg trade costs of goods shipped from i to j, and w_i represent the wage rate in location i. Zero profit then implies that equilibrium output of each variety is equal to a constant $x_{im_i} = (\sigma-1)A_iF_i$. This constant equilibrium output combined with labor market clearing then implies that the measure of varieties M_i is given by $M_i = L_i/(\sigma F_i)$.

2.2.3 Goods trade

Given the setup of the model, bilateral trade flows between locations i, j are equal to total expenditure on consumption goods m_i produced in i, presented in (2.3), aggregated across varieties. This is given by

$$X_{ij} = p_{ij}c_{ij} = \sum_{m_i} \left(\frac{p_{ijm_i}}{P_j}\right)^{1-\sigma} w_j L_j$$

where P_j is the price index dual to (2.2). Iceberg trade costs imply the landed price in location j is equal to the factory gate price scaled by the trade costs, or $p_{ij} = p_i\tau_{ij}$. Identical technology across firms within a country implies $\sum_{m_i} p_{ijm_i} = M_i p_{ij}$ where M_i is the measure of varieties. Taken together with labor market clearing, these imply we can rewrite bilateral trade flows as

$$X_{ij} = \left(\frac{\tau_{ij}}{\Pi_i P_j}\right)^{1-\sigma} w_i L_i w_j L_j \tag{2.4}$$

where Π_i and P_j are the "multilateral resistance" terms of Anderson and van Wincoop (2003), defined as

$$P_j^{1-\sigma} = \sum_i \left(\frac{\tau_{ij}}{\Pi_i}\right)^{1-\sigma} w_i L_i \tag{2.5}$$

$$\Pi_i^{1-\sigma} = \sum_j \left(\frac{\tau_{ij}}{P_j}\right)^{1-\sigma} w_j L_j. \tag{2.6}$$

To close out the model, factor prices are pinned down by the above system and the labor market clearing condition, which yields

$$w_i = B\left(\frac{A_i}{\Pi_i}\right)^{\frac{\sigma-1}{\sigma}} \tag{2.7}$$

where $B = f^{-1/\sigma}\sigma^{-1}(1-\sigma)^{\frac{\sigma-1}{\sigma}}$. Together, (2.4)–(2.7) form a system of equations for trade flows, the multilateral resistance terms, and factor prices. As a final step, we rewrite this system using the "exact hat algebra" of Dekle et al. (2007), where levels are rewritten into changes of the form

$\hat{Z} = Z'/Z$. (2.4)–(2.7) then become

$$\hat{X}_{ij} = \left(\frac{\hat{\tau}_{ij}}{\hat{\Pi}_i \hat{P}_j}\right)^{1-\sigma} \hat{w}_i \hat{L}_i \hat{w}_j \hat{L}_j \tag{2.8}$$

$$\hat{P}_j^{1-\sigma} = \sum_i s_{ij}^X \left(\frac{\hat{\tau}_{ij}}{\hat{\Pi}_i}\right)^{1-\sigma} \hat{w}_i \hat{L}_i \tag{2.9}$$

$$\hat{\Pi}_i^{1-\sigma} = \sum_j s_{ij}^M \left(\frac{\hat{\tau}_{ij}}{\hat{P}_j}\right)^{1-\sigma} \hat{w}_j \hat{L}_j \tag{2.10}$$

$$\hat{w}_i = \left(\frac{\hat{A}_i}{\hat{\Pi}_i}\right)^{\frac{\sigma-1}{\sigma}} \tag{2.11}$$

where s_{ij}^X is i's share of exports sent to j, or i's export share, and s_{ij}^M is j's share of imports sourced from i, or j's import share. The trade shares are defined respectively as

$$s_{ij}^X \equiv \frac{X_{ij}}{w_i L_i} = \left(\frac{\tau_{ij}}{\Pi_i P_j}\right)^{1-\sigma} w_j L_j$$

$$s_{ij}^M \equiv \frac{X_{ij}}{w_j L_j} = \left(\frac{\tau_{ij}}{\Pi_i P_j}\right)^{1-\sigma} w_i L_i.$$

2.2.4 Migration

To model migration in this framework, we first assume there is an initial allocation of labor in each location denoted by L_i^0. Workers are geographically mobile, and can choose to stay in their starting region or move to another region to maximize utility. Given our specification of preferences in (2.1), the indirect utility for a worker beginning in region i is

$$V_i = b_i P_i^{-\alpha} r_i^{\alpha-1} w_i$$

which is increasing in the location amenity and wage rate, and decreasing in the price index and housing rental rate. A worker beginning in region i can also choose to move to another region j. The indirect utility for a worker moving from i to j is

$$V_{ij} = b_j P_j^{-\alpha} r_j^{\alpha-1} w_j \tilde{v}_{ij}$$

where $\tilde{v}_{ij} = \mu_{ij}^{-1}(L_i^0/L_{ij})^\beta$ is the disutility from moving which is a function of the ratio of migration flows L_{ij} to the starting population of i (a congestion externality), and a migration cost $\mu_{ij} > 1$. In equilibrium, the marginal worker is indifferent between remaining in region i and moving to j, i.e. $V_i = V_{ij}$ for all i. Exploiting this, we can write the bilateral migration flows as

$$L_{ij} = \left(\frac{\mu_{ij}}{\Lambda_j V_i}\right)^{-\frac{1}{\beta}} L_j L_i^0 \tag{2.12}$$

where $V_i = \left(\sum_j \left(\frac{w_j}{P_j} b_j\right)^{1/\beta} \mu_{ij}^{-1/\beta}\right)^\beta = \left(\sum_j (\mu_{ij}\Lambda_j)^{-1/\beta} L_j\right)^\beta$ and $\Lambda_j^{1/\beta} = \sum_i (\mu_{ij} V_i)^{-1/\beta} L_i^0$. V_i and Λ_j are exact analogues to the multilateral resistance terms (2.5) and (2.6) in goods trade gravity. The outward resistance V_i is a function of prices, wages, regional amenities, and migration costs across destination regions, and the inward resistance Λ_j is a function of the same across origin regions, and represent general equilibrium *migration* cost indices as (2.9) and (2.10) are trade cost indices. Finally, we use the "exact hat algebra" from before and rewrite (2.12) in changes as

$$\frac{\hat{L}_{ij}}{\hat{L}_j} = \left(\frac{\hat{\mu}_{ij}}{\hat{\Lambda}_j \hat{V}_i}\right)^{-\frac{1}{\beta}} \tag{2.13}$$

where $\hat{\Lambda}_j^{1/\beta}, \hat{V}_i^{1/\beta}$ are defined similarly to (2.9) and (2.10) as functions of changes in the populations, migration costs, the other resistance term, and migration shares.

2.3 Data

To empirically investigate EIA formation and parameterize our model, we combine data from a number of sources. Data on EIA participation come from the NSF-Kellogg Institute Database on Economic Integration Agreements.[2] This database tracks every bilateral pair of 195 countries over the period 1950-2020 and contains information about every country's participation in an EIA with each possible trading partner. In particular, from these data we obtain a set of binary indicators equal to 1 if the country-pair is a member of a certain type of EIA in a given year (where our types consist of non-reciprocal PTAs, reciprocal PTAs, FTAs, customs unions, common markets, and economic unions). Bilateral trade data measured nominally in US dollars for 258 countries over the period 1962-2019 is sourced from UN Comtrade and the International Trade and Production

[2]These data were constructed by Scott Baier and Jeffrey Bergstrand and is provided here.

Database for Estimation (ITPD-E) created by Borchert et al. (2021), with own-country trade flows constructed by Baier and Standaert (2024).[3]

We also draw on CEPII's gravity database and the World Bank's World Development Indicators (WDI) database for additional country and country-pair economic variables.[4] CEPII gravity data contain a number of important country and bilateral gravity variables such as distance, contiguity, nominal GDP and GDP per capita, and population for 252 countries over the period 1948-2020 (Conte et al., 2022). The WDI data provide more detailed information on a larger range of economic and development variables that may be important factors of EIA participation and formation, like employment in agriculture, industrial, or services sectors.

Finally, since the theory delivers structural gravity of migration, we need data on migration flows. These come from Abel and Cohen (2019), who use migration stock data from the World Bank and United Nations to estimate five-year bilateral migration flows over the period 1990-2015 for 200 countries, measured in the number of people moving from some country i to another country j. This process is necessary as most countries do not report migration flows, but does suffer the drawback of only providing a migration flow variable every 5 years over the period. While this reduces the length (and frequency) of the panel we can leverage, we do not believe this introduces any bias in the estimation, and likely provides more intertemporal variation in EIA membership.

We merge bilateral EIA, trade, gravity, WDI, and migration data together at the country-pair and year level. Due to constraints imposed by the availability of global migration flows, the panel covers 185 country-pairs over the years 1990-2015, with migration data only being available in the years 1990, 1995, 2000, 2005, 2010, and 2015. We use all available years when possible, and restrict the sample to 5 year increments only when using the migration data. With these data now in hand, we proceed to outline our empirical strategies.

2.4 Empirical Methodology

2.4.1 Structural gravity of migration

As discussed in section 2.2.4, the model delivers a structural gravity equation for bilateral migration flows. Assuming the expected value of bilateral migration flows is given by (2.12) and

[3] The ITPD-E can be found here.

[4] The gravity data is provided on CEPII's website here, and the development data is provided on the World Bank's website here.

generalizing to a panel setting with the addition of a time subscript t, we can write observed migration flows between destinations i, j in period t as

$$L_{ijt} = \exp\left(-\frac{1}{\beta}\ln(\mu_{ijt}) - \frac{1}{\beta}\ln(\Lambda_{jt}) - \frac{1}{\beta}\ln(V_{it}) + \ln(L_{jt}) + \ln(L^0_{it})\right) v_{ijt}$$

where v_{ijt} is an error term with mean zero and variance a function of the migration cost vector, and migration flows L_{ijt} are estimates of the number of migrant transitions from origin i to destination j during a time interval.[5] We will find it advantageous to further rewrite bilateral migration flows as

$$L_{ijt} = \exp\left(\gamma EIA_{ijt} + \alpha_{ij} + \delta_{it} + \delta_{jt}\right) v_{ijt} \tag{2.14}$$

where α_{ij} is a country-pair fixed effect, δ_{it} is an exporter-year fixed effect, and δ_{jt} is an importer-year fixed effect. As discussed in Baier and Bergstrand (2007), exporter- and importer-year fixed effects are sufficient to account for multilateral resistance terms in goods trade gravity estimation. In our migration setting, Λ_{jt}, V_{it} are multilateral resistance terms and as such, will be absorbed by the same set of fixed effects.[6] We add the country-pair fixed effect α_{ij} to absorb time-invariant bilateral characteristics contained within the bilateral migration cost μ_{ijt}, leaving EIA membership as a time-varying shock to migration costs that allows us to identify the effect on bilateral migration flows.

We suspect the relationship between EIA membership and migration flows is non-linear. In particular, the effect of EIA formation between two countries on bilateral migration flows is likely heterogeneous depending on the distance between the countries. A reduction in migration costs via deeper EIA formation may lead to increased migration for many countries, but *less* migration for contiguous countries due to the possibility of cross-border commuting. To account for this possibility, we consider rewriting (2.14) as

$$\begin{aligned} L_{ijt} = \exp\Big(&\gamma_1 EIA_{ijt} \times CONTIG_{ij} + \gamma_2 EIA_{ijt} \times CLOSE_{ij} + \gamma_3 EIA_{ijt} \times MID_{ij} + \\ &\gamma_4 EIA_{ijt} \times FAR_{ij} + \alpha_{ij} + \delta_{it} + \delta_{jt}\Big) v_{ijt} \end{aligned} \tag{2.15}$$

where we interact the EIA variable with a contiguity indicator and binned distance indicators.

[5]As mentioned in section 2.3, these flows are estimated from migration stock data over 5-year time intervals.
[6]An added benefit of δ_{it}, δ_{jt} is the absorption of population stock variables L_{jt}, L^0_{it}.

Formally, $CONTIG_{ij}$ is an indicator equal to 1 if ij share a border, $CLOSE_{ij}$ is equal to 1 if the countries do not share a border but the bilateral distance less than or equal to 1,500 miles, MID_{ij} is equal to 1 if the bilateral distance is in the interval $(1500, 3000]$, and FAR_{ij} is equal to 1 if the bilateral distance is greater than 3,000 miles. This specification permits differential effects of EIA formation on migration flows depending on the distance of the countries.

Finally, we define three separate EIA variables to disentangle heterogeneous effects across agreement types. We bin nonreciprocal and reciprocal preferential trade agreements (PTAs) together into PTA_{ijt}, free trade agreements in FTA_{ijt}, and customs unions, common markets, and economic unions into EIA_deep_{ijt}. Thus, we can capture differences in the relationship between distance and EIA-induced migration effects by agreement depth. This is advantageous because we suspect deeper agreements will have a stronger relationship with migration than more shallow agreements that do not generate as much trade nor shock migration costs explicitly. For robustness, we also consider a separate indicator for each of our six agreement types, as well as alternative sets of fixed effects. An idea of the relationship between agreement formation and migration flows and how it is shaped by distance is informative about the ultimate welfare effects of EIAs and thus determinants for EIA formation.

2.4.2 Probability of EIA formation

The structure of goods trade gravity and migration gravity in (2.4) and (2.12) imply a set of variables related to changes in trade and migration flows, such as economic size or geography. EIA membership induces changes in trade and migration flows, and the magnitude of these changes are related to the welfare gains of an agreement. Therefore, the fundamental factors defining trade and migration flows may act as determinants for agreement formation. A simple way to think about this relationship is to estimate the probability of agreement formation as a function of these model primitives. Formally, we consider estimating

$$EIA_{ijt} = Z_{ijt}\beta + \varepsilon_{ijt} \tag{2.16}$$

where Z_{ijt} is a matrix of six different economic variables: the contiguity indicator, log distance, the sum of i and j's log GDP and log GDP per capita, and the difference between i and j's log GDP and log GDP per capita. Contiguity and distance are strongly related to trade costs, and we showed in the previous section they shape how migration flows (and thus welfare gains) respond to EIA formation. The sum and difference of GDP measures capture the economic size of the country-pair, as well as the economic size of a country *relative* to its trading partner. EIA_{ijt} is equal to 1 if i, j have an agreement of a certain type in time t, and we construct six different EIA indicators for the six main types of agreement: non-reciprocal PTA, reciprocal PTA, FTA, customs union, common market, and economic union.

We consider estimating (2.16) across the six agreement types with OLS (yielding a linear probability model), probit, and logit. We also consider alternative specifications involving exporter and importer fixed effects that soak up other unobserved country-level economic characteristics that may contribute to EIA formation. Before turning to estimation, we report summary statistics for the key economic variables included in Table 2.1, broken up by agreement type. From this, some simple trends emerge. For example, average distance of trading partners is decreasing in the depth of the agreement, and fraction of trading partners that share borders is increasing in agreement depth. Meanwhile, partners in more shallow agreements have lower GDP sums and higher GDP differences on average, while countries involved in deeper agreements are significantly more similar in terms of income level and economic size.

Figure B1 in the appendix displays kernel densities of the selected economic characteristics by agreement type. We can verify many of the simple insights from Table 2.1 graphically – most notably, that countries with deeper agreements are often closer, more likely to share a border, and more likely to be larger and similar in terms of income levels. Non-reciprocal PTAs stand out as having a significantly larger mass of country-pairs with substantial differences in total and per capita income, suggesting these types of EIAs are largely made between developed and less developed countries, and may act as a form of development aid.

2.4.3 Random forest

As we will see below, while the simple approach outlined above is informative, there are limits to its usefulness. In this section, we propose a different approach to investigate the importance of certain economic characteristics in the formation of EIAs. Here, we model EIA formation as a classification problem, and use a random forest of classification trees to estimate the importance of economic characteristics in determining the class (or EIA) certain country-pairs choose. We then train the model on a subset of data, predict out-of-sample EIA formation based on economic data from that year, and compare predictions to realized EIA formation data. We first describe the basics of the methodology before laying out how we evaluate model fit and variable importance, and then outline the prediction exercise before moving on to discuss results.[7]

2.4.3.1 Basics of the model

The basic element of a random forest is a decision tree. We model our decision problem as a *classification* problem, where a country-pair decides to form an EIA of a certain type. This binary classification is modeled as a response to a number of independent "predictor" variables. We consider a separate binary classification problem for each type of EIA using the same set of predictor, or input, variables to more clearly identify heterogeneity across agreement types.

The basic classification tree approach recursively partitions the input space, and then defines a local model in each resulting region. This can be represented by a tree, with one "leaf" per region. Each leaf stems from a split, whereby a function chooses the best feature (or predictor variable of the input space p), as well as the best value for that feature, to minimize classification cost.[8] A leaf (or node) can then be further split into subsequent leaves, resulting in a number of terminal nodes. At each node (whether at the beginning, middle, or end of the tree) there is a distribution of classes that satisfies the series of restrictions placed by previous splits.

Classification trees are useful for several reasons: easy to interpret, insensitive to monotonic transformations, automatic variable selection, robustness, and scalability. They present intuitive graphical depictions, provided the tree is not grown too deep. However, a single tree is often unstable – small changes in input data can have large effects on the structure of a tree, and can result in vastly different split decisions, tree depth, and number/structure of nodes. In other words, they are a *high variance* estimator, which is a problematic property.

A common and intuitive way to reduce the variance of the classification tree method is to grow multiple trees on different subsets of the data and average across estimates. However, re-running the same algorithm on different subsets often results in highly correlated predictors, which limits variance reduction. The random forest technique decorrelates the predictors by tree by growing each tree (1) on a different subset of data (as before) but importantly, *also* (2) using a randomly selected

[7] A more formal treatment of tree-based methods and random forests is available in Murphy (2012). For the purposes of this paper, we provide a high-level overview.

[8] Minimizing classification cost can be thought of as minimizing mean square error in an OLS regression – the goal is to minimize the distance between predicted and actual class.

subset of predictor/input variables – of the predictor space p, each tree is grown using $m < p$ of the predictors, where m are selected randomly from p. Thus, every tree grown has a different choice set of variables when making a split decision which can allow different variables to emerge in earlier nodes when they would otherwise be masked by other variables. The one drawback of this approach is that multiple tree methods lose nice interpretability properties of simple classification trees – we cannot represent the model graphically, and must rely one some post-processed summary measures. In the following subsection, we describe the key measures used to evaluate the results and fit of our random forest.

2.4.3.2 Interpretation of variable importance

To evaluate the results of the random forest estimation, we rely on a number of summary statistics unique to prediction-based machine learning models and in particular tree-based methods. In this subsection, we provide short descriptions of these statistics and their interpretation, which helps us understand the "importance" of the independent, or input variables, in classifying the response variable – which in our case, is membership in a certain type of EIA.

- *Mean minimum depth* is a measure of how early in a tree a variable is chosen to split on (on average) – the smaller the number (and closer to 1), the more often the variable appears in an early node and therefore the more "important" it is as a predictor of the dependent's class (with 1 being the root, or first node/split and therefore the most important of the randomly selected predictor subset m of that tree); the larger the number, the deeper in the tree it shows up, indicating other variables are being chosen for earlier nodes instead.

- *Number of nodes* is a count of how many times, across the 500 trees, the variable is selected to split the tree. Higher number here means the variable is chosen more often to split the tree, and therefore the more important the variable is as a predictor of classification. One drawback to this measure is that binary indicator variables (like contiguity) have fewer possible values and mechanically will be chosen at fewer nodes, while continuous variables with a large domain will be chosen much more often due to the larger range of possible values to select as a split value.

- *Number of trees* tells us in how many of the 500 trees grown the variable is chosen at least once as a node to split on. Given the small number of predictors p in this random forest exercise, we do not believe this is as informative a measure as others. However, if this is less than 500, then the variable is not being chosen as a predictor in several instances when it is contained in the choice set, and thus is a less important predictor.

- *Number of times a root* is a count of how many times the variable is chosen as the *first* split, or node, in a tree. A higher number here indicates that when this predictor is included in the random set of predictors when growing a specific tree within the forest, it is often chosen as the first node. This indicates the variable is very important for classification, and is tied closely to the mean minimum depth measure – the more often a variable is a root, the closer to 1 the mean minimum depth will be.

The next two "fit statistics" measure the quality of a split, primarily through a measure of classification cost or error. They are also often referred to as "node impurity" measures, where purity refers to how much of a single class is partitioned into a node following a split.

- *Mean accuracy decrease* is associated with how the removal of the variable changes the misclassification rate. Formally, it measures "by how much the classification accuracy decreases when training the model without this feature" – a lower number suggests removing the variable does not impede accurate classification very much, while a larger number suggests removing the variable results in more misclassification, and thus the variable is a more influential factor.

- *Mean Gini decrease* is associated with the Gini index, which is another measure of the error rate. It measures by how much node purity decreases when training the model without this feature, where node purity refers to the percentage of a single class contained in the node or leaf following a split. Important variables will have a higher number (indicating removal significantly alters classification accuracy and results in less pure nodes), while less important variables will have a smaller number.

Something to note about the mean accuracy and Gini decrease measures is that while they attempt to measure a similar thing (classification error), the Gini measure is more sensitive to changes in class probability. The Gini measure will prefer splits that result in **pure** nodes, i.e. the node contains only one class, while the misclassification rate is invariant to node purity. Before turning to discuss empirical results, we first outline the structure of the prediction exercise we use to evaluate the fit of the random forest model.

2.4.3.3 Prediction exercise

The primary advantage of the random forest model is its accuracy at predicting out-of-sample classifications. To leverage this advantage, we use the model to conduct a prediction exercise to evaluate how well we can predict EIA formation using the suite of independent variables we consider in previous sections. Below, we outline the basic steps of this exercise and how we will evaluate prediction accuracy.

The basic operation consists of training a random forest on a subset of the bilateral EIA panel data. As before, our primary dependent (response) variable is an indicator variable for specific type of EIA. Since we consider 6 broad types of agreement, we have six distinct response variables. Our independent (input/predictor) variables consist of log distance, contiguity indicator, sum of logged GDP and logged GDP per capita, and difference in logged GDP and logged GDP per capita. For each subset of data, six random forests are trained: one for each for the six types of EIA (non-reciprocal PTA, PTA, FTA, customs union, common market, and economic union). Then, I use the trained model to predict the value of the pairwise EIA indicator for the next chronological year after the training subset. This procedure is looped repeatedly, with each loop adding one year to the training data and predicting the EIA indicator for the following year. For example, the first loop would train the model on data ranging 1990-1995, and then predict the EIA indicator in 1996 (given data on the predictor variables from that year); the next loop would train on 1990-1996 and predict 1997. The loop runs until training on 1990-2014 and testing on 2015.

At each iteration, we compute a prediction accuracy for each of the six random forests. Once the algorithm has iterated through each subset of training and test data, we average across iterations to arrive at a mean prediction accuracy for each of the six response variables we predict. Formally, for each response variable (and thus for each random forest estimated within one iteration of training and testing data subsets) the prediction accuracy of newly formed EIAs of a certain type is an average across country-pairs ij of

$$p_{ij} = 1 - EIA_{ij,t-1} \times \begin{cases} 1 & \text{if} \quad EIA_{ij,t} = pred_{ij,t} \\ 0 & \text{if} \quad EIA_{ij,t} \neq pred_{ij,t} \end{cases} \tag{2.17}$$

and the over-prediction rate is an average of

$$o_{ij} = (1 - EIA_{ij,t-1}) \times pred_{ij,t} \times (1 - EIA_{ij,t}) \tag{2.18}$$

where $EIA_{ij,t}, EIA_{ij,t-1}, pred_{ij,t} \in \{0,1\}$, t denotes the testing year, and $t-1$ denotes the last period the training data subset. For each iteration, we compute p_{ij} and o_{ij} for all pairs of i,j and average across all observations. This yields six measures – one for each dependent variable and respective forest – which we then average across all iterations of the training and testing subsets, or all combinations of $\{t-1, t\}$. In the end, we will get an average percentage of new EIAs the model correctly predicts, and an average percentage of new EIAs the model predicts that do not occur in the data. With an understanding of our empirical methodology, we now turn to present the results of estimation.

2.5 Empirical Results

In this section we present results from the empirical strategies outlined in the previous section. We first discuss the partial equilibrium estimation of migration gravity, before turning to formally investigate the determinants of EIA formation using both regression and classification approaches.

2.5.1 Structural gravity of migration

We estimate (2.15) using both log-linear ordinary least squares (OLS) and Poisson psuedo-maximum likelihood (PPML), and report the results in Table 2.2.[9] Standard errors are clustered at the country-pair level. PTAs (reciprocal and non-reciprocal) and FTAs do not seem to have very large or consistent effects on migration flows, though for the "mid" distance bin, there are positive and statistically significant effects across specifications. This is consistent with the idea that migration flows may respond to shallow EIAs through an indirect channel, as illustrated by model.

Among deeper EIAs such as customs unions, common markets, and economic unions, a strong relationship emerges. For the "close" and "mid" distance bins, partnership in these agree-

[9] PPML for structural gravity is a widely-adopted approach to addressing common issues in trade data; see Silva and Tenreyro (2006) for more information.

ments has strong migration-creating effects. However, for country-pairs sharing borders, or more than 3,000 miles away, the effect of EIA partnership on migration is muted. In the OLS specification, migration flows *decrease* for adjacent countries following an EIA. This is also intuitive – deep EIAs liberalize labor flows significantly, and for individuals living near borders this offers cross-border commuting as an alternative to migration. Taken together, there is a distinct "U-shaped" effect of EIAs on migration by distance, with EIAs having little effect on migration flows on contiguous countries but on non-contiguous countries within 3,000 miles, and then a more muted effect beyond. This relationship is further emphasized in Table 2.3, where we consider the same specification but with only the deep EIA interaction variables. The largest post-EIA migration gains are made by countries close, but *not* adjacent, while the changes in migration flows are more muted for those either more distant or sharing a border.

For robustness, we also consider estimating gravity models with the full suite of six EIA indicators, no distance interactions, and alternative sets of fixed effects – the results of this are presented in Table B1 in Appendix B. We see large migration creation across agreements, with the largest and most robust effects among the deeper agreements. In columns (3) and (6) which have the same set of theoretically implied FEs as above, we cannot identify migration effects from non-reciprocal PTAs and free trade agreements, with reciprocal PTAs have small but positive and statistically significant effects on migration flows. Altogether, our results here suggest a strong relationship between EIA formation and migration flows that is robust across a range of agreement types and heterogeneous in magnitude across agreement type and bilateral distance.

2.5.2 Probability of EIA formation

With an understanding of the relationship between EIA partnership and bilateral migration flows, we now turn to examining more directly the economic determinants of EIA formation. As discussed in the previous section, we estimate (2.16) via OLS, probit, and logit. The results of the OLS estimation are reported in Table 2.4, with each column representing a different binary agreement indicator. These naive results suggest some significant heterogeneity in the economic factors associated with partnership across agreement types.

For example, non-reciprocal PTA membership is *more* likely the further the distance between the country-pair – the exact opposite relationship between every other EIA and distance, where more distant countries are more likely to sign agreements. In a similar vein, contiguity lowers the likelihood of non-reciprocal PTAs, while increasing the likelihood of most other agreement types. Further, non-reciprocal PTA membership is also more probable in the absolute difference between trading partner GDP measures – suggesting these types of agreements are not only made between more distant countries, but also largely made between countries of very different economic development levels. Deeper agreements seem to be signed more often among countries of a similar economic size, as the estimated coefficients on the absolute differences in GDP and GDP per capita are negative. Contiguity also matters much more for deeper agreements. Across the board, it seems country-pairs with larger sums of GDP and GDP per capita are more likely to sign agreements, suggesting richer countries participate more often in EIAs than less-developed countries.

We also estimate (2.16) using probit and logit; the results are reported in Table 2.5. The relationships we outline above are robust to this change in specification, though there are some differences (such as in the estimated coefficients on contiguity). Note that the estimates reported in Table 2.5 are not marginal effects, and are not directly comparable with the OLS estimates. As a final step, we estimate the OLS specification with importer and exporter fixed effects, as well as importer-year and exporter-year fixed effects to account for unobserved country-level heterogeneity that may influence agreement formation. We report the results of these estimations in Table B2 in Appendix B.[10] The addition of the fixed effects improves the model fit by soaking up unobserved heterogeneity, and does impact some of the estimated coefficients, such as those on contiguity, but largely the results are robust to these alternative specifications.

For the large part, across all specifications, it is clear we are explaining very little of the probability of EIA partnership across agreement types, and some of the identified effects are not robust across specifications. Thus, while we are identifying some significant effects of these economic variables on EIA probability that shed some light on the determinants of EIA formation, it is hard to argue we have a full picture. In the next section, we consider an alternative approach to see if this limited set of economic characteristics is indeed insufficient to explain EIA partnership, or if this is a model limitation.

[10] Note that the sum of log GDP is absorbed the importer-year and exporter-year fixed effects in panel (b). We believe that the aggregate GDP sums are perfectly collinear with these fixed effects for some country pairs due to limited variation, and thus they are omitted.

2.5.3 Random forest: importance and prediction

In this section, we estimate many random forests and present the results in two distinct ways. As discussed in section 2.4.3, we model EIA formation as a classification problem and use the set of covariates from the previous section as the predictor space p. Our empirical approach has two distinct dimensions. First, we choose a small subset of the full panel to implement a less-efficient random forest algorithm that delivers richer post-processed summary statistics on variable importance. The choice of the subset is somewhat arbitrary, but necessary for computational efficiency. For our purposes, we sample all country-pairs in the year 2000 for this exercise. Second, we conduct the iterated prediction exercise outlined in section 2.4.3.3, which involves using the full panel over the period 1990-2015 but in chunks as we estimate the model on a set of training data, use the subsequent year as test data to predict the EIA variable, then compute accuracy measures defined in (2.17) and (2.18) by observation. We loop this procedure over a range of training and testing data subsets, starting with a training set of 1990-1995 and a testing set of 1996, and ending with a training set of 1990-2014 and a testing set of 2015. Accuracy measures are averaged over bilateral observations and the 20 iterations of testing and training subsamples.

Table 2.6 reports measures summarizing the variable importance results obtained by estimating the random forest model on all country-pairs during the year 2000 by broad agreement type. Figure 2.1 presents accuracy and Gini decrease measures from the 4th and 5th columns of Table 2.6 in graphical form, which are the two measures of misclassification error. From these results, we can immediately see heterogeneity in the importance of our included economic variables across agreement type. For example, contiguity is not important in predicting non-reciprocal PTAs – the mean minimum depth is very large, suggesting the variable is not being chosen to split trees until much later in the tree; the Gini decrease is very small and close to zero, suggesting the removal of the contiguity variable does not contribute much to additional classification error. Meanwhile, income sums and differences are the most important features for non-reciprocal PTAs – difference in GDP per capita and sum of GDP have the smallest minimum depths, are the most frequent roots, and contibute to the largest Gini decreases.

As we move towards deeper agreements, predictor importance changes in significant ways – distance matters quite a bit more for all other agreement types than non-reciprocal PTAs; income differences matter less for deeper agreements than for non-reciprocal PTAs. Among the deepest agreements (common markets and economic unions), the sum of GDP and GDP per capita emerge as the most important predictors along with distance – all three have the lowest minimum depth, the highest number of nodes, the largest Gini decreases from variable removal, and are frequent tree roots. One interesting result among economic unions is the relative unimportance of contiguity across measures like minimum depth and Gini decrease, but it surprisingly appears as a tree root (i.e. the first split) more frequently than every variable besides distance. This is also surprising when compared to common markets, where distance and sum of GDP and GDP per capita are still most important, but contiguity is now across the board the least important predictor of common market membership, and not only rarely shows up as a root but also appears in only 316 out of 500 trees.

In summary, there are some interesting patterns that emerge. Sum of country-pair GDP tends to be an important predictor of agreement classification across most agreement types. Regardless of depth of agreement, richer countries tend to sign more of them. Difference in country-pair GDP and GDP per capita are relatively more important for non-reciprocal PTAs and PTAs than deeper agreements, which is consistent with the regression results discussed in section 2.5.2, suggesting in particular non-reciprocal PTAs may be disproportionately signed between countries of vastly different income and development levels.[11] On the other hand, distance and contiguity do not matter much for non-reciprocal PTAs (also consistent with our regression results), which suggests these agreements are often signed by more distant country-pairs.

As agreements deepen, the importance of distance increases. This is intuitive to us, as deeper agreements start to liberalize capital and labor flows and harmonize other forms of policies that make more sense for partners in closer proximity. The sum of per capita GDP also emerges as the most important variable in the deepest agreements while difference in GDP and per capita GDP are among the least important, suggesting deeper agreements are favored by countries with *similar* income levels as well as *higher* income levels. We think income similarity is important so resulting post-EIA trade or migration flows are not too unidirectional, which we believe would yield lopsided welfare gains (and perhaps even welfare losses). The importance of GDP and GDP per capita across the spectrum of agreement types seems to indicate that regardless of the specific shape of agreements, likelihood of EIA participation is increasing in development level.

[11] We posit these one-way PTAs act as a form of development aid from richer to poorer countries, but this is just one possible explanation.

To complete our random forest analysis, we conduct the prediction exercise outlined above by estimating the model on a rolling subsample and predicting out-of-sample EIA formation. We report averages of p_{ij} and o_{ij} from (2.17) and (2.18) across all observations within an iteration, then across the 20 iterations for each of our EIA variables in Table 2.7. The first row displays the average percentage of *new* EIAs correctly predicted by the model, or 0 to 1 changes from t to $t + 1$ in the respective EIA indicator, which is an average of p_{ij} across observations and iterations. The second row displays the average percentage of new EIAs predicted by the model that did not occur in the data, which is an average of o_{ij} across observations and iterations.

From Table 2.7, we see that the random forest correctly predicts new EIA formation for deeper agreements 98 to 99 percent of the time on average over testing years 1996 to 2015. It correctly predicts new FTA formation 94 percent of the time, new reciprocal PTA formation about 95 percent of the time, and new non-reciprocal PTA formation 82 percent of the time on average. In the second row, we see the over-prediction rate is minuscule for the deeper agreements, but larger for non-reciprocal PTAs at 0.13 percent. We find contiguity, distance, and the sums and differences in bilateral GDP and GDP per capita do a remarkable job correctly predicting deeper EIA formation over the sample period. These economic characteristics correctly predict shallow agreements like PTAs and FTAs with slightly less success, but have some significant error in predicting non-reciprocal PTAs. While suggesting we are missing a piece of the story for non-reciprocal PTAs, the high prediction accuracy for most other agreement types suggest the importance of economic forces in influencing the formation of new EIAs. In the next section, we outline how we plan to take the model in section 2.2 directly to the data to estimate general equilibrium trade and migration effects of EIAs and characterize the welfare gains across agreement type.

2.6 Model Simulation

In this section, we present a simple sketch of a counterfactual exercise using the framework outlined in section 2.2 and our collected data on trade flows, migration flows, gravity variables, and EIA membership. This discussion lays the groundwork for the next stages of this research.

Anderson et al. (2018) illustrate a simple way to estimate general equilibrium comparative statics of gravity models that relies on convenient theoretical properties of the PPML estimator, where the directional fixed effects in a structural gravity estimation perfectly approximate the multilateral resistance terms (Arvis & Shepherd, 2013; Fally, 2015). The basic algorithm proceeds as follows: estimate a baseline structural gravity model using PPML to obtain the baseline multilateral resistance terms, define a counterfactual and estimate a structural gravity model using the counterfactual trade cost vector, then obtain the new multilateral resistance terms while not allowing income or expenditure to change. Then, estimate the general equilibrium effects by allowing prices, income and expenditure to change, re-estimating structural gravity and iterating until convergence. This allows us to ultimately compute changes in welfare from counterfactual changes in trade costs.[12]

Our approach is very similar to the basic algorithm outlined above, with the addition of structural gravity of migration and its respective parts: vectors of migration costs, multilateral resistance terms, and populations. Counterfactual changes in trade and migration costs are determined by EIA membership, with only deeper EIAs providing explicit counterfactual changes in migration costs. However, migration flows can change indirectly from the general equilibrium effects of trade creation, whereby even shallow EIAs change the indirect utility of living in a specific region or moving to a new region. Since migration decisions hinge off an indifference condition in indirect utility, any change to indirect utility (related or unrelated to migration costs) will impact optimal bilateral migration flows.

This approach will permit us to ask a few questions of interest. First, what are the welfare effects of EIAs when you permit a migration response as opposed to when there is no migration? We suspect even the indirect migration effects will alter the welfare effect of EIAs, suggesting the importance of accounting for migration when assessing the impact of agreement formation. Second, we can construct a variety of counterfactual EIAs for a set of countries and compare welfare gains across depth of agreement. For example, how different would the effects of a deeper EIA between the U.S., Canada, and Mexico be compared to observed effects of NAFTA or USMCA? Will the observed EIA be associated with the largest welfare gain? We suspect so, but if not there are interesting follow-up questions. Finally, this approach will allow us to revisit the fundamental question of Baier and Bergstrand (2004) and consider counterfactual *economic characteristics* of countries and their trading partners, and evaluate how these shocks impact welfare effects across EIA types. If we increased the distance between two countries, which agreement type would generate the largest welfare gains? How similar in income and sectoral employment shares must countries be for an economic union to generate the largest gains? This exercise in particular will provide insight

[12]This approach is also described in more detail in Yotov et al. (2016).

into the importance of specific economic characteristics in determining the welfare effects of EIAs, and thus agreement formation.

2.7 Conclusion

In this paper, we set out to investigate the economic determinants that influence the formation of EIAs, paying special attention to heterogeneity across broad types of agreements and the role of migration flows. We present a spatial model of trade that delivers structural gravity of goods trade and migration, which we can use to estimate the partial and general equilibrium trade and migration effects of EIAs. We provide novel evidence on the relationship between EIA formation and migration flows, documenting larger migration effects for deeper EIAs and a "U-shaped" relationship between the effect and bilateral distance that suggests non-adjacent but geographically close countries see the largest migration gains from deeper EIAs. We also conduct an exploratory empirical exercise to investigate the link between some simple economic variables and EIA formation, with the main takeaway being that new EIA agreements (particularly deeper ones) can be very accurately predicted with just a small set of gravity and country-pair income variables, though there is some significant heterogeneity in the contribution of these variables towards agreement formation by agreement type. We believe these results suggest the importance of understanding the economic determinants of EIAs that may be much broader in nature than the classical forces of trade creation and diversion. We conclude with an outline of how we intend to leverage the model in section 2.2 to estimate general equilibrium trade and migration effects of EIA formation with goals of quantitatively evaluating welfare effects across agreement depth, and quantitatively evaluating the impact of changing fundamental economic characteristics on post-EIA welfare gains.

Table 2.1: Summary statistics

Variable	Mean	SD	Min	p25	Median	p75	Max
EIA: No Agreement							
dist	8591.4	4459.347	10.479	5083.627	8108.696	11905.556	19904.447
contig	0	0.094	0	0	0	0	1
sumlogGDP	31.6	3.588	19.631	29.078	31.463	33.978	46.928
difflogGDP	2.6	1.909	0	1.038	2.199	3.704	12.315
sumlogGDPcap	2.2	1.616	0	0.873	1.834	3.102	9.396
difflogGDPcap	1.7	1.201	0	0.67	1.431	2.411	6.931
EIA: Non-Reciprocal PTA							
dist	7983.4	3756.191	169.526	5114.49	7577.717	10145.654	19586.182
contig	0	0.041	0	0	0	0	1
sumlogGDP	34.9	2.821	24.83	32.914	34.798	36.749	45.539
difflogGDP	3.8	2.28	0	1.966	3.64	5.374	12.315
sumlogGDPcap	3.2	1.7	0	1.912	3.094	4.377	9.262
difflogGDPcap	2.6	1.337	0	1.576	2.622	3.626	6.707
EIA: Preferential Trade Agreement							
dist	3772.5	2927.509	105.181	1641.642	2804.291	5413.075	19711.857
contig	0.1	0.326	0	0	0	0	1
sumlogGDP	33	3.752	22.63	30.083	33.267	35.85	44.933
difflogGDP	2.3	1.709	0	0.876	1.905	3.246	8.799
sumlogGDPcap	2.7	1.641	0	1.459	2.542	3.793	8.597
difflogGDPcap	1.5	1.194	0	0.509	1.137	2.244	6.247
EIA: Free Trade Agreement							
dist	3379.3	3335.766	59.617	1298.833	2188.659	3612.088	19079.875
contig	0.1	0.299	0	0	0	0	1
sumlogGDP	36	3.254	23.215	34.169	36.339	38.258	44.919
difflogGDP	1.9	1.458	0	0.744	1.651	2.782	7.622
sumlogGDPcap	4.2	1.948	0	2.928	4.502	5.557	9.233
difflogGDPcap	1.3	0.938	0	0.51	1.152	1.937	5.097

Table 2.1: Summary statistics (cont.)

Variable	Mean	SD	Min	p25	Median	p75	Max
EIA: Customs Union							
dist	1280.7	902.207	131.692	561.14	1109.901	1799.45	6621.323
contig	0.2	0.404	0	0	0	0	1
sumlogGDP	33.4	4.921	21.386	29.156	34.052	37.747	42.734
difflogGDP	1.9	1.481	0.001	0.763	1.579	2.693	7.853
sumlogGDPcap	3.7	1.954	0.003	2.176	3.872	5.149	8.853
difflogGDPcap	0.8	0.651	0	0.276	0.607	1.155	3.914
EIA: Common Market							
dist	1487.8	825.09	59.617	884.611	1342.89	2022.635	4882.096
contig	0.1	0.294	0	0	0	0	1
sumlogGDP	37.9	2.314	29.676	36.38	37.995	39.535	43.9
difflogGDP	1.9	1.346	0	0.719	1.761	2.701	6.212
sumlogGDPcap	6.4	1.18	0.023	5.849	6.502	7.083	9.355
difflogGDPcap	0.8	0.552	0	0.268	0.665	1.132	2.875
EIA: Economic Union							
dist	1386.3	754.283	59.617	822.552	1297.695	1887.733	3766.31
contig	0.2	0.426	0	0	0	0	1
sumlogGDP	36.7	4.055	26.895	32.951	37.704	39.87	43.849
difflogGDP	1.7	1.246	0.001	0.636	1.472	2.343	6.082
sumlogGDPcap	5.3	2.658	0.001	2.504	6.582	7.196	8.925
difflogGDPcap	0.6	0.587	0	0.196	0.513	0.83	4.124

Table 2.2: Gravity of migration and distance bin interactions: all EIAs

Dependent Variables:	log(mig_rate)	mig_rate
Model:	(1)	(2)
	OLS	PPML
Variables		
contig × pta	0.0092	0.0126
	(0.0638)	(0.0674)
close × pta	-0.0154	-0.0558
	(0.0402)	(0.0897)
mid × pta	0.1123***	0.3571***
	(0.0337)	(0.0671)
far × pta	0.0267	0.0516
	(0.0183)	(0.0446)
contig × fta	-0.0274	-0.1078
	(0.0718)	(0.0738)
close × fta	-0.0356	-0.2237**
	(0.0361)	(0.0991)
mid × fta	0.0710*	0.1056*
	(0.0415)	(0.0613)
far × fta	0.0362	0.0796
	(0.0231)	(0.0651)
contig × eia_deep	-0.2241***	0.0518
	(0.0847)	(0.1321)
close × eia_deep	0.2258***	0.3357**
	(0.0514)	(0.1594)
mid × eia_deep	0.6057***	0.5265***
	(0.0577)	(0.1156)
far × eia_deep	0.2861**	-0.0829
	(0.1141)	(0.1242)
pair FEs	Yes	Yes
exporter-year	Yes	Yes
importer-year	Yes	Yes
Observations	53,688	57,590
R^2	0.80813	0.92815

Note: contig = 1 if country-pair contiguous; close = 1 if contig = 0 and distance $\in [0, 1500]$; mid = 1 if distance $\in (1500, 3000]$; far = 1 if distance > 3000. pta = 1 if EIA is non-reciprocal or reciprocal PTA; fta = 1 if EIA is FTA; eia_deep = 1 if EIA is customs union, common market, or economic union. Clustered (pairwise) standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.3: Gravity of migration and distance bin interactions: deep EIAs

Dependent Variables:	log(mig_rate)	mig_rate
Model:	(1)	(2)
	OLS	PPML
Variables		
contig × eia_deep	-0.2226***	0.1088
	(0.0698)	(0.0988)
close × eia_deep	0.2384***	0.4908***
	(0.0400)	(0.1339)
mid × eia_deep	0.5299***	0.3259***
	(0.0471)	(0.1061)
far × eia_deep	0.2507**	-0.1626*
	(0.1136)	(0.0952)
pair FEs	Yes	Yes
exporter-year	Yes	Yes
importer-year	Yes	Yes
Observations	53,733	57,639
R^2	0.80809	0.92813

Note: contig = 1 if country-pair contiguous; close = 1 if contig = 0 and distance $\in [0, 1500]$; mid = 1 if distance $\in (1500, 3000]$; far = 1 if distance > 3000. eia_deep = 1 if EIA is customs union, common market, or economic union. Clustered (pairwise) standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.4: Probability of EIA by type: OLS

Dependent:	nrpta	pta	fta	customs	common	union
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Constant	-0.5789***	0.4285***	0.3696***	0.1839***	0.1985***	0.1005***
	(0.0050)	(0.0049)	(0.0048)	(0.0034)	(0.0024)	(0.0019)
logdist	0.0088***	-0.0428***	-0.0591***	-0.0168***	-0.0274***	-0.0126***
	(0.0004)	(0.0004)	(0.0005)	(0.0003)	(0.0003)	(0.0002)
contig	-0.0413***	0.0946***	0.0554***	0.0369***	-0.0133***	0.0515***
	(0.0016)	(0.0038)	(0.0039)	(0.0024)	(0.0022)	(0.0025)
sumlogGDP	0.0126***	-0.0005***	0.0056***	-0.0010***	0.0012***	0.0004***
	(0.0001)	(0.0001)	(0.0001)	(0.0000)	(0.0000)	(0.0000)
sumlogGDPcap	0.0204***	0.0009***	0.0093***	0.0026***	0.0088***	0.0033***
	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
difflogGDP	0.0200***	-0.0004***	-0.0048***	0.0000	-0.0010***	-0.0007***
	(0.0002)	(0.0001)	(0.0001)	(0.0000)	(0.0000)	(0.0000)
difflogGDPcap	0.0474***	-0.0011***	-0.0059***	-0.0020***	-0.0047***	-0.0026***
	(0.0003)	(0.0001)	(0.0001)	(0.0000)	(0.0001)	(0.0001)
Observations	804,374	804,374	804,374	804,374	804,374	804,374
R^2	0.1247	0.0423	0.1074	0.0421	0.0897	0.0476
Adjusted R^2	0.1247	0.0423	0.1074	0.0421	0.0897	0.0476

Note: Robust standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.5: Probability of EIA by type: probit and logit

Dependent:	nrpta	pta	fta	customs	common	union
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): Probit						
Constant	-6.1690***	1.9095***	0.0153	4.1232***	0.6187***	1.0996***
	(0.0378)	(0.0408)	(0.0423)	(0.0830)	(0.0875)	(0.0890)
logdist	0.0683***	-0.4520***	-0.6032***	-0.6380***	-0.8109***	-0.5705***
	(0.0031)	(0.0034)	(0.0036)	(0.0069)	(0.0076)	(0.0074)
contig	-0.7036***	0.1665***	-0.1907***	-0.0276	-0.8512***	0.0877***
	(0.0318)	(0.0145)	(0.0154)	(0.0221)	(0.0290)	(0.0230)
sumlogGDP	0.0867***	0.0030***	0.0961***	-0.0471***	0.0675***	0.0298***
	(0.0007)	(0.0009)	(0.0011)	(0.0021)	(0.0024)	(0.0024)
sumlogGDPcap	0.1469***	0.0142***	0.0730***	0.1230***	0.3516***	0.1349***
	(0.0013)	(0.0017)	(0.0018)	(0.0038)	(0.0050)	(0.0042)
difflogGDP	0.1008***	-0.0065***	-0.0815***	-0.0372***	-0.0642***	-0.0978***
	(0.0010)	(0.0015)	(0.0019)	(0.0042)	(0.0042)	(0.0051)
difflogGDPcap	0.2864***	0.0075***	-0.0517***	-0.1687***	-0.2534***	-0.3553***
	(0.0016)	(0.0023)	(0.0027)	(0.0072)	(0.0090)	(0.0104)
Observations	804,374	804,374	804,374	804,374	804,374	804,374
Squared Correlation	0.134	0.029	0.106	0.056	0.242	0.085
Pseudo R^2	0.192	0.099	0.267	0.318	0.542	0.385
Panel (b): Logit						
Constant	-11.2411***	4.4024***	-0.1011	10.4250***	0.8388***	1.9021***
	(0.0713)	(0.0850)	(0.0856)	(0.1790)	(0.1807)	(0.2035)
logdist	0.1400***	-0.9019***	-1.1566***	-1.4043***	-1.6088***	-1.1745***
	(0.0059)	(0.0070)	(0.0072)	(0.0150)	(0.0149)	(0.0161)
contig	-1.5499***	0.2043***	-0.4621***	-0.2578***	-1.9897***	0.0772
	(0.0758)	(0.0268)	(0.0284)	(0.0476)	(0.0563)	(0.0489)
sumlogGDP	0.1549***	0.0005	0.1938***	-0.1324***	0.1421***	0.0751***
	(0.0014)	(0.0021)	(0.0023)	(0.0049)	(0.0049)	(0.0057)
sumlogGDPcap	0.2754***	-0.0014	0.1230***	0.3098***	0.7380***	0.3241***
	(0.0025)	(0.0037)	(0.0038)	(0.0089)	(0.0103)	(0.0101)
difflogGDP	0.1812***	-0.0198***	-0.1664***	-0.0708***	-0.1244***	-0.2075***
	(0.0018)	(0.0035)	(0.0040)	(0.0099)	(0.0084)	(0.0117)
difflogGDPcap	0.5378***	0.0036	-0.0836***	-0.4517***	-0.4740***	-0.9205***
	(0.0030)	(0.0051)	(0.0058)	(0.0184)	(0.0186)	(0.0257)
Observations	804,374	804,374	804,374	804,374	804,374	804,374
Squared Correlation	0.1322	0.0286	0.0973	0.0573	0.2336	0.0847
Pseudo R^2	0.1870	0.0917	0.2487	0.2980	0.5260	0.3679

Note: Robust standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.6: Random forest variable importance measures

variable	min_depth	no_nodes	acc_dec	gini_dec	no_trees	times_root
EIA: Non-reciprocal PTA						
contig	7.556	863	0.0001	2.910	444	0
difflogGDP	1.400	278,201	−0.023	751.759	500	95
difflogGDPcap	1.198	278,896	−0.022	862.572	500	140
logdist	2.244	275,703	−0.020	624.069	500	39
sumlogGDP	1.110	278,516	−0.022	894.667	500	160
sumlogGDPcap	1.596	279,917	−0.025	769.672	500	66
EIA: Reciprocal PTA						
contig	2.230	1,930	0.004	24.742	500	129
difflogGDP	1.858	51,976	0.018	280.577	500	83
difflogGDPcap	1.708	52,236	0.022	293.706	500	91
logdist	1.144	54,475	0.037	442.034	500	160
sumlogGDP	1.860	55,487	0.026	367.277	500	34
sumlogGDPcap	2.194	53,394	0.022	320.886	500	3
EIA: Free Trade Agreement						
contig	2.776	2,477	0.003	30.028	500	81
difflogGDP	2.518	47,855	0.015	264.272	500	0
difflogGDPcap	2.202	49,073	0.021	314.033	500	36
logdist	1.152	52,832	0.036	653.789	500	149
sumlogGDP	1.452	50,025	0.026	373.071	500	97
sumlogGDPcap	1.280	50,538	0.026	397.521	500	137
EIA: Customs Union						
contig	3.536	1,122	0.001	5.134	499	126
difflogGDP	2.218	12,023	0.004	60.187	500	0
difflogGDPcap	1.962	11,909	0.004	59.524	500	72
logdist	1.074	13,730	0.006	80.873	500	180
sumlogGDP	1.782	13,393	0.005	72.195	500	47
sumlogGDPcap	1.962	12,251	0.004	62.507	500	75

Table 2.6: Random forest variable importance measures (cont.)

variable	min_depth	no_nodes	acc_dec	gini_dec	no_trees	times_root
EIA: Common Market						
contig	6.061	443	0.0001	2.783	316	5
difflogGDP	2.374	4,773	0.001	30.789	500	33
difflogGDPcap	1.956	5,325	0.002	38.744	500	65
logdist	1.388	6,208	0.003	74.925	500	104
sumlogGDP	1.472	5,678	0.002	44.874	500	127
sumlogGDPcap	1.108	6,424	0.003	82.960	500	166
EIA: Economic Union						
contig	2.252	1,007	0.001	14.517	500	137
difflogGDP	2.568	10,233	0.003	54.717	500	0
difflogGDPcap	2.084	10,893	0.005	71.144	500	42
logdist	1.156	11,634	0.006	96.139	500	153
sumlogGDP	1.678	12,149	0.005	96.159	500	55
sumlogGDPcap	1.360	11,612	0.005	99.842	500	113

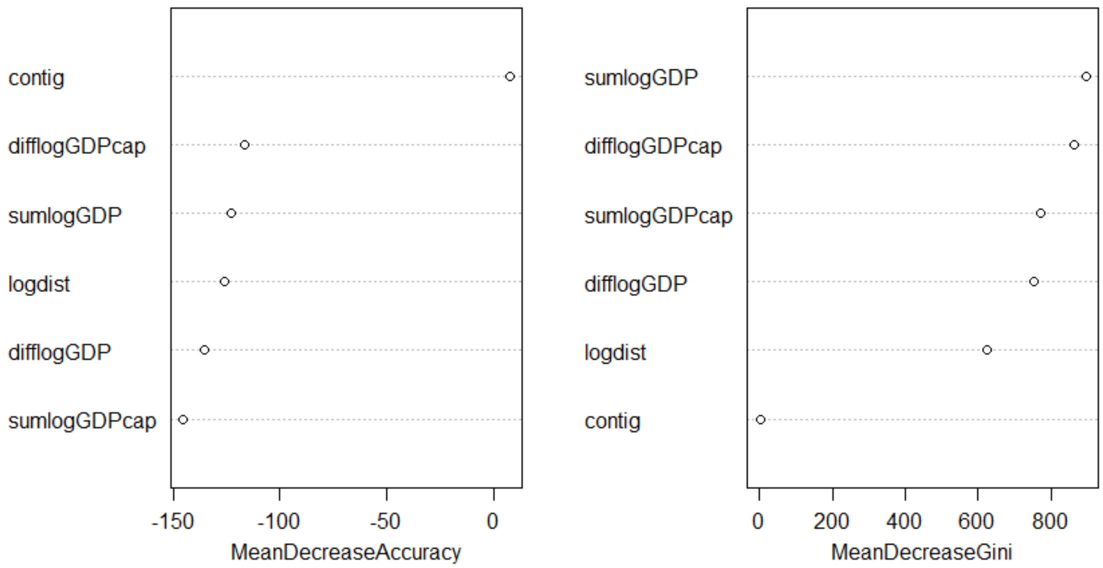
Table 2.7: Random forest predictive fit

Dependent:	nrpta (1)	pta (2)	fta (3)	customs (4)	common (5)	union (6)
New EIA accuracy	0.824	0.948	0.940	0.994	0.984	0.990
New EIA over-prediction	0.013	0.001	0.001	0.0001	0.001	0.00001

Note: For each of 20 iterations, one random forest is trained for each of the six dependent variables, and then used to predict the dependent in the next year out-of-sample. Prediction and realized data are used to compute (2.17) and (2.18), which are then averaged over (1) all observations within each iteration, and then (2) across all 20 iterations with different test and training datasets.

Figure 2.1: Random forest variable importance measures

(a) Non-reciprocal PTAs



(b) Reciprocal PTAs

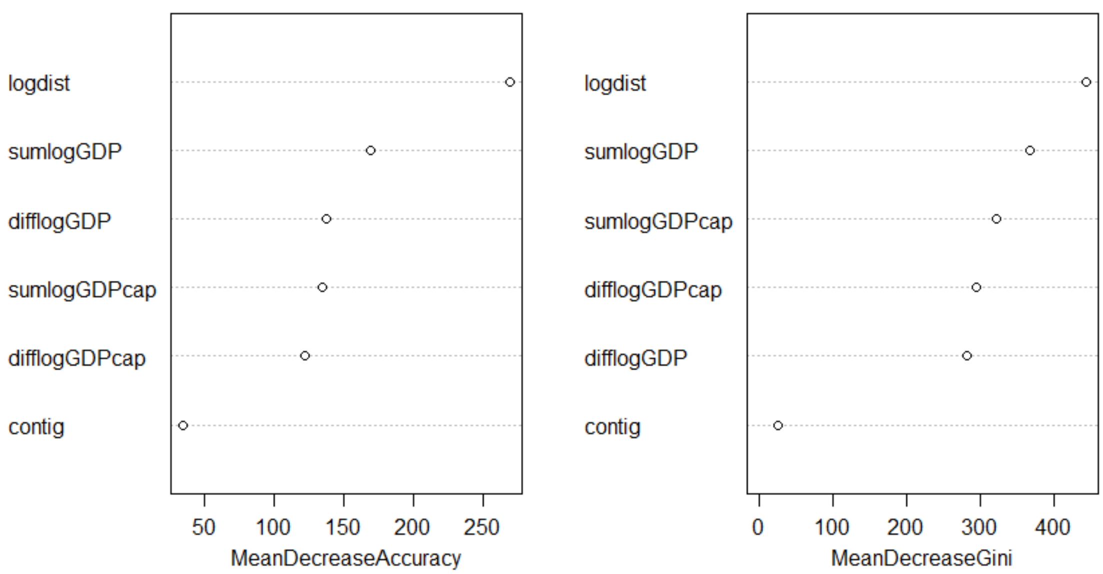
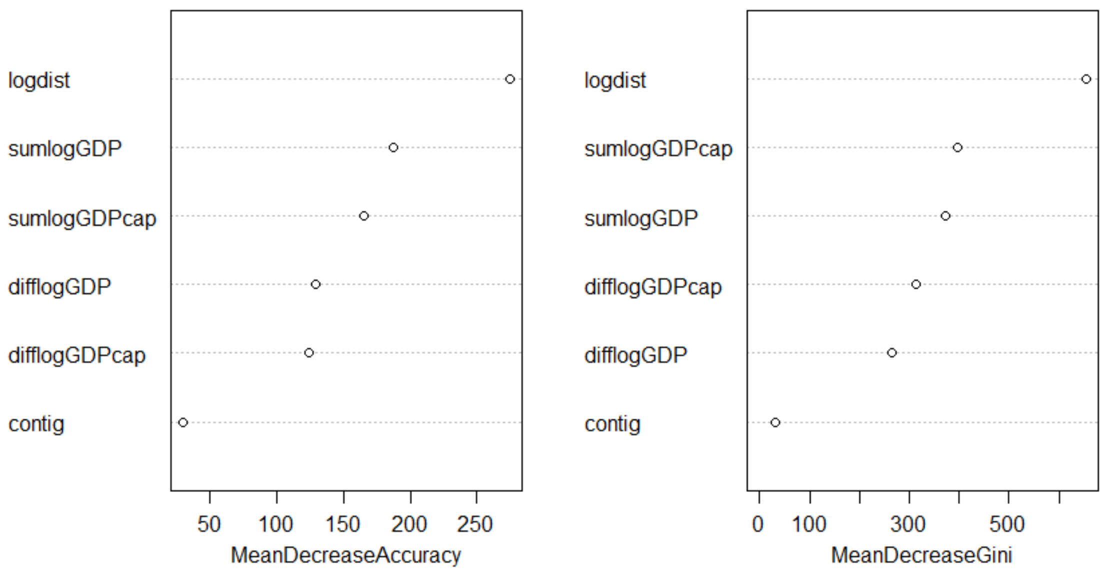


Figure 2.1: Random forest importance measures (cont.)

(c) FTAs



(d) Customs Unions

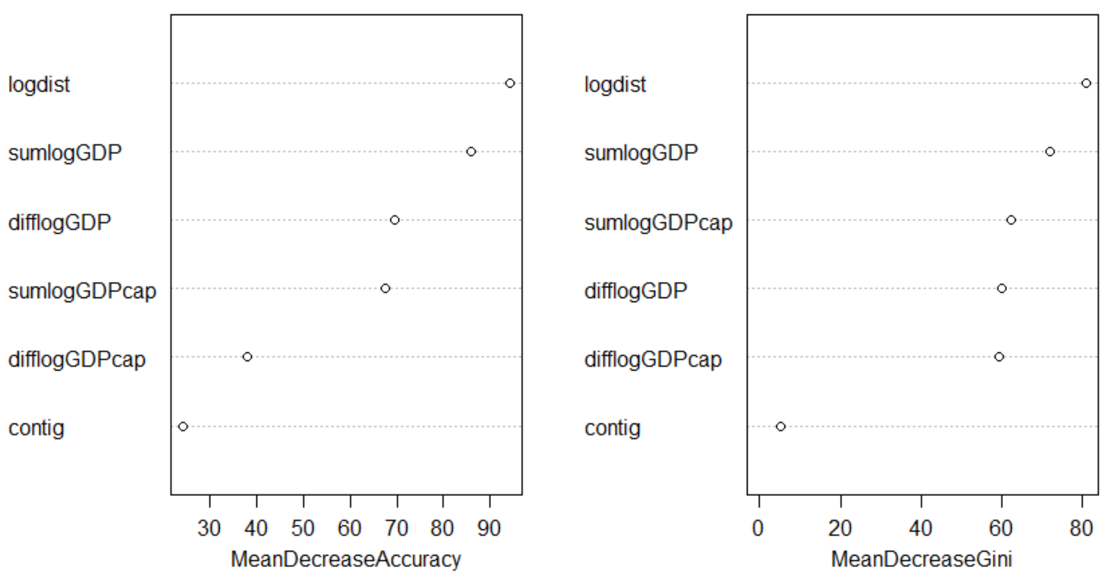
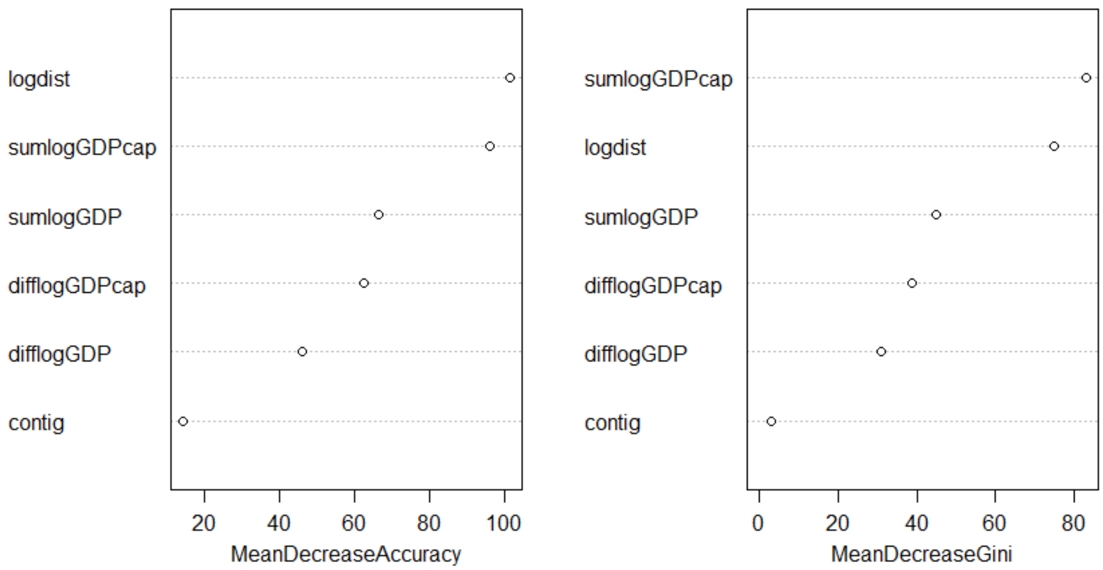
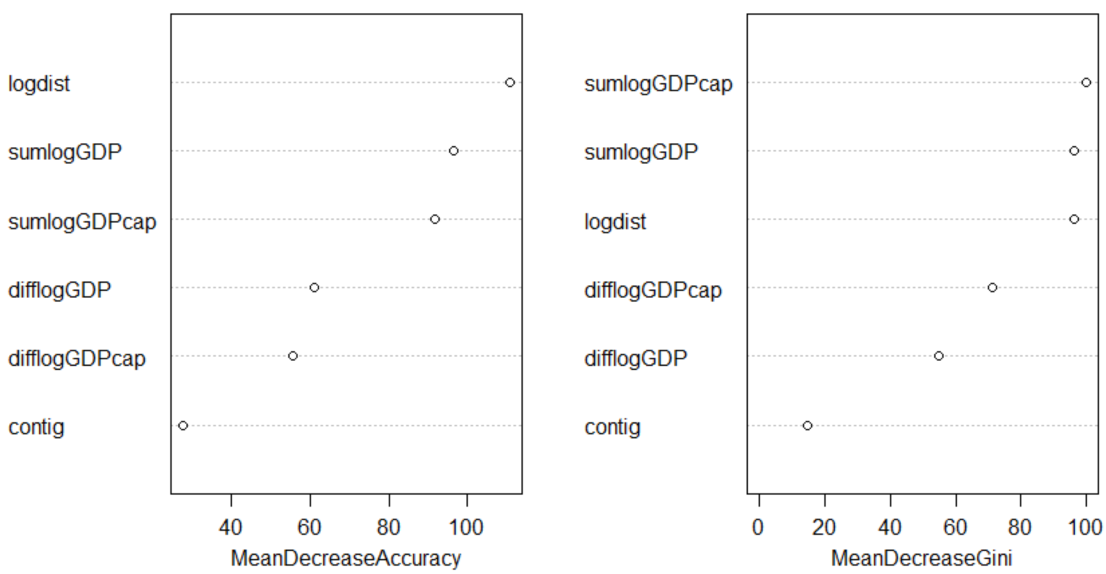


Figure 2.1: Random forest importance measures (cont.)

(e) Common Markets



(f) Economic Unions



Appendices

Appendix A Supplementary Results for "Temporary Trade Barriers and Trade Growth"

Table A1: Average growth rates by treatment status and destination

	All products		Same HS4	
Treatment	volume	share	volume	share
Argentina				
0	18.007	14.809	19.637	16.075
1	26.003	22.858	26.003	22.858
Australia				
0	10.405	9.947	12.254	11.928
1	26.673	19.758	26.673	19.758
Brazil				
0	16.713	15.510	18.406	16.947
1	35.964	26.383	35.964	26.383
Canada				
0	16.089	11.971	17.653	13.752
1	32.200	24.338	32.200	24.338
Colombia				
0	17.386	15.370	18.717	15.994
1	28.722	20.397	28.722	20.397
E.U.				
0	9.669	9.137	12.242	11.472
1	30.612	25.315	30.612	25.315
India				
0	19.531	11.532	21.594	12.761
1	31.232	18.513	31.232	18.513
Mexico				
0	3.023	15.079	4.191	17.277
1	28.958	26.399	28.958	26.399
Turkey				
0	15.510	14.479	16.452	15.225
1	46.705	41.940	46.705	41.940
U.S.				
0	11.302	9.814	12.861	11.353
1	29.050	26.857	29.050	26.857

Table A2: Cox proportional hazard

Dependent Variable:	AD_{ijt}
g_{ijt}	0.001*
	(0.001)
Observations	7,450
R^2	0.0004
Max. Possible R^2	0.853
Log Likelihood	−7,140.973
Wald Test	3.310* (df = 1)
LR Test	3.306* (df = 1)
Score (Logrank) Test	3.308* (df = 1)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table A3: Effect of AD investigation in the focal market, export data

Dependent:	Level			Growth		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): Quantity						
AD	-0.2528***	-0.3078***	-0.1536***	-8.376***	-7.841***	-7.689***
	(0.0479)	(0.0432)	(0.0554)	(0.5654)	(0.6151)	(0.6242)
Observations	269,921	269,921	269,921	276,149	276,149	276,149
R^2	0.85199	0.86808	0.89736	0.14769	0.14877	0.18198
Within R^2	0.00040	0.00064	0.00012	0.00028	0.00021	0.00020
Panel (c): Unit Value						
ad_init	-2.17×10^{-5}	0.0104	-0.0155	0.0023	0.0007	0.0010
	(0.0147)	(0.0143)	(0.0200)	(0.0029)	(0.0031)	(0.0032)
Observations	269,921	269,921	269,921	235,952	235,952	235,952
R^2	0.91407	0.92343	0.93689	0.13683	0.13744	0.16094
Within R^2	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Fixed effects						
ISO-year	✓	✓	–	✓	✓	–
HS-ISO	HS-6	HS-6	HS-6	–	HS-2	–
HS-year	HS-2	HS-4	–	HS-4	HS-4	HS-4
HS-year-ISO	–	–	HS-4	–	–	HS-2

Note: Estimates of (4) and (5) using Chinese export data to the 10 focal destinations. Standard errors in parentheses are clustered at the HS6-ISO level; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: Effect of AD investigation in the focal market, WLS

Dependent:	Level			Growth		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): Quantity						
AD	-0.4196***	-0.4960***	-0.3483***	-12.08***	-11.31***	-10.53***
	(0.0448)	(0.0410)	(0.0470)	(0.6075)	(0.6384)	(0.6482)
Observations	277,612	277,612	277,612	282,332	282,332	282,332
R^2	0.84721	0.85831	0.89736	0.14365	0.14533	0.19271
Within R^2	0.00122	0.00174	0.00069	0.00072	0.00054	0.00047
Panel (b): Import share						
AD	-0.0135***	-0.0198***	-0.0058	-10.23***	-9.347***	-9.193***
	(0.0047)	(0.0047)	(0.0056)	(0.4964)	(0.5279)	(0.5373)
Observations	387,481	387,481	387,481	293,838	293,838	293,838
R^2	0.68303	0.69738	0.74684	0.11061	0.11186	0.14637
Within R^2	0.00009	0.00020	0.00001	0.00069	0.00049	0.00047
Panel (c): Unit Value						
AD	0.1091***	0.1170***	0.0665***	0.0284***	0.0225***	0.0187***
	(0.0167)	(0.0168)	(0.0208)	(0.0037)	(0.0039)	(0.0040)
Observations	277,612	277,612	277,612	243,043	243,043	243,043
R^2	0.95541	0.95699	0.96962	0.10983	0.11142	0.16871
Within R^2	0.00021	0.00024	0.00007	0.00010	0.00005	0.00003
Fixed effects						
ISO-year	✓	✓	–	✓	✓	–
HS-ISO	HS-6	HS-6	HS-6	–	HS-2	–
HS-year	HS-2	HS-4	–	HS-4	HS-4	HS-4
HS-year-ISO	–	–	HS-4	–	–	HS-2

Note: Standard errors in parentheses are clustered at the HS6-ISO level; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: Effect of AD investigation in the focal market, SA (2021)

Dependent:	Level			Growth		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): Quantity						
ATT	-1.219***	-1.141***	-0.8373***	-11.48***	-9.978***	-9.598***
	(0.0578)	(0.0552)	(0.0676)	(0.5971)	(0.6212)	(0.6362)
Observations	277,611	277,611	277,611	282,331	282,331	282,331
R^2	0.84806	0.85884	0.89734	0.14400	0.14572	0.19295
Within R^2	0.00894	0.00800	0.00402	0.00196	0.00189	0.00171
Panel (b): Import share						
ATT	-0.0710***	-0.0737***	-0.0563***	-9.635***	-8.199***	-8.357***
	(0.0058)	(0.0059)	(0.0073)	(0.4839)	(0.5098)	(0.5210)
Observations	387,480	387,480	387,480	293,837	293,837	293,837
R^2	0.68500	0.69910	0.74820	0.11025	0.11155	0.14593
Within R^2	0.00459	0.00436	0.00295	0.00173	0.00164	0.00156
Panel (c): Unit Value						
ATT	0.1939***	0.1805***	0.0852***	0.0259***	0.0186***	0.0165***
	(0.0216)	(0.0218)	(0.0260)	(0.0039)	(0.0040)	(0.0041)
Observations	277,611	277,611	277,611	243,042	243,042	243,042
R^2	0.95611	0.95766	0.97007	0.11030	0.11185	0.16912
Within R^2	0.00192	0.00182	0.00108	0.00067	0.00066	0.00061
Fixed effects						
ISO-year	✓	✓	–	✓	✓	–
HS-ISO	HS-6	HS-6	HS-6	–	HS-2	–
HS-year	HS-2	HS-4	–	HS-4	HS-4	HS-4
HS-year-ISO	–	–	HS-4	–	–	HS-2

Note: Standard errors in parentheses are clustered at the HS6-ISO level; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A6: Effect of AD investigation in the focal market, other developing targets

Dependent:	Level			Growth		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): Quantity						
AD	-0.2486***	-0.2346***	-0.1265	-17.03***	-17.05***	-17.57***
	(0.0905)	(0.0875)	(0.1034)	(1.353)	(1.395)	(1.528)
Observations	194,979	194,979	194,979	216,780	216,780	216,780
R^2	0.79559	0.79979	0.86812	0.06852	0.06935	0.15862
Within R^2	0.00010	0.00009	0.00002	0.00027	0.00025	0.00025
Panel (b): Import share						
AD	-0.0056	-0.0058	-0.0090*	-16.66***	-16.80***	-17.24***
	(0.0047)	(0.0047)	(0.0052)	(1.288)	(1.325)	(1.453)
Observations	407,609	407,609	407,609	223,660	223,660	223,660
R^2	0.58589	0.58973	0.66765	0.05236	0.05301	0.14125
Within R^2	0.00004	0.00004	0.00009	0.00027	0.00025	0.00025
Panel (c): Unit Value						
AD	0.0247	0.0357	-0.0016	0.0204***	0.0132*	0.0134
	(0.0281)	(0.0288)	(0.0386)	(0.0074)	(0.0075)	(0.0085)
Observations	194,979	194,979	194,979	151,869	151,869	151,869
R^2	0.95607	0.95666	0.97121	0.10038	0.10155	0.20883
Within R^2	0.00001	0.00001	0.00000	0.00001	0.00000	0.00000
Fixed effects						
ISO$_o$-ISO$_d$-year	✓	✓	–	✓	✓	–
HS-ISO$_o$-ISO$_d$	HS-6	HS-6	HS-6	–	HS-2	–
HS-year	HS-2	HS-4	–	HS-4	HS-4	HS-4
HS-year-ISO$_o$-ISO$_d$	–	–	HS-4	–	–	HS-2

Note: AD cases of top 10 petitioners targeting India, Indonesia, Malaysia, and Thailand. Standard errors in parentheses are clustered at the HS6-ISO$_o$-ISO$_d$ level; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A7: Effect of AD investigation in the focal market, developed targets

Dependent:	Level			Growth		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): Quantity						
AD	0.0702	0.0783	-0.0131	-6.116***	-6.149***	-5.514***
	(0.0519)	(0.0522)	(0.0684)	(0.6871)	(0.6945)	(0.7283)
Observations	1,231,133	1,231,133	1,231,133	1,291,499	1,291,499	1,291,499
R^2	0.86123	0.86419	0.90389	0.03596	0.03635	0.09644
Within R^2	0.00001	0.00001	0.00000	0.00002	0.00002	0.00002
Panel (b): Unit Value						
AD	-0.0554***	-0.0437**	-0.0026	-0.0005	-0.0015	-0.0017
	(0.0180)	(0.0179)	(0.0228)	(0.0042)	(0.0043)	(0.0044)
Observations	1,231,133	1,231,133	1,231,133	1,028,191	1,028,191	1,028,191
R^2	0.92221	0.92464	0.94721	0.03102	0.03140	0.08863
Within R^2	0.00001	0.00001	0.00000	0.00000	0.00000	0.00000
Fixed effects						
ISO_o-ISO_d-year	✓	✓	–	✓	✓	–
HS-ISO_o-ISO_d	HS-6	HS-6	HS-6	–	HS-2	–
HS-year	HS-2	HS-4	–	HS-4	HS-4	HS-4
HS-year-ISO_o-ISO_d	–	–	HS-4	–	–	HS-2

Note: estimates using export data from and AD cases targeting the United States, the European Union, Japan, and South Korea. Standard errors in parentheses are clustered at the HS6-ISO_o-ISO_d level; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A8: Effect of AD investigation in the focal market, successful cases

Dependent:	Level			Growth		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): Quantity						
AD	-0.5703***	-0.6279***	-0.4241***	-11.55***	-10.62***	-9.872***
	(0.0522)	(0.0478)	(0.0541)	(0.6744)	(0.7079)	(0.7186)
Observations	272,215	272,215	272,215	276,998	276,998	276,998
R^2	0.84596	0.85713	0.89652	0.14539	0.14714	0.19397
Within R^2	0.00163	0.00202	0.00076	0.00039	0.00029	0.00026
Panel (b): Import share						
AD	-0.0264***	-0.0333***	-0.0157**	-9.652***	-8.735***	-8.443***
	(0.0054)	(0.0054)	(0.0064)	(0.5738)	(0.6049)	(0.6160)
Observations	381,089	381,089	381,089	288,397	288,397	288,397
R^2	0.68217	0.69650	0.74658	0.11149	0.11285	0.14794
Within R^2	0.00031	0.00050	0.00009	0.00033	0.00024	0.00023
Panel (c): Unit Value						
AD	0.1308***	0.1335***	0.0852***	0.0327***	0.0245***	0.0210***
	(0.0182)	(0.0183)	(0.0230)	(0.0042)	(0.0044)	(0.0046)
Observations	272,215	272,215	272,215	238,177	238,177	238,177
R^2	0.95535	0.95694	0.96962	0.10974	0.11126	0.16846
Within R^2	0.00032	0.00033	0.00011	0.00005	0.00002	0.00002
Fixed effects						
ISO-year	✓	✓	–	✓	✓	–
HS-ISO	HS-6	HS-6	HS-6	–	HS-2	–
HS-year	HS-2	HS-4	–	HS-4	HS-4	HS-4
HS-year-ISO	–	–	HS-4	–	–	HS-2

Note: Using investigation date of successful cases as treatment. Standard errors in parentheses are clustered at the HS6-ISO level; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A9: Effect of AD investigation in the focal market, log-differenced growth rates

Dependent:	OLS		WLS		SA (2021)	
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): Quantity						
AD	-0.1684***	-0.1526***	-0.1681***	-0.1529***	-0.1603***	-0.1373***
	(0.0094)	(0.0096)	(0.0092)	(0.0094)	(0.0091)	(0.0094)
Observations	243,043	243,043	243,043	243,043	243,042	243,042
R^2	0.09668	0.09867	0.09697	0.09897	0.09798	0.10001
Within R^2	0.00054	0.00040	0.00074	0.00054	0.00198	0.00189
Panel (b): Import share						
AD	-0.1240***	-0.1127***	-0.1236***	-0.1133***	-0.1176***	-0.1024***
	(0.0070)	(0.0072)	(0.0068)	(0.0069)	(0.0068)	(0.0070)
Observations	259,695	259,695	259,695	259,695	259,694	259,694
R^2	0.05370	0.05546	0.05436	0.05614	0.05492	0.05672
Within R^2	0.00047	0.00034	0.00066	0.00048	0.00176	0.00167
Fixed effects						
ISO-year	✓	✓	✓	✓	✓	✓
HS-ISO	–	HS-2	–	HS-2	–	HS-2
HS-year	HS-4	HS-4	HS-4	HS-4	HS-4	HS-4

Note: Unit value omitted. Standard errors in parentheses are clustered at the HS6-ISO level; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A10: Correlation matrix of AD activity between all filers and 10 focal destinations

	IND	USA	EUN	BRA	TUR	ARG	MEX	COL	AUS	CAN
ARG	0.03	0.07	0.14	0.29	0.13	1.00	0.08	0.23	0.05	0.12
AUS	0.05	0.23	0.12	0.10	-0.01	0.05	0.13	0.10	1.00	0.21
BRA	0.07	0.15	0.21	1.00	0.23	0.29	0.16	0.18	0.10	0.11
CAN	-0.01	0.33	0.21	0.11	0.04	0.12	0.18	0.15	0.21	1.00
COL	0.03	0.07	0.18	0.18	0.04	0.23	0.15	1.00	0.10	0.15
EUN	0.13	0.28	1.00	0.21	0.17	0.14	0.19	0.18	0.12	0.21
IDN	0.09	0.22	0.16	0.14	0.13	0.06	0.15	0.02	0.07	0.27
IND	1.00	0.07	0.13	0.07	0.05	0.03	0.02	0.03	0.05	-0.01
ISR	0.01	0.02	0.01	-0.01	0.06	0.04	0.03	-0.01	0.03	0.02
JAM	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
JPN	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
KOR	0.08	0.03	0.04	0.04	0.00	0.04	0.03	0.05	0.04	-0.01
MEX	0.02	0.20	0.19	0.16	0.12	0.08	1.00	0.15	0.13	0.18
MYS	0.06	0.19	0.21	0.11	0.16	-0.01	0.23	0.02	0.08	0.21
NZL	-0.01	0.07	0.03	0.00	0.00	0.00	0.05	0.04	0.00	0.04
PAK	0.01	0.20	0.24	0.03	0.01	0.08	0.19	0.07	0.09	0.13
PER	0.00	0.02	0.03	0.14	0.06	0.17	0.08	0.29	0.02	0.03
PHL	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
RUS	0.12	0.13	0.16	0.17	0.10	0.07	0.04	0.04	-0.01	0.04
THA	0.06	0.29	0.25	0.11	0.12	0.07	0.23	0.04	0.14	0.31
TTO	-0.01	0.08	-0.01	-0.01	0.04	0.12	-0.01	0.11	0.19	0.18
TUR	0.05	0.11	0.17	0.23	1.00	0.13	0.12	0.04	-0.01	0.04
UKR	0.01	0.03	0.06	0.04	0.01	0.01	0.02	0.04	-0.01	-0.01
URY	0.00	0.00	0.00	0.00	0.08	0.00	0.00	0.00	0.00	0.00
USA	0.07	1.00	0.28	0.15	0.11	0.07	0.20	0.07	0.23	0.33
ZAF	0.06	0.13	0.09	0.11	0.13	0.09	0.10	0.14	0.15	0.08

Note: 17 columns representing non-focal destinations omitted for table display, but included in the empirical analysis.

Table A11: Effect of AD investigation in third markets, WLS

Dependent:	Level			Growth		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): Quantity						
AD	0.1044***	0.1292***	0.1094**	-1.170*	-1.250**	-1.223*
	(0.0379)	(0.0389)	(0.0447)	(0.6189)	(0.6216)	(0.6312)
s_i^{AD}				0.4078	0.4130	0.4139
				(0.2795)	(0.2808)	(0.2858)
Observations	2,173,522	2,173,522	2,173,522	2,412,041	2,412,041	2,412,041
R^2	0.84625	0.85367	0.90407	0.08316	0.08512	0.15476
Within R^2	0.00021	0.00024	0.00023	0.00001	0.00001	0.00001
Panel (b): Unit Value						
AD	-0.0014	0.0090	0.0142	0.0028	0.0029	0.0026
	(0.0175)	(0.0171)	(0.0198)	(0.0038)	(0.0038)	(0.0039)
s_i^{AD}				0.0017	0.0017	0.0019
				(0.0015)	(0.0015)	(0.0016)
Observations	2,173,522	2,173,522	2,173,522	1,715,749	1,715,749	1,715,749
R^2	0.90158	0.90757	0.93741	0.09234	0.09448	0.14710
Within R^2	0.00000	0.00000	0.00001	0.00001	0.00001	0.00001
Fixed effects						
ISO-year	✓	✓	–	✓	✓	–
HS-ISO	HS-6	HS-6	HS-6	–	HS-2	–
HS-year	HS-2	HS-4	–	HS-4	HS-4	HS-4
HS-year-ISO	–	–	HS-4	–	–	HS-2

Note: s_i^{AD} is standardized. Standard errors in parentheses are clustered at the HS6 level; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A12: Effect of AD investigation in third markets, SA (2021)

Dependent:	Level			Growth		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): Quantity						
AD	0.0435	0.1095**	0.0296	-0.9873	-1.074*	-1.046
	(0.0462)	(0.0434)	(0.0469)	(0.6494)	(0.6518)	(0.6627)
$AD \times s_i^{AD}$	0.0243	0.0133	0.0243	-1.442**	-1.448**	-1.353**
	(0.0262)	(0.0288)	(0.0266)	(0.6132)	(0.6175)	(0.6287)
s_i^{AD}				1.233***	1.249***	1.188**
				(0.4761)	(0.4799)	(0.4865)
Observations	2,173,522	2,173,522	2,173,522	2,412,041	2,412,041	2,412,041
R^2	0.84615	0.85339	0.85930	0.08347	0.08541	0.15494
Within R^2	0.00273	0.00161	0.00292	0.00116	0.00116	0.00123
Panel (b): Unit Value						
AD	0.0034	-0.0048	0.0051	0.0056	0.0058	0.0055
	(0.0207)	(0.0200)	(0.0208)	(0.0045)	(0.0045)	(0.0047)
$AD \times s_i^{AD}$	0.0158	0.0130	0.0156	0.0027	0.0028	0.0027
	(0.0119)	(0.0140)	(0.0120)	(0.0036)	(0.0037)	(0.0038)
s_i^{AD}				-0.0013	-0.0013	-0.0011
				(0.0029)	(0.0030)	(0.0031)
Observations	2,173,522	2,173,522	2,173,522	1,715,749	1,715,749	1,715,749
R^2	0.90362	0.90918	0.90951	0.08861	0.09066	0.14218
Within R^2	0.00289	0.00148	0.00299	0.00088	0.00088	0.00091
Fixed effects						
ISO-year	✓	✓	–	✓	✓	–
HS-ISO	HS-6	HS-6	HS-6	–	HS-2	–
HS-year	HS-2	HS-4	–	HS-4	HS-4	HS-4
HS-year-ISO	–	–	HS-4	–	–	HS-2

Note: s_i^{AD} is standardized. Standard errors in parentheses are clustered at the HS6 level; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A13: Effect of AD investigation in third markets, other developing targets

Dependent:	Level			Growth		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): Quantity						
AD	0.0992**	0.1312***	0.1541**	-1.922**	-1.995**	-2.033**
	(0.0433)	(0.0474)	(0.0603)	(0.8526)	(0.8594)	(0.8757)
s_i^{AD}				0.4077	0.3770	0.3903
				(0.2584)	(0.2594)	(0.2631)
Observations	1,357,111	1,357,111	1,357,111	1,633,000	1,633,000	1,633,000
R^2	0.76933	0.77193	0.86707	0.04211	0.04391	0.11359
Within R^2	0.00010	0.00013	0.00023	0.00002	0.00002	0.00002
Panel (b): Unit Value						
AD	-0.0183	-0.0151	-0.0155	-0.0030	-0.0030	-0.0032
	(0.0157)	(0.0162)	(0.0201)	(0.0024)	(0.0025)	(0.0026)
s_i^{AD}				-0.0007	-0.0005	-0.0004
				(0.0010)	(0.0010)	(0.0010)
Observations	1,357,111	1,357,111	1,357,111	928,175	928,175	928,175
R^2	0.82099	0.82334	0.89834	0.07314	0.07634	0.14355
Within R^2	0.00002	0.00001	0.00001	0.00000	0.00000	0.00000
Fixed effects						
ISO$_o$-ISO$_d$-year	✓	✓	–	✓	✓	–
HS-ISO$_o$-ISO$_d$	HS-6	HS-6	HS-6	–	HS-2	–
HS-year	HS-2	HS-4	–	HS-4	HS-4	HS-4
HS-year-ISO$_o$-ISO$_d$	–	–	HS-4	–	–	HS-2

Note: estimates using export data from and AD cases of top 10 petitioners targeting India, Indonesia, Malaysia, and Thailand. s_i^{AD} is standardized. Standard errors in parentheses are clustered at the HS6 level; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A14: Effect of AD investigation in third markets, developed targets

Dependent:	Level			Growth		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): Quantity						
AD	0.0690**	0.1044***	0.1115***	-2.106***	-2.070***	-2.082***
	(0.0343)	(0.0340)	(0.0414)	(0.4668)	(0.4683)	(0.4722)
s_i^{AD}				0.2127*	0.2269*	0.2270*
				(0.1213)	(0.1248)	(0.1272)
Observations	4,003,394	4,003,394	4,003,394	4,489,988	4,489,988	4,489,988
R^2	0.85927	0.86130	0.91323	0.03800	0.03875	0.07492
Within R^2	0.00004	0.00008	0.00008	0.00002	0.00002	0.00002
Panel (b): Unit Value						
AD	-0.0171	-0.0169	-0.0006	-0.0044	-0.0043	-0.0044*
	(0.0207)	(0.0213)	(0.0224)	(0.0028)	(0.0027)	(0.0027)
s_i^{AD}				0.0002	0.0002	0.0003
				(0.0008)	(0.0007)	(0.0007)
Observations	4,003,394	4,003,394	4,003,394	3,054,627	3,054,627	3,054,627
R^2	0.91723	0.91905	0.94970	0.02542	0.02686	0.06272
Within R^2	0.00001	0.00001	0.00000	0.00000	0.00000	0.00000
Fixed effects						
ISO_o-ISO_d-year	✓	✓	–	✓	✓	–
HS-ISO_o-ISO_d	HS-6	HS-6	HS-6	–	HS-2	–
HS-year	HS-2	HS-4	–	HS-4	HS-4	HS-4
HS-year-ISO_o-ISO_d	–	–	HS-4	–	–	HS-2

Note: estimates using export data from and AD cases targeting the United States, the European Union, Japan, and South Korea. s_i^{AD} is standardized. Standard errors in parentheses are clustered at the HS6 level; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A15: Effect of AD investigation in third markets, successful cases

Dependent:	Level			Growth		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): Quantity						
AD	0.1058**	0.1614***	0.1430***	-1.107*	-1.177*	-1.141*
	(0.0412)	(0.0399)	(0.0454)	(0.6655)	(0.6682)	(0.6796)
s_i^{AD}				0.4335	0.4350	0.4375
				(0.2739)	(0.2756)	(0.2819)
Observations	1,951,238	1,951,238	1,951,238	2,162,986	2,162,986	2,162,986
R^2	0.84416	0.85186	0.90251	0.08061	0.08260	0.15500
Within R^2	0.00021	0.00037	0.00038	0.00001	0.00001	0.00001
Panel (b): Unit Value						
AD	-0.0139	-0.0192	-0.0141	-0.0002	-0.0002	-0.0004
	(0.0183)	(0.0175)	(0.0204)	(0.0042)	(0.0042)	(0.0044)
s_i^{AD}				0.0012	0.0013	0.0014
				(0.0015)	(0.0015)	(0.0016)
Observations	1,951,238	1,951,238	1,951,238	1,540,974	1,540,974	1,540,974
R^2	0.90685	0.91260	0.93916	0.09466	0.09684	0.15193
Within R^2	0.00001	0.00002	0.00001	0.00000	0.00000	0.00000
Fixed effects						
ISO-year	✓	✓	–	✓	✓	–
HS-ISO	HS-6	HS-6	HS-6	–	HS-2	–
HS-year	HS-2	HS-4	–	HS-4	HS-4	HS-4
HS-year-ISO	–	–	HS-4	–	–	HS-2

Note: Using investigation year for successful cases as treatment. s_i^{AD} is standardized. Standard errors in parentheses are clustered at the HS6 level; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

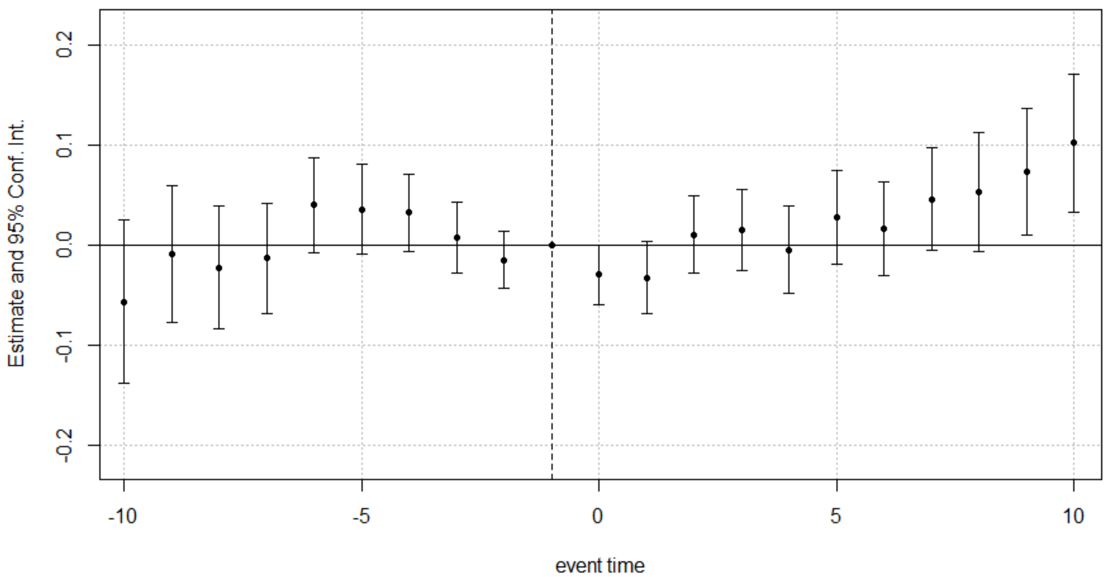
Table A16: Effect of AD investigation in third markets, log-differenced growth rates

Dependent:	Quantity Growth		
Model:	(1)	(2)	(3)
Panel (a): OLS			
AD	-0.0078	-0.0084	-0.0082
	(0.0057)	(0.0057)	(0.0058)
s_i^{AD}	0.0043*	0.0043*	0.0041
	(0.0025)	(0.0026)	(0.0026)
Observations	1,715,749	1,715,749	1,715,749
R^2	0.08574	0.08918	0.15812
Within R^2	0.00001	0.00001	0.00001
Panel (a): WLS			
AD	-0.0073	-0.0079	-0.0078
	(0.0057)	(0.0057)	(0.0058)
s_i^{AD}	0.0039	0.0039	0.0038
	(0.0026)	(0.0026)	(0.0027)
Observations	1,715,749	1,715,749	1,715,749
R^2	0.08689	0.09036	0.15941
Within R^2	0.00001	0.00001	0.00001
Panel (c): SA (2021)			
AD	-0.0050	-0.0055	-0.0060
	(0.0066)	(0.0066)	(0.0067)
s_i^{AD}	0.0031	0.0031	0.0031
	(0.0026)	(0.0026)	(0.0027)
Observations	1,715,748	1,715,748	1,715,748
R^2	0.08628	0.08972	0.15866
Within R^2	0.00060	0.00060	0.00064
Fixed effects			
ISO-year	✓	✓	–
HS-ISO	–	HS-2	–
HS-year	HS-4	HS-4	HS-4
HS-year-ISO	–	–	HS-2

Note: s_i^{AD} is standardized. Standard errors in parentheses are clustered at the HS6 level; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure A1: Effect of AD on unit value in focal markets, export data

(a) log import volume



(b) Growth in import volume

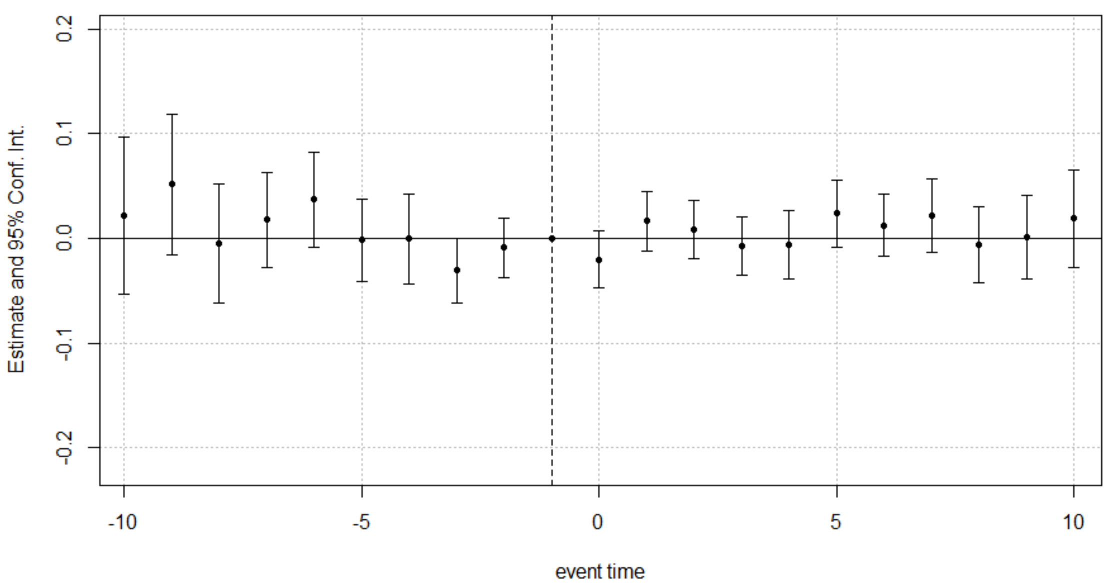
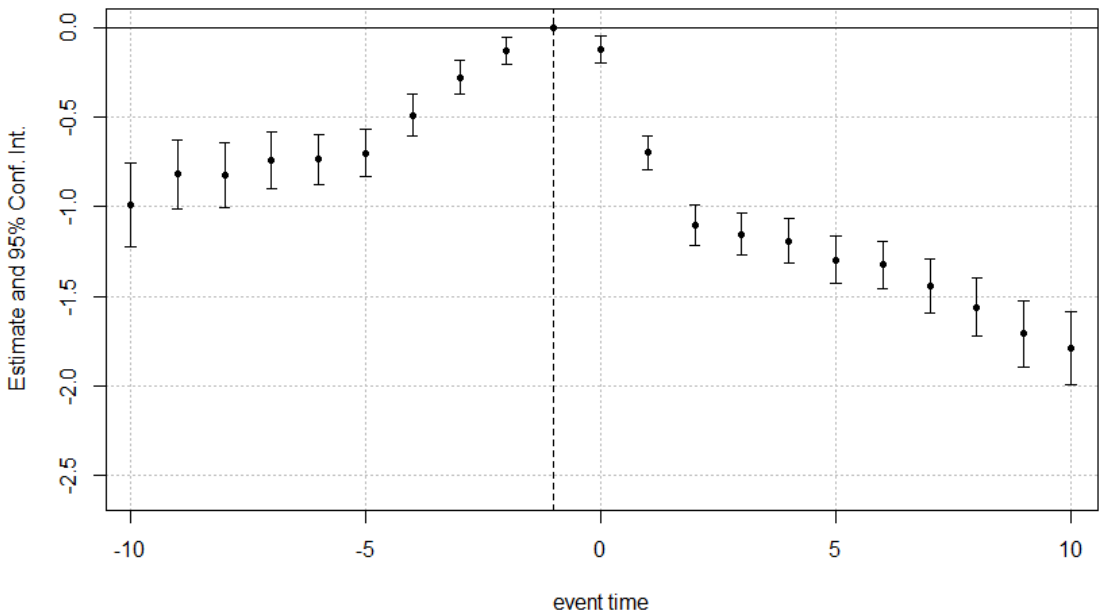


Figure A2: Effect of AD on import volume in focal markets, WLS

(a) log import volume



(b) Growth in import volume

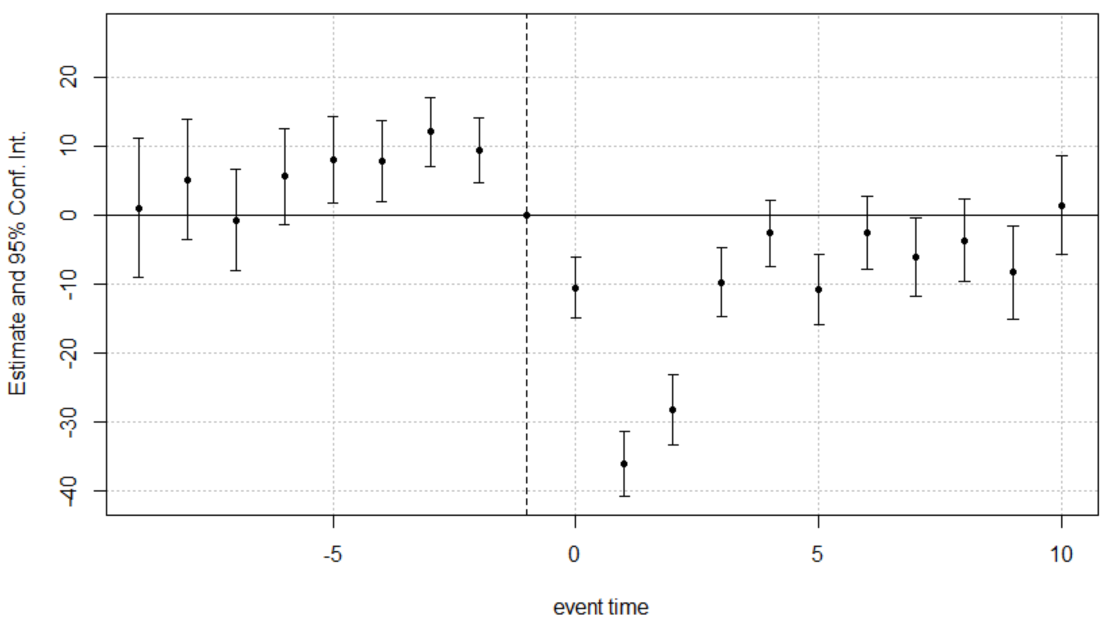
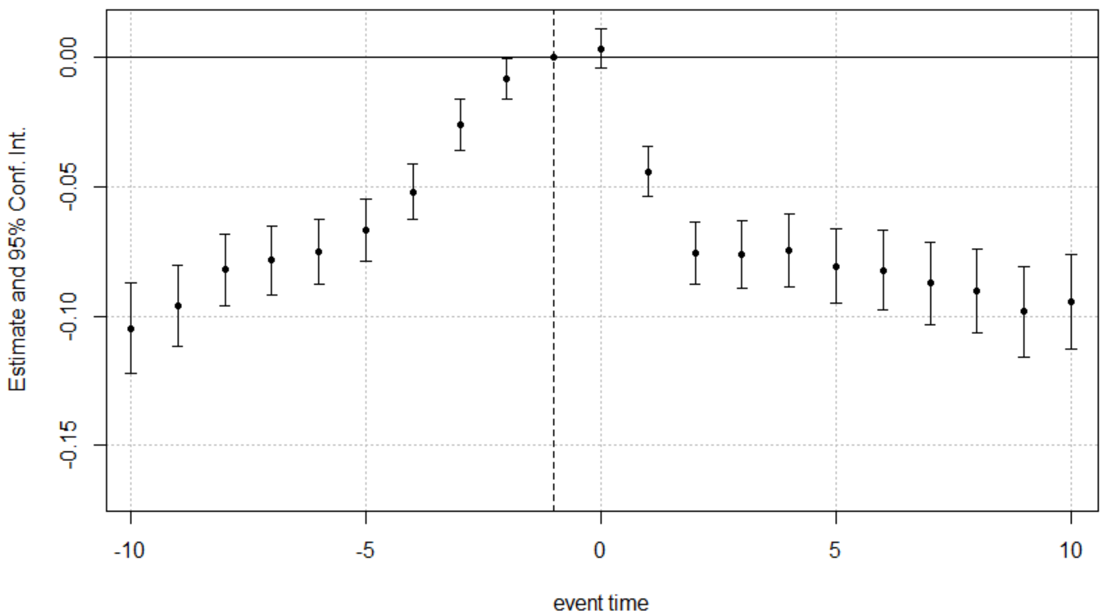


Figure A3: Effect of AD on import share in focal markets, WLS

(a) Import share



(b) Growth in import share

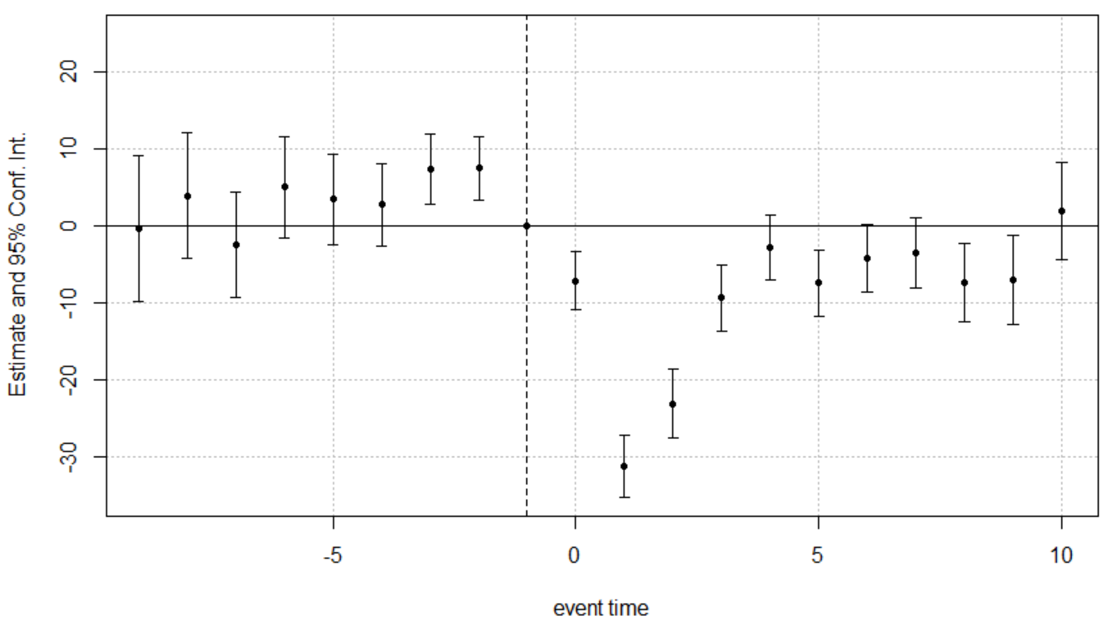
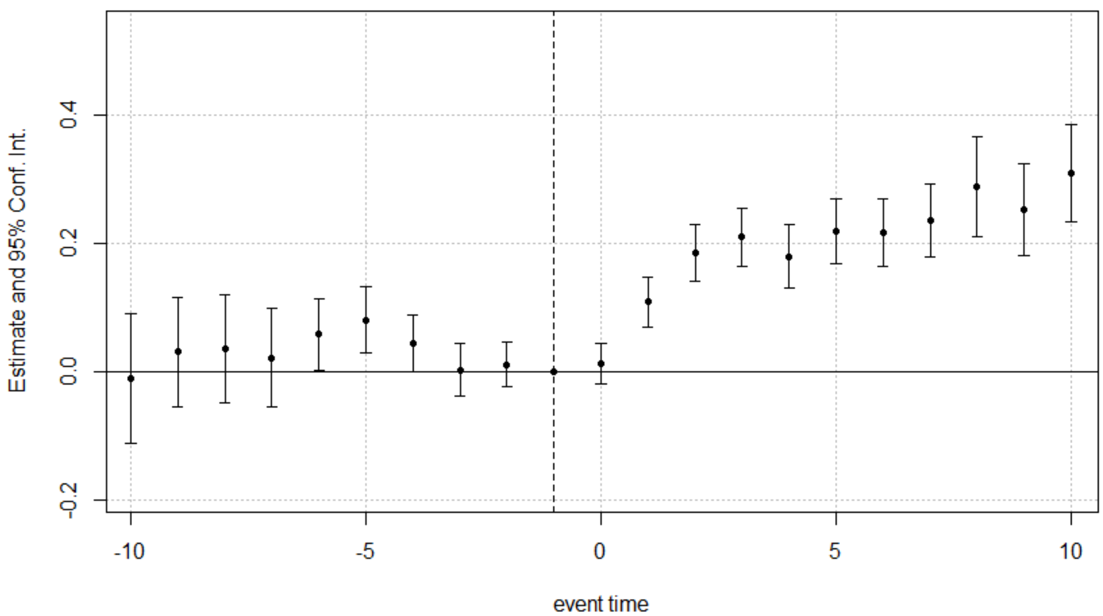


Figure A4: Effect of AD on unit value in focal markets, WLS

(a) Log unit value



(b) Growth in unit value

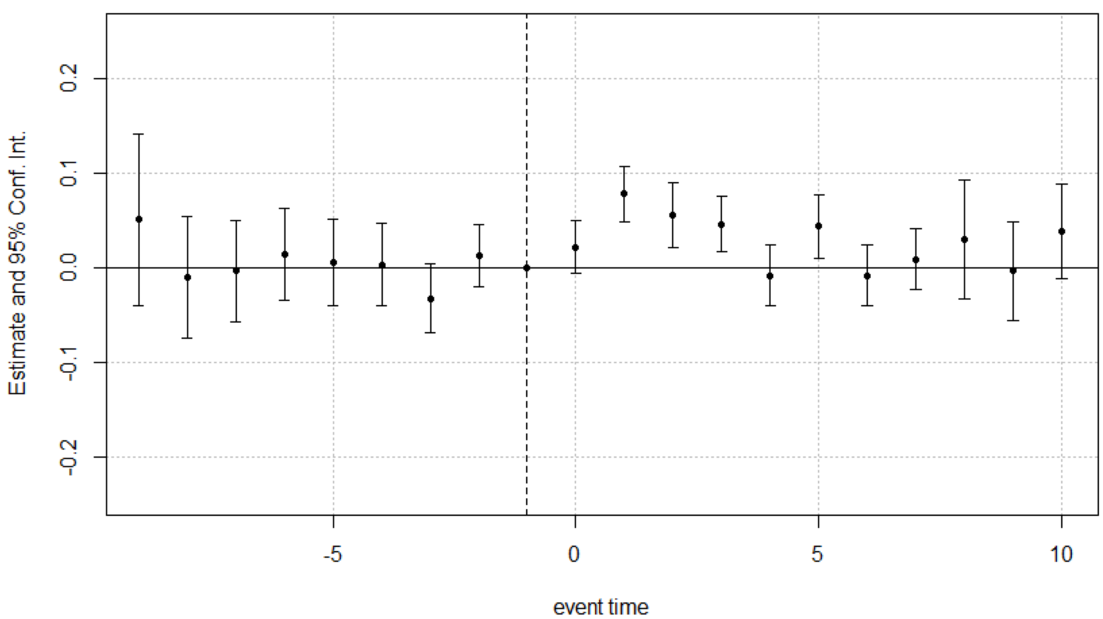


Figure A5: Effect of AD on import volume in focal markets, SA (2021)

(a) log import volume

(b) Growth in import volume

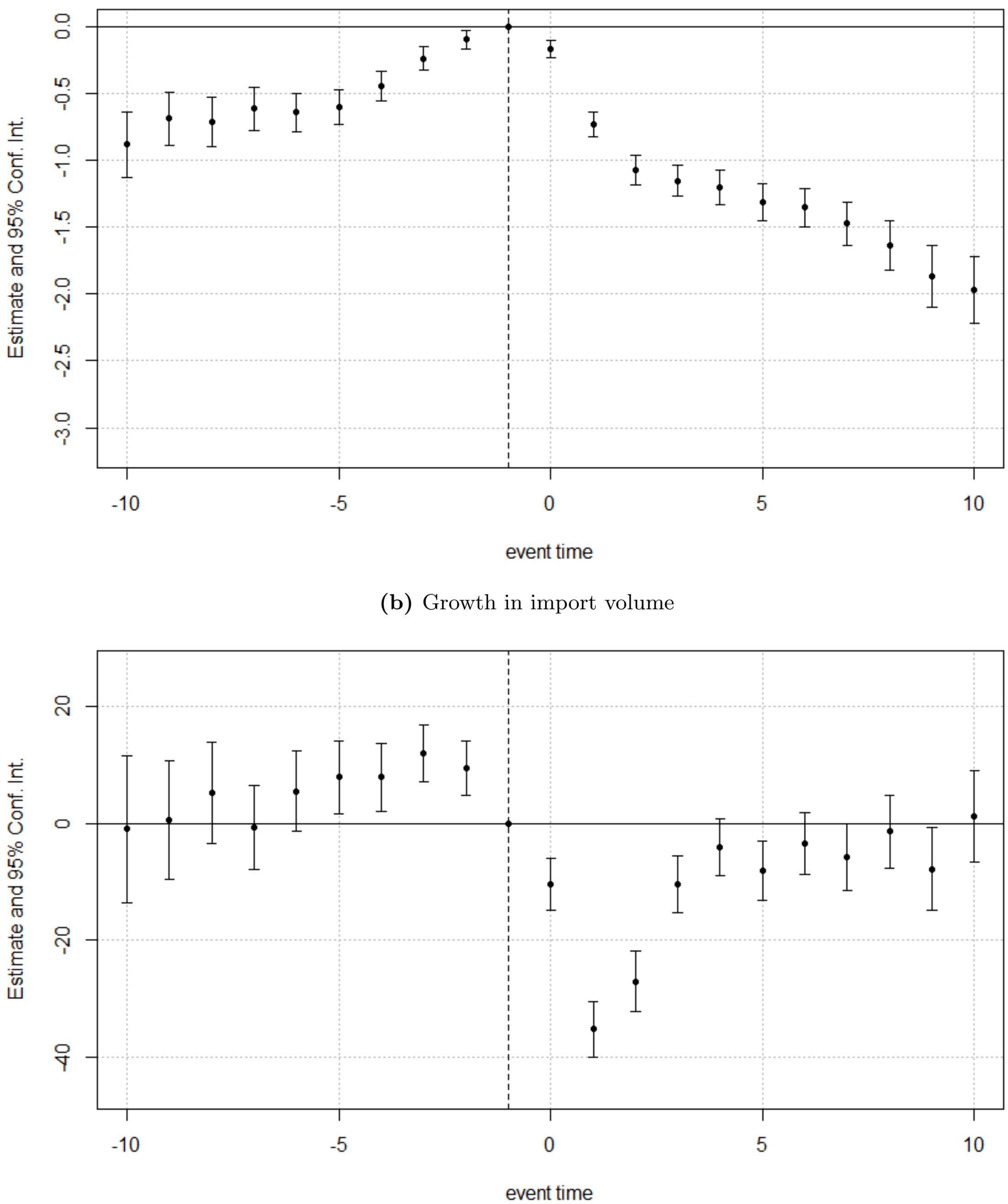
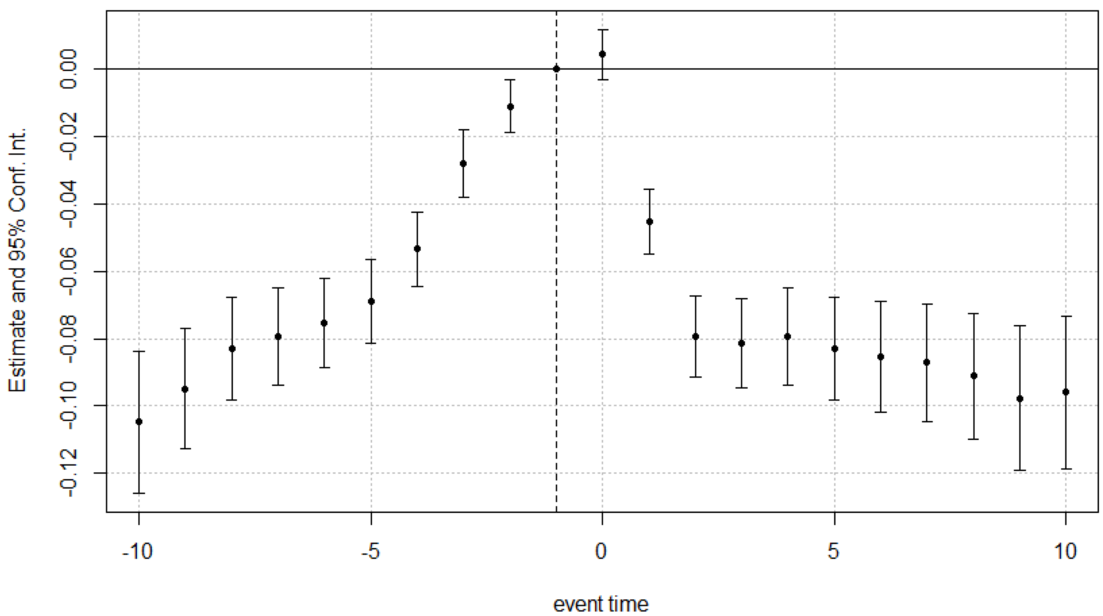


Figure A6: Effect of AD on import share in focal markets, SA (2021)

(a) Import share



(b) Growth in import share

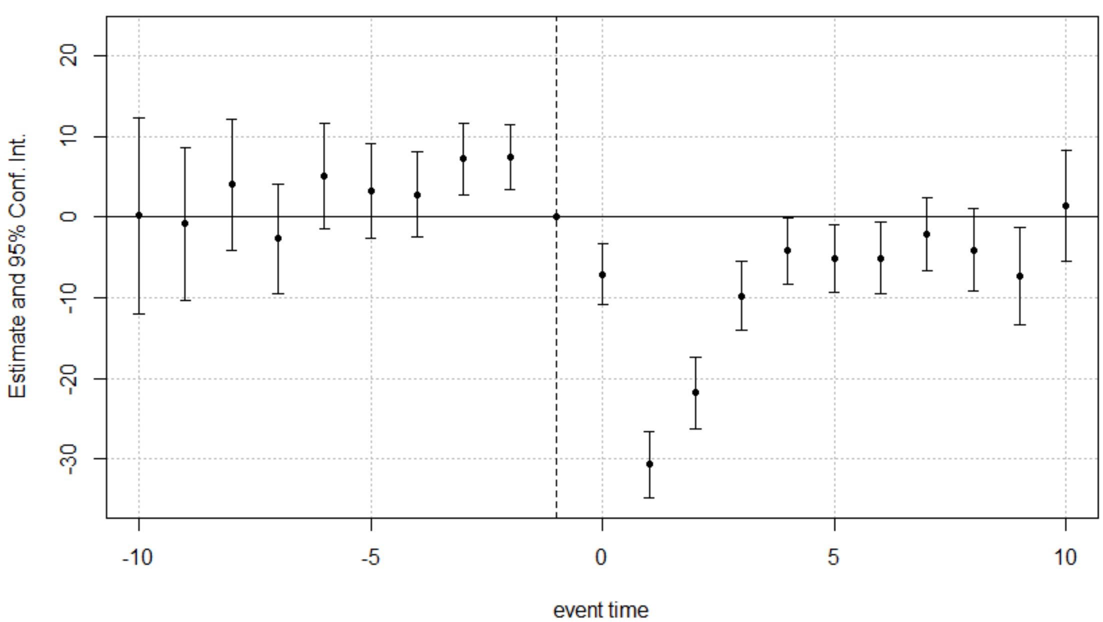
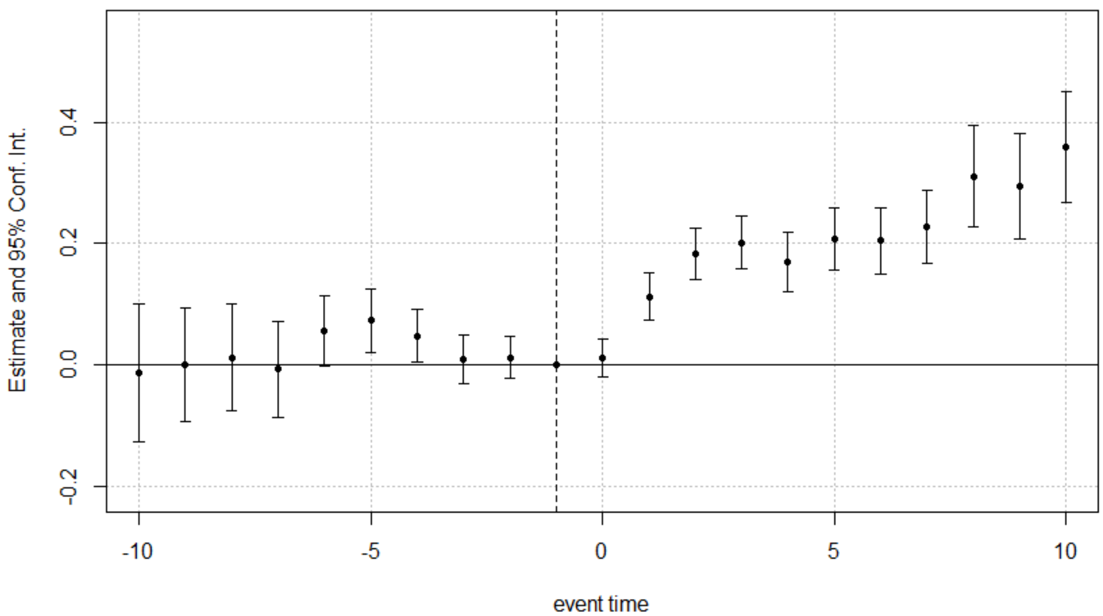


Figure A7: Effect of AD on unit value in focal markets, SA (2021)

(a) Log unit value



(b) Growth in unit value

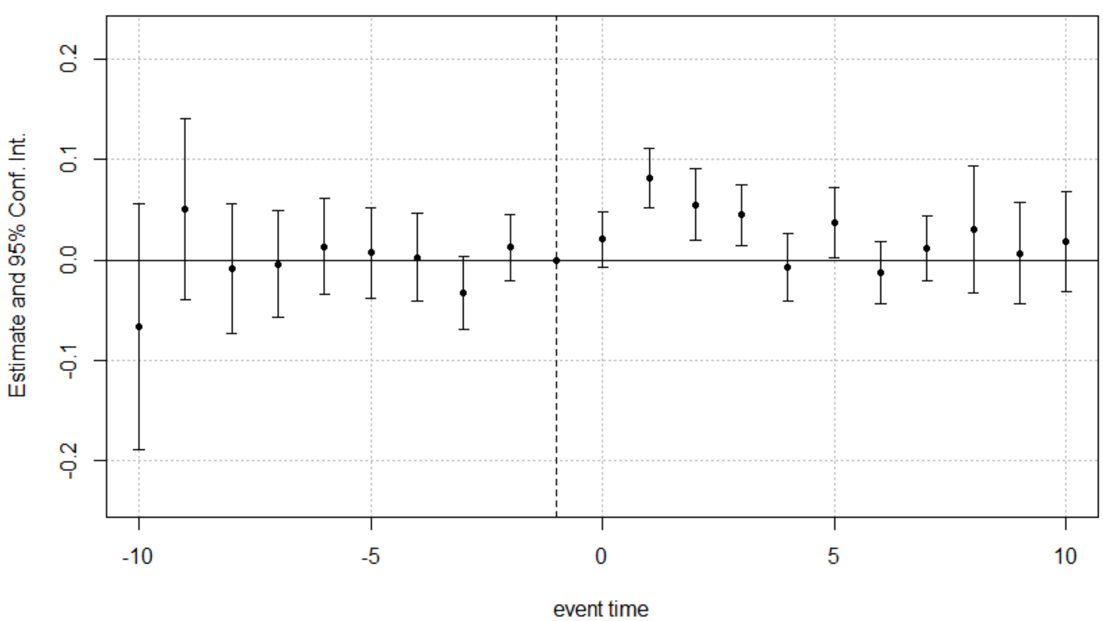
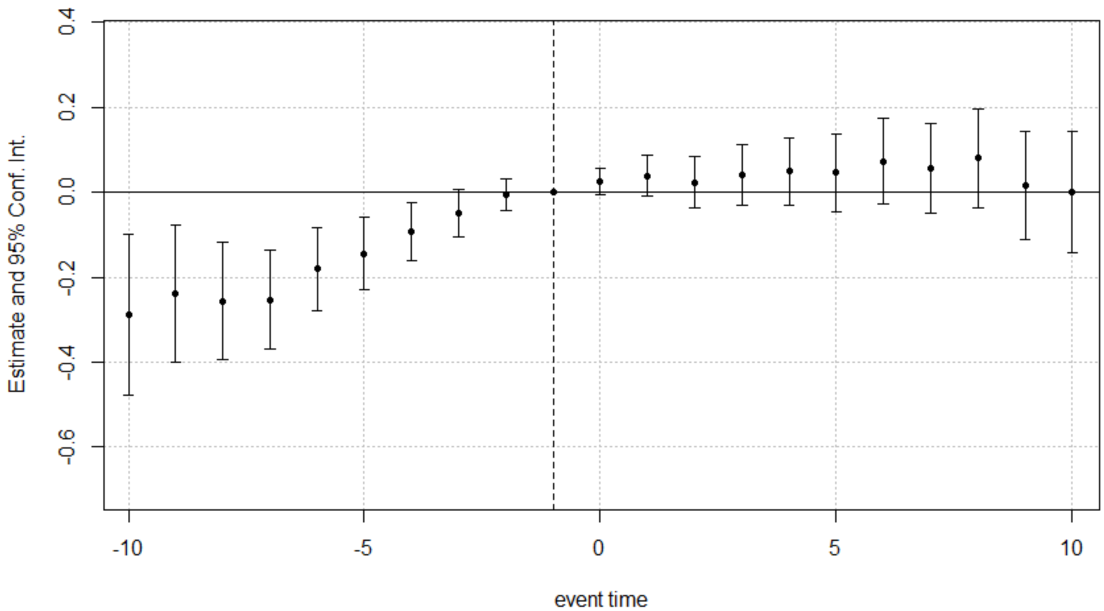


Figure A8: Effect of AD on import volume in non-target markets, WLS

(a) Log import volume



(b) Growth in import volume

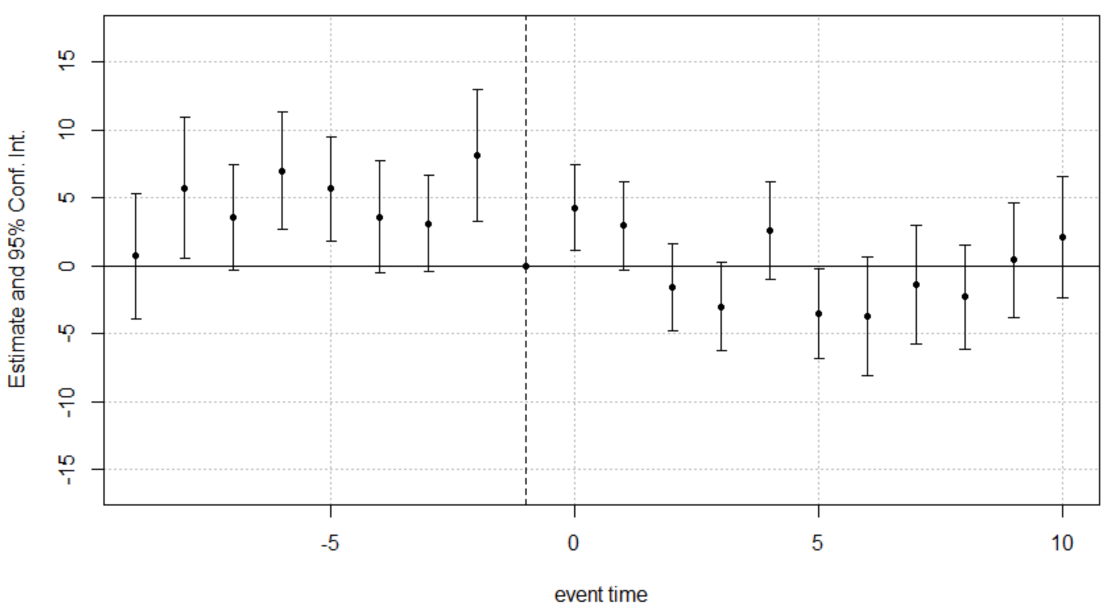
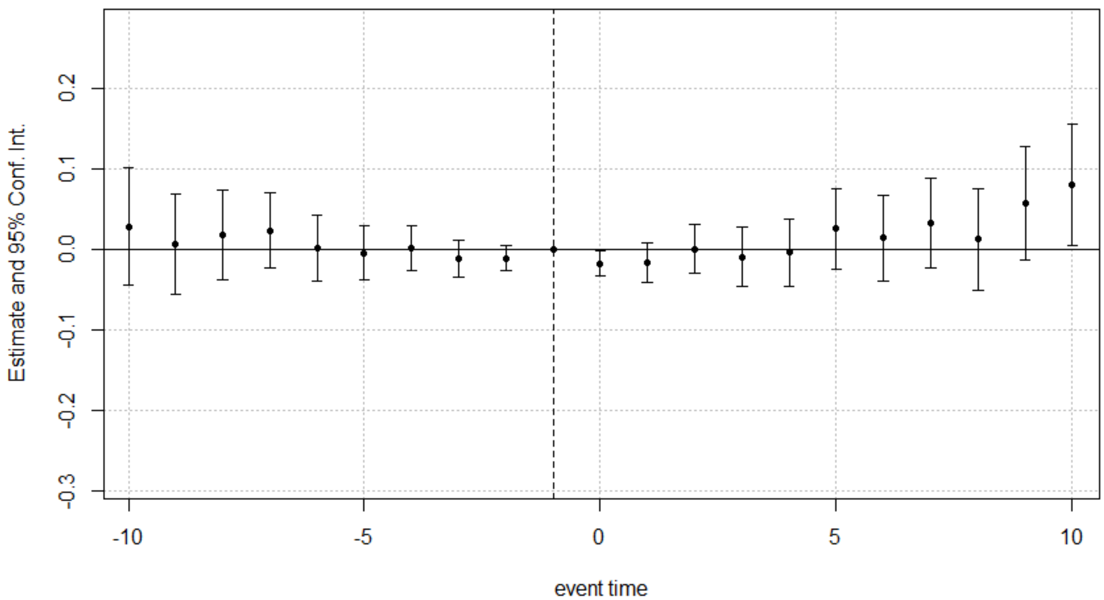


Figure A9: Effect of AD on unit value in non-target markets, WLS

(a) Log unit value



(b) Growth in unit value

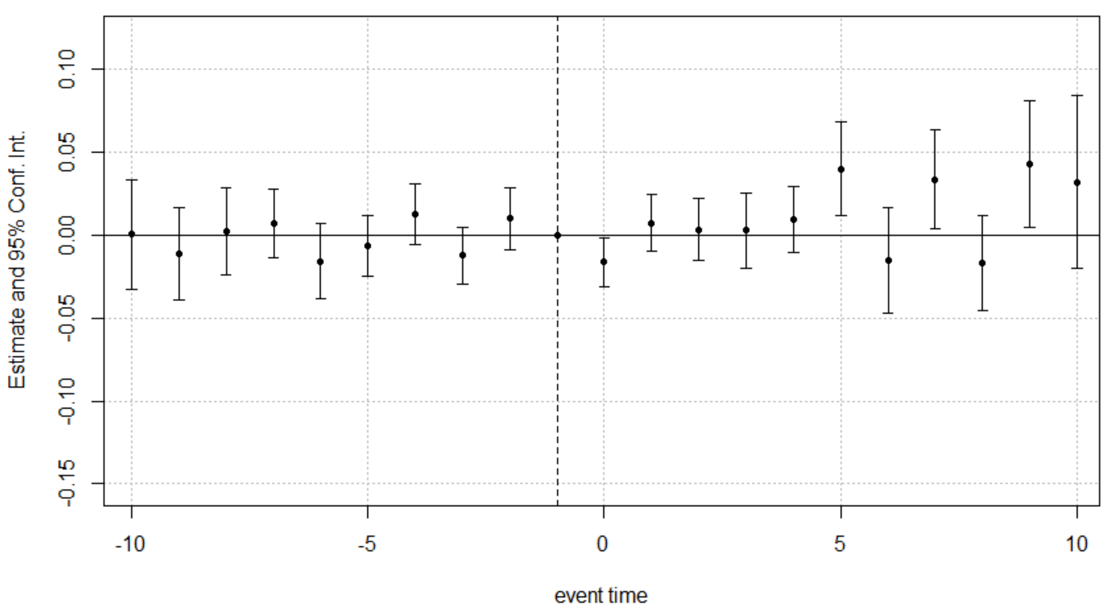
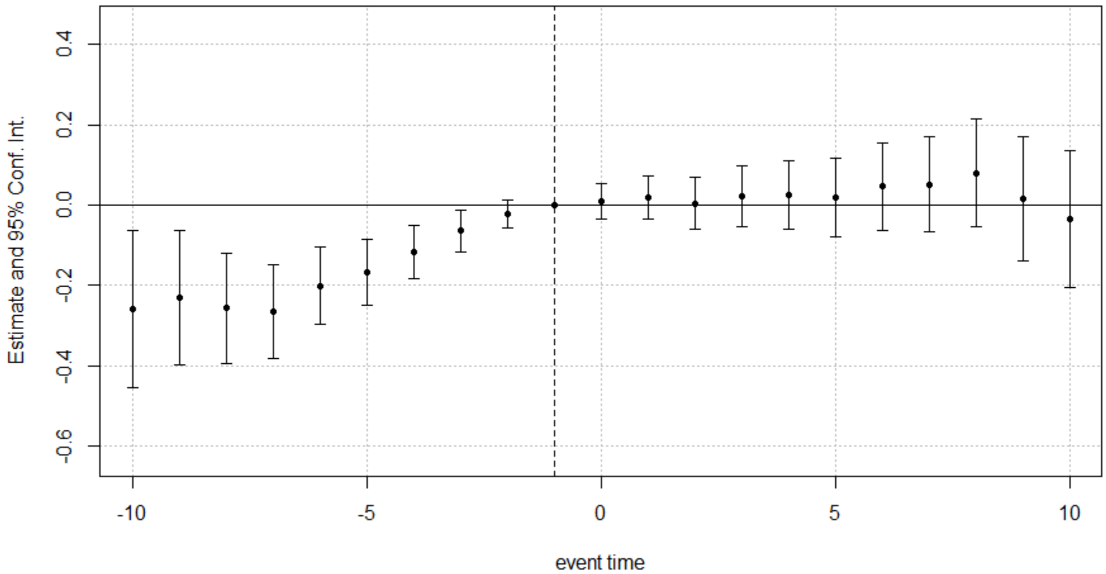


Figure A10: Effect of AD on import volume in non-target markets, SA (2021)

(a) Log import volume



(b) Growth in import volume

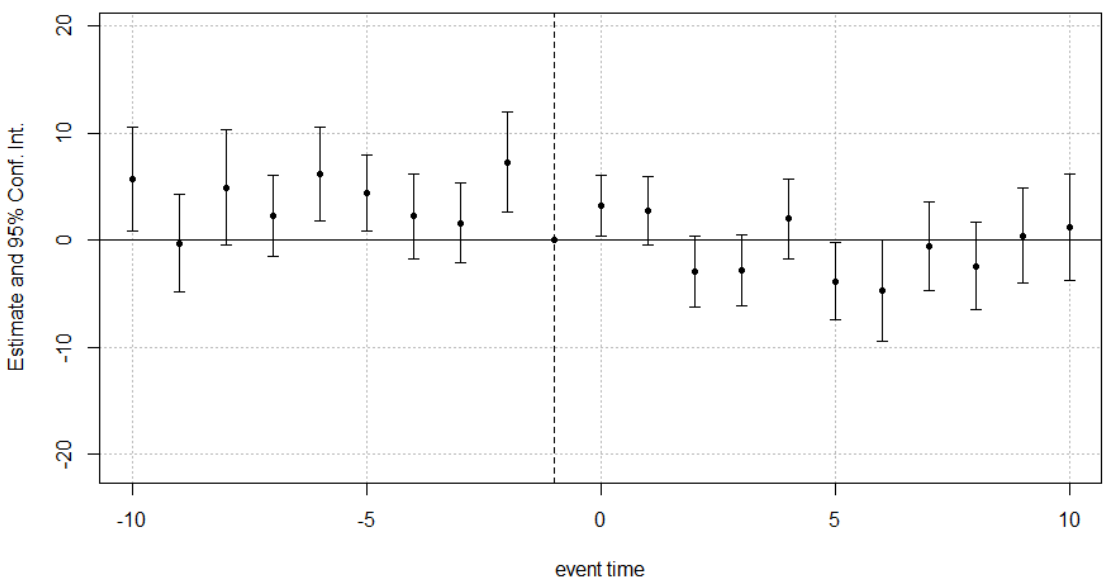
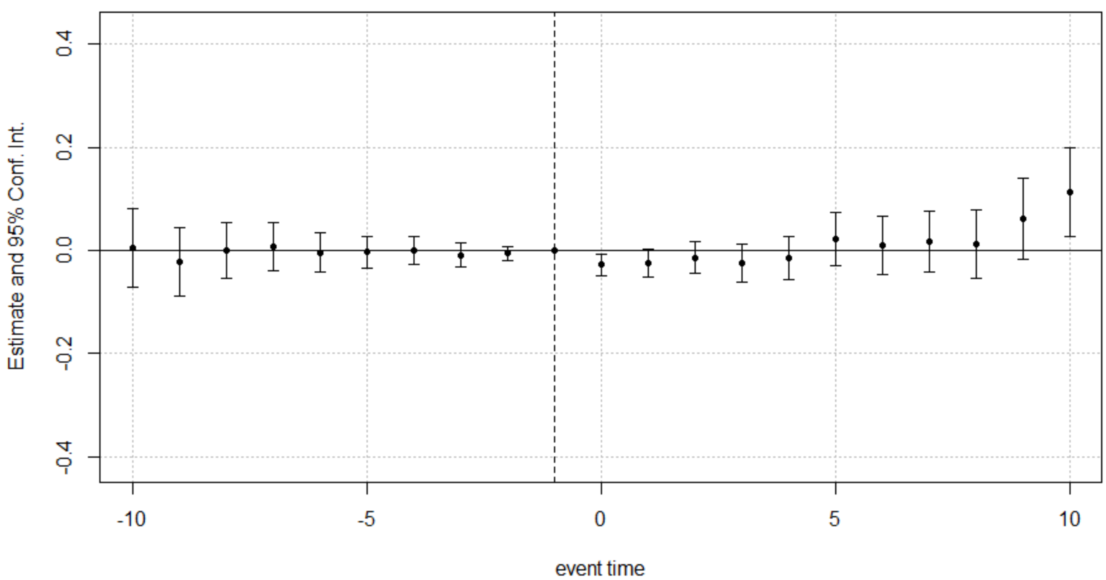
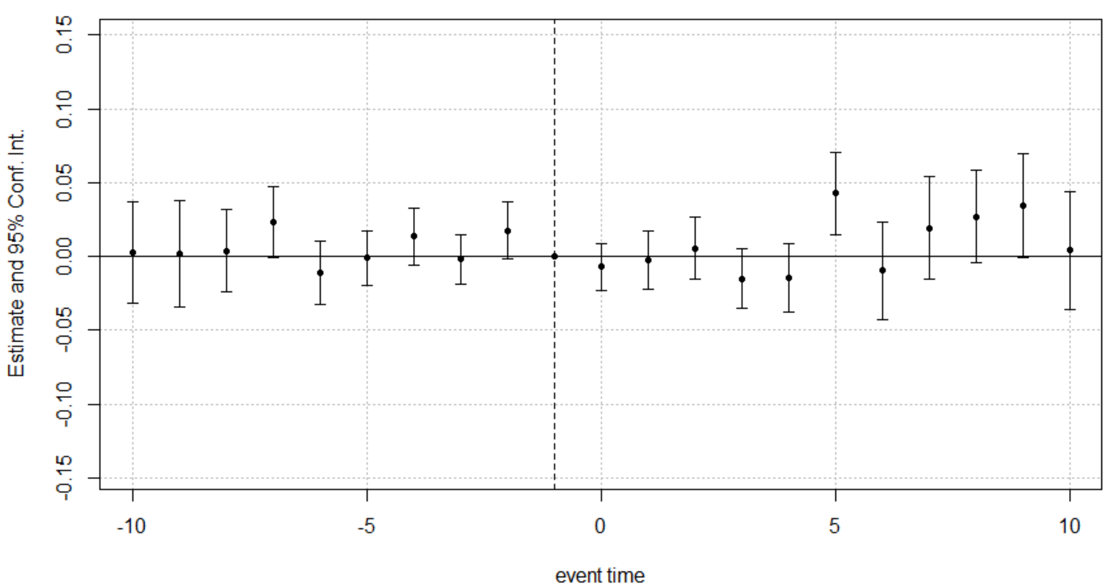


Figure A11: Effect of AD on unit value in non-target markets, SA (2021)

(a) Log unit value



(b) Growth in unit value



Appendix B Supplementary Results for "Economic Determinants of EIA Formation"

Table B1: Gravity of migration, no interactions

Dependent Variables:	log(mig_rate)			mig_rate		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	PPML	PPML	PPML
logdist	-1.343***	-1.339***		-1.290***	-1.321***	
	(0.0190)	(0.0420)		(0.0522)	(0.1224)	
nonrecip_pta	0.1415***	0.1673**	0.0220	-0.1792	-0.1183	0.0272
	(0.0477)	(0.0716)	(0.0188)	(0.1100)	(0.1763)	(0.0314)
pta	0.5035***	0.5104***	0.0824***	0.8428***	0.7715***	0.1650***
	(0.0440)	(0.0667)	(0.0203)	(0.1238)	(0.1874)	(0.0454)
fta	0.5271***	0.5124***	0.0283	1.579***	1.141***	-0.0095
	(0.0483)	(0.0692)	(0.0193)	(0.1480)	(0.2383)	(0.0391)
custun	0.8772***	0.8741***	0.3133***	1.545***	1.215***	0.0741
	(0.1068)	(0.1667)	(0.0504)	(0.3075)	(0.3624)	(0.0865)
commkt	0.3141***	0.5167***	0.3705***	0.5846*	0.7076	0.4288***
	(0.0801)	(0.1013)	(0.0426)	(0.3390)	(0.5028)	(0.0814)
econun	0.5585***	0.7518***	0.2076***	1.007***	0.9499***	0.2602***
	(0.0920)	(0.1292)	(0.0407)	(0.2177)	(0.3495)	(0.0819)
exporter-year FE	Yes		Yes	Yes		Yes
importer-year FE	Yes		Yes	Yes		Yes
exporter FE		Yes			Yes	
importer FE		Yes			Yes	
pair FE			Yes			Yes
Fit statistics						
Observations	53,688	53,688	54,071	195,894	195,906	58,083
Squared Correlation	0.71153	0.69248	0.98111	0.74724	0.67149	0.99567

Robust standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

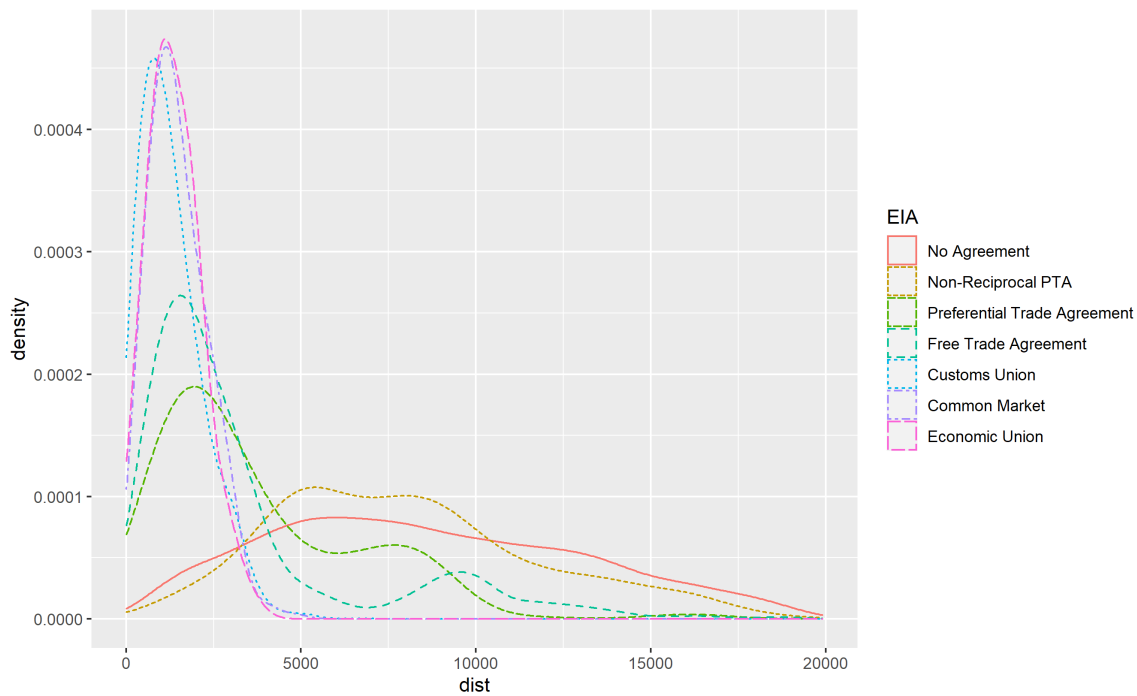
Table B2: Probability of EIA by type: importer and exporter fixed effects

Dependent:	nrpta	pta	fta	customs	common	union
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): Importer and Exporter Fixed Effects						
logdist	0.0334***	-0.0467***	-0.0710***	-0.0229***	-0.0235***	-0.0103***
	(0.0004)	(0.0005)	(0.0006)	(0.0004)	(0.0003)	(0.0002)
contig	0.0048***	0.0791***	0.0454***	0.0274***	-0.0069***	0.0525***
	(0.0018)	(0.0036)	(0.0038)	(0.0024)	(0.0020)	(0.0024)
sumlogGDP	0.0007***	-0.0031***	0.0122***	0.0009***	0.0049***	0.0040***
	(0.0002)	(0.0003)	(0.0002)	(0.0001)	(0.0001)	(0.0001)
sumlogGDPcap	0.0257***	0.0253***	0.0033***	0.0001	-0.0024***	-0.0032***
	(0.0003)	(0.0004)	(0.0003)	(0.0001)	(0.0001)	(0.0001)
difflogGDP	0.0175***	-0.0001	-0.0078***	-0.0018***	-0.0029***	-0.0021***
	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
difflogGDPcap	0.0478***	0.0079***	-0.0032***	-0.0006***	-0.0119***	-0.0071***
	(0.0003)	(0.0002)	(0.0002)	(0.0001)	(0.0002)	(0.0001)
Observations	804,374	804,374	804,374	804,374	804,374	804,374
R^2	0.6055	0.1233	0.1676	0.1038	0.1737	0.1071
Within R^2	0.1022	0.0539	0.0894	0.0500	0.0576	0.0436
Panel (b): Importer-Year and Exporter-Year Fixed Effects						
logdist	0.0340***	-0.0469***	-0.0710***	-0.0230***	-0.0235***	-0.0104***
	(0.0004)	(0.0005)	(0.0005)	(0.0004)	(0.0002)	(0.0002)
contig	0.0053***	0.0796***	0.0456***	0.0271***	-0.0071***	0.0527***
	(0.0019)	(0.0036)	(0.0037)	(0.0023)	(0.0019)	(0.0023)
sumlogGDPcap	0.0407***	0.0338***	0.0001	-0.0020***	-0.0190***	-0.0098***
	(0.0004)	(0.0004)	(0.0003)	(0.0001)	(0.0002)	(0.0002)
difflogGDP	0.0168***	0.0003**	-0.0078***	-0.0019***	-0.0029***	-0.0018***
	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
difflogGDPcap	0.0546***	0.0126***	-0.0049***	-0.0016***	-0.0198***	-0.0102***
	(0.0003)	(0.0002)	(0.0003)	(0.0001)	(0.0002)	(0.0002)
Observations	804,374	804,374	804,374	804,374	804,374	804,374
R^2	0.6903	0.1853	0.2188	0.1264	0.2530	0.1597
Within R^2	0.1209	0.0534	0.0870	0.0517	0.0710	0.0474

Robust standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure B1: Summary statistics: kernel density plots

(a) Distance by EIA type



(b) Contiguity by EIA type

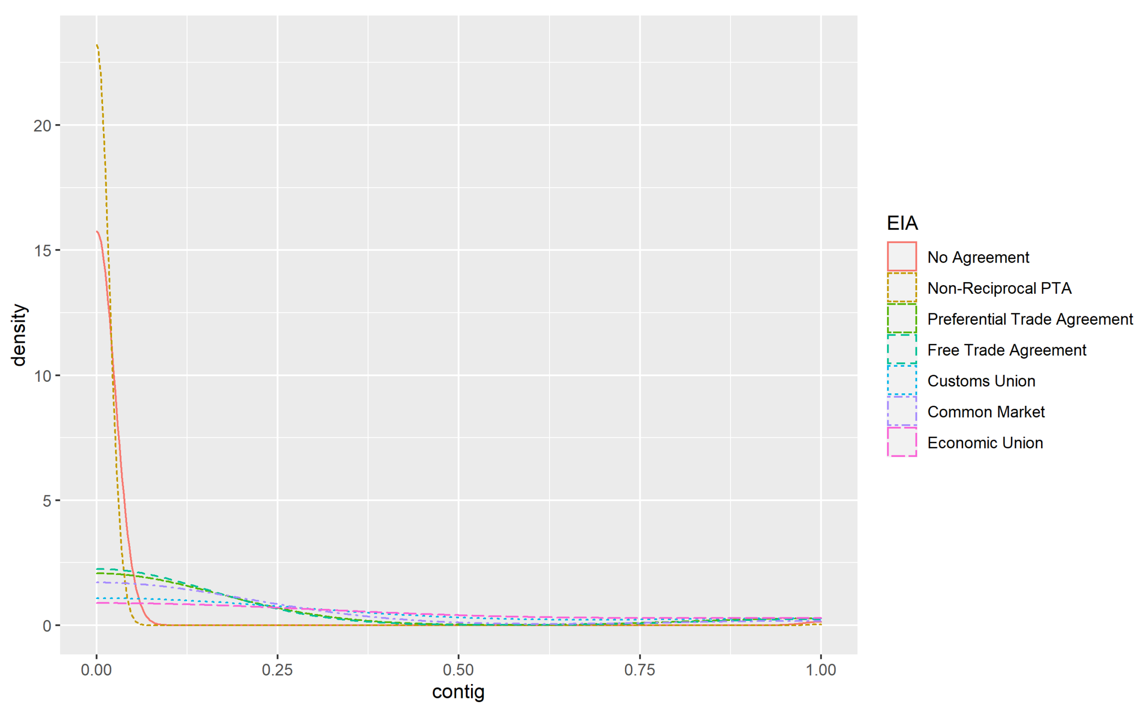
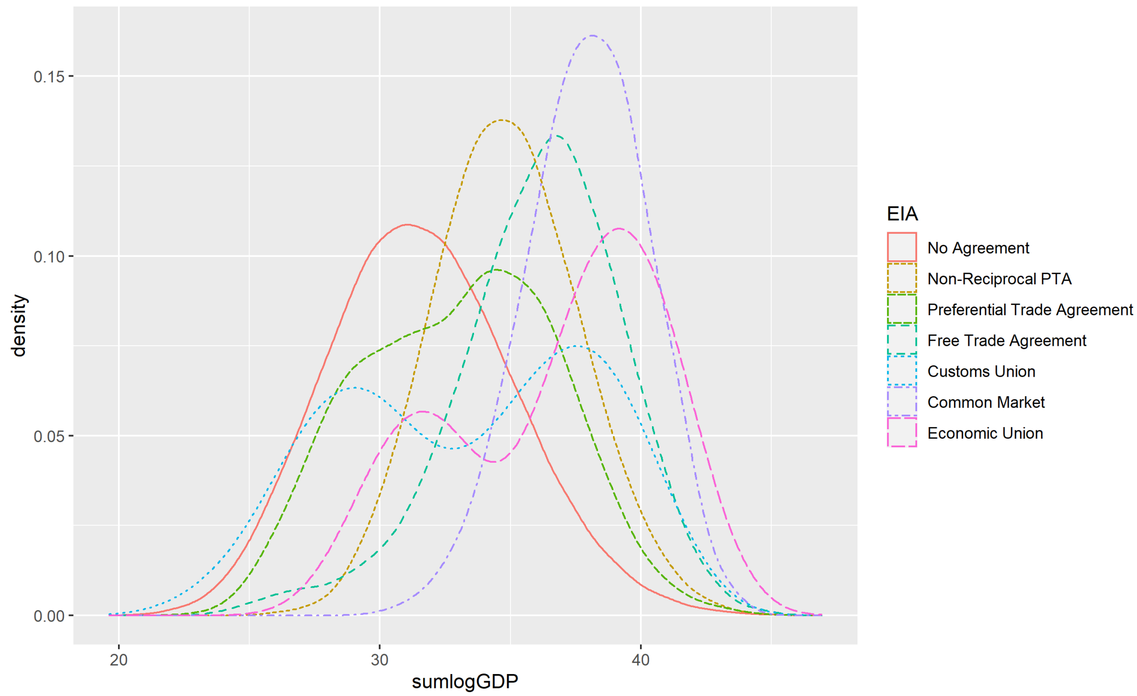


Figure B1: Summary statistics: kernel density plots (cont.)

(c) Sum of importer and exporter log GDP by EIA type



(d) Absolute difference between importer and exporter log GDP by EIA type

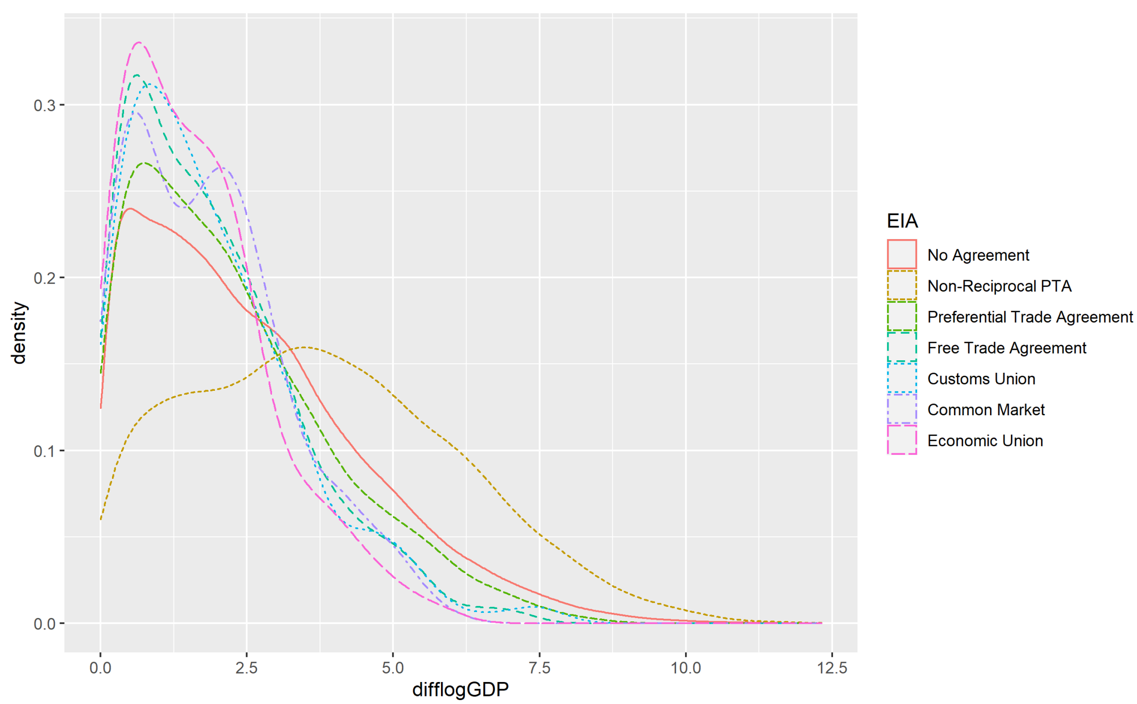
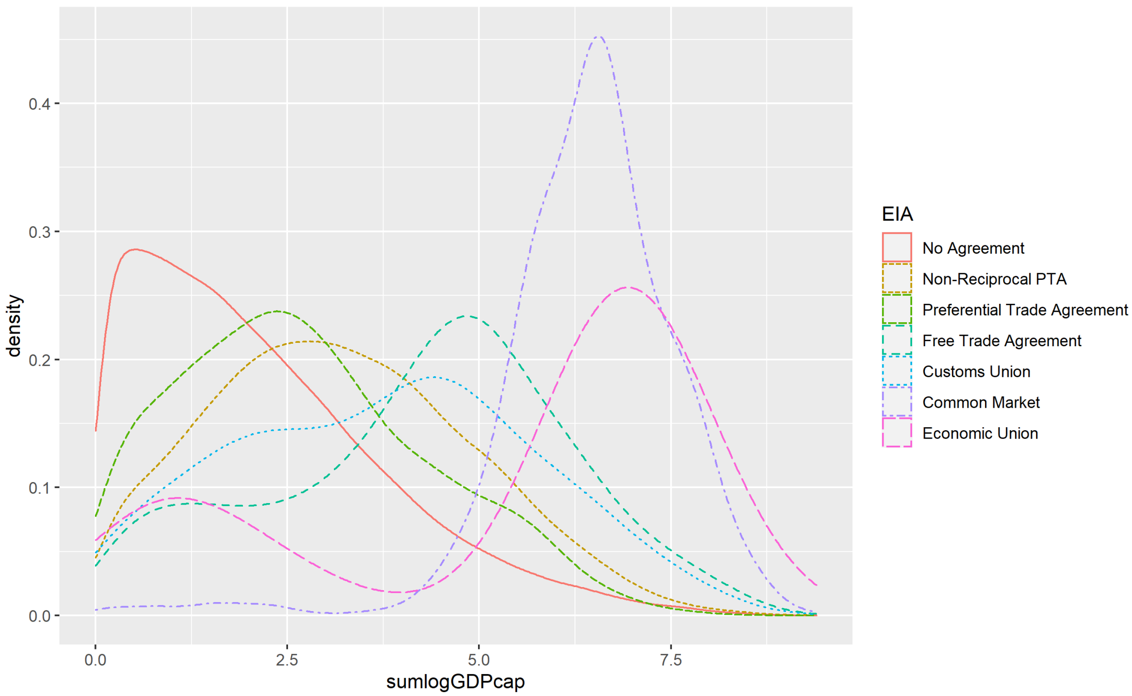
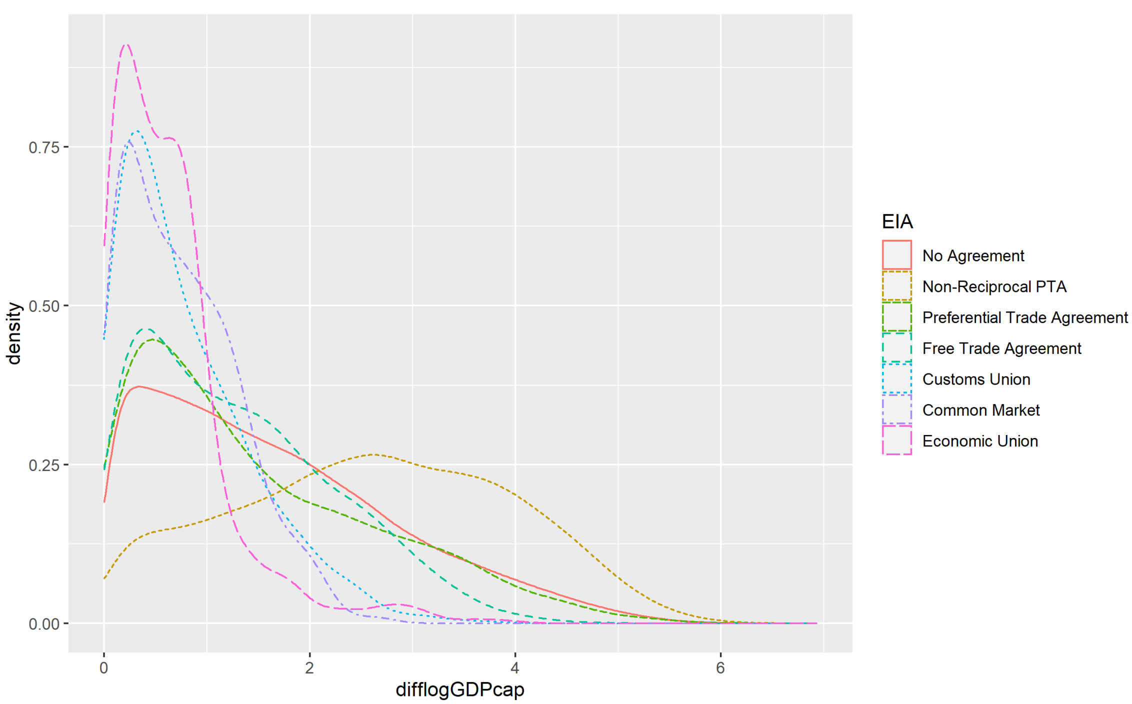


Figure B1: Summary statistics: kernel density plots (cont.)

(e) Sum of importer and exporter log GDP per capita by EIA type



(f) Absolute difference between importer and exporter log GDP per capita by EIA type



Bibliography

Abel, G. J., & Cohen, J. E. (2019). Bilateral international migration flow estimates for 200 countries. *Scientific Data*, *6*(1), 82.

Aggarwal, A. (2004). Macro economic determinants of antidumping: A comparative analysis of developed and developing countries. *World Development*, *32*(6), 1043–1057.

Albornoz, F., Brambilla, I., & Ornelas, E. (2021). *Firm export responses to tariff hikes*. Centre for Economic Performance.

Alfaro-Urena, A., Castro-Vincenzi, J., Fanelli, S., & Morales, E. (2023). Firm export dynamics in interdependent markets.

Amiti, M., Redding, S. J., & Weinstein, D. E. (2019). The impact of the 2018 tariffs on prices and welfare. *Journal of Economic Perspectives*, *33*(4), 187–210.

Amiti, M., Redding, S. J., & Weinstein, D. E. (2020). Who's paying for the US tariffs? a longer-term perspective. *AEA Papers and Proceedings*, *110*, 541–546.

Anderson, J. E., Larch, M., & Yotov, Y. V. (2018). GEPPML: General equilibrium analysis with PPML. *The World Economy*, *41*(10), 2750–2782.

Anderson, J. E., & van Wincoop, E. (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review*, *93*(1), 170–192.

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

Antràs, P., & Staiger, R. W. (2012). Offshoring and the role of trade agreements. *American Economic Review*, *102*(7), 3140–3183.

Arvis, J.-F., & Shepherd, B. (2013). The poisson quasi-maximum likelihood estimator: A solution to the 'adding up' problem in gravity models. *Applied Economics Letters*, *20*(6), 515–519.

Bagwell, K., & Staiger, R. W. (1990). A theory of managed trade. *The American Economic Review*, *80*(4), 779–795.

Bagwell, K., & Staiger, R. W. (2011). What do trade negotiators negotiate about? empirical evidence from the world trade organization. *American Economic Review*, *101*(4), 1238–1273.

Baier, S. L., & Bergstrand, J. H. (2004). Economic determinants of free trade agreements. *Journal of International Economics*, *64*(1), 29–63.

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

Baier, S. L., Bergstrand, J. H., & Clance, M. W. (2018). Heterogeneous effects of economic integration agreements. *Journal of Development Economics*, *135*, 587–608.

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

Baier, S. L., & Regmi, N. R. (2023). Using machine learning to capture heterogeneity in trade agreements. *Open Economies Review*, *34*(4), 863–894.

Baier, S. L., & Standaert, S. (2024). Gravity, globalization and time-varying heterogeneity. *European Economic Review*, *163*, 104679.

Baier, S. L., Yotov, Y. V., & Zylkin, T. (2019). On the widely differing effects of free trade agreements: Lessons from twenty years of trade integration. *Journal of International Economics*, *116*, 206–226.

Bao, X., Blonigen, B. A., & Yu, Z. (2021). *Cross-product and cross-market adjustments within multiproduct firms: Evidence from antidumping actions* (No. 29521). National Bureau of Economic Research.

Baylis, K., & Perloff, J. M. (2010). Trade diversion from tomato suspension agreements. *Canadian Journal of Economics*, *43*(1), 127–151.

Besedeš, T., & Prusa, T. J. (2017). The hazardous effects of antidumping. *Economic Inquiry*, *55*(1), 9–30.

Blanchard, E. J., Bown, C. P., & Johnson, R. C. (2016). *Global supply chains and trade policy* (Working Paper No. 21883). National Bureau of Economic Research.

Blonigen, B. A., & Bown, C. P. (2003). Antidumping and retaliation threats. *Journal of International Economics*, *60*(2), 249–273.

Blonigen, B. A., & Haynes, S. E. (2002). Antidumping investigations and the pass-through of antidumping duties and exchange rates. *American Economic Review*, *92*(4), 1044–1061.

Blonigen, B. A., & Park, J.-H. (2004). Dynamic pricing in the presence of antidumping policy: Theory and evidence. *American Economic Review*, *94*(1), 24.

Blonigen, B. A., & Prusa, T. J. (2016). Chapter 3 – dumping and antidumping duties. In K. Bagwell & R. W. Staiger (Eds.), *Handbook of commerical policy* (pp. 107–159, Vol. 1). North-Holland.

Borchert, I., Larch, M., Shikher, S., & Yotov, Y. V. (2021). The international trade and production database for estimation (ITPD-e). *International Economics*, *166*, 140–166.

Bown, C. P. (2011). Taking stock of antidumping, safeguards and countervailing duties, 1990–2009. *The World Economy*, *34*(12), 1955–1998.

Bown, C. P., Cieszkowsky, M., Erbahar, A., & Signoret, J. (2020). *Temporary trade barriers database*. Washington DC: World Bank.

Bown, C. P., Conconi, P., Erbahar, A., & Trimarchi, L. (2023). Politically motivated trade protection.

Bown, C. P., & Crowley, M. A. (2007). Trade deflection and trade depression. *Journal of International Economics*, *72*(1), 176–201.

Bown, C. P., & Crowley, M. A. (2010). China's export growth and the China safeguard: Threats to the world trading system? *Canadian Journal of Economics*, *43*(4), 1353–1388.

Bown, C. P., & Crowley, M. A. (2013a). Import protection, business cycles, and exchange rates: Evidence from the great recession. *Journal of International Economics*, *90*(1), 50–64.

Bown, C. P., & Crowley, M. A. (2013b). Self-enforcing trade agreements: Evidence from time-varying trade policy. *American Economic Review*, *103*(2), 1071–1090.

Bown, C. P., & Crowley, M. A. (2016). Chapter 1 - the empirical landscape of trade policy. In K. Bagwell & R. W. Staiger (Eds.), *Handbook of commercial policy* (pp. 3–108, Vol. 1). North-Holland.

Bown, C. P., & Tovar, P. (2011). Trade liberalization, antidumping, and safeguards: Evidence from india's tariff reform. *Journal of Development Economics*, *96*(1), 115–125.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32.

Breinlich, H., Leromain, E., Novy, D., & Sampson, T. (2022). Import liberalization as export destruction? evidence from the united states. *SSRN Electronic Journal*.

Broda, C., Limao, N., & Weinstein, D. E. (2008). Optimal tariffs and market power: The evidence. *American Economic Review*, *98*(5), 2032–2065.

Callaway, B., Goodman-Bacon, A., & Sant'Anna, P. H. C. (2021). Difference-in-differences with a continuous treatment.

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230.

Chandra, P. (2016). Impact of temporary trade barriers: Evidence from China. *China Economic Review*, *38*, 24–48.

Conconi, P., Facchini, G., & Zanardi, M. (2014). Policymakers' horizon and trade reforms: The protectionist effect of elections. *Journal of International Economics*, *94*(1), 102–118.

Conte, M., Cotterlaz, P., & Mayer, T. (2022). *The CEPII gravity database*. CEPII.

Crowley, M. A. (2011). Cyclical dumping and US antidumping protection: 1980-2001. *SSRN Electronic Journal*.

Davis, S. J., Haltiwanger, J. C., & Schuh, S. (1998). *Job creation and destruction*. The MIT Press.

Dekle, R., Eaton, J., & Kortum, S. (2007). Unbalanced trade. *American Economic Review*, *97*(2), 351–355.

Durling, J. P., & Prusa, T. J. (2006). The trade effects associated with an antidumping epidemic: The hot-rolled steel market, 1996–2001. *European Journal of Political Economy*, *22*(3), 675–695.

Egger, P., & Larch, M. (2008). Interdependent preferential trade agreement memberships: An empirical analysis. *Journal of international Economics*, *76*(2), 384–399.

Egger, P., Larch, M., Staub, K. E., & Winkelmann, R. (2011). The trade effects of endogenous preferential trade agreements. *American Economic Journal: Economic Policy*, *3*(3), 113–43.

Egger, P., & Nelson, D. (2011). How bad is antidumping? evidence from panel data. *The Review of Economics and Statistics*, *4*(93), 1374–1390.

Fajgelbaum, P., Goldberg, P. K., Kennedy, P. J., Khandelwal, A., & Taglioni, D. (2023). The US-china trade war and global reallocations.

Fally, T. (2015). Structural gravity and fixed effects. *Journal of International Economics*, *97*(1), 76–85.

Feinberg, R. M. (1989). Exchange rates and "unfair trade". *The Review of Economics and Statistics*, *71*(4), 704.

Feinberg, R. M., & Hirsch, B. T. (1989). Industry rent seeking and the filing of 'unfair trade' complaints. *International Journal of Industrial Organization*, *7*(3), 325–340.

Felbermayr, G., & Sandkamp, A. (2020). The trade effects of anti-dumping duties: Firm-level evidence from china. *European Economic Review*, *122*, 103367.

Furceri, D., Ostry, J. D., Papageorgiou, C., & Wibaux, P. (2021). Retaliatory temporary trade barriers: New facts and patterns. *Journal of Policy Modeling*, *43*(4), 873–891.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277.

Grossman, G. M. (2016, January 1). Chapter 7 - the purpose of trade agreements. In K. Bagwell & R. W. Staiger (Eds.), *Handbook of commercial policy* (pp. 379–434, Vol. 1). North-Holland.

Hillberry, R., & McCalman, P. (2016). Import dynamics and demands for protection. *Canadian Journal of Economics*, (3), 29.

Hinz, J. (2023). The ties that bind: Geopolitical motivations for economic integration. *Review of World Economics*, *159*(1), 51–100.

Hoai, N. T., Toan, N. T., & Van, P. H. (2017). Trade diversion as firm adjustment to trade policy: Evidence from EU anti-dumping duties on Vietnamese footwear. *The World Economy*, *40*(6), 1128–1154.

Jabbour, L., Tao, Z., Vanino, E., & Zhang, Y. (2019). The good, the bad and the ugly: Chinese imports, european union anti-dumping measures and firm performance. *Journal of International Economics*, *117*, 1–20.

Knetter, M. M., & Prusa, T. J. (2003). Macroeconomic factors and antidumping filings: Evidence from four countries. *Journal of International Economics*, *61*(1), 1–17.

Lasagni, A. (2000). Does country-targeted anti-dumping policy by the EU create trade diversion? *Journal of World Trade*, *34*(4).

Limão, N. (2016, January 1). Chapter 6 - preferential trade agreements. In K. Bagwell & R. W. Staiger (Eds.), *Handbook of commercial policy* (pp. 279–367, Vol. 1). North-Holland.

Lu, Y., Tao, Z., & Zhang, Y. (2013). How do exporters respond to antidumping investigations? *Journal of International Economics*, *91*(2), 290–300.

Maggi, G., & Ossa, R. (2021). The political economy of deep integration. *Annual Review of Economics*, *13*, 19–38.

Mattoo, A., Mulabdic, A., & Ruta, M. (2022). Trade creation and trade diversion in deep agreements. *Canadian Journal of Economics/Revue canadienne d'économique*, *55*(3), 1598–1637.

McLaren, J. (2016, January 1). Chapter 2 - the political economy of commercial policy. In K. Bagwell & R. W. Staiger (Eds.), *Handbook of commercial policy* (pp. 109–159, Vol. 1). North-Holland.

Monte, F., Redding, S. J., & Rossi-Hansberg, E. (2018). Commuting, migration, and local employment elasticities. *American Economic Review*, *108*(12), 3855–3890.

Morales, E., Sheu, G., & Zahler, A. (2019). Extended gravity. *The Review of Economic Studies*, *86*(6), 2668–2712.

Murphy, K. P. (2012). *Machine learning: A probabilistic perspective*. MIT Press.

Orefice, G. (2015). International migration and trade agreements: The new role of PTAs. *Canadian Journal of Economics/Revue canadienne d'économique*, *48*(1), 310–334.

Osnago, A., Rocha, N., & Ruta, M. (2019). Deep trade agreements and vertical FDI: The devil is in the details. *Canadian Journal of Economics/Revue canadienne d'économique*, *52*(4), 1558–1599.

Pierce, J. R. (2011). Plant-level responses to antidumping duties: Evidence from U.S. manufacturers. *Journal of International Economics*, *85*(2), 222–233.

Prusa, T. J. (1997). The trade effects of U.S. antidumping actions. In R. C. Feenstra (Ed.), *The effect of u.s. trade and protection* (pp. 191–214). Chicago: University of Chicago Press.

Prusa, T. J. (2001). On the spread and impact of anti-dumping. *Canadian Journal of Economics*, *34*(3), 591–611.

Prusa, T. J., & Skeath, S. (2002). The economic and strategic motives for antidumping filings. *Weltwirtschaftliches Archiv*, *138*(3), 389–413.

Prusa, T. J., Teh, R., & Zhu, M. (2022). PTAs and the incidence of antidumping disputes. *Journal of International Economics*, *138*, 103630.

Reynolds, K. M. (2006). Subsidizing rent-seeking: Antidumping protection and the byrd amendment. *Journal of International Economics*, *70*(2), 490–502.

Sandkamp, A. (2020). The trade effects of antidumping duties: Evidence from the 2004 EU enlargement. *Journal of International Economics*, *123*, 103307.

Silva, J. M. C. S., & Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, *88*(4), 641–658.

Staiger, R. W., Wolak, F. A., Litan, R. E., Katz, M. L., & Waverman, L. (1994). Measuring industry-specific protection: Antidumping in the united states. *Brookings Papers on Economic Activity: Microeconomics*, *1994*, 51–118.

Steinbach, S., & Khederlarian, A. (2022). Pre-trends and trade effects of temporary trade barriers. *SSRN Electronic Journal*.

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175–199.

USITC. (2015). Antidumping and countervailing duty handbook.

Vandenbussche, H., & Zanardi, M. (2010). The chilling trade effects of antidumping proliferation. *European Economic Review*, *54*(6), 760–777.

Viner, J. (1950). Full employment at whatever cost. *The Quarterly Journal of Economics*, *64*(3), 385–407.

Yotov, Y. V., Piermartini, R., Monteiro, J.-A., & Larch, M. (2016). *An advanced guide to trade policy analysis: The structural gravity model*. WTO iLibrary.