Empirical Essays on International Trade

Myriam Alejandra Gomez Cardenas

The University of Sheffield

Department of Economics

Supervisors:

Dr Antonio Navas

Professor Karl Taylor

A thesis submitted in partial fulfilment of
the requirements for the degree of
Doctor of Philosophy

September 2022

Acknowledgements

This thesis is dedicated to the loving memory of my Mom and Grandpa, who always believed in me and unconditionally supported me at all times. It is also dedicated to my Dad, who always encouraged me to dream big, and to my whole family for their support during this journey.

I want to express my greatest gratitude to Professor Karl Taylor for his strong commitment, valuable insights, and meticulous feedback. I want to equally express my gratitude to Dr Antonio Navas for his guidance, insightful discussions, and career advice.

Aside from my supervisors, I want to thank my Viva examiners, Dr Enrico Vanino and Dr Francesco Serti, for their valuable comments and suggestions. I would also like to thank Dr Juan Paez-Farrell and Professor Alberto Montagnoli for their feedback during the Confirmation Review, and Dr Kanchana Nadarajah and Professor Arne Risa Hole for their insights.

I want to acknowledge the informal mentoring given by Dr Justyna Drobnik-Rogers, Dr Tuong Vu, Emilia Skulska, and Raymundo Rodriguez and thank them for keeping me focused during the critical stages of the thesis. To my aunt, Rosa Cardenas, thank you for helping me proofread the thesis. I would also like to express my appreciation to Pille Kustala for mentoring me all these years.

I want to thank my friends and peers from the PhD journey who were always there during this rollercoaster experience. Special thanks go to Andira Hernandez, Belle and Basile Marquier, Bingxue Wang, Camille Allard, Clara Franco, Cyriac Kodath, Emily Barker, Itzelle Medina, Mahvish Faran, Miroslava Quiroga, Mustabsar Awais, and Soizic van Geersdaele.

This extends to my friends back in Mexico who always supported and encouraged me throughout the entire process: Alejandra Felix, Francisco Pacheco, Kenia Perez, Luis Sainz, Raymundo Rodriguez, and Susana Alanis.

I also want to express my gratitude to my former bosses for encouraging and supporting me to pursue a PhD: Dr Jose Antonio Gonzalez, Dr Qi Zhang, Dr Thuy Duong Tran, Ana Laura Gomez, Gerardo Vera, and Romina Burillo. Finally, I gratefully acknowledge the funding received from the Mexican government through the Consejo Nacional de Ciencia y Tecnología (CONACYT).

Abstract

This thesis aims to contribute to the empirical trade literature on global value chains from the perspective of an emerging economy: Mexico. This country represents an interesting case study from an international point of view as the country is strongly integrated with developed economies through bilateral trade and FDI. Therefore, this thesis aims to identify different trade mechanisms, such as FDI, trade processing, and technology embedded in imports. First, we investigate the effects of FDI and bilateral trade on business cycle comovements. Our findings confirm that bilateral trade strongly impacts the transmission of business cycles. Moreover, these results also reveal the importance of FDI on business cycle comovements. Second, we examine the relationship between new imported varieties and new exported varieties. Our results show that importing new varieties from a source country constitutes an important determinant for exporting new varieties to that same country. Third, we analyze the impact of new imported intermediate inputs and new imported capital goods on exports of new varieties. Our findings confirm that new imported intermediate inputs strongly impact exports of new varieties. Furthermore, these results also shed light on the importance of new imported capital goods on exports of new varieties.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

For the past decades, we have seen a rising phenomenon called global value chains (GVCs), where firms fragment their production and offshore production stages to other countries. The emergence of these GVCs can be explained by a series of conditions, such as trade liberalization, lower investment barriers, lower transportation costs, and technological advances in telecommunications. An interesting feature is that not only developed economies are integrated into GVCs, but also emerging and developing economies.

According to the OECD, Mexico has a significant participation in GVCs through the export activity driven by processing trade of imported intermediate inputs (OECD 2013). Therefore, Mexico represents an interesting case study from an international perspective. Furthermore, the country constitutes an emerging economy that is strongly integrated with developed economies through bilateral trade and vertical Foreign Direct Investment (FDI).[1]

Mexico's main trading partners and FDI investors constitute major players in the trade arena; these partner countries include the United States, Canada, China, and Japan. Moreover, the country offers two main competitive advantages to investment partners. First, Mexico offers low trade costs; for example, low transportation costs due to geographical proximity to the United States.[2] The country also possesses a vast network of free trade agreements and preferential trade agreements, which translates into lower, or even, zero tariffs for imported goods including intermediate inputs. Second, relative factor endowment differences; in other words, Mexico offers low wages, which is especially attractive to multinational firms during the production stages requiring labor-intensive activities (Bergin et al. 2009). Thus, this country represents a strategic destination to allocate investments, especially in the form of vertical FDI.

[1] Vertical FDI can be defined as the allocation of part of the production chain to an affiliate located abroad. This concept differs from horizontal FDI, which is the case where the affiliate located abroad replicates the production process aiming to target that foreign market (Helpman 1984, Helpman et al. 2004).

[2] Clark & Van Wincoop (2001) explain that distance represents an important trade barrier as it translates into transportation and communication costs. In the case of Mexico, a significant amount of manufacturing plants specializing in assembly activities are located in the border with the United States (Bergin et al. 2009). In particular, Hanson (1997) identifies the six largest U.S.-Mexican border-city pairs, which are San Diego-Tijuana, Imperial County-Mexicali, El Paso-Ciudad Juarez, Laredo-Nuevo Laredo, McAllen-Reynosa, and Brownsville-Matamoros.

Trade and FDI policies are often integrated, hence the importance of studying both. In terms of Foreign Direct Investment, the Mexican Government grants fiscal stimuli to multinational firms settling in the country through the "Maquiladora Program". As Burstein et al. (2008) documented, this vertical form of FDI inflows can be traced since the mid-1960's when the "Maquiladora Program" was created by the official authorities to tackle high unemployment rates in the northern region of the country.

As Bergin et al. (2009) explain, the Maquiladora Program consisted of allocating part of the production stages of foreign-owned firms (e.g., assembly of final goods) to the northern region of Mexico. Burstein et al. (2008) mention that one of the most important incentives that the Mexican government offers to foreign-owned firms is a tariff exemption for imported inputs and equipment, conditional on their re-export, after a transformation process takes place in the production chain. Today, this program is still in force under the name of "Manufacturing, Maquila and Export Service Industry Program", also known as the IMMEX Program.

Mexican trade policy is determined endogenously by policymakers, as suggested by Kandilov & Leblebicioğlu (2012). In fact, Mexico has gone through a trade liberalization process that intensively took place in the mid-1980s. In this decade, Mexico became a member of the General Agreement on Tariffs and Trade (GATT); this represented a steppingstone in the trade liberalization process. The second most important milestone was the enforcement of the North American Free Trade Agreement (NAFTA) in 1994. The NAFTA agreement set the ground rules for Mexico's upcoming free trade agreements with other countries. Today, Mexico possesses a vast network of 15 free trade agreements and six preferential trade agreements in the Latin-American region.[3]

This trade liberalization process has also led to significant changes in the productive structure of the country. De Hoyos & Iacovone (2013) identify four main channels through which trade reforms have impacted productivity in Mexico; these channels are competition, intermediate inputs, exports, and FDI. The authors also explain that the NAFTA agreement led to a productivity increase of Mexican plants through two channels: an increase in import competition and a wider access to imported intermediate inputs.

As a result of these trade policies, the country shifted from an import substitution regime to a liberalized trade regime. According to the World Integrated Trade Solution (WITS), Mexico reported an export value of 56.0 billion dollars and an import value of 48.8 billion dollars in 1993; this is the year before the NAFTA agreement entered into force. In comparison, the country reported an export value of 358.6 billion dollars and an import value of 402.3 billion dollars in 2016. Thus, the country has experienced a dramatical growth in trade since these trade liberalization policies entered into force.

[3]According to the Foreign Trade Information System of the Organization of American States, Mexico has the following free trade agreements in force: North American Free Trade Agreement (1994), Colombia (1995), European Union (2000), Chile (1999), Israel (2001), European Free Trade Association (2001), Uruguay (2004), Japan (2005), Bolivia (2010), Peru (2012), Central America (2013), Panama (2015), Pacific Alliance (2016), Comprehensive and Progressive Agreement for Trans-Pacific Partnership (2018), and the United States-Mexico-Canada Agreement (2020).

Also, Mexico has the following preferential trade agreements in force: Ecuador (1993), Paraguay (1993), Brazil (2002), Mercosur auto sector agreement (2003), Mercosur (2006), and Argentina (2007).

1.2 Aims and Motivation

This thesis aims to contribute to the existing empirical trade literature on global value chains from the perspective of an emerging economy: Mexico. Thus, we aim to identify different mechanisms, such as FDI, trade processing, and technology embedded in imports. Furthermore, this thesis allows us to obtain a comprehensive overview of these mechanisms within Mexico at the state level but also at the country level with other partner countries.

Through three empirical chapters, we aim to address the following research questions related to key trade policy components. In the first empirical chapter, we are interested in exploring the effects of FDI on the transmission of business cycles. Then, we move to the second empirical chapter aiming to analyze the impact of importing new varieties on exports of new varieties. We also examine whether importing new varieties from a source country has an impact on exports of new varieties to that same country. Finally, in the third empirical chapter, we are interested in studying the effects of importing new intermediate inputs and new capital goods on exports of new varieties.

The first empirical chapter is motivated by the limited literature on the impact of FDI on business cycle comovements. To the best of our knowledge, only Hsu et al. (2011) and Jansen & Stokman (2014) study the role of FDI on business cycle synchronizations from the perspective of developed countries. However, this relationship has not been previously explored from the perspective of developing countries. Another interesting motivation of this chapter is to use the recent Hamilton regression filter to detrend GDP time series instead of using the traditional Hodrick-Prescott and Baxter-King filters common in the business cycle comovements literature.

The second empirical chapter is motivated by the increasing literature on the effects of imported varieties on exports (Aristei et al. 2013, Bas & Strauss-Kahn 2014, Castellani & Fassio 2019, Feng et al. 2016, Lo Turco & Maggioni 2013, Navas et al. 2020, Xu & Mao 2018). Nonetheless, two main features have not been explored in-depth: trade complementarities and new traded varieties. First, it remains unclear whether importing products from a specific country has an impact on exports to that same country. Second, the approach of new varieties is quite novel (Castellani & Fassio 2019), and it remains unclear what is the impact of importing new varieties on exports of new varieties from the perspective of an emerging economy.

The third empirical chapter aims to refine the disaggregation level of the second empirical chapter by disentangling new imported varieties into intermediate inputs and capital goods. Moreover, this chapter is motivated by the increasing literature on the importance of imported intermediate inputs on exports (Aristei et al. 2013, Castellani & Fassio 2019, Feng et al. 2016, Lo Turco & Maggioni 2013, Navas et al. 2020). However, despite this extensive literature, the role of imported capital goods has not been widely explored, except for Damijan et al. (2014). Furthermore, the link between imported capital goods and exports focusing on new varieties remains unexplored for emerging economies.

1.3 Structure and Content

The overall structure of this thesis takes the form of five chapters comprising an introduction, three empirical chapters, and a conclusion. Chapter 1 provides an introduction to the entire thesis. The empirical chapters constitute independent studies exploring trade-related topics from the perspective of an emerging economy: Mexico. Thus, Chapter 2 examines the effects of FDI and bilateral trade on business cycle comovements. Chapter 3 studies the relationship between new imported varieties and new exported varieties focusing on trade complementarities at the country level. Chapter 4 explores the impact of new imported intermediate inputs and new imported capital goods on exports of new varieties. Finally, Chapter 5 offers an overall conclusion of the entire thesis. Now, we proceed to present a brief overview of each of the empirical chapters.

1.3.1 FDI, Trade, and Business Cycle Comovements

The empirical trade literature has identified several key determinants of business cycle comovements. Within the most popular determinants, we can find bilateral trade, productive structure, intra-industry trade, production sharing, distance, borders, and multinational firms, among others (Burstein et al. 2008, Calderón et al. 2007, Clark & Van Wincoop 2001, Di Giovanni & Levchenko 2010, Imbs 2004, Kleinert et al. 2015, Zlate 2016). However, the relevance of FDI has been somehow neglected from the literature, except for Hsu et al. (2011) and Jansen & Stokman (2014), who base their analyses on developed countries. Nonetheless, the role of FDI on business cycle comovements has not been explored for emerging economies.

The main contribution of this first empirical chapter is to examine the role of FDI and bilateral trade on the transmission of business cycles from the perspective of an emerging economy. Another contribution is that we implement a novel filtering technique in our analysis: the Hamilton regression filter. This filter offers important advantages over the traditional Hodrick-Prescott and Baxter-King filters commonly used in the business cycle comovements literature.

Our estimation sample consists of paired observations between 47 major partner countries and all 32 Mexican states using sub-periods to create a time element for the variables included in our model. Our results confirm that bilateral trade has a strong and positive effect on the transmission of business cycles from partner countries to Mexican states. However, what stands out from our analysis is that Foreign Direct Investment also has a positive and significant effect on business cycle comovements for an emerging economy.

From a policy perspective, it may be worth promoting FDI inflows from emerging economies to southern states in Mexico to boost these states' economic performance. On the other hand, it is worth maintaining and reinforcing FDI inflows from traditional partner countries to northern states in Mexico.

1.3.2 New Traded Varieties and Source Countries: Evidence of Trade Complementarities

A strand of the trade literature focuses on the relationship between imported varieties and exports (Aristei et al. 2013, Bas & Strauss-Kahn 2014, Castellani & Fassio 2019, Feng et al. 2016, Lo Turco & Maggioni 2013, Navas et al. 2020, Xu & Mao 2018). However, the literature leaves aside the effects of importing new products from a specific country on exports to that same country.

The main contribution of this second empirical chapter is to analyze the link between new imported varieties and exports of new varieties by revealing a trade complementarity effect at the country level. In other words, we are interested in analyzing to what extent imports of new products from a specific country lead to an increase in exports of new products to that same country. To perform this analysis, we exploit the bilateral component of our database at the country level. Furthermore, we perform this analysis from the perspective of an emerging economy using a recent period compared to the existing literature.

We start our analysis by using a decomposition exercise of the annual growth of traded varieties at the product level. This decomposition exercise aims to identify new, continuing, and withdrawn varieties. This chapter focuses on new varieties; thus, we use two criteria to define a new variety in line with Colantone & Crinò (2014). First, we consider a new variety when a product is traded (i.e., exported or imported) with a partner country for the first time. Second, we consider a new variety when a product is introduced to the Harmonized System classification. Next, we aggregate the number of new varieties at the industry level. Thus, our estimation sample consists of 74,240 new traded varieties belonging to the manufacturing sector over the time period from 2005 to 2016.

Then, we employ three empirical strategies. First, we use a fixed effects logit model to estimate the probability of exporting new varieties based on the number of new imported varieties. Next, we use a fixed effects negative binomial model to analyze the impact of new imported varieties on the number of new exported varieties. Finally, we employ linear fixed effects models to examine the impact of new imported varieties on the export share of new varieties at the extensive and intensive margins; the aim of these last models is to detect a trade complementarity effect with source countries.

Our findings suggest that importing new varieties has a positive and strong effect on the probability of exporting new varieties, on the number of new exported varieties, and on the export shares of new varieties to the same source countries. Thus, we provide empirical evidence of a degree of trade complementarity between new imports and new exports at the country level from the perspective of an emerging economy.

From a policy perspective, we find that Mexico imports new varieties from new partner countries that do not benefit from free or preferential trade agreements. Therefore, it may be worth exploring an expansion of free trade agreements, or preferential trade agreements focused on identified sectors, to other geographical regions.

1.3.3 New Imported Inputs, New Export Varieties: Capital Matters

Another strand of the literature focuses on the role of imported intermediate inputs on exports (Aristei et al. 2013, Castellani & Fassio 2019, Feng et al. 2016, Lo Turco & Maggioni 2013, Navas et al. 2020). However, to date, the role of imported capital goods has not been widely explored in the literature, except for Damijan et al. (2014). Additionally, the trade literature has not yet explored the link between imported capital goods and exports focusing on new varieties from the perspective of an emerging economy.

The main contribution of this third empirical chapter is to disentangle the effects of importing new intermediate inputs and new capital goods on exports of new varieties. Thus, we shed light on the importance of new imported capital goods on exports of new varieties from the perspective of an emerging economy. An interesting feature of our analysis is that we focus on new exported varieties, which has not been considered in the literature.

We begin our analysis by identifying new, withdrawn, and continuing products from the entire universe of manufacturing goods traded by Mexico comprising from 2003 to 2016. Consistent with the second empirical chapter and with Colantone & Crinò (2014), we focus on new varieties based on two criteria. We consider a new variety when a product is traded with a partner country for the first time or when a product is introduced to the Harmonized System classification.

Then, we disentangle these new varieties into intermediate, capital, and consumption goods by using the UN Broad Economic Categories (BEC), which is aligned to the empirical literature (Arkolakis et al. 2008, Damijan et al. 2014, Dean et al. 2011, Feng et al. 2016, Koopman et al. 2012). From these end-use categories, we focus only on mutually exclusive categories of intermediate and capital goods. Next, we aggregate these new intermediate inputs and new capital goods at the industry level. Therefore, our estimation sample is conformed by 68,727 new traded varieties belonging to the manufacturing sector over the period from 2005 to 2016.

We employ four empirical strategies to our estimation sample. First, we use a fixed effects negative binomial model to evaluate the number of new exported varieties. Next, we employ a fixed effects logit model to study the probability of exporting new varieties. Then, we use a linear fixed effects model to estimate the export value of new varieties. Finally, we use a log-first difference estimator to examine the net change in the number of new exported varieties.

Our results confirm a strong and positive relationship between new imported intermediate inputs and new exported varieties. More interestingly, our results also suggest that new imported capital goods have a positive and statistically strong effect on the probability of exporting new varieties, on the export value of new varieties, and on the net change of new exported varieties. Thus, we show that new imported capital goods play an important role on exports of new varieties from the perspective of an emerging economy.

From a policy perspective, it may be worth negotiating free or preferential trade agreements to import new intermediate inputs and capital goods. By doing this, Mexican firms could benefit from spillover effects, such as technology transfers due to a learning by importing mechanism.

Chapter 2

FDI, Trade, and Business Cycle Comovements

2.1 Introduction

Mexico constitutes an emerging economy that is strongly integrated with developed economies through trade and vertical Foreign Direct Investment (FDI). In particular, the Mexican economy is strongly intertwined with the United States through FDI. An example of this relationship is the high concentration of manufacturing plants across the U.S.-Mexican border paired cities, as identified by (Hanson 1997). Thus, the economy of these paired cities located on the Mexican side of the border highly relies on manufacturing activities performed by foreign affiliates.

Besides the strong ties to the United States, Mexican policymakers have recently adopted a market diversification strategy. Therefore, Mexico has opened up to trade with new countries by negotiating several bilateral and multilateral trade agreements. These new trade agreements and the attraction of FDI have paved Mexico's way to integrate into global value chains (GVCs). Moreover, as we have seen in the past years, emerging countries, such as Mexico, have increased their participation in GVCs, making this country an interesting case study.

Furthermore, Mexico exhibits significant heterogeneity among states. In this regard, states in the northern and central regions display higher GDP levels compared to the rest of the country. In particular, the economy of northern states highly relies on trade and manufacturing activities carried out by foreign affiliates. Therefore, these states are more exposed to business cycle comovements from the United States and other partner countries. For this reason, it is worth exploring the link between FDI and business cycle comovements. In other words, we are interested in examining to what extent partner countries can transmit business cycle fluctuations to Mexican states via FDI.

Regarding the literature, a vast number of studies analyze the role of the international economy on business cycle synchronizations. Some of the key determinants examined in the empirical literature are bilateral trade (Imbs 2004, Calderón et al. 2007, Di Giovanni & Levchenko 2010), productive structure (Clark & Van Wincoop 2001, Imbs 2004, Calderón et al. 2007), intra-industry trade (Imbs 2004, Calderón et al. 2007, Di Giovanni & Levchenko 2010), production sharing (Burstein et al. 2008, Bergin et al. 2009, Zlate 2016), borders and distance (Clark & Van Wincoop 2001), and multinational firms (Boehm et al. 2019, Kleinert et al. 2015), among others. However, little attention has been given to the role of Foreign Direct Investment (FDI) as a key determinant of business cycle comovements in the empirical literature (Hsu et al. 2011, Jansen & Stokman 2014).

The main contribution of this chapter is to investigate the impact of FDI inflows on business cycle comovements for an emerging economy. The previous literature mainly focuses on cross-country studies using developed countries to examine the relationship between FDI and business cycle comovements (Hsu et al. 2011, Jansen & Stokman 2014). In contrast, we now provide an analysis for an emerging economy using state level data. This state level of disaggregation allows us to better understand the relationship between FDI and business cycle synchronizations by accounting for differences across Mexican states. Our study also incorporates a novel filtering technique: the Hamilton regression filter. This filter presents advantages compared to commonly used filtering techniques in the business cycle comovements literature, such as the Hodrick–Prescott (HP) and Baxter-King filters.

The methodology employed is a linear fixed effects model. Our estimation sample consists of paired observations between 47 major partner countries and all 32 Mexican states using sub-periods to create a time element for the variables employed in our econometric model. Our findings suggest that FDI constitutes an important determinant of business cycle transmissions from partner countries to Mexican states. Our results hold after performing a series of robustness checks, such as discarding zero-value observations, employing an alternative measure of FDI, using an alternative specification of the filtering technique, excluding influential investors, using different sub-samples based on income profiles, and exploring the dynamic effect of the dependent variable. We also run Two-Stage Least Squares (2SLS) regressions employing landlocked status, continent, colonizer, and FDI status as instrumental variables to discard potential endogeneity. After running all these robustness checks, we can conclude that our results hold throughout all the different specifications.

The rest of the study is organized as follows: Section 2.2 reviews the existing literature. Section 2.3 describes the data. Section 2.4 exhibits the descriptive statistics. Section 2.5 explains the methodology. Section 2.6 shows the results. Section 2.7 presents the robustness analysis. Section 2.8 concludes.

2.2 Literature Review

This section encompasses a revision of the empirical trade literature on key determinants of business cycle comovements. These determinants include: bilateral trade, Foreign Direct Investment, productive structure, multinational firms, distance, and borders, among others. The papers included in this subsection correspond to the most recent studies on trade determinants and business cycle comovements. A description of the data, methodology, filtering techniques, and findings for each of the key papers is included. Detailed information regarding the filtering techniques is provided in the Appendix.

2.2.1 Bilateral Trade

We start with bilateral trade, which is agreed to play a major role in business cycle comovements between country-pairs (Imbs 2004, Di Giovanni & Levchenko 2010). Imbs (2004) suggests that bilateral trade is strongly correlated to high levels of business cycle synchronizations. The author employs a dataset on bilateral trade, financial integration, and specialization structure variables for U.S. states and 24 countries over the decades 1980s and 1990s. The estimation strategy consists of a Three-Stage Least Squares (3SLS) approach; this approach is useful for isolating the causal effects of variables that are highly intertwined (i.e., in this case, the variables intertwined are trade, financial links, sector specialization, and business cycle comovements). To measure business cycle comovements, the author uses quarterly and annual GDP data, as well as the Baxter-King bandpass filter to detrend these time series. The endogenous variables used in the simultaneous system of equations are bilateral trade intensity, which consists of bilateral trade normalized by output; bilateral financial integration; and a specialization index to measure the similarity in industry specialization. The author also incorporates control variables, which include distance, border, common language, and GDP per capita. In this chapter, we also use bilateral trade as one of our explanatory variables. However, our analysis is focused on the role of Foreign Direct Investment on business cycle comovements.

Likewise, Di Giovanni & Levchenko (2010) suggest that sector pairs experiencing higher levels of bilateral trade tend to display stronger business cycle correlations, especially in sectors that are highly dependent on intermediate inputs. Furthermore, they suggest that the relationship between bilateral trade and business cycle synchronizations is stronger for North-North countries, than for South-South and North-South countries. To perform the analysis, the authors use a database on manufacturing production and trade variables at the sector level for 55 countries, including developed and developing economies, over the period 1970-1999. The empirical technique consists of linear regressions with fixed effects. The authors measure business cycle synchronizations with the correlations of real GDP growth rates. The main explanatory variable is bilateral trade intensity measured in two main forms: bilateral trade at the sector level normalized by output and bilateral trade at the sector level normalized by the sum of total trade in the two countries. It is also worth highlighting that this paper uses a fine level of disaggregation, which is at the sector level. Furthermore, the authors employ a panel approach, which is not very common in the business cycle comovements literature. As mentioned before, we also use bilateral trade as one of the explanatory variables in this chapter. An interesting feature that we also explore is the relationship between bilateral trade and business cycle synchronization taking into account the different levels of development between country and state pairs; this level of disaggregation at the state level has not been explored in the literature for a developing country.

The results in Di Giovanni & Levchenko (2010) are in line with Calderón et al. (2007), who also agree that the relationship between bilateral trade and business cycle comovements is stronger among developed countries. Although Calderón et al. (2007) also find that this relationship is also positive and statistically significant among developing countries, albeit the magnitude of the impact is smaller compared to developed country pairs. To show this, the authors use trade data for 147 countries, including developed and developing countries over the period 1960-1999. The authors employ two types of analysis: cross-section regressions using the average values of annual data for the whole period, as well as panel regressions with country fixed effects. The authors measure business cycle comovements with GDP logged data using four filtering techniques: quadratic trend, first-differenced series, Baxter-King, and HP filters. The main explanatory variable is bilateral trade intensity; this variable has two alternative measurements: one of them constitutes the average of bilateral trade normalized by trade, and the other one is the average of bilateral trade normalized by output. Moreover, the authors include a dissimilarity index variable to measure the similarities in the production structure, as well as an intra-industry trade variable measured as the absolute value of the trade balance over total trade. We can observe three interesting features in Calderón et al. (2007). First, the authors include both developed and developing countries in their dataset. Second, they exploit the panel structure of their dataset by taking into account the time element. Third, they explore the relationship between trade and business cycle comovements from the perspective of different combinations of income profile paired countries. In this chapter, we reach a higher disaggregation level, where we consider country-state paired observations from the perspective of a developing country. Furthermore, we also include panel regressions and we divide our sample into four sub-samples based on income profiles of partner countries and Mexican states. Different from these authors who employ first difference, HP filter, and Baxter-King filters, we focus on the novel Hamilton regression filter.

Due to the importance of bilateral trade on business cycle synchronizations between country pairs, several empirical studies include this key variable despite their focus is on other determinants, such as border and distance (Clark & Van Wincoop 2001), Foreign Direct Investment (Hsu et al. 2011, Jansen & Stokman 2014), and foreign affiliates (Kleinert et al. 2015). This leads us to the following determinants of business cycle comovements.

2.2.2 Other Trade Determinants

We start this subsection by exploring the impact of border and distance followed by productive structures, intra-industry trade, production sharing, and trade openness. Distance and borders also play a key role for determining business cycle comovements. Clark & Van Wincoop (2001) explain that distance represents an important trade barrier due to an associated increase in transportation and communication costs. Thus, larger distances between partner countries tend to lower business cycle correlations. In terms of the border effect, the authors suggest that a national border might discourage trade volumes between countries compared to trade volumes among regions. Therefore, borders can negatively impact business cycles among countries. To investigate this, the authors use data on sector specialization, trade, monetary and fiscal policy, as well as distance and border for 14 EU countries and 9 U.S. regions over the period 1964-1997. The empirical strategy consists of cross-section regressions. Business cycle correlations are measured as detrended annual GDP growth rates using the Baxter-King and

HP filters; these measures employ the average values of two sub-sample periods: 1964-1980 and 1981-1997. The main explanatory variables are border measured by a dummy variable; and distance measured as the weight of population average times the log of the distance between pairs of capital cities. These authors also include bilateral trade intensity, which is measured as bilateral trade normalized by GDP. Furthermore, a dissimilarity index is also incorporated to measure the similarity of industry specialization, which leads us to the next determinant.

The impact of the productive structure has also been studied as a mechanism influencing business cycle comovements. Recalling the work of Clark & Van Wincoop (2001) and Imbs (2004), the authors also found that countries with a similar productive structure tend to display stronger patterns of business cycle synchronizations.

Intra-industry trade constitutes another important determinant of business cycle comovements. Imbs (2004) supports the relevance of trade on business cycle comovements and adds that an important amount of trade is constituted by intra-industry trade. Calderón et al. (2007) show that trade integration has a higher impact on business cycles between countries that trade intensively in intermediate goods. Di Giovanni & Levchenko (2010) point out that an increase in bilateral trade within sector pairs intensively using intermediate inputs might lead to stronger comovements.

Production sharing has also been examined as a mechanism leading to business cycle synchronizations.[1] We identify, three case studies focusing on the transmission of business cycles from the United States to Mexico through production sharing. Burstein et al. (2008) conclude that the size of the effect of production-sharing intensity is comparable to the effect of trade volume on correlations of manufacturing output. In other words, the authors show that countries closely engaging in production-sharing tend to display higher correlations of manufacturing output. Bergin et al. (2009) claim that the Mexican offshoring industry can represent a mechanism through which the United States transmits volatility to the Mexican business cycle. Also, Zlate (2016) provides evidence of a positive relationship between the share of export offshoring and business cycle comovements.

Another trade-related variable explored in the business cycle comovement literature is trade openness. Caselli et al. (2020) suggest that in the event of a sector-shock transmission, trade openness can alleviate GDP volatility through a diversification strategy of partner countries.

2.2.3 Multinational Firms

A recent strand of the trade literature focuses on the role of firms on business cycle comovements. Gabaix (2011) introduced the term "granular hypothesis"; this term suggests that shocks received by large individual firms can lead to aggregate fluctuations that impact the economic performance of a country. These individual large firms (e.g., top 50 or top 100 firms in a country) act as grains within an economy. Di Giovanni & Levchenko (2012) support the granular hypothesis by providing empirical evidence showing that shocks experienced by large

[1] Burstein et al. (2008) define production sharing as "trade in intermediate goods that are part of vertically integrated production networks that cross international borders".

trading firms lead to GDP volatility within a country.

These multinational firms possess features that differentiate them from the rest of the firms. Alviarez et al. (2017) list three ways in which multinational firms do not behave in a similar manner as domestic firms and mention how their presence varies from one country to another; these authors explain that multinational firms are active in different sectors within an economy compared to domestic firms; parent firms and their affiliates are larger in size than domestic firms; finally, countries diverge in the presence of multinational firms and in the composition of source countries that own foreign affiliates. On top of this, Kurz & Senses (2016) mention that firms with trading activity are heterogeneous in terms of the number and types of varieties traded, as well as on the amount and characteristics of their trade partners.

Garetto et al. (2016) argue that multinational firms constitute major actors in the global economy and explain their evolution over time (e.g., an affiliate starting with horizontal sales activity tends to expand towards vertical and export-platform sales over time). Di Giovanni et al. (2017) point out that trade flows are dominated by only a few large firms; these authors even show that multinational firms have the potential to significantly contribute to business cycle comovements by analyzing the top 100 firms in France; the authors also conclude that firms display stronger correlations with countries through both trade and multinational linkages. Cravino & Levchenko (2017) stress the importance of the interdependence between source countries and their foreign affiliates as a channel for business cycle comovements; they also mention that the most integrated countries have a higher propensity to be influenced by foreign shocks. Finally, Boehm et al. (2019) show that the elasticity of substitution between imported and domestic inputs plays a key role in the transmission of shocks from source to destination countries.

Another branch of the literature narrows the focus to country performance during economic recessions. Gopinath & Neiman (2014) identify that during a crisis, a contraction of imported intermediate inputs has a significant impact on the country's productivity, which leads to welfare losses. The authors also point out that a few key firms and sectors have the potential to concentrate an important share of total trade within a country. In a similar vein, Alviarez et al. (2017) show that multinational firms experienced a similar size collapse in sales during the Great Recession. Sandqvist (2017) also shows that business cycle correlations may significantly increase during economic recessions. The author identifies this increase is especially true for the manufacturing sector; thus, this sector tends to become more synchronized during economic contractions than during expansions.

Finally, Kleinert et al. (2015) show that the presence of foreign affiliates has a positive and significant impact on business cycle comovements between source countries and regions where these firms are located. In order for this to be the case, the authors explain that two conditions need to be met; first, foreign affiliates must account for a significant share of GDP in the regions established; second, the parent and foreign affiliate should display a positive correlation of value-added growth. To perform this analysis, the authors use a firm-level dataset for 21 regions in France and 162 partner countries over the period 1990-2006. The methodology employed consists of cross-section regressions with country and region fixed effects. The dependent variable is business cycle correlations between French regions and partner countries; this variable is measured by the annual GDP growth rates in natural logs and HP-filtered. The main explanatory variable is the share of employment by foreign affiliates established in French regions. The control variables are bilateral trade, a dissimilarity index based on the difference of sectoral export shares between country and region pairs, and an intra-industry trade measure. A main critique of this paper is that the authors did not exploit the time dimension of their database by using a panel regression approach, instead they use a cross-section approach for a specific year (i.e., 2004). This chapter uses a similar approach to Kleinert et al. (2015); nonetheless, our chapter differs in several ways. First, our chapter focuses on the role of FDI instead of the presence of foreign affiliates measured with the share of employment. Second, we use the perspective of a developing country. Third, we employ the recent Hamilton regression filter. Fourth, we exploit the panel structure of our database and use linear fixed effects regressions. As mentioned before, this chapter focuses on the impact of Foreign Direct Investment on business cycle comovements, which leads us to the next subsection.

2.2.4 Foreign Direct Investment

The empirical trade literature has given little attention to the role of Foreign Direct Investment inflows as a determinant of business cycle comovements. Jansen & Stokman (2014) suggest that countries with stronger FDI linkages tend to display more synchronized business cycles, especially for the years following 1995, when a boost of FDI flows was globally experienced. The authors employ a dataset for 12 developed economies over the period 1982-2007. The methodologies used are cross-section regressions and Two-Stage Least Square (2SLS) regressions. To measure the business cycle comovements, the authors use quarterly and annual growth rates of real GDP. This data was detrended using the Baxter-King and HP filters. The explanatory variables are FDI and bilateral trade. FDI is calculated as the sum of the inward and outward FDI stocks of country pairs as a percentage of the reporting country's GDP. On the other hand, bilateral trade is calculated in a similar manner as FDI linkages, which is bilateral trade of country pairs normalized by country's GDP. The authors analyze different estimation samples due to differences in the pace of FDI activity: the mean values of the whole estimation sample (1982-2007) and the mean values of two sub-samples (1982-1994 and 1995-2007). In this chapter, we also explore the link between FDI and business cycle comovements; however, we focus on a developing country. We extend the analysis by exploiting the panel structure of our database, which is something not exploited in Jansen & Stokman (2014).

Hsu et al. (2011) suggest that FDI constitutes an important channel for the transmission of business cycles across countries; moreover, they also suggest that FDI plays an important role in the diffusion of technology and financial investment. The authors use a dataset conformed by 77 developed country-pairs over the period 1988-2002. The empirical approach consists of a comparison between single equation estimations with fixed effects and with random effects, as well as simultaneous equation estimations using cross-sectional data against the results of the error component three-stage least squares (EC3SLS). The authors suggest that this last methodology displays more reliable estimates as this method controls for a potential endogeneity bias and also presents more coefficients that are statistically significant. To measure business cycle correlations, the authors use GDP annual growth rates measured in natural logs, as first-difference, and Hodrick-Prescott-filtered data. The system of equations include three endogenous variables: bilateral trade, FDI, and a dissimilarity index to measure the similarity in industry specialization. Bilateral trade is measured as the sum of the exports and imports made between a country-pair over the total amount of exports and imports of that country-pair. FDI has no standard measure according to these authors; therefore, they propose an index similar to the bilateral trade index; this index represents the sum of inward and outward FDI flows of country pairs as a percentage of the sum of total FDI by that country pair. In our chapter, we also explore the relationship between FDI and business cycle comovements. However, our analysis focuses on the perspective of a developing country. An interesting feature of this chapter is that the unit of observation is state-country pairs, which is not common in the literature. Finally, we employ the recent Hamilton regression filter as the filtering technique in our baseline regression.

2.2.5 Contributions to the Literature

The empirical trade literature has identified several key determinants of business cycle comovements. Within the most popular determinants we can find bilateral trade, productive structure, intra-industry trade, production sharing, distance, borders, and multinational firms, among others; however, little attention has been given to the relevance of FDI as a determinant of business cycle comovements especially from the perspective of a developing country.

In fact, a source country has the potential to transmit fluctuations to the business cycle of host states through FDI. On the one hand, FDI promotes a demand for skilled workers in host states; this situation leads to a skilled premium, where wages are higher for skilled workers employed in foreign affiliates compared to those employed in domestic firms in those states. On the other hand, experiencing an economic downturn in the source country can lead to a significant reduction of FDI inflows to host states; this situation could lead to layoffs, or even, closure of manufacturing plants in host states; thus, impacting the local economy.

In this chapter, we focus on the role of FDI on business cycle comovements. We also incorporate key trade determinants in the analysis as control variables (i.e., a bilateral trade index and a dissimilarity index). The analysis constitutes a case study of a developing country, which displays heterogeneous degrees of economic development among states. This level of disaggregation has not been studied before for a developing country since the few available studies center on the role of FDI using country-level data for developed economies. The methodology employed consists of a linear fixed effects model. This analysis also incorporates the recent Hamilton regression filter.

2.3 Data

The database in this chapter is the result of the compilation of several datasets retrieved from Mexican authorities and international organizations. This database consists of FDI and trade variables for 47 major partner countries and all 32 Mexican states over the period spanning from 1999 to 2016.[2] [3]

We employ GDP annual data of partner countries and Mexican states to construct the dependent variable (i.e., business cycle correlations). We use GDP at purchasing power parity (PPP) in constant international dollars for country-level data, which is retrieved from the World Bank Development Indicators. On the other hand, Mexican state GDP data is taken from the National Institute of Statistics and Geography (INEGI); this state data is available as a percentage of total GDP, which we then converted it into GDP (PPP) in constant international dollars by multiplying each state percentage times the annual Mexican GDP at PPP.

FDI data is reported by the Mexican Ministry of Economy. The Ministry reports annual data on the value of FDI inflows expressed in millions of U.S. dollars. These FDI inflows include greenfield investments, as well as mergers and acquisitions.[4] Furthermore, these FDI inflows represent net inflows of investment (i.e., new investment inflows less disinvestment). FDI data is only available for Mexico's top 50 partner countries, which account for 98% of total inflows captured by Mexico.[5] The remaining 2% of FDI inflows is clustered in a category named "rest of the world". It is worth mentioning that this FDI data is reported in three sets of datasets, each one comprised only of two dimensions. In other words, one dataset includes country and state level data; another dataset includes country and sector level data; and a third dataset includes state and sector level data. However, a limitation of this data is that there is not a unique FDI dataset available at a finer level of disaggregation, which comprises all of the three dimensions including country, state, and sector level data. Finally, it is worth mentioning that we employ annual FDI inflows instead of FDI stocks as it offers more flexibility in terms of variable construction for the panel specification.

The main explanatory variable in our analysis is FDI. To construct this variable, we use data on FDI inflows from the source country c to the host Mexican state r over FDI inflows

[2]The 47 countries encompassed in the estimation sample are: Argentina, Australia, Austria, Belgium, Belize, Brazil, Canada, Chile, China, Colombia, Costa Rica, Czech Republic, Denmark, Ecuador, El Salvador, Finland, France, Germany, Guatemala, Hong Kong, India, Indonesia, Ireland, Israel, Italy, Japan, South Korea, Luxembourg, Malaysia, the Netherlands, New Zealand, Nicaragua, Norway, Panama, Peru, the Philippines, Poland, Portugal, Russian Federation, Singapore, South Africa, Spain, Sweden, Switzerland, United Kingdom, United States, and Uruguay.

[3]The 32 Mexican states included in the estimation sample are: Aguascalientes, Baja California, Baja California Sur, Campeche, Chihuahua, Chiapas, Coahuila, Colima, Durango, Guerrero, Guanajuato, Hidalgo, Jalisco, Estado de Mexico, Mexico City, Michoacan, Morelos, Nayarit, Nuevo Leon, Oaxaca, Puebla, Queretaro, Quintana Roo, Sinaloa, San Luis Potosi, Sonora, Tabasco, Tamaulipas, Tlaxcala, Veracruz, Yucatan, and Zacatecas.

[4]According to the OECD Benchmark Definition of Foreign Direct Investment, greenfield investments refer to the creation of a subsidiary in a foreign country. On the other hand, mergers and acquisitions refer to the case where an existing firm is taken over, fully or partially, by other firms.

[5]The original FDI dataset also included Taiwan, Puerto Rico, and Venezuela. However, these three countries were dropped from the estimation sample because the World Bank does not list Taiwan as a separate country. In the case of Puerto Rico and Venezuela, these countries do not report their GDP for the years 2013-2016 and 2015-2016, respectively. Thus, we ended up with an estimation sample composed by 47 countries.

received at that Mexican state r. For this purpose, we select FDI inflows reported by country c and state r. This selected measure constitutes the aggregation of all the sectors in the economy: agricultural, extraction, manufacturing, and services.[6]

Mexican trade data is retrieved from two main national sources: the Mexican Ministry of Economy and the INEGI. The first source releases data on annual exports and imports at the product level (HS 8-digits) by destination or source country.[7] The second source releases aggregate trade data at the country level; this source also reports annual exports by state and sector (NAICS 3-digits), but does not report annual imports; these export sectors only include mining and manufacturing activities.

Furthermore, annual data on country trade flows is accessible from two sources: the World Trade Organization (WTO) and the UN Comtrade. The first source releases trade flows at sectoral level according to the Standard International Trade Classification (SITC) Revision 3. The sectors included in this study are: fuels; non-ferrous metals, ores, and other minerals; food; textiles; clothing; chemicals; iron and steel; and machinery and transport equipment. On the other hand, the UN Comtrade releases trade flows at the country level.

Data on landlocked status, continent, colonizer, as well as latitude and longitude of capital cities is taken from the Geo CEPII Database (Calderón et al. 2007, Imbs 2004, Navas et al. 2020). On the other hand, latitude and longitude data of Mexican cities can be retrieved from the World Cities Database. Bilateral distance is computed using the great-circle distance formula to calculate the distance between capital cities of partner countries and Mexican state capital cities (Jansen & Stokman 2014, Navas et al. 2020). The border variable between partner countries and Mexican states is a construct dummy variable; this variable is equal to 1 if a Mexican state shares a border with a partner country, and zero, otherwise (Clark & Van Wincoop 2001, Imbs 2004, Calderón et al. 2007, Jansen & Stokman 2014, Kleinert et al. 2015).[8] We also constructed a free trade agreement (FTA) variable employing data from the Foreign Trade Information System of the Organization of American States (OAS) records. This FTA variable is a dummy equal to 1 if a partner country has a free trade agreement with Mexico; and zero, otherwise (Jansen & Stokman 2014).

Supplementary information on standard codes for countries (i.e., ISO 3166) and Mexican subdivisions (i.e., ISO 3166-2) is sourced from the International Organization for Standardization (ISO). In addition to these standard codes, the Mexican Ministry of Economy also possesses its own country codes. Annual exchange rates used to translate reported values in Mexican pesos to U.S. dollars are obtained from the OECD. The correspondence table (version 2017) between the Harmonized System codes (2012) and the NAICS codes (2013) is sourced from the INEGI.

[6]Sector-level data is consistently reported at NAICS 2- to 6-digit levels. The acronym NAICS stands for North American Industry Classification System. This coding system was developed by the national statistical agencies of Canada, the United States, and Mexico to classify industries.

[7]Mexican data on bilateral trade is reported at the 8-digit Harmonized System (HS) level, which is the finest level of disaggregation corresponding to products.

[8]Mexican states sharing a border with the U.S. are Baja California, Sonora, Chihuahua, Coahuila, Nuevo Leon, and Tamaulipas. Mexican states sharing a border with Guatemala are Campeche, Chiapas, and Tabasco. Mexican states sharing a border with Belize are Campeche and Quintana Roo.

2.4 Descriptive Statistics

2.4.1 Variable Description

Dependent Variable

The dependent variable consists of GDP growth rate correlations between country c and Mexican state r in period t. In a similar manner as Calderón et al. (2007) and Hsu et al. (2011), we create a time element of the dependent variable by splitting the data into 5-year time spans that we call periods t. Therefore, we calculate the linear correlation coefficients for the following periods: 1996-1999, 2000-2004, 2005-2009, and 2010-2014.[9]

As a matter of context, the empirical literature on business cycle comovements have used different filtering techniques to separate time series into the trend and the cyclical component (Clark & Van Wincoop 2001, Imbs 2004, Calderón et al. 2007, Burstein et al. 2008, Hsu et al. 2011, Jansen & Stokman 2014, Kleinert et al. 2015, Zlate 2016).[10] However, trade economists have not achieved a consensus yet on which of these filtering techniques is the most appropriate for business cycle comovements.[11]

In this study, we calculate the dependent variable using Hamilton regression logged and filtered data. We were inclined to use the novel Hamilton regression filter as it is perceived to offer advantages over the commonly used HP filter, which is criticized of incurring on spurious cycles (Canova 1998, Hamilton 2017). Nonetheless, we also present the results after using other filtering techniques in the Appendix section (i.e., these filtering techniques consist of first differences, the Baxter-King filter, frequency domain filter, and the Hodrick-Prescott filter). Additionally, we present an alternative specification of the Hamilton regression filter using standardized data as part of the Robustness Analysis.

The raw version of the dependent variable is calculated in line with Kleinert et al. (2015):

$$\rho_{crt} = corr\left(\frac{GDP_{c,t} - GDP_{c,t-1}}{GDP_{c,t-1}}, \frac{GDP_{r,t} - GDP_{r,t-1}}{GDP_{r,t-1}}\right), \tag{2.1}$$

where c stands for country, r represents Mexican state, and t is the period. Thus, the dependent variable corresponds to the GDP growth rate correlations between 47 major partner countries and all 32 Mexican states over period t. Furthermore, the dependent variable employed in the baseline specification consists of 5-year GDP growth rate correlations of country-state paired observations using logged and Hamilton filtered data.

[9]The initial year of the first period (i.e., 1995) is not included due to data limitations. Furthermore, we dropped the last two years of our original dataset (i.e., 2015 and 2016) because calculating linear correlation coefficients with only two observations is inaccurate.

[10]Clark & Van Wincoop (2001) and Jansen & Stokman (2014) employ the Baxter-King and HP filters. Imbs (2004) uses the Baxter-King filter. Calderón et al. (2007) employ the first difference, Baxter-King, and HP filters. Burstein et al. (2008) and Hsu et al. (2011) use the first difference and HP filters. Kleinert et al. (2015) and Zlate (2016) employ the HP filter.

[11]Table 2.16, presented in the Appendix section, summarizes the most common filtering techniques employed in the business cycle comovements literature. In addition, the Appendix provides a formal derivation of these filtering techniques.

Main Explanatory Variable

The main explanatory variable is constructed as a ratio of FDI inflows from country c to Mexican state r in period t over FDI inflows captured by state r in period t (i.e., $pFDI_{crt}$). The previous studies focusing on the impact of FDI on business cycle comovements measure this variable in a similar manner as bilateral trade intensity; thus, two measures of this variable have been used in the literature: FDI as a percentage of GDP (Jansen & Stokman 2014) and an index of bilateral FDI (Hsu et al. 2011). This chapter employs a similar approach to Hsu et al. (2011) to measure the influence of partner countries on state FDI inflows.[12] Thus, the main explanatory variable in this chapter is defined as follows:

$$pFDI_{crt} = \frac{FDI_{crt}}{FDI_{rt}}, \tag{2.2}$$

where FDI_{crt} stands for FDI inflows from source country c to Mexican state r in period t, and the denominator FDI_{rt} corresponds to FDI inflows received in Mexican state r in period t. This explanatory variable uses mean values of country-state paired observations over 5-year periods. Furthermore, this FDI measure includes aggregate values of all sectors in the economy: agricultural, extractive, manufacturing, and services.

Control Variables

The bilateral trade intensity index is calculated following Frankel & Rose (1998). This index proxies trade integration and has been employed in other studies examining the impact of trade on business cycle comovements (Clark & Van Wincoop 2001, Imbs 2004, Kleinert et al. 2015). Thus, the variable is constructed as a ratio of total bilateral trade over the sum of the GDP produced by country c and Mexican state r in period t:

$$BT_{crt} = \frac{expo_{crt} + impo_{crt}}{GDP_{ct} + GDP_{rt}}, \tag{2.3}$$

where $expo_{crt}$ stands for exports by Mexican state r to country c in period t. On the other hand, $impo_{crt}$ represents imports by Mexican state r from country c in period t. It is worth noting that we encountered some data limitations. In fact, the INEGI releases data on exports by Mexican state, but does not report data on imports by Mexican state due to potential confidentiality issues. Thus, we assume that $expo\_SH_{rt} \equiv impo\_SH_{rt}$, where the export share $expo\_SH_{rt}$ is equivalent to the import share $impo\_SH_{rt}$.[13] The export share can be calculated as the following ratio:

$$expo\_SH_{rt} = \frac{expo_{rt}}{\Sigma(expo_{rt})}, \tag{2.4}$$

[12] In subsection 2.7.2 Alternative Measure of FDI of the Robustness Analysis section, we show that our results are similar if we use an alternative specification of FDI, which is inspired in Jansen & Stokman (2014); this alternative measure corresponds to FDI inflows from country c to state r as a share of state GDP.

[13] We make this assumption based on the aggregate exports and imports at the country level spanning from 1993 to 2016; these trade time series are reported by the INEGI. We can observe that on average, the bulk of exports represents about 97% of total imports at the country level.

where the export share $expo\_SH_{rt}$ is expressed as state exports ($expo_{rt}$) weighted over total exports (i.e., $\Sigma(expo_{rt})$); this share is a number between 0 and 1. Therefore, the index of bilateral trade intensity can be constructed as follows:

$$BT_{crt} = \frac{(expo\_SH_{rt})(expo_{ct} + impo_{ct})}{GDP_{ct} + GDP_{rt}}, \tag{2.5}$$

where $expo\_SH_{rt}$ is the export share of Mexican state r in period t multiplied by Mexico's total trade ($expo_{ct} + impo_{ct}$), weighted over the sum of the GDP of country c and Mexican state r in period t. This control variable uses mean values of country-state paired observations over 5-year periods. In terms of the interpretation, a higher value of the index denotes a higher intensity of trade between a partner country and a Mexican state. On the other hand, a lower value of the index denotes a lower intensity of trade between the country-state pairs.

Regarding additional data limitations, this trade variable only includes extractive and manufacturing sectors.[14] Also, the periods belonging to the time spans 1996-1999 and 2000-2004 were dropped from the estimation sample due to data limitations at the state-level (i.e., trade data at the state-level is only available from 2003).

We also use a dissimilarity index as part of our control variables. The main objective of this index is to reveal the comparative advantage in terms of export composition. This index is calculated following Kleinert et al. (2015):

$$DISSIM_{crt} = \sum_i \left| \frac{expo^i_{ct}}{expo_{ct}} - \frac{expo^i_{rt}}{expo_{rt}} \right|, \tag{2.6}$$

where $\frac{expo^i_{ct}}{expo_{ct}}$ is the share of exports by country c belonging to sector i in period t over total exports by country c in period t. On the other hand, $\frac{expo^i_{rt}}{expo_{rt}}$ stands for the share of exports by Mexican state r belonging to sector i in period t over total exports by Mexican state r in period t. Then, we take the absolute value of the difference between the two shares (i.e., $\frac{expo^i_{ct}}{expo_{ct}}$ and $\frac{expo^i_{rt}}{expo_{rt}}$). Finally, we take the sum of the sectors i. This control variable uses mean values of country-state paired observations over 5-year periods. In terms of the interpretation, a higher value of the index can be seen as a greater difference of export structures between the partner country c and the Mexican state r, and vice versa. Thus, more similar export structures lead to stronger business cycle correlations.

To construct this variable, we also experienced two data limitations. First, this index includes the extractive and manufacturing sectors only.[15] Second, the periods belonging to the time spans 1996-1999 and 2000-2004 were dropped from the estimation sample due to data limitations at the state-level (i.e., trade data at the state-level is only available from 2003).

[14] Due to the nature of trade data, services are not included; moreover, data on agricultural exports by Mexican state is not available.

[15] Due to limited data, we matched WTO sectors (i.e., merchandise aggregates in SITC Revision 3) to corresponding NAICS-3 digit sectors. Thus, we included the following sectors: fuels; non-ferrous metals, ores, and other minerals; food; textiles; clothing; chemicals; iron and steel; and machinery and transport equipment.

2.4.2 Summary Statistics

Table 2.1 reports the summary statistics of the analyzed variables that correspond to the estimation sample. In a similar manner as Calderón et al. (2007) and Hsu et al. (2011), we construct a panel data using 5-year periods to create the time element for the dependent variable (i.e., GDP growth rate correlations). This estimation sample covers different sectors depending on the data availability.[16]

Table 2.1: Summary Statistics

Variables	Labels	(1) N	(2) Mean	(3) Std.Dev.
Correlation of GDP growth rates (Raw)	ρ_{crt}^{Raw}	3,008	0.4541	0.4590
Correlation of GDP growth rates (Ham)	ρ_{crt}^{Ham}	3,008	0.5368	0.4680
Correlation of GDP growth rates (Ham-STD)	$\rho_{crt}^{Ham-STD}$	3,008	0.5270	0.4705
FDI inflows as a state proportion	$pFDI_{crt}$	3,008	0.0213	0.0848
FDI inflows as a share of GDP	$sFDI_{crt}$	3,008	0.0003	0.0018
Bilateral Trade Index	BT_{crt}	3,008	0.0001	0.0003
Dissimilarity Index	$DISSIM_{crt}$	3,008	1.1714	0.4823

Notes: This estimation sample is composed of country-state paired observations using 5-year GDP growth rate correlations to create the time element of the dependent variable. The explanatory variables constitute the mean values of country-state paired observations over 5-year periods. These periods include 2005-2009 and 2010-2014. Due to data restrictions of the control variables, the periods belonging to the time spans 1996-1999 and 2000-2004 were dropped from the estimation sample.

The average GDP growth rate correlation between country c and Mexican state r using Hamilton logged and filtered data is 53.7%. This mean correlation is comparable in size to Hsu et al. (2011) for developed countries, who found a mean correlation of 46.5% using first-differenced data. Nonetheless, our mean correlation is dramatically larger than in Calderón et al. (2007) for cross-countries including developed and developing countries, who found a mean correlation of 3.7% using first-differenced data and 5.9% using HP-filtered data.

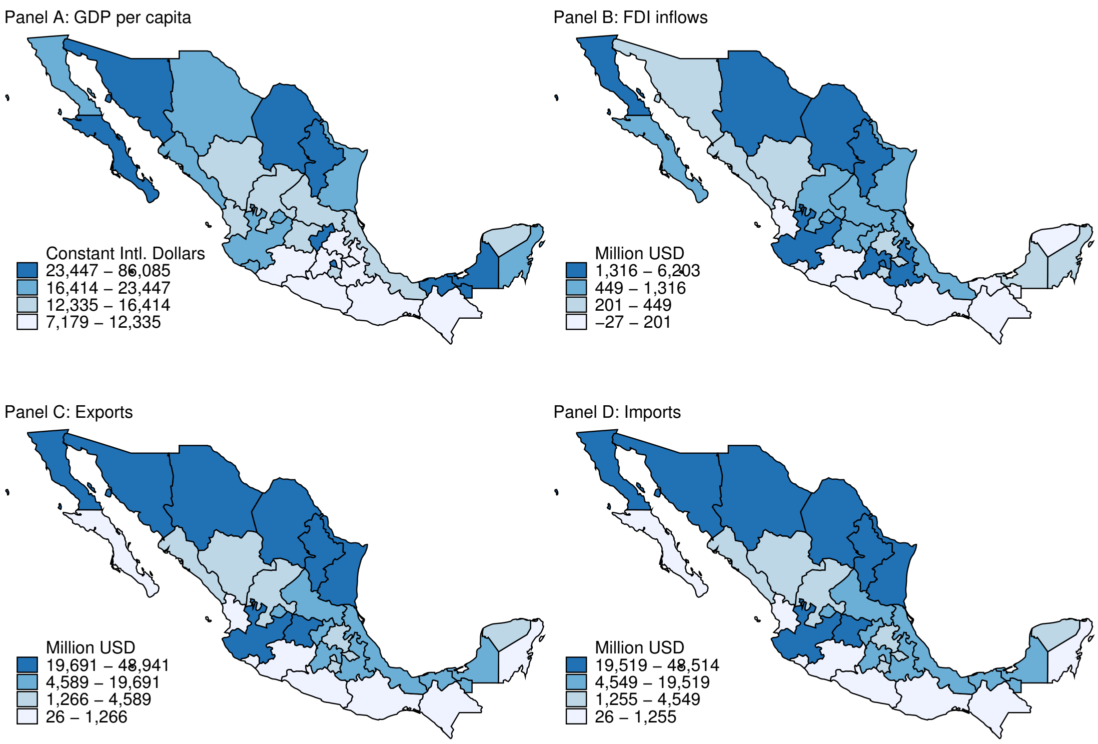
In terms of the mean value of our main explanatory variable (i.e., FDI inflows from source country c to host Mexican state r over FDI inflows received in Mexican state r) is 2.1% for Mexican states, compared to 0.8% in Hsu et al. (2011) for 77 paired developed countries, and to 12.2% in Jansen & Stokman (2014) for 8 developed countries. Furthermore, the mean value of our alternative measure of FDI (i.e., FDI as a share of state GDP) is 0.03% for Mexico, which is comparable in size to Kleinert et al. (2015), who report an average of 0.03% for their main explanatory variable (i.e., foreign employment share) for the whole sample. Furthermore, bilateral trade has an average value of 0.01% for Mexico, which is consistent with Kleinert et al. (2015), who report an average of 0.02% for the whole sample. Nonetheless, this bilateral trade mean value for Mexico is smaller compared to Hsu et al. (2011), who report a mean value of 5.5%. Also, the mean value of the dissimilarity index for Mexico (1.17) is comparable to Kleinert et al. (2015), who report a mean value of 1.07.

[16]The independent variables are constructed using different sectors: FDI inflows include the agricultural, extractive, manufacturing, and service sectors; while bilateral trade and the dissimilarity index only comprise the extractive and manufacturing sectors; these last two variables encounter data limitations due to the nature of trade data that excludes services. Also, export data at the state-level do not report information for the agricultural sector.

2.4.3 Economic Indicators

Figure 2.1 illustrates important economic indicators of Mexican states: GDP per capita, FDI inflows, exports, and imports.[17] Dark blue states perform better in these economic indicators, while lighter blue states exhibit a lower performance. Panel A displays the GDP per capita by state measured at purchasing power parity (PPP) in constant international dollars.[18] From this map, Campeche is the state with the largest GDP per capita (86,085 USD), followed by Mexico City (42,782 dollars), Nuevo Leon (31,183 dollars), Tabasco (28,510 dollars), and Sonora (25,665 dollars). It is worth mentioning that Nuevo Leon and Sonora are two states located in the northern region.

Figure 2.1: Economic Indicators



Notes: This figure displays key economic indicators of Mexican states. Panel A exhibits GDP per capita by state measured at purchasing power parity (PPP) in constant international dollars. Panel B shows FDI inflows captured by state measured in million U.S. dollars. Panel C presents exports by state measured in million U.S. dollars. Panel D displays imports by state measured in million U.S. dollars. The reference year for these maps is 2016.

Panel B exhibits FDI inflows received by state. We can observe that Mexico City concentrates a significant amount of FDI inflows (6.2 billion USD), followed by Nuevo Leon (3.1 billion USD), Estado de Mexico (2.4 billion USD), Jalisco (2.0 billion USD), and Chihuahua (1.9 billion USD). From this figure, we can identify that two of the major FDI recipients, Nuevo Leon and Chihuahua, are located in the northern region. Also, some other important recipients are located in the central region of the country.

Panel C shows exports by state. The state with the largest export activity is Chihuahua

[17] A map of the administrative divisions in Mexico with state names is available in the Appendix section.

[18] We use PPP in constant international dollars in this map to be consistent with the construction of the dependent variable, which is GDP growth rate correlations.

(48.9 billion USD), followed by Coahuila (42.6 billion USD), Baja California (40.6 billion USD), Nuevo Leon (36.3 billion USD), and Tamaulipas (28.2 billion USD). We can now observe a clear pattern showing that all top five states with export activity are located in the northern region sharing a border with the United States.

Finally, Panel D displays imports by state.[19] The state with the largest import activity is also Chihuahua (48.5 billion USD), followed by Coahuila (42.1 billion USD), Baja California (40.2 billion USD), Nuevo Leon (36.0 billion USD), and Tamaulipas (27.9 billion USD). We can observe that export figures are slightly larger than imports.

We can conclude that these maps suggest the presence of heterogeneity among states. In fact, we can observe that states located in the northern and central regions perform better in most of these economic indicators, while states in the southern region display a lower performance. In terms of GDP per capita, we can notice two exceptions to this pattern: Tabasco and Campeche; these two states are located in the southern region and are among the top five states with larger GDP per capita; a plausible explanation is that these two states specialize in the oil and gas industry. Regarding other economic indicators, we can observe FDI figures are relatively small compared to exports. Furthermore, we can identify a clear trade specialization pattern, where states located in the northern region exhibit larger exports and imports compared to other regions. As an overall conclusion, we can suspect that FDI and bilateral trade may have an effect on business cycle comovements at the state level.

2.4.4 FDI Inflows

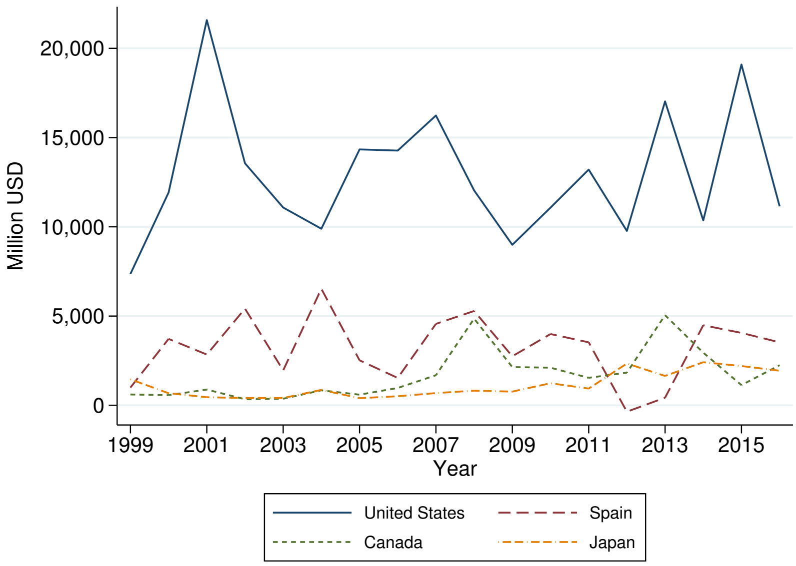
Figure 2.2 exhibits FDI inflows captured by Mexico over the period 1999-2016. This figure displays FDI inflows from Mexico's four major investors. We can observe that the United States is by far Mexico's top investment partner. Moreover, it is possible to notice that all of the time series report a significant decline of FDI inflows that coincides with the Great Recession of 2008-2009. In the case of the United States, the country experienced an early decline in its economy that can be explained by the subprime mortgage market issues in 2007.

Figure 2.3 displays FDI inflows sourced by Mexico's four major investment partners in 2016: the United States (10.6 billion USD), Spain (3.0 billion USD), Germany (2.6 billion USD), and Canada (2.2 billion USD). From this figure, we can observe some patterns of investment allocation in Mexican states. For instance, the United States invests a significant amount of FDI in states located in the northern border; however, this country also invests in two other states located in the central region: Mexico City and Queretaro. A plausible explanation to invest in these states is that Mexico City represents a natural spot to allocate headquarters of multinational firms; on othe other hand, Queretaro has become highly industrialized in the past few decades.

We can also observe that Germany allocates a significant proportion of FDI in the central

[19] As mentioned before, the INEGI does not report import data at the state level. Thus, we assume the export share is equivalent to the import share. Therefore, we multiply the import value in U.S. dollars at the national level times the export share to obtain an approximate of the imports at the state level.

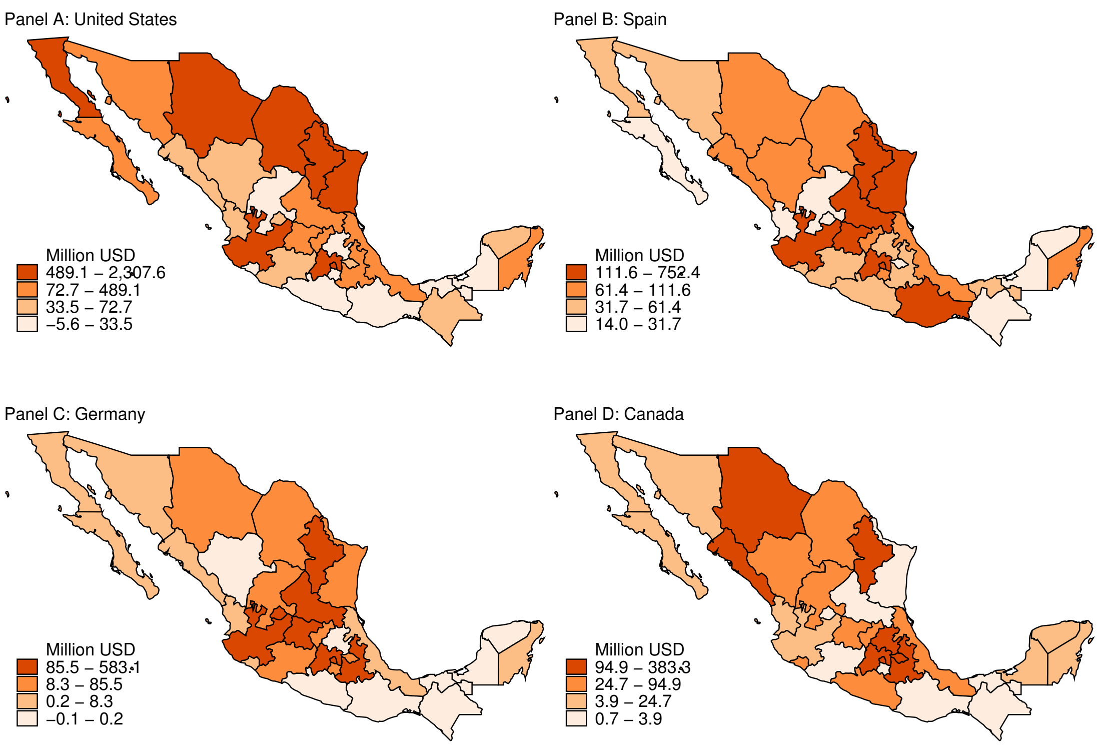
Figure 2.2: FDI Inflows



Source: This figure exhibits FDI inflows from Mexico's four major investors over the period 1999-2016.

region. Also, Canada concentrates FDI inflows in a few Mexican states specializing in mining activities; these states are located in the northwestern and central regions. Canada also invests in some tourist destinations located in the Pacific coastline, as well as in Mexico City. In the case of Spain, we cannot identify a clear pattern of geographical concentration.

Figure 2.3: Distribution of FDI Inflows

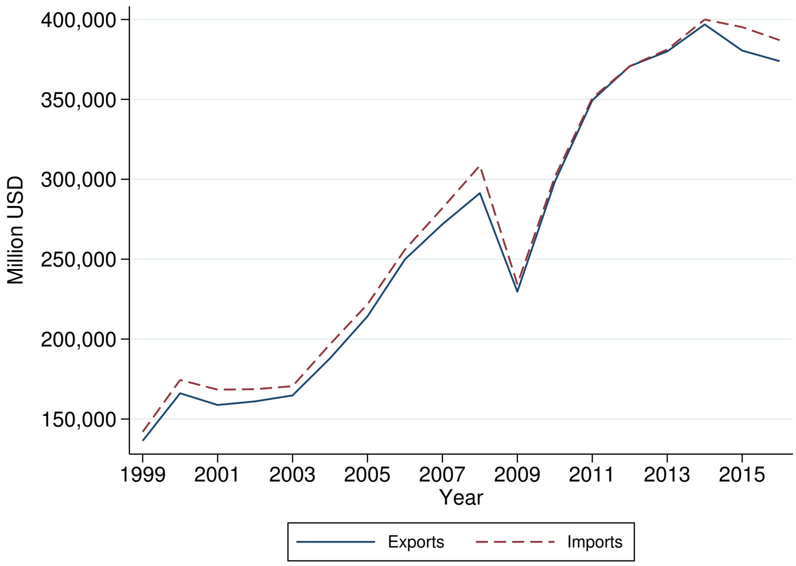


Source: These maps show the geographical distribution at the state level of FDI inflows from Mexico's four major investment partners. The reference year is 2016.

2.4.5 Trade

We now move on to Mexican total exports and imports during the period spanning from 1999 to 2016. Figure 2.4 shows an overall increase of both imports and exports, except for the sharp drop experienced during the Great Recession of 2008-2009. Another interesting trend that we can detect is that the value of exports is comparable in size to the value of imports. On average, the bulk of exports represent about 97% of imports.

Figure 2.4: Bilateral Trade



Source: This figure presents Mexican exports and imports over the period 1999-2016.

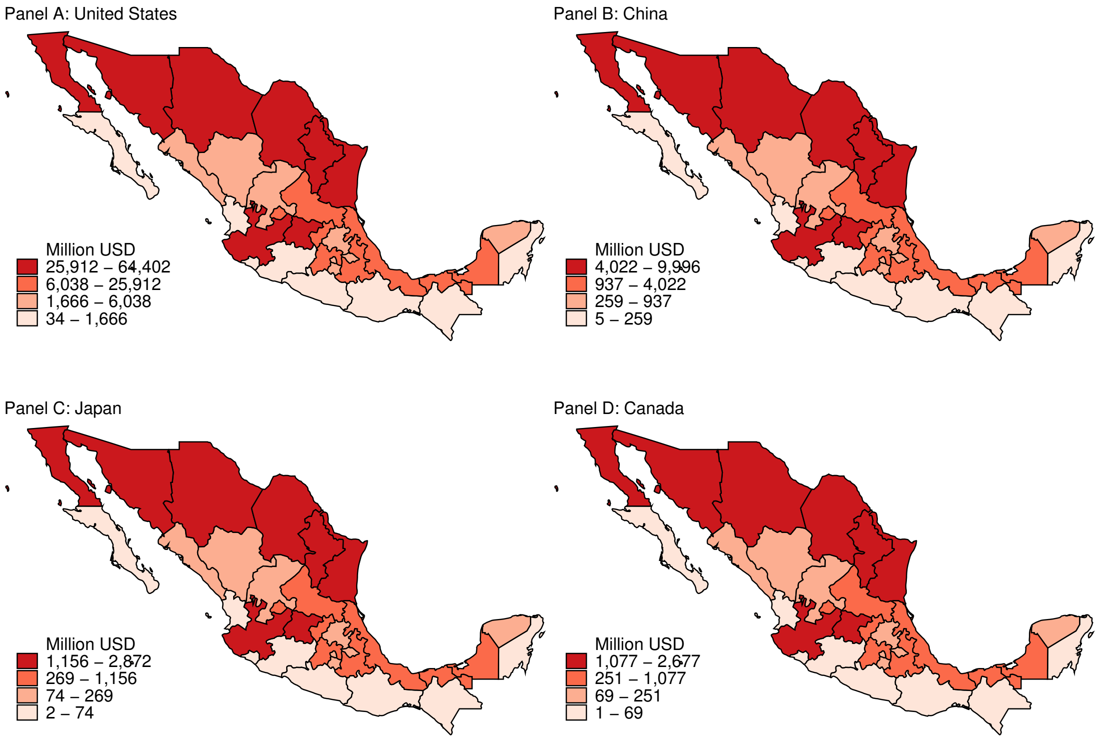
Figure 2.5 exhibits the distribution of total trade (i.e., the sum of imports and exports) of Mexico's four major trade partners in 2016: the United States (482.7 billion USD), China (74.9 billion USD), Japan (21.5 billion USD), and Canada (20.1 billion USD).[20] This figure displays a clear pattern among trade partners. Two of Mexico's top trade partners are located in the North American region: the United States and Canada. The U.S. constitutes by far Mexico's major trade partner for several reasons. Some of these motivations are that both countries share a border; benefited from the former North American Free Trade Agreement (NAFTA) for around 20 years; and recently negotiated the United States-Mexico-Canada Agreement (USMCA). On the other hand, Canada also benefited from free trade with Mexico under the NAFTA agreement; negotiated the new USMCA agreement; and is also a member of the Comprehensive and Progressive Agreement for Trans-Pacific Partnership (TPP11).[21]

However, a surprising outcome is that two Asian countries are among Mexico's top trade partners: China and Japan. This is surprising due to the remoteness of these two countries leading to high transportation costs. Also, Mexico does not possess a free trade agreement with China. This means that imported goods from China pay tariffs. Despite these tariffs,

[20] We display the distribution of total trade as this variable was useful to construct the bilateral trade index, which is one of the independent variables.

[21] The member countries of the Comprehensive and Progressive Agreement for Trans-Pacific Partnership are: Australia, Brunei, Canada, Chile, Japan, Malaysia, Mexico, New Zealand, Peru, Singapore, and Vietnam.

Figure 2.5: Distribution of Total Trade



Source: These maps display the geographical distribution at the state level of total trade with Mexico's four major trade partners. The reference year is 2016.

it seems that firms still perceive profitable to trade with this Asian country. Overall, we can observe that both Asian countries display strong trade ties with the Mexican states located in the northern and central regions.

Finally, we can also notice that states intensively trading with these four partner countries are located in the northern and central regions of the country. A plausible explanation for the northern states is that these share a border with the United States. In terms of the states located in the central region, we can identify an important concentration of large firms in Jalisco, Guanajuato, Puebla, and Mexico City.[22]

[22]The Appendix section exhibits a map with the geographical distribution of firms.

2.5 Methodology

The methodology employed consists of a linear fixed effects model using country-state paired data formed by 47 major partner countries and all 32 Mexican states over 5-year periods. The aim of the study is to examine the effects of FDI and trade linkages on business cycle comovements. The baseline regression equation follows the methodology proposed by Kleinert et al. (2015); although, the main explanatory variable used in this chapter is FDI inflows as a proportion of country c in state r in period t.[23] Another difference is that we now include a time dimension, similar to Calderón et al. (2007) and Hsu et al. (2011), to exploit the panel structure of our database.

Thus, the baseline regression equation is defined as follows:

$$\rho_{crt} = \beta_1 \, pFDI_{crt} + \beta_2 \, BT_{crt} + \beta_3 \, DISSIM_{crt} + \nu_c + \nu_r + \nu_t + \varepsilon_{crt}, \tag{2.7}$$

where ρ_{crt} corresponds to a vector of GDP growth rate correlations between country c and Mexican state r in period t. We constructed the dependent variable using Hamilton regression logged filtered data. The main explanatory variable is $pFDI_{crt}$, which stands for FDI inflows from country c to Mexican state r in period t as a ratio of FDI received by the host state r in period t.

From the literature review, we motivate the inclusion of the following two control variables: bilateral trade index (BT_{crt}) and a dissimilarity index between country-state export structures ($DISSIM_{crt}$). We also include country (ν_c), state (ν_r), and period (ν_t) fixed effects, as well as a disturbance term (ε_{crt}). Fixed effects are used to allow for potential endogeneity, due to an omitted variable bias. Thus, state fixed effects focus on the within-state variation, while country fixed effects focus on the within-country variation. These fixed effects are helpful in controlling for certain unobservable characteristics that could also influence the allocation of FDI and trade flows.

2.6 Results

Table 2.2 presents the baseline results for Eq.(2.7). The methodology employed consists of a linear fixed effects model. The dependent variable corresponds to the vector of GDP growth rate correlations using Hamilton logged filtered data of country-state paired observations over 5-year periods. All the right-hand side variables are expressed in mean values. Finally, we also include country, state, and period fixed effects to control for potential omitted variable bias (i.e., unobservable characteristics).

In column (1), we start by exploring the impact of FDI on business cycle comovements; this regression includes country, state, and period fixed effects. We can observe that FDI has

[23] Kleinert et al. (2015) use share of employment by foreign affiliates of country c in region r as the main explanatory variable. Moreover, the authors employ cross-section regressions.

Table 2.2: FDI and Business Cycle Comovements

Dependent Variable ρ_{crt}	(1) Ham_GDP_5	(2) Ham_GDP_5	(3) Ham_GDP_5	(4) Ham_GDP_5	(5) Ham_GDP_5	(6) Ham_GDP_5
FDI as a Proportion	0.3404**			0.3669**	0.3497**	0.3735**
	(0.1431)			(0.1450)	(0.1430)	(0.1449)
Bilateral Trade Index		231.5735***		240.0178***		221.5138***
		(58.4496)		(59.7621)		(58.5798)
Dissimilarity Index			0.0958**		0.0967**	0.0908**
			(0.0392)		(0.0390)	(0.0391)
Observations	3,008	3,008	3,008	3,008	3,008	3,008
R-squared	0.159	0.161	0.162	0.162	0.163	0.166
Country-State Pairs	1,504	1,504	1,504	1,504	1,504	1,504
Country FE	YES	YES	YES	YES	YES	YES
State FE	YES	YES	YES	YES	YES	YES
Period FE	YES	YES	YES	YES	YES	YES

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

Notes: This table constitutes the baseline to analyze the impact of FDI and trade on business cycle comovements between partner country c and Mexican state r in time period t. The table reports the estimates from Eq.(2.7) using a linear fixed effects model. The estimation sample is composed of country-state paired observations over 5-year periods spanning from 2005 to 2014. The dependent variable corresponds to 5-year GDP growth rate correlations between country c and Mexican state r in period t using Hamilton regression logged filtered data. The main explanatory variable is defined as FDI inflows from country c to Mexican state r in period t as a proportion of FDI inflows received in Mexican state r in period t. All independent variables are expressed in mean values of 5-year periods. All regressions include country, state, and period fixed effects.

a positive and statistically significant effect on business cycle synchornizations for a developing country. These results for a developing country are aligned to the empirical literature for developed countries (Hsu et al. 2011, Jansen & Stokman 2014). Moving on to the interpretation, our results suggest that a unit increase in FDI raises the business cycle comovements between partner countries and Mexican states by about 0.003.

In column (2), we proceed to investigate the impact of bilateral trade on business cycle comovements; this column also includes the full set of country, state, and period fixed effects. These results suggest that bilateral trade has a positive and strong effect on the synchronization of business cycles. This is in line with the empirical literature (Clark & Van Wincoop 2001, Imbs 2004, Calderón et al. 2007, Hsu et al. 2011, Jansen & Stokman 2014, Kleinert et al. 2015). Our results suggest that a unit increase in the bilateral trade index raises the business cycle comovements between partner countries and Mexican states by about 2.32.

In column (3), we now study the impact of having a different export structure between countries and Mexican states. Our results suggest that a dissimilar export structure has a positive and statistically significant effect on business cycle comovements. This dissimilarity in export structures could indicate export complementarities between countries and Mexican states. Columns (4)-(5) include the main explanatory variable (i.e., FDI inflows) and incorporate control variables separately. From these columns, we can notice that FDI remains positive and statistically significant in both columns.

Finally, column (6) represents our baseline specification where we include all the independent variables in our regression equation, as well as the full set of country, state, and period fixed effects. We can observe that FDI remains positive and statistically significant once we include all control variables. Our results suggest that a unit increase in FDI raises the business cycle comovements between partner countries and Mexican states by about 0.004.

Compared to the literature, our results are aligned to Hsu et al. (2011), where FDI has a positive and statistically significant effect on business cycle synchronizations for a single equation estimation employing a linear fixed effects model. Our results are also aligned to Jansen & Stokman (2014); although, the authors employ pooled cross-section regressions and their results are sensitive to sub-periods.

2.7 Robustness Analysis

In this section, we present a series of robustness checks. We begin by dealing with zero-value observations. Then, we use an alternative measure of FDI, which we define as FDI as a share of GDP. After that, we use an alternative specification of the filtering technique employed to calculate the dependent variable. Next, we exclude influential investment countries from our estimation sample. Then, we divide our sample into four sub-samples based on income groups of countries and states (i.e., countries are divided into high-income OECD countries and low- and middle-income countries, while Mexican states are divided in high-income states and low- and middle-income states). Later, we explore the dynamic effect of the dependent variable. Finally, we use a Two-Stage Least Squares model to discard potential endogeneity issues.

2.7.1 Zero-Value Observations

This subsection deals with the issue of having an excess of zero-value observations in the database. This issue is common in trade and FDI datasets, as some countries may not trade or invest in all the possible partner countries. In this specific study, not all Mexican states capture trade transactions nor investments from all partner countries. The main concern with zero-value observations is that these may lead to distorted estimates for semi-log models. Thus, to deal with zero-value observations we follow a standard method in the trade literature, which consists of discarding zero-value observations as in Kleinert et al. (2015).[24]

Table 2.3 shows the results after excluding the observations that do not report FDI values due to a lack of investment from certain partner countries to specific Mexican states. We can notice a significant reduction of around 65% of the observations compared to the baseline results reported in Table 2.2. Columns (1)-(2) and (4) show that the magnitude of the coefficients of FDI are slightly smaller after discarding zero-value observations compared to Table 2.2. These slightly smaller coefficients are consistent with Kleinert et al. (2015), who also observe a reduction on the magnitude of the coefficients once they restricted their sample by approximately 80% of the observations. Overall, we can observe that our results hold and FDI remains positive and statistically significant after discarding zero-value observations.

[24]Another interesting approach to deal with zero-value observations is to use the Poisson Pseudo-Maximum-Likelihood (PPML) estimator proposed by Santos Silva & Tenreyro (2006). Nonetheless, the PPML estimator is helpful when the dependent variable contains a significant amount of zero-value observations, which is not the case of our dependent variable (i.e., GDP growth rate correlations).

Table 2.3: Zero-Value Observations

Dependent Variable ρ_{crt}	(1) Ham_GDP5	(2) Ham_GDP5	(3) Ham_GDP5	(4) Ham_GDP5
FDI as a Proportion	0.3373**	0.3639**	0.3500**	0.3708**
	(0.1469)	(0.1494)	(0.1467)	(0.1490)
Bilateral Trade Index		244.6254***		197.2583**
		(90.1369)		(90.8002)
Dissimilarity Index			0.1261***	0.1203***
			(0.0422)	(0.0424)
Observations	1,955	1,955	1,955	1,955
R-squared	0.228	0.231	0.238	0.240
Country-State Pairs	1,116	1,116	1,116	1,116
Country FE	YES	YES	YES	YES
State FE	YES	YES	YES	YES
Period FE	YES	YES	YES	YES

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

Notes: This table reports the results after discarding zero-value observations from the estimation sample. This sample is composed of country-state paired observations over 5-year periods spanning from 2005 to 2014. The methodology employed consists of a linear fixed effects model. The dependent variable corresponds to 5-year GDP growth rate correlations between country c and Mexican state r in period t using Hamilton regression logged filtered data. The main explanatory variable is defined as FDI inflows from country c to Mexican state r in period t as a proportion of FDI inflows received in Mexican state r in period t. All independent variables correspond to 5-year average values. All regressions include country, state, and period fixed effects.

2.7.2 Alternative Measure of FDI

We proceed to estimate Eq.(2.7) employing FDI as a share of state GDP along the lines of Jansen & Stokman (2014) and combined with the time element proposed by Calderón et al. (2007) and Hsu et al. (2011):

$$sFDI_{crt} = \frac{FDI_{crt}}{GDP_{rt}}, \tag{2.8}$$

where FDI_{crt} denotes FDI inflows from country c to state r in period t weighted over GDP_{rt}, which stands for state GDP in period t. Just as before, the dependent variable is the vector of GDP growth rate correlations of country-state paired data over period t using Hamilton logged filtered data.

Table 2.4 reports the results for this alternative measure of FDI. We can notice that the magnitude and the precision of the coefficients of FDI as a share of GDP are larger compared to the baseline results presented in Table 2.2. Despite these differences, FDI remains positive and strongly significant throughout the different columns. These results suggest that a unit increase in the share of FDI to GDP raises the business cycle comovements between partner countries and Mexican states by roughly 0.19.

Compared to the literature, the magnitude of the coefficient using this alternative measure of FDI is larger than in Jansen & Stokman (2014). However, the literature has shown that these results are sensitive to the methodology and period employed.

Table 2.4: Alternative Measure of FDI

Dependent Variable	(1)	(2)	(3)	(4)
ρ_{crt}	Ham_GDP5	Ham_GDP5	Ham_GDP5	Ham_GDP5
FDI as a share of GDP	18.0771***	19.1429***	18.0031***	18.9910***
	(5.7714)	(5.7047)	(5.6783)	(5.6198)
Bilateral Trade Index		238.4776***		219.9931***
		(60.0693)		(58.9862)
Dissimilarity Index			0.0957**	0.0897**
			(0.0390)	(0.0391)
Observations	3,008	3,008	3,008	3,008
R-squared	0.159	0.162	0.163	0.166
Country-State Pairs	1,504	1,504	1,504	1,504
Country FE	YES	YES	YES	YES
State FE	YES	YES	YES	YES
Period FE	YES	YES	YES	YES

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

Notes: This table reports the results after using an alternative measure of FDI as the main explanatory variable. This main explanatory variable is now defined as FDI inflows from country c to Mexican state r in period t as a share of Mexican state r GDP in period t. The estimation sample is composed of country-state paired observations over 5-year periods spanning from 2005 to 2014. The methodology employed consists of a linear fixed effects model. The dependent variable corresponds to 5-year GDP growth rate correlations between country c and Mexican state r in period t using Hamilton regression logged filtered data. All independent variables correspond to 5-year average values. All regressions include country, state, and period fixed effects.

2.7.3 Alternative Filtering Technique Specification

In this sub-section, we use a different specification of the GDP time series data employed before applying the Hamilton regression filter. Instead of using logged data of these GDP time series, we now use standardized data. After this process, we apply the Hamilton regression filter to the standardized data.

Table 2.5 displays the results for this alternative specification of the Hamilton regression filter using standardized data. In column (1), we can observe FDI is positive and strongly statistically significant. Compared to the baseline table, the magnitude of the FDI coefficient is larger now. Columns (2)-(4) include control variables in a step wise manner. After including all the control variables, we can notice that FDI remains positive and strongly significant. In terms of the magnitude of the coefficient, FDI has now a larger coefficient compared to the baseline results exhibited in Table 2.2.

2.7.4 Influential Investors

Next, we analyze our results by excluding influential investors from our estimation sample: the United States, Spain, Germany, and Canada. Thus, we want to explore whether our results hold even if we do not include these countries in our sample. We employ linear regressions with fixed effects, where the dependent variable corresponds to GDP growth rate correlations using Hamilton regression logged filtered data.

Table 2.6 displays the results for the estimation sample excluding the United States, Spain, Germany, and Canada; each column excludes only one country. All columns, except for column (2), suggest that even if we exclude the United States, Germany, and Canada from our

Table 2.5: Alternative Filtering Technique Specification

Dependent Variable ρ_{crt}	(1) Ham_STD5	(2) Ham_STD5	(3) Ham_STD5	(4) Ham_STD5
FDI as a Proportion	0.3836***	0.4096***	0.3928***	0.4161***
	(0.1424)	(0.1443)	(0.1424)	(0.1442)
Bilateral Trade Index		235.1046***		216.7850***
		(59.6770)		(58.5965)
Dissimilarity Index			0.0957**	0.0899**
			(0.0388)	(0.0389)
Observations	3,008	3,008	3,008	3,008
R-squared	0.140	0.143	0.144	0.147
Country-State Pairs	1,504	1,504	1,504	1,504
Country FE	YES	YES	YES	YES
State FE	YES	YES	YES	YES
Period FE	YES	YES	YES	YES

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

Notes: This table reports the results after employing an alternative specification of the Hamilton regression approach to filter data. The estimation sample is composed of country-state paired observations over 5-year periods spanning from 2005 to 2014. The methodology employed consists of a linear fixed effects model. The dependent variable corresponds to 5-year GDP growth rate correlations between country c and Mexican state r in period t using the Hamilton regression filter on standardized data. The main explanatory variable is defined as FDI inflows from country c to Mexican state r in period t as a proportion of FDI inflows received in Mexican state r in period t. All independent variables correspond to 5-year average values. All regressions include country, state, and period fixed effects.

Table 2.6: Influential Investors

Dependent Variable ρ_{crt}	(1) Excl USA	(2) Excl ESP	(3) Excl DEU	(4) Excl CAN
FDI as a Proportion	0.6425***	0.1855	0.3633**	0.4739***
	(0.1848)	(0.1324)	(0.1508)	(0.1630)
Bilateral Trade Index	246.1459***	245.1850***	229.3743***	232.3778***
	(68.1493)	(59.0777)	(59.1824)	(60.8004)
Dissimilarity Index	0.0911**	0.0869**	0.0946**	0.0943**
	(0.0398)	(0.0396)	(0.0397)	(0.0397)
Observations	2,944	2,944	2,944	2,944
R-squared	0.168	0.159	0.172	0.166
Country-State Pairs	1,472	1,472	1,472	1,472
Country FE	YES	YES	YES	YES
State FE	YES	YES	YES	YES
Period FE	YES	YES	YES	YES

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

Notes: This table reports the results after excluding influential investors from the estimation sample. Column (1) reports the results after excluding the United States; column (2) excludes Spain; column (3) excludes Germany; and column (4) excludes Canada. The estimation sample is composed of country-state paired observations over 5-year periods spanning from 2005 to 2014. The methodology employed consists of a linear fixed effects model. The dependent variable corresponds to 5-year GDP growth rate correlations between country c and Mexican state r in period t using Hamilton regression logged filtered data. The main explanatory variable is defined as FDI inflows from country c to Mexican state r in period t as a proportion of FDI inflows received in Mexican state r in period t. All independent variables correspond to 5-year average values. All regressions include country, state, and period fixed effects.

estimation sample, FDI remains positive and statistically significant as in our baseline results reported in Table 2.2.

Interestingly, column (2) shows that once we exclude Spain from the estimation sample, the magnitude of the FDI coefficient becomes insignificant. This suggests that FDI has no effect on business cycle comovements when we exclude Spain from the sample. We could interpret this as Spain mainly driving the effects of FDI on business cycle synchronizations. This might seem a bit counterintuitive, as we would expect the United States to be the main driver of this effect. A plausible explanation might be that Spain is intensively investing in a larger number of Mexican states during the examined periods as shown by Figure 2.3. The higher level of FDI inflows may be motivated by similarities in tastes and preferences due to historical linkages and common language. Thus, Mexico represents an interesting entry point to the Latin American market for Spain.

In terms of the control variables, bilateral trade and the dissimilarity index remain positive and statistically significant as in our baseline results presented in Table 2.2.

2.7.5 Income Profile

Similar to Calderón et al. (2007), we divide our sample into four sub-samples based on the income profile of partner countries and Mexican states. Our two categories of partner countries are high-income OECD countries and low- and middle-income countries.[25] On the other hand, we define two categories of Mexican states: high-income states and low- and middle-income states.[26] We are now interested in analyzing whether the relationship between FDI and business cycle comovements is stronger or weaker depending on the income profile of paired country-states.

Table 2.7 reports the results of our sub-samples based on income groups. The methodology used consists of linear fixed effects regressions. The dependent variable corresponds to GDP growth rate correlations of country-state paired observations. In column (1), we can notice that FDI has no impact on business cycle synchronizations for the combination of high-income OECD countries and high-income states. Turning now to column (2), we can observe that FDI has a positive, albeit weakly significant impact on business cycle comovements for our combination of high-income OECD countries and low- and middle-income Mexican states. A plausible explanation may be that low- and middle-income states are starting to attract FDI inflows from high-income OECD countries.

Interestingly, column (3) shows that FDI has a positive and statistically significant effect

[25] High-income countries are defined by the OECD as member countries with a GNI per capita income above 12,236 U.S. dollars in 2016; thus, our sub-sample of high-income OECD countries includes Australia, Austria, Belgium, Canada, Chile, Czech Republic, Denmark, Finland, France, Germany, Hong Kong, Ireland, Israel, Italy, Japan, South Korea, Luxembourg, Netherlands, New Zealand, Norway, Panama, Poland, Portugal, Singapore, Spain, Sweden, Switzerland, United Kingdom, United States, and Uruguay.

The rest of the countries in our sample are included in the low- and middle-income country sub-sample; therefore, our sub-sample of low- and middle-income countries comprised of Argentina, Belize, Brazil, China, Colombia, Costa Rica, Ecuador, El Salvador, Guatemala, India, Indonesia, Malaysia, Nicaragua, Peru, Philippines, Russia, and South Africa.

[26] We define high-income states as those with a GDP per capita larger than 10,000 U.S. dollars in 2016; thus, our sub-sample of high-income Mexican states includes Baja California Sur, Campeche, Mexico City, Coahuila, Nuevo Leon, Queretaro, Sonora, and Tabasco.

The rest of the states in our sample are included in the low- and middle-income state sub-sample; therefore, our sub-sample of low- and middle-income Mexican states comprised of Aguascalientes, Baja California, Chiapas, Chihuahua, Colima, Durango, Estado de Mexico, Guanajuato, Guerrero, Hidalgo, Jalisco, Michoacan, Morelos, Nayarit, Oaxaca, Puebla, Quintana Roo, San Luis Potosi, Sinaloa, Tamaulipas, Tlaxcala, Veracruz, Yucatan, and Zacatecas.

Table 2.7: Income Profile

Dependent Variable ρ_{crt}	(1) INDc - INDr	(2) INDc - DEVr	(3) DEVc - INDr	(4) DEVc - DEVr
FDI as a Proportion	0.5894	0.2951*	1.2001**	-1.1217
	(0.3700)	(0.1582)	(0.5570)	(0.9861)
Bilateral Trade Index	-244.4893	-22.1480	499.4358***	262.8436***
	(149.4229)	(158.0382)	(162.2786)	(91.5837)
Dissimilarity Index	0.4712**	0.0688	-0.0972	0.1094
	(0.1946)	(0.0432)	(0.2947)	(0.0879)
Observations	480	1,440	272	816
R-squared	0.168	0.192	0.125	0.167
Country-State Pairs	240	720	136	408
Country FE	YES	YES	YES	YES
State FE	YES	YES	YES	YES
Period FE	YES	YES	YES	YES

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

Notes: This table reports the results for the different combinations of sub-samples based on income levels. Column (1) displays the results for paired high-income OECD countries c (IND_c) and high-income Mexican states r (IND_r). Column (2) shows the results for paired high-income OECD countries c (IND_c) and low- and middle-income Mexican states r (DEV_r). Column (3) exhibits the results for paired low- and middle-income countries c (DEV_C) and high-income Mexican states r (IND_r). Column (4) reports the results for paired low- and middle-income countries c (DEV_c) and low- and middle-income Mexican states r (DEV_r). The estimation sample is composed of country-state paired observations over 5-year periods spanning from 2005 to 2014. The methodology employed consists of a linear fixed effects model. The dependent variable corresponds to 5-year GDP growth rate correlations between country c and Mexican state r in period t using Hamilton regression logged filtered data. The main explanatory variable is defined as FDI inflows from country c to Mexican state r in period t as a proportion of FDI inflows received in Mexican state r in period t. All independent variables correspond to 5-year average values. All regressions include country, state, and period fixed effects.

on the transmission of business cycles between low- and middle-income countries and high-income Mexican states. This effect between developing countries and developed states could be driven by the increasing participation of developing countries in global value chains. Within Mexico, developed states, mainly located in the border with the United States, offer a strategic geographical advantage, as well as skilled-labor, which attract more FDI inflows.

These results suggest that a unit increase in FDI inflows reduces the business cycle comovements between low- and middle-income countries and high-income Mexican states by about 0.01. In column (4), we can notice that FDI has no impact on business cycle comovements between low- and middle-income countries and low- and middle-income states.

In terms of bilateral trade, our results are in line with Calderón et al. (2007), who also found a positive relationship between bilateral trade and business cycle comovements between pairs of developed countries.[27] Our results suggest that most of the bilateral trade occurs between developing countries and developing states, but especially with developed Mexican states. A plausible explanation for this trade concentration between developing countries and Mexican states is their participation in global value chains.

[27]We are not able to compare our results for the relationship between FDI and business cycle comovements using the different pairwise combinations of income profiles because FDI was not considered in Calderón et al. (2007).

2.7.6 Dynamic Effects

We now proceed to investigate the potential dynamic effect of GDP_{ct-1} affecting GDP_{rt} through FDI and trade. For this specification, the dependent variable corresponds to the GDP growth rate correlations between country c in period $t-1$ and Mexican state r in period t using Hamilton logged and filtered data. The methodology consists of a linear fixed effects model. All regressions include country, state, and period fixed effects.

Table 2.8: Dynamic Effect

Dependent Variable $\rho(GDP_{ct-1}, GDP_{rt})$	(1) Ham_GDP5_	(2) Ham_GDP5_	(3) Ham_GDP5_	(4) Ham_GDP5_
FDI as a Proportion	1.0040***	1.0331***	0.9909***	1.0226***
	(0.2838)	(0.2872)	(0.2819)	(0.2854)
Bilateral Trade Index		263.9828**		293.4666**
		(120.4066)		(118.8682)
Dissimilarity Index			-0.1367***	-0.1446***
			(0.0444)	(0.0446)
Observations	3,008	3,008	3,008	3,008
R-squared	0.059	0.061	0.063	0.066
Country-State Pairs	1,504	1,504	1,504	1,504
Country FE	YES	YES	YES	YES
State FE	YES	YES	YES	YES
Period FE	YES	YES	YES	YES

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

Notes: This table investigates the potential dynamic effect. The estimation sample is composed of country-state paired observations over 5-year periods spanning from 2005 to 2014. The methodology employed consists of a linear fixed effects model. The dependent variable corresponds to 5-year GDP correlations between country c in period $t-1$ and Mexican state r in period t; both series were logged and detrended using the Hamilton regression filter. The main explanatory variable is defined as FDI inflows from country c to Mexican state r in period t as a proportion of FDI inflows received in Mexican state r in period t. All independent variables correspond to 5-year average values. All regressions include country, state, and period fixed effects.

Table 2.8 examines the potential dynamic effect. In column (1), we can notice FDI has a positive and strong statistically effect on the transmission of the business cycle from country c in the previous period $t-1$ to Mexican state r in period t; in other words, the business cycle propagates with a lagged effect. In columns (2)-(3), we can observe the effect of FDI remains positive and strongly significant. Column (4) also shows FDI has a positive and strong impact on the transmission of business cycles with a lagged effect. Furthermore, bilateral trade also have a positive and strong impact on the propagation of business cycles with a lag.

2.7.7 Two-Stage Regressions

A potential problem with our empirical model is that of endogeneity arising from reverse causality and an omitted variable bias. To tackle this potential endogeneity problem, we make use of a Two-Stage Least Squares (2SLS) model. In a similar vein as Jansen & Stokman (2014), we employ similar instrumental variables; therefore, we use having a free trade agreement (FTA), as well as gravity variables, such as landlocked status, common continent, and common colonizer. The aim of using instrumental variables is that these instruments do not impact business cycle comovements directly, but rather have an indirect impact through FDI inflows.

Regarding the model specification of these 2SLS regressions, it is worth mentioning that due to the time-invariant nature of the instruments, we cannot include country fixed effects as these observations will drop out; despite this, we include state and period fixed effects in our specification. Furthermore, we do not include bilateral trade and the dissimilarity index as these control variables may be endogenous as well.

Table 2.9 reports the results of our 2SLS regressions. These regressions include landlocked status, continent, colonizer, and FTA as instrumental variables. The top panel reports the second stage of the IV regressions and the bottom panel exhibits the first stage of the regressions.

Table 2.9: 2SLS Regressions

	(1)	(2)	(3)	(4)	(5)
	Second Stage	Second Stage	Second Stage	Second Stage	Second Stage
Second Stage Regressions	Ham_GDP5	Ham_GDP5	Ham_GDP5	Ham_GDP5	Ham_GDP5
FDI as a Proportion	-3.2112***	0.5144	-0.4340	3.1459***	0.4246**
	(0.8733)	(0.5055)	(0.5618)	(0.5909)	(0.1678)
Observations	3,008	3,008	3,008	3,008	3,008
State FE	YES	YES	YES	YES	YES
Period FE	YES	YES	YES	YES	YES
Adj. R2	0.059	0.452	0.439	0.174	0.453
Underidentification stat.	40.863	20.404	77.217	75.525	100.279
Prob underident. stat.	0.000	0.000	0.000	0.000	0.000
Weak identification stat.	55.131	20.615	86.334	80.370	27.370
Endogeneity F-test	19.007	0.314	1.483	35.525	0.328
Prob endogeneity test	0.000	0.576	0.223	0.000	0.567
Hansen J-statistic					45.721
Prob Hansen J-stat.					0.000
	First Stage	First Stage	First Stage	First Stage	First Stage
First Stage Regressions	FDI_p_5	FDI_p_5	FDI_p_5	FDI_p_5	FDI_p_5
Landlocked	-0.0192***				-0.0284***
	(0.0026)				(0.0030)
Continent		0.0271***			0.0996***
		(0.0060)			(0.0149)
Colonizer			-0.0261***		-0.1125***
			(0.0028)		(0.0148)
FTA				0.0286***	0.0331***
				(0.0032)	(0.0037)

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

Notes: This table reports the results after employing a Two-Stage Least Squares (2SLS) model. The estimation sample is composed of country-state paired observations over 5-year periods spanning from 2005 to 2014. The instrumental variables employed are landlocked status, continent, colonizer, and free trade agreement. The top panel shows the second stage IV regressions, while the bottom panel exhibits the first stage regressions. The dependent variable on the top panel corresponds to GDP growth rate correlations between country c and Mexican state r in period t using Hamilton regression logged filtered data.

From the first stage, we individually evaluate the impact of each of the instrumental variables on FDI inflows; the last column shows the impact of all the instruments on FDI in the first stage. Our results from columns (1) and (3) suggest that being landlocked and having a common colonizer have a negative and strong significant impact on FDI inflows. On the other hand, columns (2) and (4) show that sharing the same continent and having a free trade agreement have a positive and strong significant effect on FDI. Finally, column (5) shows that being landlocked, locating in the same continent, having a common colonizer, and having a free trade agreement have a strong statistically significant effect on FDI inflows.

For the second stage, we focus now on column (5), which incorporates all the instruments. We can observe that FDI has a positive and statistically significant effect on business cycle comovements. These results suggest that a unit increase in FDI raises the business cycle comovements between partner countries and Mexican states by about 0.004. In comparison to the model where exogeneity was assumed exhibited in Table 2.2, the results are similar.

We also report a series of tests. We can confirm that we do not face any issue of under-identification. Also, our F-statistic for the weak identification test is larger than the critical values; thus, our instruments have good explanatory power for the endogenous variable. Then, we use an endogeneity test to test whether our suspected endogenous variable can be treated as exogenous. This diagnostic test shows that the regressor can be treated as exogenous in column (5).

Finally, we also report the Hansen J-statistic for overidentification of the instruments; this statistic tests that the instruments are not correlated with the error term of the model. The null hypothesis is that the instruments are valid. In this case, we reject the null hypothesis and conclude that the instrumental variables are not valid instruments as these may be correlated with the error term. This suggests that we need to take these results with caution as our estimates may still be biased.[28]

2.8 Conclusions

The previous empirical literature has extensively analyzed the relationship between different trade determinants and business cycle comovements; however, little attention has been given to FDI, especially for developing countries. The main contribution of this chapter is to examine the effects of FDI inflows on business cycle comovements for an emerging economy. Our estimation sample consists of paired observations between 47 major partner countries and all 32 Mexican states using sub-periods to create a time element for the variables comprised in our model specification.

We employ a linear fixed effects model. As part of the methodology, we use the recent Hamilton regression filter to detrend GDP time series used to construct the dependent variable. Our results suggest that FDI inflows have a positive and statistically significant impact on business cycle comovements for an emerging economy. Furthermore, we confirm that trade has a positive and strong significant effect on business cycles synchronizations.

We also perform a series of robustness checks, which includes discarding zero-value observations, using an alternative measure of FDI, excluding influential investors, employing different sub-samples based on income groups, and exploring the dynamic effect of the dependent variable. We also make use of 2SLS regressions, employing landlocked status, continent, colonizer,

[28] We also present the 2SLS estimates employing distance and border as instrumental variables in the Appendix section; nonetheless, these two instruments are problematic because of their impact on business cycle synchronizations as shown by Clark & Van Wincoop (2001).

and free trade agreement as instrumental variables to discard potential endogeneity. Overall, we can conclude that FDI holds for all specifications.

From a policy perspective, the state disaggregation level used in this chapter allows us to gain a better understanding of the relationship between FDI inflows and the transmission of business cycles for an emerging economy with notable differences across states. This level of disaggregation is especially important in designing public policies aiming to boost the economic development for those states exhibiting growth rates below the national average through trade and FDI. Our results imply that FDI inflows are important determinants of business cycles synchronizations.

Furthermore, we can observe that Spain is mainly driving the impact of FDI on business cycle comovements. A plausible explanation may be that Spain seeks partner countries that share historical linkages and a common language, which facilitates establishing affiliates in Mexican states. Furthermore, Mexican states offer an interesting entry point to the United States and to the Latin American markets. Thus, a policy recommendation would be to continue diversifying markets and investment partners so that in case of an economic downturn of the Spanish economy, the transmission of business cycles smooths as Mexican states can rely on other partner countries.

Based on state heterogeneity, policymakers could explore the idea of creating industry clusters in Mexican states located in the southern region to promote their economic growth. Furthermore, public policies aiming to attract FDI to these southern states could target on emerging economies, as we can observe there is an increasing participation of these countries in global value chains. An advantage offered by Mexican southern states is their proximity to the Latin American market. Along with this, it would be interesting to expand the free trade agreement network to other emerging economies to promote and increase trade of intermediate inputs. As a result, Mexican states located in the southern region could benefit from FDI and trade spillover effects and could better integrate to global value chains. On the other hand, policymakers can keep public policies designed to maintain and reinforce FDI inflows in the northern states of Mexico.

Finally, the chapter presents certain limitations regarding data accessibility. These limitations include a restricted sample of partner countries investing in Mexican states, albeit this restricted sample represents 98% of FDI inflows. In terms of trade, accessing firm-level data was not possible due to confidentiality issues. Also, data on imports at the state level is not available due to the same confidentiality reasons; the authorities argue that large firms could be traced down with trade data at the state level.

In terms of methodology, a linear fixed effects model presents some drawbacks since we used 5-year periods to construct the time element of the dependent variable, which is GDP growth rate correlations between country-state paired observations. Therefore, a sample size of five paired observations is rather small to obtain meaningful correlations for the dependent variable. Furthermore, due to data limitations on the control variables, we ended up with a short panel; nonetheless, this panel estimation approach is preferred over a cross-section approach.

2.9 Appendix

2.9.1 Countries and States

Table 2.10: List of Partner Countries

Label	Name	Label	Name	Label	Name	Label	Name
ARG	Argentina	CZE	Czech Republic	IRL	Ireland	PER	Peru
AUS	Australia	DEU	Germany	ISR	Israel	PHL	Philippines (The)
AUT	Austria	DNK	Denmark	ITA	Italy	POL	Poland
BEL	Belgium	ECU	Ecuador	JPN	Japan	PRT	Portugal
BLZ	Belize	ESP	Spain	KOR	Korea (The Republic of)	RUS	Russian Federation
BRA	Brazil	FIN	Finland	LUX	Luxembourg	SGP	Singapore
CAN	Canada	FRA	France	MYS	Malaysia	SLV	El Salvador
CHE	Switzerland	GBR	United Kingdom	NIC	Nicaragua	SWE	Sweden
CHL	Chile	GTM	Guatemala	NLD	Netherlands (The)	URY	Uruguay
CHN	China	HKG	Hong Kong	NOR	Norway	USA	United States of America
COL	Colombia	IDN	Indonesia	NZL	New Zealand	ZAF	South Africa
CRI	Costa Rica	IND	India	PAN	Panama		

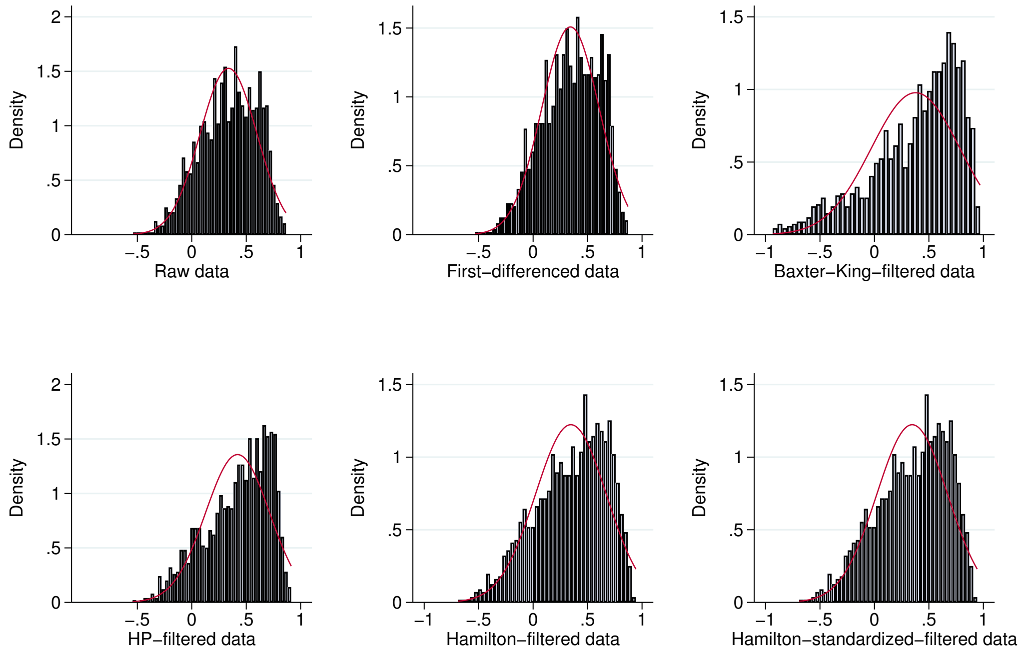
Figure 2.6: Map of the Administrative Divisions in Mexico

Table 2.11: List of Mexican States

Label	Name	Label	Name	Label	Name	Label	Name
MX-AGU	Aguascalientes	MX-COL	Colima	MX-MOR	Morelos	MX-SLP	San Luis Potosi
MX-BCN	Baja California	MX-DUR	Durango	MX-NAY	Nayarit	MX-SON	Sonora
MX-BCS	Baja California Sur	MX-GRO	Guerrero	MX-NLE	Nuevo Leon	MX-TAB	Tabasco
MX-CAM	Campeche	MX-GUA	Guanajuato	MX-OAX	Oaxaca	MX-TAM	Tamaulipas
MX-CHH	Chihuahua	MX-HID	Hidalgo	MX-PUE	Puebla	MX-TLA	Tlaxcala
MX-CHP	Chiapas	MX-JAL	Jalisco	MX-QUE	Queretaro	MX-VER	Veracruz
MX-CMX	Ciudad de Mexico	MX-MEX	Estado de Mexico	MX-ROO	Quintana Roo	MX-YUC	Yucatan
MX-COA	Coahuila	MX-MIC	Michoacan	MX-SIN	Sinaloa	MX-ZAC	Zacatecas

2.9.2 Additional Descriptive Statistics

Figure 2.7: Distribution of Business Cycle Comovements



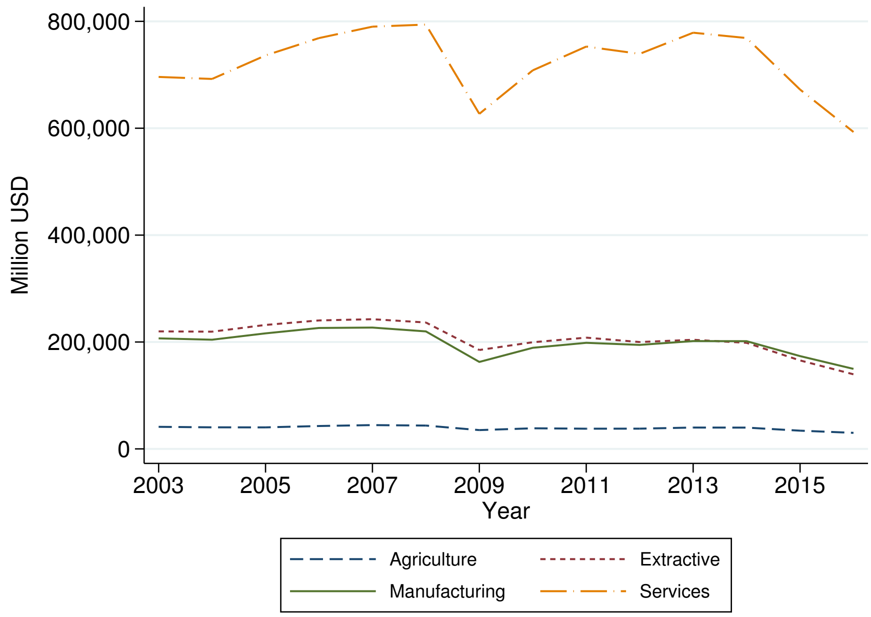
Notes: These distributions are defined across country-state paired observations over the period 1999-2016.

Table 2.12: Pairwise Correlations

Variables	(1) Raw_GD5	(2) D1_GDP	(3) BK_GDP	(4) FD_GDP	(5) HP_GDP	(6) Ham_GDP	(7) Ham_STD	(8) pFDI	(9) sFDI	(10) BT	(11) DISSIM
Correlation of GDP growth rates (Raw)	1.000										
Correlation of GDP growth rates (D1)	1.000	1.000									
	(0.000)										
Correlation of GDP growth rates (BK)	0.314	0.314	1.000								
	(0.000)	(0.000)									
Correlation of GDP growth rates (FD)	0.606	0.609	0.535	1.000							
	(0.000)	(0.000)	(0.000)								
Correlation of GDP growth rates (HP)	0.656	0.660	0.484	0.914	1.000						
	(0.000)	(0.000)	(0.000)	(0.000)							
Correlation of GDP growth rates (Ham)	0.469	0.470	0.582	0.634	0.620	1.000					
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)						
Correlation of GDP growth rates (Ham-STD)	0.477	0.478	0.591	0.631	0.625	0.991	1.000				
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)					
FDI inflows as a state proportion (pFDI)	0.023	0.023	0.026	0.023	0.023	0.043	0.044	1.000			
	(0.209)	(0.205)	(0.155)	(0.204)	(0.199)	(0.017)	(0.017)				
FDI inflows as a share of GDP (sFDI)	0.012	0.012	-0.024	-0.007	-0.007	0.008	0.008	0.783	1.000		
	(0.510)	(0.505)	(0.191)	(0.694)	(0.688)	(0.674)	(0.647)	(0.000)			
Bilateral Trade Index (BT)	0.011	0.011	0.148	0.085	0.072	0.128	0.116	0.313	0.250	1.000	
	(0.560)	(0.560)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		
Dissimilarity Index (DISSIM)	-0.110	-0.108	-0.045	-0.014	-0.006	-0.065	-0.080	-0.066	-0.054	0.017	1.000
	(0.000)	(0.000)	(0.014)	(0.452)	(0.722)	(0.000)	(0.000)	(0.000)	(0.003)	(0.363)	

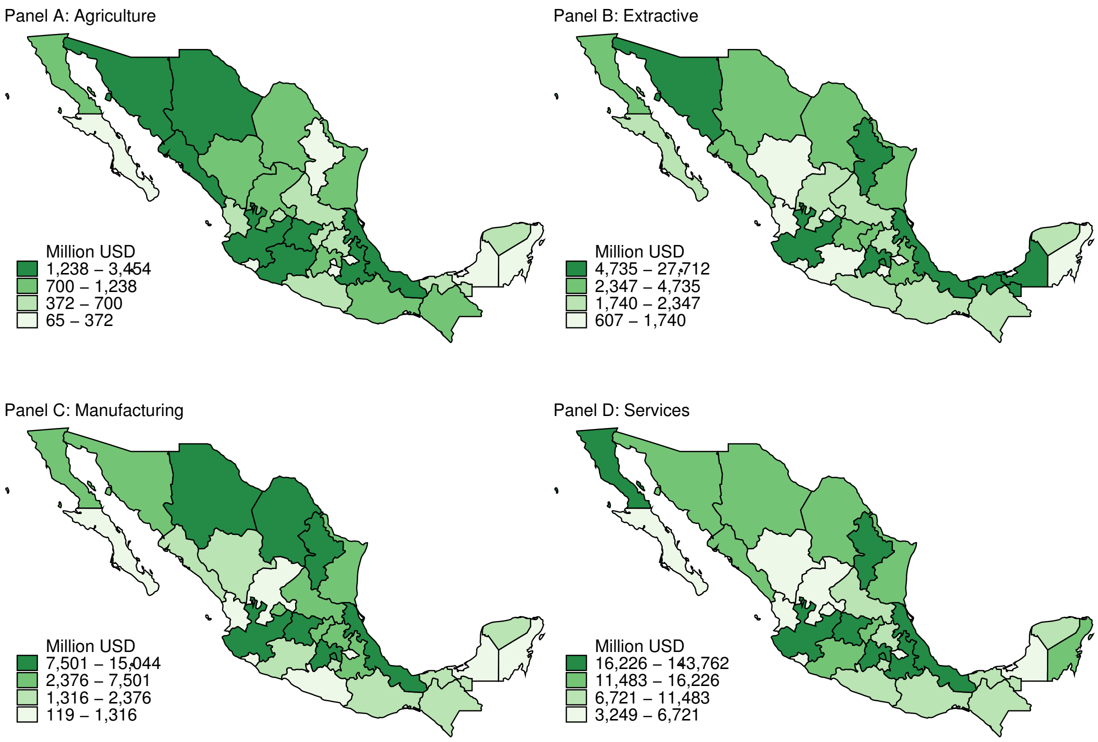
Notes: This table exhibits the paired correlations between the different measures of the dependent variable and the explanatory variables of our estimation sample.

Figure 2.8: GDP by Sector



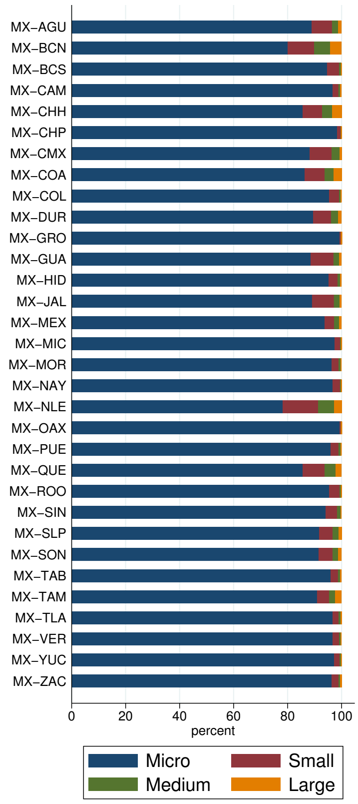
Source: This figure presents GDP by sector over the period 2003-2016. This period is slightly shorter due to access limitations in terms of data disaggregation by sector (i.e., NAICS-2 digit level).

Figure 2.9: Distribution of GDP



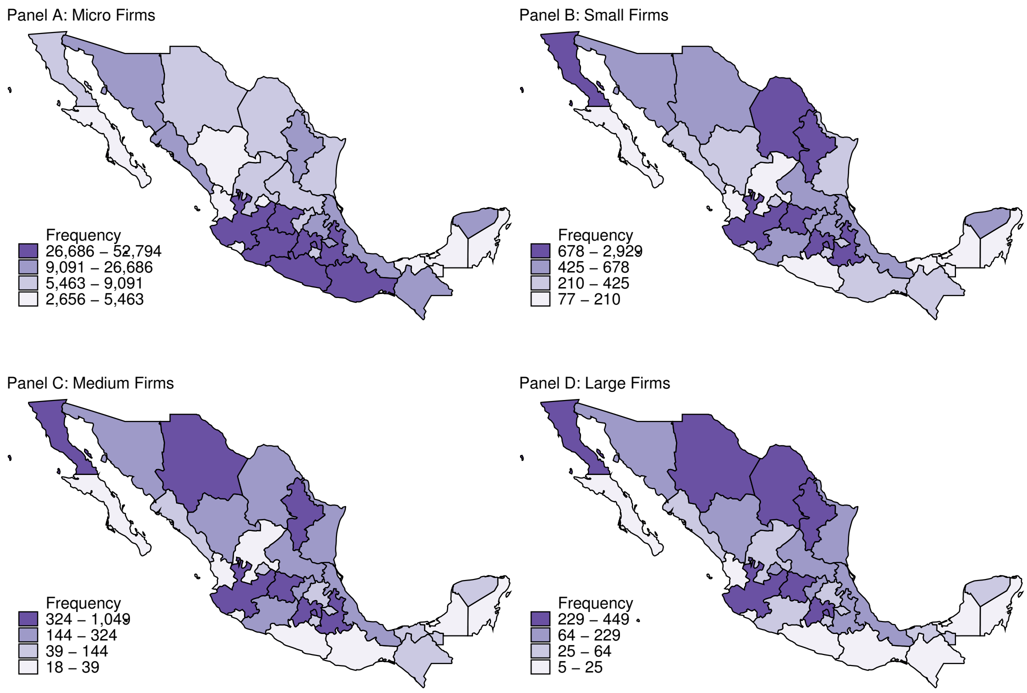
Source: These maps display the geographical distribution at the state level of GDP by sector. The reference year is 2016.

Figure 2.10: Firm Size



Source: This figure exhibits the firm size breakdown by state. This figure uses data from the DENUE dataset sourced from the INEGI. The reference year is 2016.

Figure 2.11: Distribution of Firms



Source: These maps show the geographical distribution at the state level of firm size. This figure employs data from the DENUE dataset sourced from the INEGI. The reference year is 2016.

2.9.3 Additional Regressions

Table 2.13: Cross-Section Regressions

Dependent Variable ρ_{cr}	(1) Raw	(2) D1	(3) BK	(4) FD	(5) HP	(6) Ham	(7) Ham-STD
FDI as a Proportion	0.0054	0.0035	-0.1872	-0.1978*	-0.1238	-0.1647	-0.1499
	(0.0917)	(0.0919)	(0.2100)	(0.1188)	(0.0892)	(0.2191)	(0.2119)
Bilateral Trade Index	42.2431**	43.2625***	52.7402**	57.2867***	47.2044***	46.5821**	50.0454**
	(16.5306)	(16.1187)	(23.8147)	(19.6299)	(15.4643)	(22.3540)	(22.3620)
Dissimilarity Index	-0.0314***	-0.0315***	-0.0171	-0.0149	-0.0218*	-0.0236	-0.0286*
	(0.0121)	(0.0120)	(0.0175)	(0.0181)	(0.0124)	(0.0165)	(0.0167)
Observations	1,504	1,504	1,504	1,504	1,504	1,504	1,504
R-squared	0.731	0.740	0.737	0.712	0.798	0.680	0.661
Country FE	YES	YES	YES	YES	YES	YES	YES
State FE	YES	YES	YES	YES	YES	YES	YES

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

Notes: This table examines the impact of FDI on business cycle comovements between partner countries c and Mexican states r employing a cross-section regression model. The estimation sample is composed by the mean values of country-state paired observations over the period 1999-2016. The dependent variable corresponds to annual GDP growth rate correlations using different measures: raw data (Raw), first difference (D1), Baxter-King filter (BK), frequency domain filter (FD), Hodrick-Prescott filter (HP), Hamilton regression filter applied to logged data (Ham), and Hamilton regression filter applied to standardized data (Ham-STD). The main explanatory variable is defined as the proportion of FDI inflows from country c to Mexican state r. All regressions include country and state fixed effects.

Table 2.14: Filtering Techniques

Dependent Variable ρ_{crt}	(1) Raw	(2) D1	(3) BK	(4) FD	(5) HP	(6) Ham-STD	(7) Ham
FDI as a Proportion	0.1975	0.1943	0.2048	0.1598	-0.0492	0.4161***	0.3735**
	(0.2401)	(0.2400)	(0.2299)	(0.2485)	(0.2201)	(0.1442)	(0.1449)
Bilateral Trade Index	-100.3997	-96.6163	421.7430***	279.2097***	344.7781***	216.7850***	221.5138***
	(102.0855)	(101.8998)	(95.0324)	(96.4589)	(85.9832)	(58.5965)	(58.5798)
Dissimilarity Index	0.0695*	0.0679*	0.0364	-0.0317	-0.0259	0.0899**	0.0908**
	(0.0384)	(0.0384)	(0.0462)	(0.0481)	(0.0419)	(0.0389)	(0.0391)
Observations	3,008	3,008	3,008	3,008	3,008	3,008	3,008
R-squared	0.392	0.399	0.036	0.475	0.509	0.147	0.166
Country-State Pairs	1,504	1,504	1,504	1,504	1,504	1,504	1,504
Country FE	YES	YES	YES	YES	YES	YES	YES
State FE	YES	YES	YES	YES	YES	YES	YES
Period FE	YES	YES	YES	YES	YES	YES	YES

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

Notes: This table reports the results after using different specifications of the dependent variable. The methodology employed consists of a linear fixed effects model. The estimation sample is composed of country-state paired observations over 5-year periods spanning from 2005 to 2014. The dependent variable corresponds to 5-year GDP growth rate correlations between country c and Mexican state r in period t using different measures: raw data (Raw), first differenced data (D1), Baxter-King filtered data (BK), frequency domain filtered data (FD), Hodrick-Prescott filtered data (HP), Hamilton regression filter data applied to standardized values (Ham-STD), and Hamilton regression filtered data (Ham). All these filtering techniques were applied to logged data. The main explanatory variable is defined as FDI inflows from country c to Mexican state r in period t as a proportion of FDI inflows received in Mexican state r in period t. All independent variables are expressed in mean values of 5-year periods. All regressions include country, state, and period fixed effects.

Table 2.15: 2SLS Regressions using an Alternative Instrument

Second Stage Regressions	(1) Second Stage Ham_GDP5	(2) Second Stage Ham_GDP5	(3) Second Stage Ham_GDP5
FDI as a Proportion	2.4811***	0.1359	0.3555**
	(0.6774)	(0.1829)	(0.1699)
Observations	3,008	3,008	3,008
State FE	YES	YES	YES
Period FE	YES	YES	YES
Adj. R2	0.288	0.454	0.454
Underidentification stat.	24.376	5.766	24.453
Prob underident. stat.	0.000	0.016	0.000
Weak identification stat.	21.859	11.645	8.276
Endogeneity F-test	22.628	0.335	0.036
Prob endogeneity test	0.000	0.563	0.850
Hansen J-statistic			25.389
Prob Hansen J-stat.			0.000
First Stage Regressions	**First Stage FDI_p_5**	**First Stage FDI_p_5**	**First Stage FDI_p_5**
Distance (in logs)	-0.0160***		-0.0082**
	(0.0034)		(0.0033)
Border		0.3125***	0.2969***
		(0.0916)	(0.0939)

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

Notes: This table reports the Two-Stage Least Squares (2SLS) regressions using distance and border as alternative instruments. The estimation sample is composed of country-state paired observations over 5-year periods spanning from 2005 to 2014. The top panel shows the second stage IV regressions, while the bottom panel exhibits the first stage regressions. The dependent variable on the top panel corresponds to GDP growth rate correlations between country c and Mexican state r in period t using Hamilton regression logged filtered data.

2.9.4 Filtering Techniques

This subsection presents the most common filtering techniques employed in the empirical literature on business cycle comovements. The aim of these techniques is to separate time series into trend and cyclical components. This subsection displays the equation form for each filtering technique.

First Order Differences

As described by Canova (1998), the first difference procedure is based on the assumptions that the resulting time series is a random walk, the cyclical component is stationary, and that these two components are uncorrelated. An additional assumption is that y_t has a unit root. The first difference is expressed as:

$$y_t = y_{t-1} + \varepsilon_t \tag{2.9}$$

the trend would be $x_t = y_{t-1}$ and c_t can be estimated as $\hat{c}_t = y_t - y_{t-1}$.

Baxter-King filter

According to StataCorp. (2017), which is based on Baxter & King (1999), the Baxter-King filter can be expressed by the following equations:

$$c_t = \sum_{j=-\infty}^{\infty} b_j y_{t-j}, \tag{2.10}$$

where c_t is the cyclical component of the time series y_t, and b_j corresponds to the coefficients of the sequence of some ideal filter.

Letting the minimum p_l and maximum p_h periods of the stochastic cycles of interest, the weights b_j in this estimation are as given by:

$$b_j = \begin{cases} \pi^{-1}(\omega_h - \omega_l) & \text{if} \quad j = 0 \\ (j\pi)^{-1}\{sin(j\omega_h) - sin(j\omega_l)\} & \text{if} \quad j \neq 0 \end{cases} \tag{2.11}$$

where $\omega_l = 2\pi/p_l$ and $\omega_h = 2\pi/p_h$ correspond to the lower and higher cutoff frequencies, respectively.

Thus, the filter estimates the cyclical component of the time series c_t in the following way:

$$c_t = \sum_{j=-q}^{+q} \hat{b}_j y_{t-j}. \tag{2.12}$$

The coefficients $\hat{b}_j$ are equal to $\hat{b}_j = b_j - \bar{b}_q$, where $\hat{b}_{-j} = \hat{b}_j$ and $\bar{b}_q$ stands for the mean of the ideal coefficients truncated at $\pm q$:

$$\bar{b}_q = (2q+1)^{-}1 \sum_{j=-q}^{q} b_j. \tag{2.13}$$

Frequency Domain filter

Following Corbae et al. (2002) and Ouliaris et al. (2014), the frequency domain filter can be defined by:

$$x_t = \Pi_2' z_t + \tilde{x}_t, \tag{2.14}$$

where x_t is an observed time series, z_t is a deterministic sequence, and $\tilde{x}_t$ is a stochastic (latent) component. The stochastic component of the data allows for the regression coefficient to vary across frequency bands:

$$\tilde{y}_t = \sum_{j=-\infty}^{\infty} \beta_j' \tilde{x}_{t-j} + \varepsilon_t = \beta(L)' \tilde{x}_t + \varepsilon_t, \tag{2.15}$$

where $\tilde{y}_t$ is a (latent) dependent variable and ε_t stands for unobserved disturbances.

Hodrick-Prescott filter

As described by Hamilton (2017), which is based on Hodrick & Prescott (1981,1997), the Hodrick-Prescott filter can be expressed as follows:

$$y_t = g_t + c_t, \tag{2.16}$$

where y_t is the natural log of a time series, g_t stands for the trend component, and c_t presents the deviations from the growth. The filter is calculated:

$$min\{g_t\}_{t-1}^{T} \left\{ \sum_{t=1}^{T} (y_t - g_t)^2 + \lambda \sum_{t=1}^{T} ((g_t - g_{t-1}) - (g_{t-1} - g_{t-2}))^2 \right\}, \tag{2.17}$$

where λ is the smoothing parameter.

Hamilton regression filter

Based on Hamilton (2017), the Hamilton regression filter is defined as the following OLS regression:

$$y_{t+h} = \beta_0 + \beta_1 y_t + \beta_2 y_{t-1} + \beta_3 y_{t-2} + \beta_4 y_{t-3} + \upsilon_{t+h}, \tag{2.18}$$

where y_{t+h} corresponds to the p=4 most recent values of the time series y_t, and h is the horizon. Thus, the residuals are:

$$\hat{\upsilon}_{t+h} = y_{t+h} - \hat{\beta}_0 - \hat{\beta}_1 y_t - \hat{\beta}_2 y_{t-1} - \hat{\beta}_3 y_{t-2} - \hat{\beta}_4 y_{t-3}. \tag{2.19}$$

2.9.5 Summary of Filtering Techniques

Table 2.16: Filtering Techniques: Advantages and Drawbacks

Filter	Description	Frequency	Advantages	Drawbacks
First difference (D1)	Detrending procedure based on first-order differencing resulting in a random walk.	D = 1 period	New values represent a move away from the original values.	Inappropriate technique in presence of autocorrelation with earlier periods.
Baxter-King filter (BK)	Band-pass filter that employs a moving average that drops observations on both ends of the time-series.	Number of periods equivalent to 3 years. • Monthly: k = 36 • Quarterly: k = 12 • Annual: k = 3	Good approximation to the optimal filter. Considers statistical features of business cycles.	Longer time-series drop observations on both ends of the original time-series.
Frequency domain filter (FD)	Band-pass filter that acts as an alternative technique to the Baxter-King filter, without dropping observations.	Oscillation periods between 1.5 and 8 years. • M: s=18 & e=96 • Q: s=6 & e=32 • A: s=2 & e=8	Does not drop observations.	Resulting time-series depends on the sample size.
Hodrick-Prescott filter (HP)	High-pass filter widely employed in the empirical literature of business cycle comovements.	The smoothing parameter is λ. • Monthly: λ = 129,600 • Quarterly: λ = 1,600 • Annual: λ = 6.25	Can be applied to non-stationary time-series.	The assumed magnitude of λ is questionable, which might generate biased estimates where the cyclical component is part of the trend; thus, displaying the dynamics of the filter itself and incurring on spurious cycles.
Hamilton regression filter (Ham)	OLS regression filtering technique that acts as an alternative to the HP filter.	Number of periods equivalent to 2 years. • Monthly: h = 24 • Quarterly: h = 8 • Annual: h = 2	Robust approach as the HP filter, without the same drawbacks.	Based on assumptions of the detrended component's characteristics. Might modify the original cyclical structure of the time-series.

Notes: This table is a compilation of different sources from the empirical literature (Canova 1998, Corbae et al. 2002, Ouliaris et al. 2014, StataCorp. 2017, Hamilton 2017, Schüler 2018).

Chapter 3

New Traded Varieties and Source Countries: Evidence of Trade Complementarities

3.1 Introduction

The trade literature has emphasized the importance of new varieties in an economy and the gains from trade (Backus et al. 1992, Arkolakis et al. 2008, Saviotti & Frenken 2008). Also, a growing literature has focused on the relationship between new imported products and exports (Aristei et al. 2013, Bas & Strauss-Kahn 2014, Feng et al. 2016, Lo Turco & Maggioni 2013, Xu & Mao 2018). However, as far as we are concerned, only Castellani & Fassio (2019) focus on the relationship between imports of new varieties and exports of new varieties.

Furthermore, we identify three mechanisms behind the relationship between new imported varieties and new exported varieties. The first mechanism is related to trade processing (Castellani & Fassio 2019). The second mechanism is linked to knowledge about the destination country (Poncet & Mayneris 2013). Finally, the third mechanism is associated with technology embedded in imported products (Hausmann & Hidalgo 2011, Colantone & Crinò 2014, Feng et al. 2016, Xu & Mao 2018). In this chapter, we aim to explore the processing trade mechanism and briefly discuss the knowledge about the destination country mechanism.

The main contribution of this chapter is to examine the impact of importing new varieties on exports of new varieties. Furthermore, this chapter also reveals a degree of trade complementarities between imports and exports of new varieties at the country level. To the best of our knowledge, this feature has not been widely explored in the literature. To perform this analysis, we exploit the bilateral trade component of our database. This database represents the compilation of several datasets including bilateral trade data on exports and imports at the product level, concordance tables to trace the evolution of products, and statistical data on countries' GDP, score for starting a business, and import tariffs, as well as trade gravity variables.

Moreover, this analysis is also important because most of the empirical studies focus on developed countries with the exception of China; however, this Asian country has different trade dynamics and firm structures compared to the rest of the developing economies. Thus, our study sheds light on the relationship between new imported varieties and exports of new varieties from the perspective of a developing country. Finally, we also analyze a more recent period in comparison to the rest of the literature, which typically focuses on data for over ten years ago.

We start our study by decomposing the annual growth of trade varieties between new, withdrawn, and continuing varieties. We use two criteria to define a new variety. First, when a product is traded (i.e., imported or exported) with a partner country for the first time. Second, by tracing the evolution of product codes over time using concordance tables. Once we identified that new varieties contribute to trade growth in a non-negligible proportion, we then center the analysis to those new varieties. Therefore, we concentrate on 74,240 new traded varieties belonging to the manufacturing sector over the period 2005-2016.

We then employ a three-fold empirical strategy on the estimation sample focused on new traded varieties. First, we use a fixed effects logit model to estimate the probability of exporting new varieties based on the number of new imported varieties. Then, we employ a fixed effects negative binomial model to examine the impact of importing new varieties on the number of new exported varieties. Finally, we use a linear fixed effects model to measure the effects of importing new varieties on the export share of new varieties at the extensive and intensive margins aiming to detect trade complementarities at the country level.

Our main findings suggest that importing new varieties constitutes a key determinant for exporting new varieties across our different specifications (i.e., on the probability, number of varieties, and on shares). Our findings also suggest that importing new varieties from a country leads to an increase of exports of new varieties to that same source country. These results hold after incorporating control variables and including a full set of country, industry, and year fixed effects.

Furthermore, we perform a series of robustness checks that involves incorporating trade gravity variables, excluding source and destination countries, excluding Mexico's main trading partner, using alternative control variables, employing an alternative methodology to the negative binomial model, dealing with zero-value observations in the dependent variable, using a log-log model, examining contemporaneous effects, increasing the lag length of the independent variables, using different combinations of fixed effects, exploring input-output linkages, examining the marginal effects by sector, and using different sub-samples. We also run two-stage regressions, where we employ applied import tariffs as our instrument, in an attempt to alleviate any potential endogeneity.

Our findings suggest that the contemporaneous effects of new imported varieties on new exported varieties are significantly larger for the number of varieties and for the export share at the intensive margin. However, reverse causality is a potential concern in this contemporaneous effects exercise. To avoid this concern, we lagged all the independent variables in our baseline specifications. Consequently, our baseline results remain statistically significant, albeit their economic significance is smaller.

The rest of the chapter is organized as follows: Section 3.2 covers the related literature.

Section 3.3 describes the data. Section 3.4 displays the descriptive statistics. Section 3.5 defines the methodology. Section 3.6 shows the results. Section 3.7 presents the robustness analysis. Finally, Section 3.8 concludes.

3.2 Literature Review

We begin this section by examining the importance of introducing new products in an economy through trade, as well as the impact of those new products on growth. Next, we define and make a distinction between two popular concepts in the trade literature: products and varieties. We then consider the different methodologies used in the empirical literature to identify new products. Later, we identify the benefits and mechanisms behind new exports. We also discuss the empirical studies centered on the gains from importing new products in an economy. Then, we focus on the available cross-country and country case studies that examine the relationship between new imported varieties and exports. Finally, we provide an overview of the gaps in the literature and the contributions of this chapter in the existing empirical literature.

3.2.1 New Products and Growth

The trade literature has reiterated that the introduction of new varieties in an economy leads to gains from trade, as consumers benefit from access to new varieties. Also, the growth effect of trading new varieties can be explained by technological spillovers. In other words, firms having access to different variaties become more productive in creating new varieties as well. One strand of the literature started by analyzing the relationship between new products introduced by trade and economic growth. One of the pioneers in this strand is the study by Rivera-Batiz & Romer (1991) that highlights the importance of promoting the flow of ideas and increasing the trade in goods to achieve economic integration. This integration has the potential to boost economic growth in the long run through increasing returns to scale in the research and development (R&D) sector, especially among developed economies. In their study, the authors developed two theoretical models with different specifications of the R&D sector (i.e., knowledge-driven specification and a lab equipment specification) as a growth engine.[1] According to the authors, trade plays an important role in the process since the exchange of goods has the potential to avoid redundant efforts in R&D activities.

Likewise, Backus et al. (1992) reveal that new imported products can be incorporated as intermediate inputs in the production chain, which leads to growth. Thus, the authors tested the relationship between the growth rate of output per employee and intra-industry trade. The

[1] In the knowledge-driven specification model, human capital and knowledge are the only factors that influence the creation of new designs. In this specification, knowledge is the result of scientific and engineering efforts, as well as of know-how accumulation. On the contrary, the lab equipment specification model relies on human capital, unskilled labor, and capital goods, but not on ideas; these ideas have in fact no direct impact on production.

results suggest that developing countries importing specialized inputs can benefit more and grow at faster rates compared to larger countries.

Another strand of the literature focuses on the product mix (i.e., new, withdrawn, and continuing products) determined by firms. The aim of these firms dealing with a product mix is to reallocate resources efficiently. This strand of the literature offers empirical studies applied to both developed and developing countries. These empirical studies emphasize the impact of firms' decisions on growth for both developed and developing countries. Nonetheless, these studies exhibit significant differences in the flexibility of adjusting the product mix and their impact on firm productivity.

Bernard et al. (2010) examine the determinants of adjustments in the product mix of manufacturing firms in the United States. The authors recorded that about half of the U.S. manufacturing firms change their product mix every five years on average. The change of products is associated with the characteristics of both firms and firm-products. The decision to incorporate a new product is positively correlated to firm productivity; on the other hand, product withdrawal is a decision related to both the firm and firm-product attributes. Therefore, new and withdrawn products can determine changes in the product scope of firms. The results suggest that adjustments in the product mix of firms may lead to a more efficient reallocation of resources.

Goldberg et al. (2010*b*) provide evidence on how adjustments in the product mix of firms can potentially contribute to growth for developing countries, using India as a case study. This growth can be explained by product additions, rather than product withdrawals in the production line. Furthermore, their results do not confirm a relationship between product mix and tariffs decline due to trade liberalization for India; the authors explain that these findings could be due to the high regulations that Indian manufacturing firms faced before the reforms (e.g., industrial licensing), which cause firms to keep product lines, even if these are not profitable.[2] Also, those firms that previously faced sunk costs to expand operations are not willing to withdraw products from the production line. Finally, the country exhibits significant wealth disparities among the population; thus, there is still a demand for old products, even if these products are considered obsolete. As the authors frame it, it is worth studying countries at different stages of development, as these countries display important differences in terms of firm size distributions and resource allocation.

3.2.2 Defining Products and Varieties

The widespread use of the term varieties is sometimes equated to products in the trade literature. Nonetheless, a product is the finest level of disaggregation of an item, often classified as a subheading or tariff line in the HS code system. On the other hand, a variety is defined as a product imported from a specific country. The term variety has been widely used in the empirical literature as the combination of subheading or tariff line and a country (Broda & Weinstein 2006, Arkolakis et al. 2008, Goldberg et al. 2009, 2010*a*,*b*, Colantone & Crinò 2014).

[2]Industrial licensing refers to regulations and restrictions to establish industries in certain sectors.

In this chapter, we define varieties as the tariff line-country combination.

3.2.3 Identification of New Products

The empirical literature includes different methodologies to identify new products. Xiang (2005) investigates the relationship between new products and the relative demand for skilled labor using product-level data for the manufacturing sector in the United States. The author identified new goods by comparing product lists over two different waves and relied exclusively on product names. These potentially new goods were classified into four groups: products that feature spelling differences, but similar names; products with identical names with clarifications in terms of product purposes; products with minor differences in names; and products with major differences in names. This last group (i.e., products with major differences in names) were classified as new products and constituted the basis of the author's analysis. A drawback of this classification is that it can be problematic when comparing several years as the author manually identified the differences in names, instead of tracing them using a systematic classification code.

Broda & Weinstein (2010) are able to undertake more detailed analysis of the impact of the creation and withdrawal of products on prices using bar code data at the household level for specific sectors (e.g., grocery, pharmacy, and mass-merchandise) in 23 cities in the United States over six years. Part of their analysis involves monitoring the three dimensions of the data: product, brand, and product group. This data allowed the authors to identify new and withdrawn products by using entry and exit rates. The downside of using this classification system is that it may become problematic when trying to compare products across countries as bar codes may not be standardized.

The classification approach used by Colantone & Crinò (2014) is arguably the best approach to identify new products. The authors examine the impact of new imported inputs on the creation of new domestic products using product-level data of imported intermediate goods and domestic products for 25 European countries over the period 1995-2007. The authors started by identifying new imported products using the Combined Nomenclature (CN) and year-to-year correspondence tables to keep track of the evolution of the codes. The authors defined two criteria to identify new imported goods: when the good is imported for the first time from a partner country, and when the code is introduced in the classification system with no previous corresponding code. In this chapter, we use this classification approach as it seems to be more systematic and consistent compared to others used in the literature.

3.2.4 Benefits and Mechanisms Behind New Exports

The trade literature explores the different benefits of exporting new products. In this regard, Isogawa et al. (2012) focus their research on product innovation, especially in the case where a product is introduced for the first time to a market; the authors call this type of innovation: "new-to-market product innovation". Their findings suggest that Japanese firms tend to experience larger sales when introducing new products. Furthermore, their results suggest that new-to-market product innovation benefits other firms from technological spillover effects (i.e., technology acquisition or technology provision).

Furthermore, Saviotti & Frenken (2008) study the link between the variety of exports and economic development. To examine this relationship, the authors make a distinction between export varieties. On the one hand, they define "related varieties" as those varieties within sectors. On the other hand, they define "unrelated varieties" as those varieties between sectors. Their results suggest that an increase in export-related varieties has an immediate impact on growth. At the same time, an increase of export-unrelated varieties also has an impact on growth, but with a time lag. The authors present evidence suggesting that countries exhibiting a low level of development at the beginning of the period, but having an ability to catch up at a fast pace, exhibited a rapid increase in their export variety.

Also, Poncet & Mayneris (2013) suggest that firms can benefit from exposure to other exporters when aiming to enter markets that present important challenges, such as the Asian market. These challenges arise from significant differences in language, culture, tastes and preferences, as well as a different business culture. Thus, these export spillover effects can be present when firms with export activity concentrate in a geographical space and share information about the export market, or even share export costs, such as participating in export promotion activities, such as international exhibitions.

Moving on to the different mechanisms behind new exports, Bahar et al. (2014) explain that the mechanism behind the ability of a country to include a new export product could be influenced by its neighbour's export basket. Thus, the authors explain that neighbour countries tend to exhibit a similar comparative advantage; nonetheless, this similarity vanishes with distance. They also explain that countries' comparative advantage can change over time depending on their absorption capacity of new technologies.

Also, Hausmann & Hidalgo (2011) claim that the mechanism behind new varieties is a concept they call capabilities; the authors define these capabilities as a large and diverse set of non-tradeable inputs. Furthermore, the authors mention two types of capabilities: one associated with countries' endowment of capabilities; and the second type, which is related to the technological requirements of products. Thus, the authors explain that countries endowed with more capabilities can produce a broader range of new products. However, these new products requiring more capabilities are less accessible to countries. Nevertheless, the authors also explain that capabilities can be traded, allowing countries to import inputs that were not previously available. In terms of complex products, the authors suggest that global value chains allow more countries to engage in the production of these more complex products; this way, countries can specialize in specific stages of production instead of producing the entire product themselves.

3.2.5 Trade Liberalization and Gains from New Imports

Arkolakis et al. (2008) provide a number of insights regarding varieties and gains from trade. The authors show that larger countries tend to have wider varieties of both intermediate and final goods. These varieties are negatively correlated to tariffs, meaning that higher tariffs are associated with lower amounts of imported varieties. The results suggest that an increase in the number of imported varieties leads to gains from trade, although these gains are small. Also, new imported varieties have a positive, albeit small impact on welfare. A plausible explanation for this is that these new varieties are imported on a small scale.

As concluded by Goldberg et al. (2009), a trade liberalization process allows firms to gain access to more and less expensive intermediate inputs. This trade liberalization process also promotes imports of products and varieties that were not previously available, which are used by manufacturing firms. In the same vein, Amiti & Konings (2007) conclude that a reduction in tariffs leads to an increase in productivity for importing firms. These importing firms benefit more from productivity gains compared to non-importing firms. The authors acknowledge that a reduction of tariffs can promote these productivity gains through a tougher import competition. Furthermore, less expensive imported inputs have a similar effect on productivity via learning by importing, access to more varieties, and quality effects.[3]

These results are analogous to those found by Topalova & Khandelwal (2011), who explain two mechanisms behind the relationship between lower tariffs and higher productivity of firms. First, increased import competition due to trade liberalization leads to more efficient firms. Second, the trade policy reform in India led to lower import tariffs translating into an increase in imported intermediate goods. Goldberg et al. (2010*a*) examine the relationship between imports of intermediate goods, the production scope of domestic firms, and trade liberalization, using firm- and product-level data for India between 1989 and 1997. The results suggest a decline in tariffs promotes the production of new goods introduced by domestic firms, through access to new varieties. Moreover, these authors also show that access to new imported intermediates allows domestic firms to increase their production scope via the inclusion of new varieties.

Furthermore, Colantone & Crinò (2014) provide evidence pointing out that new imported products have a strong and positive impact on the creation of new domestic products using a set of 25 European countries. This positive impact can be explained by the combination of two mechanisms: scale effects, which expands the range of intermediate goods available, and access to superior varieties with quality-adjusted prices. These mechanisms result in benefits from trade as countries gain access to more and better-quality intermediate varieties.

This strand of the literature also has some sub-branches providing more information about new goods. Xiang (2005) shows empirical evidence revealing that industries investing more in research and development (e.g., chemical, machinery, electronics, and transportation equipment) tend to create more new goods. Broda & Weinstein (2010) provide evidence that the creation of new products is procyclical, meaning that new products are introduced in the market during economic expansions.

[3]Learning by importing refers to the ability of firms to improve their productivity after starting to import inputs from abroad. This mechanism is further discussed in Chapter 4.

3.2.6 New Imported Varieties and Exports

Import and export activities are two important intertwined strategies used by firms. Therefore, different studies focus on the relationship between new imported varieties and exports. This branch of the empirical literature is composed of cross-country studies, as well as country case studies focusing on developed countries. These studies converge in providing evidence on the strong relationship between new imported varieties and exports. It is worth emphasizing that this relationship has only been tested for one developing country (i.e., China), but has not been extended to other developing countries. The motivation of this chapter is to examine the link between new imported varieties and exports of new varieties for a developing country. As far as we know, none of the other papers have explored the bilateral trade component, which this chapter aims to analyze through trade complementarities.

We proceed to review a couple of cross-country studies. First, Aristei et al. (2013) examine the two-way relationship between the export and import activity of firms. To test this relationship, the authors used a probit model employing firm-level data in the manufacturing sector for 27 countries located in Eastern Europe and Central Asia from 2002 to 2008. The empirical model includes a vector of control variables that is comprised of productivity, firm size, ownership structure, share of white collars, and product innovation, as well as country and sector fixed effects. All independent variables have three lags due to the nature of the survey (i.e., World Bank Business Environment and Enterprise Performance Survey), which is administered every three years. The authors show that the analyzed relationship holds only in one direction: the import activity has a positive effect on the probability of exporting. This result can be supported by the idea that the import activity allows firms to boost their productivity and to innovate. This paper constitutes one of the few studies focused on the two-way relationship between export and import activities using firm level data. Despite its important contribution, this study does not distinguish whether importing from a specific country impacts the probability of exporting to that same country. In this chapter, we move one step further by exploiting this bilateral feature and revealing some degree of trade complementarities. Also, we expand the sample size in terms of countries and period length.

Guarascio & Pianta (2017) examine the relationship between new products, exports, and profits using a Three Stage Least Squares (3SLS) methodology for six E.U. economies (i.e., Germany, France, Italy, Spain, the Netherlands, and the U.K.) over the period 1995-2011. The 3SLS approach was employed by the authors to explore simultaneous relationships among variables and allow for endogeneity. The instrument used is the growth rate of value added. This methodology is also useful when controlling for all potential sources of heterogeneity affecting these relationships. The results suggest that the interdependence between the three key factors studied (i.e., the inclusion of new products, export growth, and increased profits) represents gains from technology for the developed countries included in the sample. Nonetheless, these gains from technology do not hold for the analyzed Southern European countries. Furthermore, technological innovations and engaging in global value chains seem to shape the firms' export patterns and profits jointly.

The empirical literature also includes a few country case studies mainly concentrating on developed economies. Lo Turco & Maggioni (2013) analyze the role of imports in raising manufacturing firms' probability of exporting using Italy as a case study. The methodology employed is a pooled probit model employing firm level data over the period 2000-2004. The authors draw two sub-samples of Italy's trade partners for the main explanatory variable; these correspond to the share of imports from both low- and high-income countries. The estimated model also includes a set of control variables, such as labor, firm average wage, firm's total factor productivity (TFP), intangible and tangible assets, as well as sector and year fixed effects. The authors claim that importing from low-income countries alone has a positive and significant effect on the export probability of Italian manufacturing firms. Nonetheless, they mention that this result should be taken cautiously as this link may only be valid in the case of Italy and might not be extrapolated to other countries. A limitation of this study is that the paper does not distinguish between the different countries for Italian exports. In this chapter, we consider this bilateral feature from the perspective of a developing country, which complements the results in Lo Turco & Maggioni (2013).

Bas & Strauss-Kahn (2014) focus on how importing intermediate varieties have the potential to raise the productivity of firms. However, the authors also explore how an increase in the number of imported inputs has the potential to contribute to an increase in the number of exported varieties using firm-level data from the French manufacturing sector between 1996 and 2005. The authors used a 2SLS approach to avoid endogeneity issues. The instrumental variable is defined as input tariffs from non-EU countries, the endogenous variable is the number of imported varieties, and the dependent variable is the number of exported varieties to the E.U. The regression equation includes firm size and firms' TFP as control variables, as well as firm and year fixed effects. All independent variables are expressed in their logarithmic form and are lagged by one year. The results suggest that imported intermediate inputs have an impact on the export scope via lower prices of intermediate goods and lower export fixed costs (i.e., firms gain access to intermediate inputs that possess technology of superior quality or sophistication). It is worth mentioning that even though the authors employed bilateral trade data, they did not quite exploit the bilateral component; in other words, the authors consider EU countries as a bloc, instead of employing data of individual countries. In contrast, this chapter examines the bilateral trade flows between Mexico and all the different partner countries; this difference allows a more comprehensive identification of bilateral trade patterns.

Feng et al. (2016) examine the importance of imported inputs on firm export activity in China. The authors used 2SLS regressions to account for potential endogeneity issues. The instrumental variable is defined as input tariffs; this instrument considers the relevance of trade liberalization, changes in terms of exchange rates, and firm decisions on imported intermediate usage. The analysis is performed on Chinese manufacturing firm-level data during the period 2002-2006. The findings suggest that Chinese firms importing more intermediate inputs tend to experience a growth in exports. The authors explain that firms that started as non-traders benefit more from the relationship between intermediate inputs and export growth. Moreover, importing upgraded intermediate inputs (i.e., embedded with higher quality and technology) can lead to more sophisticated exported goods. Finally, local private firms importing more intermediate inputs experience larger export growth compared to their foreign counterparts located in China.

Xu & Mao (2018) study the relationship between imported intermediate inputs and the export quality of firms using China as a case study. The authors used linear regressions with fixed effects employing firm- and product-level data for the manufacturing sector over the period 2000-2007. The empirical specification includes a vector of control variables comprising firm size, average wage, firm profit, credit constraint, government subsidy, exchange rates, and firm ownership, as well as firm and year fixed effects. The results suggest that imported intermediate inputs have the potential to boost the export quality of manufacturing firms. The authors claim that quality embedded in imported intermediate goods, quality of institutions, and market share reallocation constitute potential determinants of export quality improvements among Chinese manufacturing firms.

Castellani & Fassio (2019) analyze the role of new imported inputs as a key determinant of new exported products for Sweden. The authors employed a negative binomial approach using data from manufacturing firms between 2001 and 2012. The dependent variable in the baseline regression equation is the number of new exported products, while the main explanatory variable is the number of new imported inputs. The authors also included control variables of ownership structure and firm-level controls (i.e., productivity, number of employees, investments, and a dummy for having at least one employee who registered a patent), as well as firm and year fixed effects. The authors claim that this negative binomial methodology is the most appropriate approach when the dependent variable is a count variable exhibiting overdispersion. The results suggest that new imported inputs constitute a key determinant of exporting new products. More specifically, an increase in the number of new imported goods leads to an increase in new products aimed to the export market. Although the authors use firm-level data, this data is not disaggregated at the product-level. Our chapter, however, does capture the effects at the product-level (i.e., HS 8-digit classification).

Navas et al. (2020) examine the indirect role of market size and geographical proximity via imports on the export patterns of Italian manufacturing firms from 2000 to 2006. The authors use two specifications of 2SLS regressions. The first specification focuses on the extensive margin, which is defined in this paper as the export status of firms (i.e., a dummy variable that equals one if the firm exports to a specific country, and zero otherwise). The second specification focuses on the intensive margin, which is defined as the total exports of a firm to a specific country. On the right-hand side of the equation, the explanatory variables were defined as firm's productivity and TFP-enhancing effect of imported intermediates. Moreover, control variables were added to the baseline specifications, such as distance, GDP, trade opening, remoteness, and market costs, as well as year-geographical and firm fixed effects.

In their second stage, Navas et al. (2020) used two instrumental variables to tackle potential endogeneity: GDP weighted by a firm's import share of each country (proxies the number of foreign intermediate varieties available) and total imports of other European countries weighted by the relative importance of a product in a firm's total imports. The authors also estimated the second specification using a conditional fixed-effects Poisson model to deal with zero-value observations. The results suggest that importing from large markets that are geographically closer may lead to higher gains in terms of productivity, which may also promote firms' export activity and increase their export value. Furthermore, the authors explain that these two determinants (i.e., market size and geographical proximity) have an indirect effect on firms' export behavior through import activity. In contrast, this chapter considers the impact of new imported varieties from a source country on the exports of new varieties to that same country.

3.2.7 Contributions to the Literature

We have seen that several studies examining the link between new imported varieties and exports focus on developed countries, but not on developing countries, except for a couple of studies that use China as a case study. These two exceptions examine the relationship between imports and exports using Chinese firm-level data. Despite the importance of this Asian economy, the results cannot be extrapolated to other developing countries due to significant differences in the firm ownership structure between China and other developing countries. We need to recall that a significant amount of Chinese firms are State-Owned Enterprises (SOEs).[4] Due to this reason, it is worth analyzing the relationship between new imported varieties and exports of new varieties from the perspective of other developing countries.

The most important contribution of our study is that we exploit the bilateral trade component of our database, which has not been widely explored in the related literature, except for Navas et al. (2020). However, the authors focused instead on the impact of firm's productivity on exports. By exploiting this bilateral trade component, we can reveal the presence of a certain degree of trade complementarity between imports and exports at the country level. In other words, we can examine to what extent importing new varieties from a specific country increases exports of new varieties to that country. Thus, we examine the impact of new imported varieties on the export share of new varieties at the extensive and intensive margins. The aim is to explore whether there is an additional effect on exports to the same source country.

Furthermore, previous studies on new imported varieties have not examined trade between asymmetric countries in terms of economic development; ergo, this study performs the analysis from the perspective of an emerging economy. It is worth highlighting that Mexico is an interesting case study because the country has strong trade ties with developed economies, especially with the United States. However, the country also trades with developing countries in the Latin American region. As part of the robustness checks, we also include a separate

[4]According to the OECD (2017), the Chinese central government owns more than 51,000 firms. This figure is significantly larger compared to a total of roughly 2,500 SOEs combined from a sample of 38 countries; this sample of countries include: Argentina, Australia, Austria, Brazil, Canada, Chile, Colombia, Costa Rica, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, India, Ireland, Israel, Italy, Japan, South Korea, Latvia, Lithuania, Mexico, Netherlands, New Zealand, Norway, Poland, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, the United Kingdom, and the United States.

analysis for high-income and low- and middle-income countries.

Finally, this study is also relevant because we use disaggregate data at the product level, and we extend the examined period. The first feature on the level disaggregation is important because it allows us to identify new varieties within industries. Regarding the examined period, the available empirical literature is based upon data from over ten years ago. Thus, it is unclear whether these results hold considering the recent trends in trade (i.e., global value chains). Therefore, we provide an updated study employing recent trade data.

3.3 Data

The data used in this chapter represents the compilation of several datasets. Bilateral data on exports and imports at the product-level was retrieved from the Mexican Ministry of Economy over the period 2003-2016. These annual datasets contain information on trade value in U.S. dollars, volume in units, and source or destination country for the universe of products encompassing the agricultural, extractive, and manufacturing sectors. This data is available at the product level using the Harmonized System (HS) classification at 8-digits of disaggregation (i.e., tariff lines), which comprises over 12,000 products. Finally, the data is reported for 231 countries and territories.

As a matter of context, the Harmonized System (HS) is an international classification for products developed by the World Customs Organization (WCO). The structure of the code starts with 21 sections that cover all the agricultural, extractive, and manufacturing sectors (e.g., Section XVII stands for "Vehicles, aircrafts, vessels and associated transport equipment"). These sections are broken down into paired digits where the first pair (HS 2-digits) corresponds to chapters of this classification (e.g., Chapter 87 – "Vehicles other than railway or tramway rolling stock, and parts and accessories thereof"). The next pair of digits denotes the headings (HS 4-digits), which represent groups within the chapter (e.g., 8702 – "Motor vehicles for the transport of ten or more persons, including the diver"). The next two digits stand for the sub-heading (HS 6-digits), which is the standard level used in the international classification of products (e.g., 8702.10 – "With compression-ignition internal combustion piston engine (diesel or semi-diesel)").

Furthermore, the sub-division beyond the HS 6-digits level is reserved for countries' national tariffs. In the case of Mexico, trade statistics are recorded at the HS 8-digits level (e.g., 8702.10.01 – "With body mounted on chassis, excluding those of Tariff items 8702.10.03 and 8702.10.05"); this number of digits constitutes the highest level of disaggregation for the tariff nomenclature in Mexico.

The Harmonized System makes a clear distinction between types of goods. Chapters 1 to 24 include agricultural goods, while Chapters 25 to 97 are comprised of non-agricultural goods. Moreover, Chapter 77 is reserved for future use in the Harmonized System classification. Chapter 98 corresponds to "project imports, laboratory chemicals, passenger's baggage, personal importation by air or post; ship stores", while Chapter 99 contains temporary modifications to the national legislation; these last two chapters are reserved for national use. In this empirical chapter, we focus on manufacturing goods; thus, we concentrate on products comprised in Chapters 25 to 97, excluding Chapter 77. The Appendix contains a table of the 21 sections of the HS nomenclature and their corresponding chapters and descriptions defined by the World Customs Organization.

Moving on to the database, it is worth mentioning that the zero-value observations may stand for trade samples.[5] The status of trade samples in Mexico is determined by the General Rules of Foreign Trade and by the Law of General Taxation regarding Imports and Exports. This status is defined as "items that, due to their quantity, weight, volume or other presentation conditions, indicate, without a doubt, that these can only be used for demonstration or to place orders". One of the requirements to fulfill the status of trade samples is that the unit price of the item should not exceed the equivalent of one U.S. dollar.

These zero-value observations represent less than 0.3% of the observations in the estimation sample. A plausible explanation for the origin of these zero-value observations, associated with trade samples, is that these zeros may have arisen from the exchange rate conversion of the reported values in Mexican Pesos (MXN) to U.S. dollars (USD), which were later truncated by the Mexico's Central Bank (i.e., no decimals are displayed in the datasets). Furthermore, after comparing these zero-value observations against their volume in units and their corresponding UN Comtrade Standard Unit Values (in U.S. dollars), we can infer that these transactions stand for trade samples.[6] To consider these goods, we assigned a small value equivalent to 0.01 U.S. dollars to these zero-value observations associated with trade samples.

Another important feature to bear in mind when working with trade data is the lack of trade transactions between some countries. As acknowledged by Santos Silva & Tenreyro (2006), Kleinert et al. (2015) and Navas et al. (2020), among others, zero-value observations are very common in trade records as not all countries trade all products with each other. As a matter of context, Mexican trade data is registered by the Ministry of the Treasury and Public Credit, via the Revenues Administration System (SAT). This trade data corresponds to all the import and export transactions made by Mexican firms. In other words, we can assume that bilateral trade in goods not reported in the trade datasets can be interpreted as true zeros (i.e., transactions did not take place). Thus, we recoded these missing values as zeros.

A steppingstone in this chapter is the identification of new, continuing, and withdrawn products. This procedure was done by tracing the evolution of HS codes between 2003 and 2016. It is important to point out that Mexican trade authorities update the Harmonized System codes every five years in compliance with the World Customs Organization (WCO) regulations. Thus, the updates that match our examined period were carried out in 2007 and 2012. We used the concordance tables TIGIE 2002-2007 and TIGIE 2007-2012 retrieved from the Integrated Foreign Trade Information System (SIICEX) to follow this code evolution.

[5]A trade sample is defined as a good that is imported exclusively for the purpose of being shown or demonstrated to request orders, but this good cannot be sold.

[6]The UN Comtrade defines the Standard Unit Value (SUV) as the median unit value of traded goods; these standard unit values are calculated at the end of each year. The Appendix contains a scatterplot reporting the zero-value observations contained in the Mexican trade datasets.

Turning now to GDP data of countries, this data was retrieved from the World Bank Development Indicators; we use this measurement at purchasing power parity (PPP) in constant international dollars. We also employ the World Bank Doing Business dataset to retrieve information on the score of starting a business of each country; this score represents the simple average of the following scores: procedures, time and cost to start and operate a business, and minimum capital requirements.[7] We consider that this variable is a good proxy for the ease of doing business and trading with the different partner countries.[8]

We also use trade gravity variables sourced from the Geo CEPII Database (Calderón et al. 2007, Imbs 2004, Navas et al. 2020). These variables include latitude and longitude of capital cities, official language, landlocked status, continent, and colonizer. We computed bilateral distance using the great-circle distance formula between Mexico City and the capital city of each partner country (Jansen & Stokman 2014, Navas et al. 2020). Moreover, we constructed the border variable as a dummy variable; this variable is equal to one if a partner country shares a border with Mexico; and zero, otherwise (Calderón et al. 2007, Clark & Van Wincoop 2001, Imbs 2004, Jansen & Stokman 2014, Kleinert et al. 2015). We also constructed the free trade agreement (FTA) variable using data from the Foreign Trade Information System of the Organization of American States (OAS) records. This FTA variable is a dummy that equals one if the partner country has a free trade agreement with Mexico; and zero, otherwise (Jansen & Stokman 2014).

As part of the Robustness Analysis, we also use input-output matrices from the World Input-Output Database (WIOD). This database was retrieved from the University of Groningen. The WIOD encompasses data for 28 EU countries and 15 major countries over the period spanning from 2000 to 2014. This input-output data is reported at the sector level under the International Standard Industrial Classification (ISIC) at a 2-digits level. It is worth mentioning that the WIOD database reports data for the agricultural, extractive, manufacturing, and service sectors. However, we only include manufacturing sectors in this chapter.[9] Along with this world input-output matrix, we employ an OECD Correspondence Table HS-ISIC to match manufacturing sectors.

Data on Most Favored Nation (MFN) and applied tariffs at different levels of the Harmonized System were sourced from the UNCTAD Trade Analysis and Information System (TRAINS) Database; this data is reported in ad valorem duties.[10] Labels on standard codes for countries (ISO 3166) were retrieved from the International Organization for Standardization (ISO). In

[7]Each of these individual indicators are measured on a scale from 0 to 100, where 0 represents the worst regulatory performance and 100 the best regulatory performance.

[8]The World Bank Doing Business dataset also includes trade-related variables considered in Navas et al. (2020): number of documents to import, cost to import in U.S. dollars per container deflated, and time required to import in days, among others. We consider these trade-related variables as alternative control variables in the Robustness Analysis section.

[9]The Appendix contains a table of the ISIC manufacturing sectors included in this chapter.

[10]MFN rates are tariffs that countries promise to impose on imports from other WTO members, unless countries are part of a preferential or free trade agreement. In practice, MFN applied tariffs are the highest rates WTO members can charge one another. On the other hand, applied tariffs are effective applied rates; these can be below MFN rates, due to preferential or free trade agreements.

Ad valorem duties are import duties expressed as a percentage of the value of the merchandise.

addition to these international standard codes, the Mexican Ministry of Economy possesses its own country codes; these codes were also used to complement trade datasets.

3.4 Descriptive Statistics

3.4.1 Decomposition Exercise

A steppingstone of this analysis is to identify new, continuing, and withdrawn varieties from the universe of manufacturing goods that Mexico traded during the period 2003-2016. In line with the standard empirical literature, we define a variety as a product-country combination. In other words, a product traded with a particular country.[11] We use two criteria to define a new variety. First, when a product is traded with a partner country for the first time. Second, by tracing the evolution of product codes over time using concordance tables provided by the Mexican authorities.

Thus, we assume a variety is new under the following circumstances. First, the tariff line is introduced to the Harmonized System in time t and does not have any previous code corresponding to it. Second, the tariff line is introduced to the Harmonized System in time t and has one or more previous codes corresponding to it, but none was traded with a particular country in any previous year. Third, the tariff line is not new to the Harmonized System but has not been traded with a particular country in any previous year. As a result, traded varieties can be counted as new only once.

We start this section with a decomposition exercise of the annual growth of exported and imported varieties between new, withdrawn, and continuing varieties following the methodology employed by Colantone & Crinò (2014):

$$\frac{\mathrm{X_{cit}\text{ - }X_{cit\text{-}1}}}{\mathrm{X_{cit\text{-}1}}} = \frac{1}{\mathrm{X_{cit\text{-}1}}}\left[\sum_{z\in New_{cit}} X^z_{cit} - \sum_{z\in Withdrawn_{cit}} X^z_{cit-1} + \sum_{z\in Continuing_{cit}} (X^z_{cit} - X^z_{cit-1})\right], \quad (3.1)$$

where c stands for partner countries, i represents the industries (HS 4-digits), and t denotes time expressed in years.[12] The superscript z represents exports or imports, respectively; and X denotes the number of traded varieties.

Table 3.1 identifies the number of traded manufacturing varieties by Mexico into new, continuing, and withdrawn. We identify a total of 227,005 new exported varieties, which account for 17.8% of the total number of exported varieties; from this number of new exported varieties, nearly 7,610 corresponds to new products (i.e., new HS 8-digit codes). On the other hand, we identify a total number of 255,087 new imported varieties, which account for 12.8% of the total number of imported varieties; from this number of new imported varieties, roughly 12,770 corresponds to new products (i.e., new HS 8-digit codes).

[11] In this chapter, we define product as a tariff line coded at HS 8-digits. By trade, we refer to products been exported or imported.

[12] After identifying new, withdrawn, and continuing varieties, we aggregate tariff lines (i.e., HS 8-digits) into industries (i.e., HS 4-digits).

Table 3.1: Identification of Traded Varieties

		Total Varieties	New Varieties	Withdrawn Varieties	Continuing Varieties
Exported Varieties	(Freq.)	1,275,607	227,005	5,019	1,043,583
Exported Varieties	*(%)*	*100.0*	*17.8*	*0.4*	*81.8*
Imported Varieties	(Freq.)	1,999,633	255,087	6,730	1,737,816
Imported Varieties	*(%)*	*100.0*	*12.8*	*0.3*	*86.9*

Notes: This table identifies the universe of traded manufacturing varieties by Mexico into new, continuing, and withdrawn. Figures displayed in odd rows represent the number of exported and imported varieties, respectively. On the other hand, figures in italics correspond to the breakdown of varieties expressed as percentages of exported and imported varieties, respectively.

In Table 3.2, we present a the decomposition exercise of the annual growth of traded varieties. The figures in this table are expressed as percentages and represent the mean values (in U.S. dollars) across industries, countries, and years; figures in italics are normalized by the growth rates presented in the first column.

Table 3.2: Decomposition of Annual Growth Rates of Traded Varieties

	Growth Rate	New Varieties	Withdrawn Varieties	Continuing Varieties
Exported Varieties	10.9	1.5	-0.2	9.7
Normalized by Growth Rates	*100.0*	*13.4*	*-2.0*	*88.5*
Imported Varieties	10.8	1.5	-0.5	9.8
Normalized by Growth Rates	*100.0*	*13.5*	*-4.8*	*91.3*

Notes: The table reports the decomposition exercise of annual growth rates of exported and imported varieties into new, continuing, and withdrawn. This decomposition exercise is based on Eq.(3.1). Figures are expressed in percentages and represent the mean values (in U.S. dollars) across industries, countries, and years. Figures in italics are normalized by the growth rates in the first column.

From this table, we can observe that new exported varieties and new imported varieties account for 13.4% and 13.5% of the average annual growth rates, respectively. The figure corresponding to new imported varieties is slightly lower compared to 17.1% for European countries in Colantone & Crinò (2014). On the other hand, the figures for withdrawn and continuing imported varieties are also lower for Mexico compared to European countries, where the average exit rate is 10.7%, and the average continuing rate is 93.6%. We cannot compare figures for exported varieties, as this was not part of the scope in Colantone & Crinò (2014); instead, the authors focused on domestic production.

We can conclude that both new exported varieties and new imported varieties constitute important features of trade that are worth examining. It is also worth noting that the exit rates are very low for Mexico; this is aligned to Goldberg et al. (2010*a*), who claim that Indian firms tend to continue adding goods to the production line, instead of discontinuing obsolete goods.

3.4.2 Variable Description

Dependent Variable

The dependent variable consists of new exported varieties to country c by industry i in year t. We employ four different measures of the dependent variable. The first measure corresponds to the probability of exporting new varieties by industry i to country c in year t. We use this dependent variable in our fixed effects logit model. We define this probability measure as:

$$Prob\_X\_NEW_{cit} > 0, \tag{3.2}$$

where $Prob\_X\_NEW_{cit}$ is a dummy variable equal to 1 if at least one new variety is exported to country c by industry i in time t; and zero, otherwise.

In a similar manner as in Castellani & Fassio (2019), the second measure is the number of new exported varieties by industry i to country c in year t. We employ this measure of the dependent variable in our fixed effects negative binomial model. Our measure in levels of the dependent variable is defined as:

$$X\_NEW_{cit} = \sum_k X\_PROD^k_{cit}, \tag{3.3}$$

where X_NEW_{cit} corresponds to the sum of new varieties k (in HS 8-digits) of exported products X_PROD to country c belonging to industry i (in HS 4-digits) in year t.

The third measure of the dependent variable stands for the export share of the number of new varieties to country c by industry i in time t. The aim is to measure new exported varieties at the extensive margin. We use this measure of the dependent variable in one of our linear fixed effects models. We calculate this extensive margin as follows:

$$ExtM\_X\_NEW_{cit} = \frac{X\_NEW_{cit}}{\sum_c (X\_NEW_{cit})}, \tag{3.4}$$

where X_NEW_{cit} represents the number of new exported varieties to country c by industry i in time t over the total number of new exported varieties by industry i in time t; this export share is expressed as a percentage.

The fourth measure denotes the export share of the value (in U.S. dollars) of new varieties to country c by industry i in time t. Now, the aim is to measure new exported varieties at the intensive margin. We employ this other measure of the dependent variable in the other linear fixed effects model. We construct this intensive margin as:

$$IntM\_X\_NEW_{cit} = \frac{X\_NEW^{USD}_{cit}}{\sum_c (X\_NEW^{USD}_{cit})}, \tag{3.5}$$

where $X\_NEW_{cit}^{USD}$ stands for the value (in U.S. dollars) of new exported varieties to country c by industry i in time t over the total value of new exported varieties (in U.S. dollars) by industry i in time t; this export share is also expressed as a percentage.

Main Explanatory Variable

The main explanatory variable is also defined in line with Castellani & Fassio (2019). Our explanatory variable corresponds to the log number of new imported varieties from country c by industry i in the previous year $t-1$.

$$ln(IM\_NEW_{cit-1}) = ln(\sum_k IM\_PROD_{cit-1}^{k}), \tag{3.6}$$

where IM_NEW_{cit-1} corresponds to the sum of new varieties k (in HS 8-digits) of imported products IM_PROD from country c belonging to industry i (in HS 4-digits) in the previous year $t-1$. To avoid the log of zero, which is undefined, we add one unit to the main explanatory variable before taking the natural logarithm.[13]

Control Variables

In terms of the control variables, we include the log of GDP at PPP in constant international dollars: $ln(\text{GDP in PPP})_{ct-1}$. It is worth mentioning that GDP is a standard control variable used in the trade literature (see, for example, Jansen & Stokman (2014)). We also include the log of the score of starting a business: $ln(\text{Starting a Business})_{ct-1}$.[14] This score is a proxy for the ease of doing business with partner countries. All control variables are expressed in natural logs and lagged by one year to tackle potential reverse causality. [15]

It is worth mentioning that although FDI plays an important role in processing trade, we did not include FDI inflows as part of the control variables as we only have data for 47 major investment countries. In fact, we need to recall that this study focuses on new varieties; therefore, one of the criteria for a product to be considered new is that the product is traded with a country for the first time. Thus, there is no FDI data available for a significant amount of new source and destination countries. Nevertheless, we include FDI inflows as a control variable in the Appendix section. However, we can notice FDI has an insignificant effect on exports of new varieties.

[13]Adding one unit and then, taking the natural logarithm is an approach often used in the empirical trade literature to deal with zero-value observations (Calderón et al. 2007). Nonetheless, we also use a Poisson Pseudo Maximum Likelihood (PPML) estimator as an alternative approach to deal with zero-value observations as part of the Robustness Analysis section.

[14]We also use other trade-related control variables considered in Navas et al. (2020) as part of the Robustness Analysis: number of documents to import, cost to import in U.S. dollars per container deflated, and time required to import in days. However, none of these variables were considered in the baseline specifications due to their relatively shorter time span covering the period 2006-2015.

[15]As mentioned before, we add one unit to the independent variables before taking the natural logarithm.

3.4.3 Summary Statistics

Our estimation sample is composed of 74,240 new varieties over the period 2005-2016. This number of new varieties are product-country combinations. These new varieties are the result of manufacturing industries introducing new products, and partner countries trading new products for the first time. These new varieties in HS 8-digits are then aggregated at the industry level (HS 4-digits); this approach of data aggregation is standard in the trade literature (see, for example, Colantone & Crinò (2014), Castellani & Fassio (2019)).

Table 3.3: Summary Statistics

Variables	Labels	(1) N	(2) Mean	(3) Std.Dev.	(4) Min	(5) Max
Probability of Exporting New Varieties	$Prob\_X\_NEW_{cit}$	890,880	0.1450	0.3522	0	1
Number of New Exported Varieties	X_NEW_{cit}	890,880	0.2114	0.6797	0	50
Export Share at the Extensive Margin	$ExtM\_X\_NEW_{cit}$	890,880	1.1570	5.3031	0	100
Export Share at the Intensive Margin	$IntM\_X\_NEW_{cit}$	890,880	1.1694	7.8977	0	100
Number of New Imported Varieties	IM_NEW_{cit}	890,880	0.2321	0.7412	0	55
ln(No. of New Imported Varieties)_cit-1	$ln(IM\_NEW)_{cit-1}$	890,880	0.1437	0.3412	0	4.0254
ln(GDP in PPP)_ct-1	$ln(GDP)\_ct-1$	890,880	25.9256	2.0616	18.9133	30.5546
ln(Starting a Business)_ct-1	$ln(Business)\_ct-1$	890,880	4.2776	0.2980	0.7885	4.6052
ln(No. Documents to Import)_ct-1	$ln(Docs2Import)_{ct-1}$	730,886	1.7892	0.4064	0.6931	3.0445
ln(Import Costs)_ct-1	$ln(Cost2Import)_{ct-1}$	730,886	7.2695	0.5945	5.9092	9.8975
ln(Time to Import)_ct-1	$ln(Time2Import)_{ct-1}$	730,886	2.7824	0.6450	1.3863	4.7622
ln(Distance)	$ln(Distance)_c$	886,224	8.9760	0.6856	6.9680	9.7742
Border	$Border_c$	886,224	0.0313	0.1740	0	1
Landlocked	$Landlocked_c$	886,224	0.1275	0.3336	0	1
Continent	$Continent_c$	886,224	0.2971	0.4570	0	1
Language	$Language_c$	886,224	0.1937	0.3952	0	1
Colonizer	$Colonizer_c$	886,224	0.1909	0.3930	0	1
Free Trade Agreement	FTA_{ct-1}	890,880	0.3775	0.4848	0	1
ln(Applied Import Tariffs)_cit-1	$ln(Applied)_{cit-1}$	888,450	1.4460	1.2148	0.0000	3.8712

Notes: The estimation sample is conformed by 74,240 new varieties over 12 years.

Table 3.3 reports the summary statistics of our empirical analysis. We can observe that the probability of exporting a new variety is 14.5%. Furthermore, we can see that the maximum number of new exported varieties over the examined period is 50 new varieties. It is worth mentioning that this measure of the dependent variable presents overdispersion around the mean.[16] In other words, the sample is concentrated on a few discrete values.

On the other hand, the main explanatory variable in its raw version is the number of new imported varieties, which ranges from zero to 55 new varieties. We also include time-variant control variables: GDP at PPP in constant international dollars and the score of starting a business.

Furthermore, we include standard trade gravity variables as part of the robustness checks: bilateral distance, common border, landlocked-status, common continent, common language, and common colonizer. In the standard trade literature, distance is a proxy for transportation costs; this variable is computed using the great-circle distance formula between Mexico City and the capital city of each partner country. It is worth mentioning that these gravity variables

[16] The number of new exported varieties variable presents a larger variance compared to the mean.

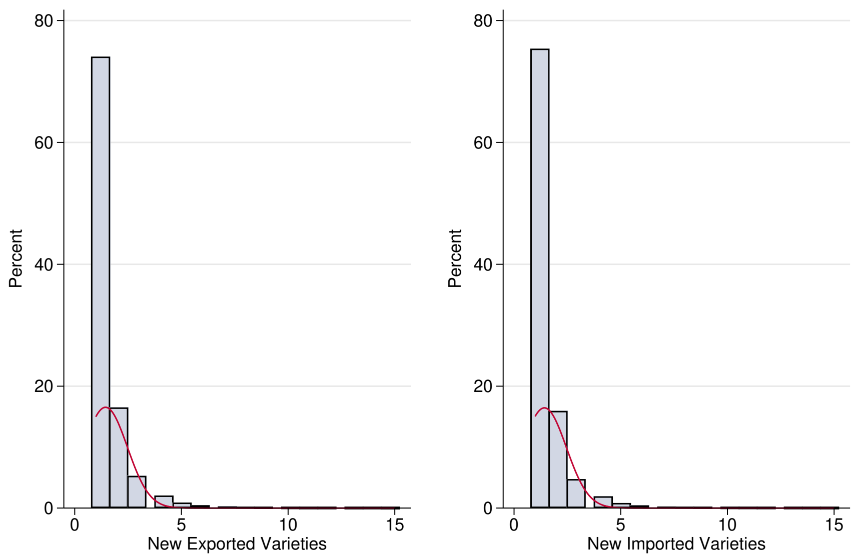
are time-invariant; thus, we cannot simultaneously include country fixed effects. Furthermore, we also run two-stage regressions with applied import tariffs as our instrumental variable.

3.4.4 Industry Distribution of New Traded Varieties

As a recap from Table 3.3, Mexican firms exported a maximum of 50 new varieties within an industry to a destination country in a specific year. In comparison, these firms imported a maximum of 55 new varieties within an industry from a source country in a specific year.[17] Figure 3.1 now exhibits the industry distribution of new traded varieties by Mexico over the examined period. For visual purposes, we truncated the scale of the horizontal axis to the 15 new varieties with the largest concentration of industries. From this figure, we can observe that about 74% of industries exported only one new variety, while 75% of industries imported only one new variety. It is worth mentioning that the frequencies dramatically drop as the number of new varieties increases.

Overall, we can notice that both distributions are right-skewed, resembling binomial distributions, with most of the industries concentrating in one to five new varieties that are both exported and imported. Compared to Castellani & Fassio (2019), the shape of these distributions looks similar to those for new traded products in Sweden; the difference is that the number of new products in the Scandinavian country is spread out.

Figure 3.1: Industry Distribution of New Traded Varieties



Notes: The left figure displays the industry distribution of new exported varieties. The right figure exhibits the industry distribution of new imported varieties.

[17] In 2007, Mexico exported 50 new varieties to the United States belonging to industry HS 8708 ("Parts and accessories of the motor vehicles of headings 87.01 to 87.05."). Likewise, the country imported 55 new varieties from the United States belonging to the same industry HS 8708 in 2007.

3.4.5 New Traded Varieties by Sector and Industry

Table 3.4 displays the number of new traded varieties by sector in 2016. We grouped these varieties by sectors according to sections of the Harmonized System classification. We can observe that the number of new imported varieties is relatively larger than the number of new exported varieties, except for the metal and transportation sectors. A plausible explanation is that these last two sectors highly rely on global value chains and are export-oriented towards the U.S. market. Furthermore, the sector that concentrates the largest number of imported varieties is the textile sector. On the other hand, the sector that exhibits the largest number of exported varieties is the machinery and electrical sector.

Table 3.4: Number of New Traded Varieties by Sector

Sector	New Exported Varieties	New Imported Varieties
25–27 Mineral Products	146	202
28–38 Chemicals and Allied Industries	1,256	1,476
39–40 Plastics and Rubbers	732	736
41–43 Raw Hides, Skins, Leather, and Furs	80	144
44–49 Wood and Wood Products	407	561
50–63 Textiles	2,243	3,649
64–67 Footwear and Headgear	905	999
68–71 Stone and Glass	337	407
72–83 Metals	1,636	1,494
84–85 Machinery and Electrical	2,975	3,348
86–89 Transportation	403	357
90–97 Miscellaneous	914	958
Total	12,034	14,331

Notes: This table shows the number of new traded varieties by sector. The reference year is 2016.

Table 3.5 exhibits the ten most common new exported varieties by industry in 2016. We can observe that most of these industries belong to the textile sector.[18] Nonetheless, the footwear, transportation, and electrical industries are also present. In contrast with Castellani & Fassio (2019), Swedish manufacturing firms concentrate their exports in new varieties related to maritime and waterway structures, parts of machinery for ships' derricks, parts, and accessories for tractors, articles of wood, and office or school supplies of plastics. This analysis is interesting as it reveals the export structure and the comparative advantage of each country.

Table 3.5: New Exported Varieties by Industry

Industry	Industry Description	New Varieties	Million USD
6403	Footwear with outer soles of rubber, plastics, or leather.	331	175.9
6204	Women's or girls' suits, ensembles, jackets, blazers, dresses, skirts.	242	129.9
6402	Other footwear with outer soles and uppers of rubber or plastics.	219	36.17
6203	Men's or boys' suits, ensembles, jackets, blazers, trousers, bib and brace overalls.	217	774.7
6302	Bed linen, table linen, toilet linen and kitchen linen.	198	5.977
6404	Footwear with outer soles of rubber, plastics, leather and uppers of textile materials.	183	16.90
8708	Parts and accessories of the motor vehicles.	151	2.197
8536	Electrical apparatus for switching or protecting electrical circuits.	148	1.516
5407	Woven fabrics of synthetic filament yarn.	142	6.821
7318	Screws, bolts, nuts, coach screws, screw hooks, rivets, cotters, washers.	129	0.130

Notes: This table displays the number of new exported varieties and corresponding export value (in U.S. dollars) by industry (HS 4-digits). The reference year is 2016.

As a counterpart, Table 3.6 displays the most common new imported varieties by industry

[18]In this chapter, sectors are defined as sections of the Harmonized System classification, while industries are defined as HS 4-digit codes.

in 2016. These new imported varieties are mainly concentrated in the textiles, footwear, and transportation sectors. A plausible explanation for the match on export and import industries is that Mexican firms intensively concentrate on intra-industry trade. It is worth mentioning that we cannot compare these common new imported varieties with Castellani & Fassio (2019) because this was not part of their scope.

Table 3.6: New Imported Varieties by Industry

Industry	Industry Description	New Varieties	Million USD
6204	Women's or girls' suits, ensembles, jackets, blazers, dresses, skirts.	500	95.72
6203	Men's or boys' suits, ensembles, jackets, blazers, trousers, bib and brace overalls.	402	138.2
5407	Woven fabrics of synthetic filament yarn.	359	182.9
6403	Footwear with outer soles of rubber, plastics, or leather.	323	152.4
6111	Babies' garments and clothing accessories, knitted or crocheted.	289	43.05
6404	Footwear with outer soles of rubber, plastics, leather and uppers of textile materials.	260	367.0
6302	Bed linen, table linen, toilet linen and kitchen linen.	253	35.48
6402	Other footwear with outer soles and uppers of rubber or plastics.	247	159.0
6209	Babies' garments and clothing accessories.	221	11.71
8708	Parts and accessories of the motor vehicles.	186	12.68

Notes: This table exhibits the number of new imported varieties and corresponding import value (in U.S. dollars) by industry (HS 4-digits). The reference year is 2016.

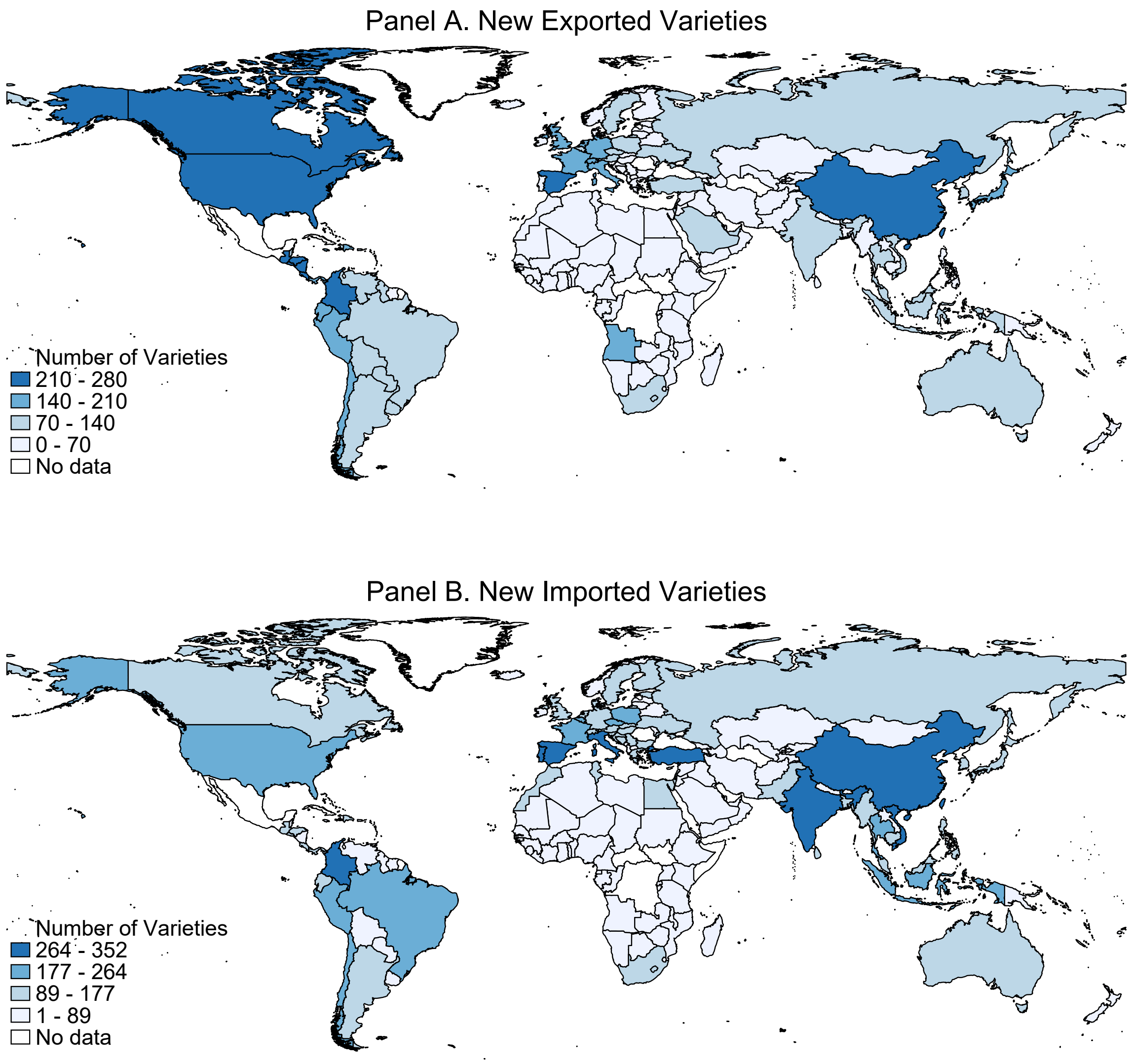
3.4.6 Source and Destination Countries of New Varieties

Figure 3.2 presents a spatial visualization of Mexico's main trade partners of new manufacturing goods in 2016. The color range displayed in these maps is determined by utilizing an equal interval classification. The purpose of this analysis is to understand which are the main source and destination countries for new traded varieties. This analysis is important because we are interested in identifying whether source and destination countries are the same (i.e., trade complementarities at the country level).

Panel A displays the destination countries for new exported varieties in 2016. We can notice that the main destinations for new exported varieties are: the United States (280 varieties), Guatemala (266 varieties), Costa Rica (256 varieties), Nicaragua (256 varieties), El Salvador (253 varieties), Spain (246 varieties), Colombia (238 varieties), Honduras (234 varieties), Canada (226 varieties), and China (212 varieties). We can conclude that most of these top destination countries are concentrated on the same continent (i.e., the Americas), except for Spain and China. A plausible explanation is that having a common language can facilitate doing business with Spain. Also, another possible explanation may be that doing business with China has become dynamic and appealing to Mexican exporters due to the large market size this Asian country represents. Moreover, as suspected, the United States represents the main destination for new goods exported by Mexico; this might be influenced by the geographical proximity and the free trade agreement between these countries (i.e., the former NAFTA agreement and the newly United States-Mexico-Canada Agreement).

On the other hand, Panel B exhibits the source countries of new imported varieties in 2016. We can observe that the main sources of new imported varieties are: Vietnam (352 varieties), followed by India (339 varieties), Spain (326 varieties), Turkey (324 varieties), China (315 varieties), Colombia (310 varieties), Portugal (299 varieties), Italy (277 varieties), Indonesia

Figure 3.2: Source and Destination Countries of New Traded Varieties



Notes: Panel A exhibits the number of new exported varieties by Mexico to the different destination countries. Panel B shows the number of new imported varieties by Mexico from the different source countries. The reference year is 2016.

(262 varieties), and Thailand (261 varieties). We can conclude that most of these source countries of new varieties are developing economies. Moreover, most of these source countries are concentrated in Asia, despite the high transportation costs and lack of trade agreements associated with these nations. It is also worth noticing that the United States is not among the top source countries of new imported varieties.

These maps reveal the presence of a certain trade complementarity effect on new varieties with Spain, Colombia, and China. These three countries constitute both main source and destination countries for new varieties traded with Mexico. Furthermore, we can observe that transportation costs and free trade agreements play an important role in the decision-making process of Mexican manufacturing firms that export new varieties; nonetheless, this does not hold when it comes to importing new varieties. Despite the tariffs associated with importing products from countries that do not benefit from free trade agreements, Mexican firms still consider more profitable to import new varieties from these Asian countries.

3.5 Methodology

Our methodology consists of a three-fold strategy, where we employ a fixed effects logit model, a fixed effects negative binomial model, and a linear fixed effects model. We present each of these methodologies in the following subsection.

3.5.1 Fixed Effects Logit Model

We first start with a logit model with fixed effects to estimate the probability of exporting new varieties as a function of the number of new imported varieties in the previous year. Our first baseline regression equation is the following:

$$Prob(X\_NEW_{cit} > 0) = \beta\, ln(IM\_NEW_{cit-1}) + \gamma X_{ct-1} + \nu_c + \nu_i + \nu_t + \varepsilon_{cit}, \tag{3.7}$$

where $Prob(X\_NEW\_cit > 0)$ is a dummy variable that stands for the probability of exporting new varieties by industry i to country c in year t; this dependent variable is equal to one if at least one new variety is exported to a country, and zero otherwise. The main explanatory variable, $ln(IM\_NEW_{cit-1})$, stands for the log number of new imported varieties by industry i from country c in year $t-1$. The vector of control variables X_{ct-1} includes GDP in PPP dollars and the score of starting a business. GDP is a standard control variable used in the trade literature (see, for example, Jansen & Stokman (2014)). We also use score of starting a business to proxy the ease of doing business with partner countries.[19]

All the right-hand side variables are expressed in natural logs and lagged by one year to avoid potential endogeneity. It is worth mentioning that as part of the Robustness Analysis, we increase the lag length of the independent variables to two and three lags, we also use different combinations of fixed effects (e.g., industry-year, country-year, and sector-year fixed effects), and we use an instrumental variable approach to control for any potential endogeneity issue in the main explanatory variable. Returning to the baseline specification, we also add a set of country (ν_c), industry (ν_i), and year (ν_t) fixed effects to all the regressions.[20] Finally, we report the average marginal effects in the Results section. It is important to mention that a drawback from including fixed effects in our specification is that the model drops observations when our variables are time-invariant.

3.5.2 Fixed Effects Negative Binomial Model

Having examined the role of importing new varieties on the probability of exporting new varieties, now we want to measure the impact of importing new varieties on the export number of new varieties. To test this relationship, we use a negative binomial model with fixed effects,

[19] We also employ other trade-related control variables as part of the Robustness Analysis.

[20] We declare the dataset to be a panel by establishing country-industry paired observations as the cross-section observations, and year as the time series observations. Thus, including both country and industry fixed effects in our regression equations is equivalent to including variety fixed effects.

given that the dependent variable is a count variable with overdispersion around the mean (i.e., the variance is larger than the mean). In other words, the sample is concentrated on a few discrete values. The Appendix includes a subsection on the negative binomial model.

This negative binomial model represents advantages over the Poisson model since it relaxes the assumption that the mean of the distribution should be equal to the variance. Nonetheless, a drawback from models combined with fixed effects is that these models drop observations where the variables are time-invariant. Thus, our second baseline regression equation is defined as:

$$X\_NEW_{cit} = \alpha + \beta\, ln(IM\_NEW_{cit-1}) + \gamma X_{ct-1} + \nu_c + \nu_i + \nu_t + \varepsilon_{cit}, \qquad (3.8)$$

where X_NEW_cit is the number of new exported varieties by industry i to country c in year t. The main explanatory variable, $ln(IM\_NEW_{cit-1})$, is the log number of new imported varieties by industry i from country c in year $t-1$. The control variables are GDP of partner countries and the score of starting a business, which we consider it as a proxy for the ease of doing business with different countries. All the independent variables are expressed in natural logs and lagged by one year to avoid potential endogeneity. We also include a full set of country, industry, and year fixed effects.

3.5.3 Linear Fixed Effects Model

As a third step, we measure the impact of importing new varieties on the export share of new varieties at the extensive and intensive margins using linear regressions with fixed effects. The definitions of the extensive and intensive margins employed in this chapter are similar to Lawless (2010); we combine these margins with the export shares aiming to investigate the trade complementarity effect at the country level.[21] In other words, we are interested in exploring whether importing new varieties from a source country has an effect on exports of new varieties to that same country. The regression equation to evaluate the number of new imported varieties on the share of new exported varieties at the extensive margin is:

$$ExtM\_X\_NEW_{cit} = \alpha + \beta\, ln(IM\_NEW_{cit-1}) + \gamma X_{ct-1} + \nu_c + \nu_i + \nu_t + \varepsilon_{cit}, \qquad (3.9)$$

where the dependent variable represents the share of the number of new exported varieties by industry i to country c in time t over the total number of new exported varieties by industry i in time t; this export share is expressed as a percentage. The main explanatory variable, $ln(IM\_NEW_{cit-1})$, is the number of new imported varieties by industry i from country c in year $t-1$. The regressions are controlled by GDP and the score of starting a business; these regressions also include a full set of country, industry, and year fixed effects. All the independent variables are expressed in their natural log form and lagged by one year.

[21] Lawless (2010) defines the extensive margin as the number of exporting firms, and the intensive margin as the average export sales. We follow the same logic and incorporate a feature that allows us to identify a trade complementarity effect; thus, we define the extensive margin as the share of the number of new exported varieties, and the intensive margin as the share of the export value of new varieties.

On the other hand, the specification to examine the number of new imported varieties on the share of new exported varieties at the intensive margin is as follows:

$$IntM\_X\_NEW_{cit} = \alpha + \beta\, ln(IM\_NEW_{cit-1}) + \gamma X_{ct-1} + \nu_c + \nu_i + \nu_t + \varepsilon_{cit}, \tag{3.10}$$

where the dependent variable denotes the share of the value (in U.S. dollars) of new exported varieties by industry i to country c in time t over the total value (in U.S. dollars) of new exported varieties by industry i in time t; this export share is also expressed as a percentage. As before, the main explanatory variable, $ln(IM\_NEW_{cit-1})$, is the number of new imported varieties by industry i from country c in year $t-1$. This regression equation also incorporates controls for GDP and score of starting a business. A full set of country, industry, and year fixed effects is included. All the independent variables are expressed in natural logs and lagged by one year.

3.6 Results

Moving on to the results, we start our analysis by employing a fixed effects logit model to estimate the probability of exporting new varieties as a function of the number of new imported varieties. Then, we use a fixed effects negative binomial model to examine the impact of importing new varieties on the number of new exported varieties. Finally, we use a linear fixed effects model to estimate the impact of importing new varieties on the export share of new varieties at the extensive and intensive margins.

3.6.1 Probability to Export New Varieties

We start by examining whether new imported varieties have an impact on the probability of exporting new varieties (i.e., $Prob(X\_NEW_{cit} > 0)$) by estimating Eq.(3.7). To do so, we transform the dependent variable into a dummy variable that equals one if at least one new variety is exported by industry i to country c in year t, and zero otherwise. For this first baseline specification, we use a logit model with fixed effects. The estimation sample is composed of 74,240 new varieties over the period 2005-2016. All the right-hand side variables are expressed in their logarithmic form and lagged by one year. Moreover, all the specifications include a full set of country, industry, and year fixed effects.

Table 3.7 reports the average marginal effects using a logit model with fixed effects. In column (1), we start our specification by only including the main explanatory variable, along with country, industry, and year fixed effects. We can observe new imported varieties is positive and strongly statistically significant. Then, we introduce the control variables in a stepwise manner. In column (2), we can observe that our main explanatory variable remains strongly significant once we include GDP in our specification; nonetheless, the magnitude of the main explanatory variable reduces dramatically. In column (3), we now include the score of starting a business as a control; we can notice that the main explanatory variable remains positive and strongly significant, albeit the magnitude is smaller compared to the first column.

Table 3.7: Probability of Exporting New Varieties

VARIABLES	(1) AME Prob_X_NEW	(2) AME Prob_X_NEW	(3) AME Prob_X_NEW	(4) AME Prob_X_NEW
ln(New Imported Varieties)_cit-1	0.0342***	0.0003***	0.0195***	0.0013***
	(0.0025)	(0.0001)	(0.0015)	(0.0001)
ln(GDP in PPP)_ct-1		0.0005***		0.0012***
		(0.0394)		(0.0407)
ln(Starting a Business)_ct-1			0.0556***	0.0033***
			(0.0315)	(0.0326)
Observations	606,648	606,648	606,648	606,648
Number of varieties	50,554	50,554	50,554	50,554
Prob Wald Chi2	0.000	0.000	0.000	0.000
Country FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table constitutes the baseline table for analyzing the probability of exporting new varieties by industry i to country c in time t. The table reports the average marginal effects of Eq.(3.7) using a logit model with fixed effects. The estimation sample is conformed by 74,240 new varieties over the period 2005-2016; however, some observations were dropped from the sample due to the nature of the methodology employed. All independent variables are expressed in natural logs and lagged by one year. All regressions include country, industry, and year fixed effects.

In column (4), we present our baseline specification, where we include all the control variables and a full set of country, industry, and year fixed effects. We can observe that our main explanatory variable, $ln(IM\_NEW_{cit-1})$, remains positive and strongly statistically significant. Moreover, both GDP and score of starting a business are positive and strongly significant, as we would expect from the trade literature. It is worth mentioning that the effect of the main explanatory variable is similar in size to GDP.

Our results suggest that a 10% increase in new imported varieties increases the probability of exporting new varieties by 0.01 percentage points. From our estimation sample, about 15% of exports were recorded as new varieties. Therefore, our results are statistically significant, albeit not economically meaningful.[22] Despite this non economically significant outcome at the country level, it may be interesting to replicate the analysis considering state heterogeneity shown in Chapter 2.[23]

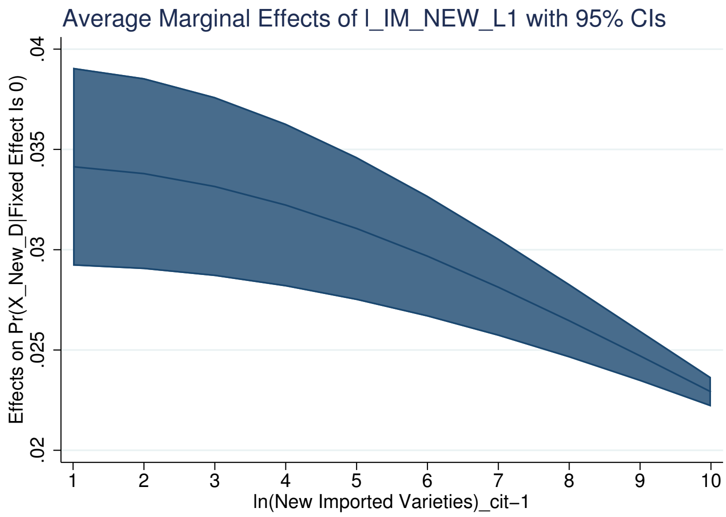
This positive and statistically significant impact of new imported varieties on the probability of exporting new varieties for Mexico is consistent with Aristei et al. (2013) for Eastern European and Central Asian countries, and with Lo Turco & Maggioni (2013) for Italy. However, our results differ from Castellani & Fassio (2019), who show that the number of new imported inputs has no impact on the probability of exporting a new product in the case of Sweden.

In Figure 3.3, we exhibit a graph of the average marginal effects of importing new varieties against the probability of exporting new varieties with a confidence interval at the 95% level. We can interpret that the effect is larger for the first new imported varieties, and then declines monotonically, but remains statistically significant.

[22] If there is a 10% increase in new imported varieties, this give us 0.1451. Thus, 100*((0.1451-0.145)/0.145) = 0.07%. This suggests there is a 0.07% increase in exports of new varieties in our sample.

[23] We could not perform this additional analysis at the state level because the National Institute of Statistics and Geography (INEGI) does not report imports at this level of disaggregation due to confidentiality issues.

Figure 3.3: Probability of Exporting New Varieties



Notes: This graph displays the average marginal effects of importing new varieties on the probability of exports of new varieties.

3.6.2 Number of New Exported Varieties

Next, we examine the relationship between the number of new imported varieties and the number of new exported varieties by estimating Eq.(3.8). For this second baseline specification, we use a negative binomial model with fixed effects, as this is the most appropriate methodology when the dependent variable is a count variable with overdispersion around the mean (i.e., the variance is larger than the mean). Just as before, the estimation sample is constituted by 74,240 new varieties over the period 2005-2016. All the right-hand side variables are expressed in logs and lagged by one year. All the specifications include country, industry, and year fixed effects.

Table 3.8 shows the coefficients of the negative binomial model with fixed effects.[24] In column (1), we start by estimating a negative binomial regression using a cross-section approach. The aim of this exercise is to include an overdispersion parameter α. If this dispersion parameter equals zero, then a Poisson model would be a better approach. On the other hand, if this dispersion parameter is significantly greater than zero, then this means that the data presents overdispersion and therefore, a negative binomial model would be a better approach. In this case, we can observe that α is larger than zero; this confirms that our data presents overdispersion; thus, the negative binomial model is the most appropriate approach for our specification.

In column (2), we now report the results of the negative binomial model with fixed effects using a panel approach.[25] We can observe new imported varieties is positive and strongly statistically significant. Then, we introduce the control variables one by one. In column (3), we can notice that the magnitude of the explanatory variable slightly reduces when we incorporate

[24]The number of observations in Table 3.8 differs from previous Table 3.7, where we use a logit fixed effect model. The likelihood function in Table 3.7 is only identified from switchers (i.e., 0 to 1, or 1 to 0), so observations always 1 or always 0 do not contribute to the likelihood function (i.e., log 1 = 0 and log 0 is undefined).

[25]Stata does not possess an option to include an overdispersion parameter α for negative binomial regressions using a panel approach. This is the reason why we use a cross-section approach in column (1), as it exhibits an overdispersion parameter.

Table 3.8: Number of New Exported Varieties

VARIABLES	(1) Cross-sec X_NEW	(2) Panel X_NEW	(3) Panel X_NEW	(4) Panel X_NEW	(5) Panel X_NEW
ln(New Imported Varieties)_cit-1	1.0287***	0.1065***	0.0977***	0.1037***	0.0955***
	(0.0070)	(0.0067)	(0.0067)	(0.0067)	(0.0067)
ln(GDP in PPP)_ct-1			0.1011***		0.0950***
			(0.0063)		(0.0063)
ln(Starting a Business)_ct-1				0.3428***	0.3156***
				(0.0239)	(0.0238)
Constant	-1.7909***	0.4273***	-2.2091***	-0.9872***	-3.3532***
	(0.0034)	(0.0160)	(0.1626)	(0.0997)	(0.1845)
Observations	890,880	607,524	607,524	607,524	607,524
Prob Wald Chi2	0.000	0.000	0.000	0.000	0.000
Alpha	2.988				
Country FE	NO	YES	YES	YES	YES
Industry FE	NO	YES	YES	YES	YES
Year FE	NO	YES	YES	YES	YES

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: This table constitutes the baseline table for examining the number of new exported varieties by industry i to country c in time t. The table reports the coefficients of Eq.(3.8) employing a negative binomial model with fixed effects. The estimation sample is conformed by 74,240 new varieties over the period 2005-2016; however, some observations were dropped from the sample due to the nature of the methodology employed. All independent variables are expressed in natural logs and lagged by one year. All panel regressions include country, industry, and year fixed effects.

GDP in our regression. In column (4), we now include the score of starting a business as a control variable; we can notice that the magnitude of the coefficient did not change substantially compared to the one presented in the second column.

In column (5), we present our baseline specification, where we include all the control variables and the full set of country, industry, and year fixed effects. We can observe that the results do not change substantially when we incorporate all the control variables. In other words, our main explanatory variable remains positive and strongly statistically significant; furthermore, the magnitude of the coefficients do not exhibit much variation. Our results suggest that a 1% increase in new imported varieties is associated with an increase in the number of new exported varieties by about 0.001.

In terms of the control variables, GDP and score of starting a business are both positive and strongly significant. Compared to the literature, the direction and magnitude of the coefficient of our main explanatory variable are analogous to Castellani & Fassio (2019) for Swedish firms.

3.6.3 The Extensive and Intensive Margins

We now evaluate how the number of new imported varieties from a source country has an impact on the export share of the number of new varieties to that same country (i.e., the extensive margin). Likewise, we evaluate how the number of new imported varieties from a source country has an impact on the export share of the value (in U.S. dollars) of new varieties to that same partner country (i.e., the intensive margin).

To perform our analysis at the extensive and intensive margins, we estimate Eq.(3.9) and Eq.(3.10), respectively. An interesting feature of this chapter, is that our results suggest some degree of trade complementarities between source and destination countries of new traded varieties at the country level. For example, if Mexico increases imports of new varieties from Spain, this may lead to an increase of Mexican exports of new varieties to Spain compared to other destinations.

The methodology employed to study the export share at the extensive and intensive margins consists of linear regressions with fixed effects. As a recap, the estimation sample is constituted by 74,240 new varieties over the period 2005-2016. All the right-hand side variables are expressed in logs and lagged by one year. Moreover, all the specifications include a full set of country, industry, and year fixed effects.

The Extensive Margin

Table 3.9 reports the results of the relationship between new imported varieties and new exported varieties at the extensive margin using linear regressions with fixed effects.[26] The dependent variable is the export share of the number of new varieties by industry i to country c in time t over the total number of new exported varieties by industry i in time t. The main explanatory variable, $ln(IM\_NEW_{cit-1})$, is the number of new imported varieties by industry i from country c in year $t-1$.

Column (1) constitutes the starting point where we only include the main explanatory variable along with the full set of country, industry, and year fixed effects. We can observe that new imported varieties is positive and strongly statistically significant. We then introduce the control variables in a stepwise manner. In column (2), we introduce GDP as the sole control variable. Then, in column (3), we include only the score of starting a business as the control variable.

Table 3.9: Export Share of New Varieties at the Extensive Margin

VARIABLES	(1) ExtM_X_NEW	(2) ExtM_X_NEW	(3) ExtM_X_NEW	(4) ExtM_X_NEW
ln(New Imported Varieties)_cit-1	0.0614***	0.0613***	0.0603***	0.0604***
	(0.0137)	(0.0137)	(0.0137)	(0.0137)
ln(GDP in PPP)_ct-1		0.0165		-0.0350
		(0.0420)		(0.0426)
ln(Starting a Business)_ct-1			0.2044***	0.2110***
			(0.0346)	(0.0351)
Constant	1.1421***	0.7172	0.3016**	1.1747
	(0.0179)	(1.0804)	(0.1418)	(1.0797)
Observations	890,880	890,880	890,880	890,880
R-squared	0.003	0.003	0.006	0.006
Number of varieties	74,240	74,240	74,240	74,240
Country FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

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

Notes: This table constitutes the baseline table for investigating the export share of the number of new varieties by industry i to country c in time t (i.e., exports at the extensive margin). This table estimates Eq.(3.9) using linear regressions with fixed effects. The estimation sample is conformed by 74,240 new varieties over the period 2005-2016. All independent variables are expressed in natural logs and lagged by one year. All regressions include country, industry, and year fixed effects.

[26]Table 3.9 has a different set of observations compared to Tables 3.7 and 3.8. An advantage of the linear fixed effects model is that it does not drop observations compared to the fixed effects logit model and fixed effects negative binomial model.

Finally, column (4) corresponds to our baseline specification, where we introduce both GDP and score of starting a business as controls. We can observe that the magnitude of the coefficients does not change dramatically compared to the first column. We can conclude that the number of new imported varieties has a positive and strong significant effect, albeit small, on the export share at the extensive margin. The results suggest that a 1% increase in new imported varieties from a country leads to an increase of 0.0006 percentage points in the export share of new varieties at the extensive margin. We can also observe that GDP has no effect on the export share at the extensive margin. In contrast, the score of starting a business has a positive and statistically significant effect.

The Intensive Margin

On the other hand, Table 3.10 shows the results of the relationship between new imported varieties and new exported varieties at the intensive margin using linear regressions with fixed effects.[27] The dependent variable corresponds to the export share of the value (in U.S. dollars) of new exported varieties by industry i to country c in time t over the total value of new exported varieties by industry i in time t. The main explanatory variable, $ln(IM\_NEW_{cit-1})$, is the number of new imported varieties by industry i from country c in the previous year.

In column (1), we start by only including our main explanatory variable along with the full set of country, industry, and year fixed effects. We can observe that new imported varieties is positive and strongly statistically significant. These results are consistent with Table 3.9, where the dependent variable is the export share at the extensive margin.

Table 3.10: Export Share of New Varieties at the Intensive Margin

	(1)	(2)	(3)	(4)
VARIABLES	IntM_X_NEW	IntM_X_NEW	IntM_X_NEW	IntM_X_NEW
ln(New Imported Varieties)_cit-1	0.0781***	0.0777***	0.0763***	0.0763***
	(0.0265)	(0.0265)	(0.0265)	(0.0265)
ln(GDP in PPP)_ct-1		0.0863		0.0001
		(0.0642)		(0.0646)
ln(Starting a Business)_ct-1			0.3531***	0.3531***
			(0.0542)	(0.0546)
Constant	1.1454***	-1.0739	-0.3065	-0.3084
	(0.0272)	(1.6503)	(0.2225)	(1.6446)
Observations	890,880	890,880	890,880	890,880
R-squared	0.002	0.002	0.005	0.005
Number of varieties	74,240	74,240	74,240	74,240
Country FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

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

Notes: This table constitutes the baseline table for examining the export share of the value (in U.S. dollars) of new varieties by industry i to country c in time t (i.e., exports at the intensive margin). This table estimates Eq.(3.10) using linear regressions with fixed effects. The estimation sample is conformed by 74,240 new varieties over the period 2005-2016. All independent variables are expressed in natural logs and lagged by one year. All regressions include country, industry, and year fixed effects.

The next step is to include the control variables one by one. In column (2), we introduce GDP as the only control variable. In column (3), we only include the score of starting a business

[27] Just as mentioned before, Table 3.10 has a different set of observations compared to Tables 3.7 and 3.8 because the linear fixed effects model does not drop observations.

as the control variable. Finally, column (4) corresponds to our baseline specification, where we incorporate both GDP and score of starting a business, along with the full set of country, industry, and year fixed effects. We can observe that the magnitude of the coefficients holds throughout the different specifications. We can conclude that the number of new imported varieties also has a positive and strong significant effect, albeit small, on the export share at the intensive margin. These results suggest that a 1% increase in new imported varieties from a country leads to an increase of roughly 0.0008 percentage points in the export share of new varieties at the intensive margin. Once again, GDP has no effect on exports at the intensive margin, while the score of starting a business has a positive and significant effect.

In summary, we can observe that there is a presence of some degree of trade complementarities between imports and exports of new varieties at the country level. We can also observe that the magnitude of the main explanatory variable is slightly larger at the intensive margin compared to the extensive margin. To the best of our knowledge, the impact of new imported varieties on the export share of new varieties at the extensive and intensive margins has not been explored before. The most closely related study is Navas et al. (2020), however, the authors examined the impact of firm's productivity on exports at the extensive and intensive margins.[28]

3.7 Robustness Analysis

This section below presents a series of robustness checks for the different methodologies employed in the study. The structure of most of the tables in this section report in their first column the average marginal effects resulting from the fixed effects logit model. The second column displays the coefficients employing a fixed effects negative binomial model. The last two columns reveal the export shares of new varieties at the extensive and intensive margins, respectively, using linear fixed effects models.

We start the robustness analysis by including trade gravity variables to our specifications. Then, we exclude countries representing both source and destination countries of new varieties from our sample. This is followed by the exclusion of Mexico's main trading partner: the United States. Next, we restrict our sample to Mexico's 50 major trading partners. This is followed by dividing the estimation sample into country sub-samples based on their income profile: high-income OECD and low- and middle-income countries. Then, we reduce our sample to the top trading partners of new varieties. Also, we use another sub-sample comprising top industries trading new varieties. After that, we use other trade-related control variables in our specifications.

We also employ an alternative methodology to the fixed effects negative binomial model: a fixed effects Poisson model. Next, we deal with zero-value observations in the dependent

[28]It is worth noting that Navas et al. (2020) define the extensive margin as the export status, which consists of a dummy variable that takes the value of one if a firm exports to a country in a given year, and zero otherwise. On the other hand, the authors define the intensive margin as the total exports of a firm to a country in a given year.

variable. Later, we use a log-log model. Then, we also look at the contemporaneous effects of importing new varieties. After that, we increase the lag length of the independent variables to avoid potential reverse causality. We also include different combinations of fixed effects to tackle a potential omitted variable bias. Next, we analyze input-output linkages across sectors. After, we examine the marginal effects by industry. Finally, we run two-stage regressions to account for potential endogeneity arising from potential omitted variables and reverse causality.

3.7.1 Trade Gravity Variables

We start by introducing standard gravity variables in our analysis; these gravity variables are distance, border, free trade agreement, landlocked-status, common continent, common language, and common colonizer. By distance, we refer to the geographical log distance between Mexico City and the capital city of each partner country calculated by the great-circle distance formula; this variable is a proxy for transportation costs. Then, with the border variable, we refer to the countries that share a border with Mexico: the United States, Guatemala, and Belize. By free trade agreement, we construct a dummy variable that equals one if the partner country has a free trade agreement with Mexico; and zero otherwise. Next, a common continent refers to partner countries located in the Americas. Common language refers to the country's official language being Spanish. Finally, a common colonizer refers to being a former colony of Spain; thus, invoking shared historical and cultural linkages.

In Table 3.11, we include these trade gravity variables in our specifications as additional controls to GDP and score of starting a business. Due to the time-invariant nature of these gravity variables, we cannot include country fixed effects as these observations will drop out; nevertheless, we include industry and year fixed effects. All the right-hand side variables are expressed in natural logs and lagged by one year. It is worth mentioning that the number of observations and varieties differs from column to column due to the nature of the logit and negative binomial methodologies combined with fixed effects, which drop time-invariant observations from the sample.

From this table, we can observe that importing new varieties has a positive and statistically strong effect on exporting new varieties throughout the different specifications (i.e., on the probability, number of varieties, and shares), which is consistent with our previous findings. Moreover, GDP and score for starting a business also have positive and statistically strong effects in all the specifications. Moving forward to the trade gravity variables, we can notice that being in the Americas and having Spanish as the official language have a positive and statistically significant effect on exporting new varieties throughout the different specifications (i.e., on the probability, number of varieties, and shares). A plausible explanation is that being on the same continent reduces transportation costs and timing. Furthermore, sharing a language facilitates doing business, which leads to an increase in traded varieties.

On the other hand, we can observe that distance, landlocked-status, and common colonizer have a negative and statistically strong effect on exporting new varieties in all the specifications (i.e., on the probability, number of varieties, and shares). We could interpret this as having a

Table 3.11: Trade Gravity Variables

VARIABLES	(1) AME Prob_X_NEW	(2) X_NEW	(3) ExtM_X_NEW	(4) IntM_X_NEW
ln(New Imported Varieties)_cit-1	0.0344***	0.2073***	0.1157***	0.1770***
	(0.0009)	(0.0059)	(0.0149)	(0.0272)
ln(GDP in PPP)_ct-1	0.0317***	0.2552***	0.1754***	0.2493***
	(0.0022)	(0.0018)	(0.0042)	(0.0058)
ln(Starting a Business)_ct-1	0.0497***	0.3909***	0.1442***	0.0991***
	(0.0146)	(0.0123)	(0.0156)	(0.0262)
ln(Distance)	-0.0649***	-0.5384***	-0.4628***	-0.1449***
	(0.0104)	(0.0081)	(0.0254)	(0.0351)
Border	-0.0658***	-0.5668***	0.2496***	0.6044***
	(0.0195)	(0.0145)	(0.0623)	(0.0868)
Free Trade Agreement	-0.0003	-0.0201***	-0.0347**	-0.0380*
	(0.0078)	(0.0064)	(0.0135)	(0.0212)
Landlocked	-0.0120***	-0.0884***	-0.0822***	-0.0961***
	(0.0119)	(0.0104)	(0.0117)	(0.0191)
Continent	0.0545***	0.3524***	0.4175***	0.8943***
	(0.0143)	(0.0111)	(0.0332)	(0.0515)
Language	0.0500***	0.4473***	0.7009***	0.5304***
	(0.0153)	(0.0117)	(0.0368)	(0.0545)
Colonizer	-0.0166***	-0.1211***	-0.2928***	-0.2873***
	(0.0153)	(0.0118)	(0.0303)	(0.0485)
Constant		-3.7869***	-0.0039	-4.6975***
		(0.0900)	(0.2380)	(0.3492)
Observations	885,708	885,708	886,224	886,224
Number of industries	1,016	1,016	1,023	1,023
Prob Wald Chi2	0.000	0.000		
R-squared			0.014	0.007
Country FE	NO	NO	NO	NO
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table includes the standard trade gravity variables as additional controls to our baseline specifications. The dependent variable in column (1) corresponds to the probability of exporting new varieties by industry i to country c in time t; this column reports the average marginal effects using a logit model with fixed effects. The dependent variable in column (2) is the number of new exported varieties by industry i to country c in time t; this column reports the coefficients using a negative binomial with fixed effects approach. The dependent variable in column (3) is the export share of the number of new varieties by industry i to country c in time t (i.e., the extensive margin). On the other hand, the dependent variable in column (4) is the export share of the value (in U.S. dollars) of new varieties by industry i to country c in time t (i.e., the intensive margin). Both columns (3) and (4) are estimated by linear regressions with fixed effects. All time-variant independent variables are expressed in natural logs and lagged by one year. All regressions include industry and year fixed effects.

greater distance between Mexico and a partner country decreases exports of new varieties; this is in line with Clark & Van Wincoop (2001), who explain that distance represents an important trade barrier due to associated transportation costs. In the case of the landlocked status, a country enclosed by land incurs in more transportation costs to ship the merchandise; thus, this landlocked status also hinders exports of new varieties. Surprisingly, having a common historical background does not positively impact exports of new varieties those countries. This could be explained by the fact that source countries of new varieties are mainly located in Asia, which do not have a common heritage with Mexico.

Furthermore, we can observe that having a free trade agreement between Mexico and a partner country has a negative and statistically significant effect on exporting new varieties measured at levels and shares, albeit an insignificant effect on the probability. We can interpret this in a similar way as having a common colonizer; in other words, we can notice that sources of new varieties are mainly Asian countries, which do not possess free trade agreements with Mexico. This means that new imported varieties from Asian countries must pay tariffs, and thus, these tariffs increase the costs of importing those new varieties. This situation could translate into a motivation for the Mexican government to negotiate free trade agreements with these identified source countries.

Finally, the only variable that we find mixed results is common border. The results suggest that sharing a border has a negative and strongly significant effect on exports of new varieties measured as a probability (column 1) and as levels (column 2); nonetheless, border has a positive and strong significant effect on exports of new varieties measured in shares (columns 3 and 4).

3.7.2 Exclusion of Source-Destination Countries

Next, we analyze our four different empirical specifications excluding Spain, Colombia, and China from our sample. The aim is to examine whether the relationship between new imported varieties and new exported varieties hold after discarding these three countries that constitute both main source and destination countries of new traded varieties according to Figure 3.2. On the one hand, Mexico, Spain, and Colombia possess strong historical and cultural linkages, including the language; thus these results are in line with the knowledge about the destination country mechanism. On the other hand, China has become an appealing market to Mexican exporters due to its large market size. With this analysis, we want to confirm that our results hold even after excluding these countries.

Table 3.12: Estimation Sample Excluding Source-Destination Countries

VARIABLES	(1) AME Ex.ESP Prob	(2) Ex.ESP X_NEW	(3) Ex.ESP ExtMarg	(4) Ex.ESP IntMarg	(5) AME Ex.COL Prob	(6) Ex.COL X_NEW	(7) Ex.COL ExtMarg	(8) Ex.COL IntMarg	(9) AME Ex.CHN Prob	(10) Ex.CHN X_NEW	(11) Ex.CHN ExtMarg	(12) Ex.CHN IntMarg
ln(IM_NEW)_cit-1	0.0041***	0.0923***	0.0583***	0.0748***	0.0004***	0.0973***	0.0544***	0.0700***	0.0002***	0.0956***	0.0634***	0.0651**
	(0.0003)	(0.0068)	(0.0137)	(0.0266)	(0.0001)	(0.0068)	(0.0135)	(0.0263)	(0.0001)	(0.0068)	(0.0137)	(0.0263)
ln(GDP)_ct-1	0.0023*	0.0926***	-0.0577	-0.0302	0.0005***	0.0940***	0.0003	0.0359	0.0003***	0.0928***	-0.0242	0.0658
	(0.0413)	(0.0064)	(0.0419)	(0.0642)	(0.0408)	(0.0063)	(0.0425)	(0.0646)	(0.0432)	(0.0066)	(0.0416)	(0.0609)
ln(Business)_ct-1	0.0119***	0.3195***	0.2199***	0.3658***	0.0012***	0.3263***	0.2123***	0.3520***	0.0006***	0.3310***	0.2101***	0.3582***
	(0.0328)	(0.0239)	(0.0349)	(0.0545)	(0.0327)	(0.0239)	(0.0350)	(0.0545)	(0.0327)	(0.0240)	(0.0350)	(0.0545)
Constant		-3.3084***	1.7063	0.4049		-3.3875***	0.2296	-1.2610		-3.3667***	0.8843	-2.0475
		(0.1855)	(1.0627)	(1.6326)		(0.1842)	(1.0771)	(1.6417)		(0.1900)	(1.0519)	(1.5402)
Observations	597,084	597,912	879,612	879,612	596,616	597,432	879,840	879,840	597,048	597,888	879,180	879,180
No. varieties	49,757	49,826	73,301	73,301	49,718	49,786	73,320	73,320	49,754	49,824	73,265	73,265
Prob Wald Chi2	0.000	0.000			0.000	0.000			0.000	0.000		
R-squared			0.006	0.005			0.005	0.005			0.006	0.006
Country FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table reports the results after excluding Spain, Colombia, and China from the estimation sample. Columns (1)-(4) present the results after excluding Spain; columns (5)-(8) exclude Colombia; and columns (9)-(12) exclude China from the sample. The dependent variable in columns (1), (5), and (9) corresponds to the probability of exporting new varieties by industry i to country c in time t; these columns report the average marginal effects using a logit model with fixed effects. The dependent variable in columns (2), (6), and (10) stands for the number of new exported varieties by industry i to country c in time t; these columns report the coefficients using a negative binomial with fixed effects approach. The dependent variable in columns (3), (7), and (11) is the export share of the number of new varieties by industry i to country c in time t (i.e., the extensive margin). On the other hand, the dependent variable in columns (4), (8), and (12) is the export share of the value (in U.S. dollars) of new varieties by industry i to country c in time t (i.e., the intensive margin). Columns (3)-(4), (7)-(8), and (11)-(12) are estimated by linear regressions with fixed effects. All independent variables are expressed in natural logs and lagged by one year. All regressions include country, industry, and year fixed effects.

In Table 3.12, we show the results for the different specifications excluding Spain, Colombia, and China from the sample. Each set of four columns excludes only one of these three countries; the first four columns exclude Spain, the following four columns exclude Colombia, and the last four columns exclude China. All regressions include GDP and score of starting a business as control variables, as well as the full set of country, industry, and year fixed effects. It is worth mentioning that the number of observations and varieties differ from each column due to the nature of the logit and negative binomial methodologies combined with fixed effects, which drop time-invariant observations.

After excluding Spain from the sample, we can observe that the results are consistent; in other words, the coefficients of the main explanatory variable remain positive and strongly significant throughout the different models employed in this chapter (i.e., on the probability, levels, and shares). We can notice similar results after excluding Colombia and China from the estimation sample by separate; the coefficients of the main explanatory variables also remain positive and strongly significant for our four different specifications. A plausible explanation for the trade complementarity effect between imports and exports of new varieties at the country level could be explained by a knowledge about the destination country mechanism, as shown in Figure 3.2 and in Table 3.11, where we introduce trade gravity variables.

3.7.3 Exclusion of the United States

As part of this analysis, we also exclude Mexico's traditional main trading partner: the United States. Table 3.13 exhibits the results of our four specifications excluding the United States from the sample. Like before, all regressions include GDP and the score of starting a business as control variables, along with the full set of country, industry, and year fixed effects.

We can observe that new imported varieties are positive and strongly statistically significant throughout the four different specifications. Thus, these results are consistent with the baseline regressions presented in section 3.6. Nonetheless, we can observe that the size of the effect becomes larger for the probability and the export shares, at both extensive and intensive margins, once we exclude the United States. This suggests that other trading partners may play more important roles in the trade of new varieties than the United States.

Table 3.13: Estimation Sample Excluding the United States

VARIABLES	(1) Excl.USA Prob_X_NEW (AME)	(2) Excl.USA X_NEW	(3) Excl.USA ExtM_X_NEW	(4) Excl.USA IntM_X_NEW
ln(New Imported Varieties)_cit-1	0.0029***	0.0933***	0.0533***	0.0801***
	(0.0002)	(0.0067)	(0.0132)	(0.0259)
ln(GDP in PPP)_ct-1	0.0021**	0.1367***	-0.0753*	-0.0728
	(0.0407)	(0.0066)	(0.0424)	(0.0641)
ln(Starting a Business)_ct-1	0.0073***	0.3149***	0.1625***	0.2745***
	(0.0326)	(0.0240)	(0.0347)	(0.0539)
Constant		-4.3833***	2.3774**	1.8428
		(0.1892)	(1.0703)	(1.6247)
Observations	598,260	599,088	881,772	881,772
Number of varieties	49,855	49,924	73,481	73,481
Prob Wald Chi2	0.000	0.000		
R-squared			0.003	0.004
Country FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table reports the results after excluding the United States from the estimation sample. The dependent variable in column (1) corresponds to the probability of exporting new varieties by industry i to country c in time t; this column reports the average marginal effects using a logit model with fixed effects. The dependent variable in column (2) stands for the number of new exported varieties by industry i to country c in time t; this column reports the coefficients using a negative binomial with fixed effects approach. The dependent variable in column (3) is the export share of the number of new varieties by industry i to country c in time t (i.e., the extensive margin). On the other hand, the dependent variable in column (4) is the export share of the value (in U.S. dollars) of new varieties by industry i to country c in time t (i.e., the intensive margin). Columns (3)-(4) are estimated by linear regressions with fixed effects. All independent variables are expressed in natural logs and lagged by one year. All regressions include country, industry, and year fixed effects.

3.7.4 Restricted Sample

Then, we restrict our sample to Mexico's major 50 trade partners.[29] This restricted sample is formed by 39,770 new varieties over the period 2005-2016. All the independent variables are in their log form and lagged by one year. All the specifications include a full set of country, industry, and year fixed effects. The reason why the number of observations and varieties differ from column to column is due to the nature of the logit and negative binomial methodologies combined with fixed effects, which drop time-invariant observations.

Table 3.14 presents the results of the different methodologies employed in the analysis. In column (1), we report the average marginal effects using a logit model with fixed effects. These results suggest that a 1% increase in new imported varieties is associated with an increase of 0.02 percentage points on the probability of exporting new varieties. Compared to the baseline regression in Table 3.7, we can observe that the impact of new imported varieties is larger in the restricted sample.

Column (2) displays the coefficients using a negative binomial model with fixed effects. The results suggest that a 1% increase in new imported varieties is associated with an increase in the number of new exported varieties by about 0.0005. The magnitude of the coefficient for

[29] Mexico's major 50 trade partners were determined based on the annual mean of total trade (i.e., the sum of imports and exports). These main trade partners are: Aruba, Argentina, Australia, Austria, Belgium, Brazil, Canada, Chile, China, Colombia, Costa Rica, Czech Republic, Dominican Republic, Ecuador, El Salvador, Finland, France, Germany, Guatemala, Honduras, Hong Kong (China), Hungary, India, Indonesia, Ireland, Israel, Italy, Japan, South Korea, Malaysia, the Netherlands, Nicaragua, Panama, Peru, the Philippines, Poland, Portugal, Puerto Rico, Russian Federation, Saudi Arabia, Singapore, Spain, Sweden, Switzerland, Taiwan, Thailand, United Kingdom, United States, Venezuela, and Vietnam.

Table 3.14: Restricted Sample to Major Trade Partners

VARIABLES	(1) AME Prob_X_NEW	(2) X_NEW	(3) ExtM_X_NEW	(4) IntM_X_NEW
ln(New Imported Varieties)_cit-1	0.0230***	0.0519***	0.0278	0.0666*
	(0.0027)	(0.0072)	(0.0203)	(0.0387)
ln(GDP in PPP)_ct-1	-0.0095	-0.0243**	-0.2729**	-0.3232*
	(0.0557)	(0.0098)	(0.1258)	(0.1924)
ln(Starting a Business)_ct-1	0.0510***	0.1212***	0.2744***	0.6169***
	(0.0438)	(0.0301)	(0.0896)	(0.1351)
Constant		0.9956***	7.9565**	7.9199
		(0.2849)	(3.2749)	(5.0476)
Observations	368,280	369,072	477,192	477,192
Number of varieties	30,690	30,756	39,766	39,766
Prob Wald Chi2	0.000	0.000		
R-squared			0.054	0.031
Country FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table focuses on the restricted sample composed by 39,770 new traded varieties over the period 2005-2016; this restricted sample corresponds to Mexico's 50 main trade partners. The dependent variable in column (1) corresponds to the probability of exporting new varieties by industry i to country c in time t; this column reports the average marginal effects using a logit model with fixed effects. The dependent variable in column (2) is the number of new exported varieties by industry i to country c in time t; this column reports the coefficients using a negative binomial with fixed effects approach. The dependent variable in column (3) is the export share of the number of new varieties by industry i to country c in time t (i.e., the extensive margin). On the other hand, the dependent variable in column (4) is the export share of the value (in U.S. dollars) of new varieties by industry i to country c in time t (i.e., the intensive margin). Both columns (3) and (4) are estimated by linear regressions with fixed effects. All independent variables are expressed in natural logs and lagged by one year. All regressions include country, industry, and year fixed effects.

the restricted sample is smaller than the coefficient in the baseline specification in Table 3.8.

Column (3) exhibits the impact of importing new varieties on the export share of new varieties at the extensive margin. We can observe that the coefficient remains positive, albeit statistically insignificant. On the other hand, column (4) shows the effect of importing new varieties on the export share of new varieties at the intensive margin. In this case, we can notice that the coefficient remains positive and statistically significant. We can infer that this trade complementarity effect is present at the intensive margin (i.e., share of the value of new exported varieties), but is not at the extensive margin (i.e., share of the number of new exported varieties for the restricted sample). Finally, these coefficients evaluated at the shares are both smaller than those on the baseline specifications in Tables 3.9 and 3.10.

3.7.5 Income Profile

In a similar vein as Lo Turco & Maggioni (2013), we now divide the estimation sample into two country sub-samples: high-income OECD countries and low- and middle-income countries. The objective now is to examine the relationship between new imported varieties and new exported varieties depending on the income profile of partner countries. All the specifications contain the same main explanatory variable, which is the number of new imported varieties by industry i from country c in the previous year $t-1$. All regressions are controlled by GDP and the score of starting a business and include a full set of country, industry, and year fixed effects. Once again, the number of observations and varieties differ from column to column due to the nature of the logit and negative binomial methodologies combined with fixed effects, which drop time-invariant observations.

Table 3.15: Income Profile

VARIABLES	(1) AME High Income Prob_X_NEW	(2) High Income X_NEW	(3) High Income ExtM_X_NEW	(4) High Income IntM_X_NEW	(5) AME Low_Middle Prob_X_NEW	(6) Low_Middle X_NEW	(7) Low_Middle ExtM_X_NEW	(8) Low_Middle IntM_X_NEW
ln(New Imported Varieties)_cit-1	0.0001***	0.0875***	0.0662***	0.0913**	0.0299***	0.1017***	0.0550***	0.0582*
	(0.0001)	(0.0097)	(0.0236)	(0.0435)	(0.0031)	(0.0093)	(0.0161)	(0.0329)
ln(GDP in PPP)_ct-1	0.0001***	0.0967***	0.3843**	-0.0649	-0.0059	0.1022***	-0.0647	-0.0540
	(0.1163)	(0.0192)	(0.1840)	(0.2705)	(0.0498)	(0.0079)	(0.0424)	(0.0649)
ln(Starting a Business)_ct-1	0.0001***	0.2884***	0.5496***	1.1426***	0.0769***	0.3276***	0.1879***	0.2534***
	(0.0973)	(0.0735)	(0.1531)	(0.2370)	(0.0350)	(0.0264)	(0.0358)	(0.0558)
Constant		-3.3024***	-11.4634**	-1.9300		-3.5712***	1.9550*	1.4195
		(0.6218)	(5.0775)	(7.4904)		(0.2285)	(1.0498)	(1.6144)
Observations	205,080	205,428	299,136	299,136	401,568	402,096	591,744	591,744
Number of varieties	17,090	17,119	24,928	24,928	33,464	33,508	49,312	49,312
Prob Wald Chi2	0.000	0.000			0.000	0.000		
R-squared			0.015	0.019			0.007	0.005
Country FE	YES	YES	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table reports the results for country sub-samples based on income levels. Columns (1)-(4) report the results for high-income OECD countries, while columns (5)-(8) show the results for low- and middle-income countries. The dependent variable in columns (1) and (5) corresponds to the probability of exporting new varieties by industry i to country c in time t; these columns report the average marginal effects using a logit model with fixed effects. The dependent variable in columns (2) and (6) is the number of new exported varieties by industry i to country c in time t; these columns report the coefficients using a negative binomial with fixed effects approach. The dependent variable in columns (3) and (7) is the export share of the number of new varieties by industry i to country c in time t (i.e., the extensive margin). On the other hand, the dependent variable in columns (4) and (8) is the export share of the value (in U.S. dollars) of new varieties by industry i to country c in time t (i.e., the intensive margin). Columns (3)-(4) and (7)-(8) are estimated by linear regressions with fixed effects. All independent variables are expressed in natural logs and lagged by one year. All regressions include country, industry, and year fixed effects.

Table 3.15 reports the results for these two country sub-samples based on income groups. The first four columns display the results for high-income OECD countries, whereas the last four columns exhibit the results for low- and middle-income countries.[30] We employ three different methodologies that we previously used.

Columns (1) and (5) report the average marginal effects using a logit model with fixed effects to analyze the impact of new imported varieties on the probability of exporting new varieties. We can observe that new imported varieties are positive and strongly statistically significant for both high-income OECD and low- and middle-income countries; although, the effect is larger for low- and middle-income countries.

Our results suggest that imports of new varieties from low- and middle-income countries experience a larger impact on Mexico's probability of exporting new varieties. A plausible explanation is that one of the criteria to identify a new variety is that a product is imported or exported for the first time with a partner country. Thus, developing countries are experiencing an increased participation in trade of new varieties. This phenomenon is aligned to the increasing integration of developing countries in global value chains, which is related to the processing

[30]According to the OECD country classification, high-income OECD countries include those members with a GNI per capita income above 12,236 U.S. dollars in 2016. These countries include: Australia, Austria, Belgium, Canada, Chile, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, South Korea, Latvia, Lithuania, Luxembourg, the Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, United Kingdom, and the United States. We classified the rest of the countries under the low- and middle-income sub-sample.

trade mechanism.

The next methodology consists of a negative binomial model with fixed effects to examine the impact of the number of new imported varieties on the number of new exported varieties by income group. Columns (2) and (6) display the coefficients for our two country sub-samples. These results also suggest that new imported varieties are positive and strongly statistically significant for both high-income OECD and low- and middle-income countries. Nonetheless, we can notice that the magnitude of the coefficients does not change much based on income groups.

Now, we examine the impact of importing new varieties on the export share of new varieties at both the extensive and intensive margins using linear regressions with fixed effects. Columns (3) and (7) show the results of the relationship at the extensive margin. We can also observe that new imported varieties are positive and strongly significant for both sub-samples of income countries at the extensive margin. Once again, the magnitude of the coefficients does not vary much based on income groups. These results are not drastically different from those in our baseline specification presented in Table 3.9.

Finally, columns (4) and (8) exhibit the results of the relationship at the intensive margin using linear regressions with fixed effects. In this case, we can now notice that importing new varieties has a positive and statistically significant effect at the intensive margin. Nonetheless, these coefficients exhibit a lower significance level compared to the baseline results presented in Table 3.10.

These findings complement the study by Lo Turco & Maggioni (2013) for Italy, where the authors argue that importing from low-income countries has a larger effect on the probability to start exporting compared to importing from high-income countries. In our study, we analyze the asymmetric trade relations with high-income and low- and middle-income countries from the perspective of a developing country. The results suggest that for a middle-income country, as in the case of Mexico, the effects of importing new varieties from low- and middle-income countries on exports of new varieties (i.e., measured as a probability and number of varieties) are larger compared to their high-income counterparts. On the contrary, the results also suggest that the effects of importing new varieties from a high-income country on exports of new varieties measured as shares are larger than for low- and middle-income countries. This suggests that the trade complementarity effect is stronger for high-income OECD countries.

3.7.6 Top Countries Trading New Varieties

We also extract a sub-sample of the top 17 trade partners of new varieties.[31] This sub-sample consists of 14,508 new varieties over the period 2005-2016. All the independent variables are expressed in their log form and lagged by one year. Moreover, all the specifications include a full set of country, industry, and year fixed effects. Again, the number of observations and varieties change in each column because the logit and negative binomial methodologies with fixed effects drop time-invariant observations from the sample.

Table 3.16: Main Trade Partners of New Varieties

VARIABLES	(1) AME Prob_X_NEW	(2) X_NEW	(3) ExtM_X_NEW	(4) IntM_X_NEW
ln(New Imported Varieties)_cit-1	0.0001***	0.0571***	0.0586	0.1554**
	(0.0001)	(0.0115)	(0.0394)	(0.0712)
ln(GDP in PPP)_ct-1	0.0001***	-0.0375***	0.8980***	0.6622**
	(0.0803)	(0.0111)	(0.1930)	(0.2949)
ln(Starting a Business)_ct-1	-0.0001	-0.0695	0.0037	0.2513
	(0.0671)	(0.0448)	(0.1686)	(0.2363)
Constant		2.0159***	-21.9613***	-16.7081**
		(0.3321)	(4.9852)	(7.7354)
Observations	138,756	139,212	174,108	174,108
Number of varieties	11,563	11,601	14,509	14,509
Prob Wald Chi2	0.000	0.000		
R-squared			0.147	0.072
Country FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table comprises exclusively Mexico's top 17 trade partners of new varieties. The dependent variable in column (1) corresponds to the probability of exporting new varieties by industry i to country c in time t; this column reports the average marginal effects using a logit model with fixed effects. The dependent variable in column (2) is the number of new exported varieties by industry i to country c in time t; this column reports the coefficients using a negative binomial with fixed effects approach. The dependent variable in column (3) is the export share of the number of new varieties by industry i to country c in time t (i.e., the extensive margin). On the other hand, the dependent variable in column (4) is the export share of the value (in U.S. dollars) of new varieties by industry i to country c in time t (i.e., the intensive margin). Both columns (3) and (4) are estimated by linear regressions with fixed effects. All independent variables are expressed in natural logs and lagged by one year. All regressions include country, industry, and year fixed effects.

Table 3.16 shows that the main independent variable (i.e., new imported varieties) is positive and statistically significant on new exported varieties measured as a probability, in levels, and as a share at the intensive margin. On the other hand, this independent variable remains positive, albeit statistically insignificant, on the export share at the extensive margin. If we compare the magnitude of the main explanatory variable to the baseline specifications, these are now smaller in size, except for the intensive margin.

[31] The sub-sample is determined by identifying on one hand, the top 10 source countries of new imports, and on the other hand, the top 10 destination countries for new exports. The reason why we end up with 17 countries is because China, Spain, and Portugal are both top source and destination countries of new varieties. The countries that conform this sub-sample are: Canada, China, Colombia, Costa Rica, El Salvador, Guatemala, Honduras, India, Indonesia, Italy, Nicaragua, Portugal, Spain, Thailand, Turkey, United States, and Vietnam.

3.7.7 Main Industries Trading New Varieties

We now extract another sub-sample of the top industries trading new varieties. This sub-sample includes a total of 609 industries belonging to the chemicals and allied industries sector (Chapters 28 to 38 of the Harmonized System), textiles sector (Chapters 50 to 63 of the Harmonized System), metals sector (Chapters 72 to 83 of the Harmonized System), and to the machinery and electrical sector (Chapters 84 and 85 of the Harmonized System) over the examined period. All independent variables are expressed in logs and lagged by one year. We also include a full set of country, industry, and year fixed effects. It is worth mentioning that the number of observations and varieties are different because the logit and negative binomial approaches combined with fixed effects drop time-invariant observations from the sample.

Table 3.17: Main Industries Trading New Varieties

VARIABLES	(1) AME Prob_X_NEW	(2) X_NEW	(3) ExtM_X_NEW	(4) IntM_X_NEW
ln(New Imported Products)_cit-1	0.0040***	0.0955***	0.0629***	0.0772**
	(0.0003)	(0.0079)	(0.0158)	(0.0322)
ln(GDP in PPP)_ct-1	0.0021	0.0974***	-0.0346	-0.0116
	(0.0505)	(0.0076)	(0.0500)	(0.0788)
ln(Starting a Business)_ct-1	0.0116***	0.3371***	0.2072***	0.3397***
	(0.0404)	(0.0290)	(0.0419)	(0.0680)
Constant		-3.4987***	1.1608	0.0317
		(0.2234)	(1.2698)	(2.0120)
Observations	379,968	380,616	551,976	551,976
Number of varieties	31,664	31,718	45,998	45,998
Prob Wald Chi2	0.000	0.000		
R-squared			0.006	0.005
Country FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table is based on a sub-sample of the main trading industries of new varieties; these industries belong to the chemicals and allied industries, textiles, metals, and to the machinery and electrical sectors. This sample is composed by 45,998 new varieties over the period 2005-2016. The dependent variable in column (1) corresponds to the probability of exporting new varieties by industry i to country c in time t; this column reports the average marginal effects using a logit model with fixed effects. The dependent variable in column (2) is the number of new exported varieties by industry i to country c in time t; this column reports the coefficients using a negative binomial with fixed effects approach. The dependent variable in column (3) is the export share of the number of new varieties by industry i to country c in time t (i.e., the extensive margin). On the other hand, the dependent variable in column (4) is the export share of the value (in U.S. dollars) of new varieties by industry i to country c in time t (i.e., the intensive margin). Both columns (3) and (4) are estimated by linear regressions with fixed effects. All independent variables are expressed in natural logs and lagged by one year. All regressions include country, industry, and year fixed effects.

Table 3.17 reports the results for this industry sub-sample. We can observe that after reducing the number of observations by almost half, our main independent variable remains positive and statistically significant through all the different specifications (i.e., on the probability, number of varieties, and shares). Furthermore, we can notice that the magnitude of the coefficients is slightly larger in this table compared to the baseline results (i.e., Tables 3.7, 3.8, 3.9, and 3.10), except for the number of varieties, where there is a small variation. We can conclude that the effects are slightly larger for those industries encompassing a large number of new varieties.

3.7.8 Alternative Control Variables

In this subsection, we incorporate other trade-related control variables in a similar manner as in Navas et al. (2020). These trade-related controls are: number of documents to import, costs to import in U.S. dollars per container deflated, and time to import in days. It is worth mentioning that we did not include these trade-related control variables in our baseline specifications because of their short time span (i.e., the data is available for the period 2006-2015).

All regressions include logged and lagged independent variables and a full set of country, industry, and year fixed effects. The sample sizes may vary from one column to another due to drawbacks of the negative binomial and logit models; these two models combined with fixed effects result in dropping observations for time-invariant variables; furthermore, the logit model drops observations for non-switchers that, therefore, do not contribute to the likelihood function. In addition, these sample sizes are smaller compared to the baseline results as the available time span for these trade-related control variables is shorter.

Table 3.18: Alternative Control Variables

VARIABLES	(1) Prob_NEW	(2) X_NEW	(3) ExtM_NEW	(4) IntM_NEW	(5) Prob_NEW	(6) X_NEW	(7) ExtM_NEW	(8) IntM_NEW	(9) Prob_NEW	(10) X_NEW	(11) ExtM_NEW	(12) IntM_NEW
ln(IM_NEW)_cit-1	0.0006***	0.0508***	0.0473***	0.0724**	0.0178***	0.0460***	0.0451***	0.0698**	0.0002***	0.0519***	0.0470***	0.0715**
	(0.0001)	(0.0075)	(0.0156)	(0.0303)	(0.0025)	(0.0075)	(0.0156)	(0.0303)	(0.0001)	(0.0076)	(0.0156)	(0.0303)
ln(GDP in PPP)_ct-1	0.0015***	0.0778***	-0.0436	0.0513	0.0259**	0.0756***	-0.0653	0.0025	0.0005***	0.0896***	-0.0246	0.0524
	(0.0502)	(0.0071)	(0.0493)	(0.0719)	(0.0505)	(0.0071)	(0.0497)	(0.0726)	(0.0510)	(0.0072)	(0.0491)	(0.0714)
ln(Import Docs)_ct-1	-0.0007***	-0.1999***	-0.2205***	-0.1217								
	(0.0288)	(0.0197)	(0.0580)	(0.0864)								
ln(Import Costs)_ct-1					-0.0668***	-0.2350***	-0.1410***	-0.2664***				
					(0.0254)	(0.0163)	(0.0419)	(0.0628)				
ln(Import Time)_ct-1									0.0002***	0.0501***	0.0882*	-0.0141
									(0.0246)	(0.0154)	(0.0486)	(0.0736)
Constant		-1.1852***	2.7220**	0.1143		0.2563	3.9155***	3.1264		-2.0003***	1.5528	-0.1015
		(0.1918)	(1.2888)	(1.8768)		(0.2314)	(1.3631)	(2.0051)		(0.1987)	(1.2832)	(1.8668)
Observations	468,706	469,732	730,886	730,886	468,706	469,732	730,886	730,886	468,706	469,732	730,886	730,886
Prob Wald Chi2	0.000	0.000			0.000	0.000			0.000	0.000		
R-squared			0.007	0.003			0.006	0.005			0.005	0.003
Number of varieties	47,366	47,471	74,240	74,240	47,366	47,471	74,240	74,240	47,366	47,471	74,240	74,240
Country FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: This table reports the results after using different trade-related control variables instead of the score of starting a business. Columns (1)-(4) report the results using the log number of documents to import as a control variable. Columns (5)-(8) exhibit the results using the log costs to import in U.S. dollars per container deflated as a control variable. Columns (9)-(12) display the results using the log time to import in days as a control variable. The dependent variable in columns (1), (5), and (9) is the probability of exporting new varieties by industry i to country c in time t; these columns report the average marginal effects using a logit model with fixed effects. The dependent variable in columns (2), (6), and (10) corresponds to the number of new exported varieties by industry i to country c in time t; these columns report the coefficients using a negative binomial with fixed effects approach. The dependent variable in columns (3), (7), and (11) is the export share of the number of new varieties by industry i to country c in time t (i.e., the extensive margin); these columns are estimated by linear fixed effects regressions. Finally, the dependent variable in columns (4), (8), and (12) is the export share of the value (in U.S. dollars) of new varieties by industry i to country c in time t (i.e., the intensive margin); these columns are estimated by linear fixed effects regressions. All independent variables are expressed in natural logs and lagged by one year. All regressions include country, industry, and year fixed effects.

Table 3.18 reports the regressions using these different trade-related control variables instead of the score of starting a business. In columns (1)-(4), we use the log number of documents to import as our control variable, along with logged GDP.[32] We can observe that new imported varieties remain positive and statistically significant through all of our specifications. Moving on to the number of documents required to import, this variable has a negative and strong impact

[32]The World Bank defines the number of documents to import as the total amount of documents per import shipment required by law or by relevant agencies, including government ministries, customs authorities, port authorities, and other related agencies.

on exports of new varieties. A plausible interpretation is that requiring more paperwork at Customs may hinder the number of new imported varieties, which could potentially be used to produce new varieties for the export market. An exception of this trend is the impact of documents required to import on export shares at the intensive margin, which is insignificant.

In columns (5)-(8), we proceed to include the log costs to import and the log of GDP as our control variables.[33] Our results are consistent with the previous set of four columns, where new imported varieties remain positive and statistically significant for the four specifications. The coefficients of our import cost variable are negative and strongly significant through all the four columns. A plausible explanation is that having high import costs may also disincentivize technology transfers of new varieties. Thus, by reducing these imports costs, we can promote more imports of new varieties.

In the last set of columns (9)-(12), we employ the log of days to import and the log of GDP as control variables.[34] Similarly, the coefficients of new imported varieties are positive and statistically significant for all the specifications. Contrary to what we could expect, days to import exhibit a positive and statistically significant impact on exports of new varieties measured as a probability, number of varieties, and as shares at the extensive margin. A plausible explanation could be that source countries of new imported varieties may be located remotely, such as the case of Asian countries; thus, this translates in more days required to import varieties from these remote locations.

Our results using different sets of control variables are consistent to our baseline results in Section 3.6. Nonetheless, the coefficients of new imported varieties are slightly smaller compared to our baseline results presented in Tables 3.8, 3.9, and 3.10.

[33]The World Bank defines import costs as those related to importing a 20-foot container of goods by sea transport through the following stages: document preparation, customs clearance and inspections, inland transport and handling, and port and terminal handling. These costs are measured in U.S. dollars per container deflated.

[34]The World Bank Doing Business dataset states that the time to import refers to the calendar days associated to import a 20-foot container of goods by sea transport through the following stages: document preparation, customs clearance and inspections, inland transport and handling, and port and terminal handling.

3.7.9 Fixed Effects Poisson Model

We now proceed to use an alternative methodology to the fixed effects negative binomial model: the fixed effects Poisson model. In Table 3.19, we examine the relationship between new imported varieties and new exported varieties using a fixed effects Poisson model. All the regressions include a full set of country, industry, and year fixed effects; control variables are included in a stepwise manner.

Table 3.19: Fixed Effects Poisson Model

VARIABLES	(1) X_NEW	(2) X_NEW	(3) X_NEW	(4) X_NEW
ln(New Imported Varieties)_cit-1	0.0562***	0.0553***	0.0549***	0.0547***
	(0.0073)	(0.0073)	(0.0073)	(0.0073)
ln(GDP in PPP)_ct-1		0.1656***		0.0496
		(0.0396)		(0.0406)
ln(Starting a Business)_ct-1			0.3615***	0.3508***
			(0.0311)	(0.0320)
Observations	607,524	607,524	607,524	607,524
Number of varieties	50,627	50,627	50,627	50,627
Prob Wald Chi2	0.000	0.000	0.000	0.000
Country FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

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

Notes: This table reports the alternative results of estimating Eq.(3.8) using a Poisson model with fixed effects. The dependent variable is the number of new exported varieties by industry i to country c in time t. The estimation sample is conformed by 74,240 new varieties over the period 2005-2016; however, some observations were dropped from the sample due to the nature of the methodology employed. All independent variables are expressed in natural logs and lagged by one year. All regressions include country, industry, and year fixed effects.

Although this methodology is not the best approach for a count dependent variable with overdispersion around the mean, it constitutes a robustness check for the consistency of our results. We can notice that the coefficients of the main explanatory variable remain positive and strongly statistically significant; however, the magnitude of the coefficients is smaller compared to the results from the negative binomial model with fixed effects presented in Table 3.8.

A drawback of the Poisson model is the restrictiveness of the assumption that the mean of the distribution should be equal to the variance.

3.7.10 Zero-Value Observations

This subsection deals with zero-value observations in the dependent variable using an alternative approach to adding one unit and then, taking the natural logarithm. This issue is common in trade datasets as some countries may not trade with all the possible partners. These zero-value observations may be problematic as these may lead to distorted estimates for semi-log models. Thus, we now employ a Poisson Pseudo-Maximum-Likelihood (PPML) estimator proposed by Santos Silva & Tenreyro (2006), to deal with zero-value observations. Information regarding the PPML estimator is provided in the Appendix.

Table 3.20: Zero-Value Observations

VARIABLES	(1) X_NEW	(2) ExtM_X_NEW	(3) IntM_X_NEW
ln(New Imported Varieties)_cit-1	0.1126**	0.1037***	0.1582***
	(0.0500)	(0.0322)	(0.0292)
ln(GDP in PPP)_ct-1	0.0462	-0.0725	-0.0362
	(0.2098)	(0.2202)	(0.2318)
ln(Starting a Business)_ct-1	0.3483*	0.2796	0.4550**
	(0.1855)	(0.2231)	(0.2304)
Constant	-3.6148	1.1420	-0.5397
	(5.2281)	(5.4600)	(5.9700)
Observations	890,364	890,364	890,364
Prob Wald Chi2	0.023	0.011	0.000
Country FE	YES	YES	YES
Industry FE	YES	YES	YES
Year FE	YES	YES	YES

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

Notes: This table reports the coefficients of Eq.(3.8)-(3.10) employing a Poisson Pseudo Maximum Likelihood (PPML) model with fixed effects; the aim is to deal with an excess of zero-value observations in the dependent variable. Column (1) estimates Eq.(3.8), where the dependent variable is a count variable representing the number of new exported varieties. Column (2) estimates Eq.(3.9), where the dependent variable is the export share of the number of new varieties (i.e., the extensive margin). Column (3) estimates Eq.(3.10), where the dependent variable is the export share of the value (in U.S. dollars) of new varieties (i.e., the intensive margin). All independent variables are expressed in natural logs and lagged by one year. All regressions include country, industry, and year fixed effects.

In Table 3.20, we report the coefficients of Eq.(3.8)-(3.10) using a PPML approach with fixed effects to deal with zero-value observations in the dependent variable. In column (1), we estimate Eq.(3.8), where the dependent variable is a count variable representing the number of new exported varieties. We can observe that the coefficient of the main explanatory variable (i.e., new imported varieties) remains positive and statistically significant; also, the magnitude of this coefficient is slightly larger compared to the baseline specification in Table 3.8.

Column (2) estimates Eq.(3.9), where the dependent variable is the share of the number of new exported varieties (i.e., the extensive margin). We also notice that our main explanatory variable remains positive and statistically significant. Just as in the previous column, the magnitude of this coefficient is larger compared to the baseline specification presented in Table 3.9.

Finally, column (3) estimates Eq.(3.10), where the dependent variable is the share of the value (in U.S. dollars) of new exported varieties (i.e., the intensive margin). We observe that our main explanatory variable remains positive and statistically significant. The magnitude of our coefficient of interest is now twice in size compared to the baseline specification exhibited in Table 3.10.

3.7.11 Log-Log Model

We also try a log-log model to study the impact of the log number of new imported varieties on the log number of new exported varieties. This time, instead of running a fixed effects negative binomial model, we run a linear fixed effects model. Table 3.21 presents the results of this log-log model. In column (1), we only include the main explanatory variable along with the full set of country, industry, and year fixed effects. We can notice that the main explanatory variable is positive and strongly statistically significant.

Table 3.21: Log-Log Model

VARIABLES	(1) l_X_NEW	(2) l_X_NEW	(3) l_X_NEW	(4) l_X_NEW
ln(New Imported Varieties)_cit-1	0.0287***	0.0287***	0.0285***	0.0285***
	(0.0014)	(0.0014)	(0.0014)	(0.0014)
ln(GDP in PPP)_ct-1		0.0154***		0.0062**
		(0.0024)		(0.0024)
ln(Starting a Business)_ct-1			0.0387***	0.0375***
			(0.0022)	(0.0022)
Constant	0.1230***	-0.2733***	-0.0362***	-0.1919***
	(0.0010)	(0.0625)	(0.0091)	(0.0618)
Observations	890,880	890,880	890,880	890,880
R-squared	0.012	0.012	0.012	0.012
Number of varieties	74,240	74,240	74,240	74,240
Country FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

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

Notes: This table estimates a log-log model using a linear fixed effects approach. In this specification, we also take the log of the dependent variable, which is the number of new exported varieties by industry i to country c in time t. The estimation sample is conformed by 74,240 new varieties over the period 2005-2016. All independent variables are expressed in natural logs and lagged by one year. All panel regressions include country, industry, and year fixed effects.

The following columns incorporate the control variables one by one. Thus, column (2) includes GDP as a control variable. Here, we can notice that the coefficient of the main explanatory variable remains the same. In column (3), we now include the score of starting a business. We can notice that the main explanatory variable remains positive and strongly significant, albeit the coefficient is slightly smaller.

In column (4), we include all the control variables. We can observe that new imported varieties remain positive and strongly statistically significant. Our results suggest that a 1% increase in the number of new imported varieties is associated with an increase of the number of new exported varieties by about 0.03%. Compared to the baseline results presented in Table 3.8, we can notice that the magnitude of the effect is larger under this log-log specification.

3.7.12 Contemporaneous Effects

We are now interested in examining the contemporaneous effects of new imported varieties on exports of new varieties. To perform this analysis, we define all the right-hand side variables in time t instead of in time $t-1$. Therefore, our main explanatory variable is now the log number of new imported varieties by industry i from country c in year t.

Table 3.22 shows the results of this contemporaneous effects of new imported varieties. In column (1), we report the average marginal effects using a logit model with fixed effects. We can notice that new imported varieties are positive and strongly statistically significant. Nonetheless, the size of the effect is smaller compared to our baseline specification, with one lag, presented in Table 3.7.

Table 3.22: Contemporaneous Effects

VARIABLES	(1) Prob_X_NEW: AME	(2) X_NEW	(3) ExtM_X_NEW	(4) IntM_X_NEW
ln(New Imported Varieties)_cit	0.0003***	0.2390***	0.0862***	0.2398***
	(0.0001)	(0.0069)	(0.0141)	(0.0332)
ln(GDP in PPP)_ct	0.0002***	0.1085***	0.0239	0.0598
	(0.0414)	(0.0064)	(0.0390)	(0.0598)
ln(Starting a Business)_ct	0.0005***	0.3345***	0.1958***	0.3304***
	(0.0356)	(0.0257)	(0.0363)	(0.0568)
Constant		-3.8032***	-0.2870	-1.7860
		(0.1910)	(0.9968)	(1.5351)
Observations	606,648	607,524	890,880	890,880
Prob Wald Chi2	0.000	0.000		
R-squared			0.007	0.012
Number of varieties	50,554	50,627	74,240	74,240
Country FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table explores the contemporaneous effects of importing new varieties. The dependent variable in column (1) corresponds to the probability of exporting new varieties by industry i to country c in time t; this column reports the average marginal effects using a logit model with fixed effects. The dependent variable in column (2) is the number of new exported varieties by industry i to country c in time t; this column reports the coefficients using a negative binomial with fixed effects approach. The dependent variable in column (3) is the export share of the number of new varieties by industry i to country c in time t (i.e., the extensive margin). On the other hand, the dependent variable in column (4) is the export share of the value (in U.S. dollars) of new varieties by industry i to country c in time t (i.e., the intensive margin). Both columns (3) and (4) are estimated by linear regressions with fixed effects. The main explanatory variable is the log number of new imported varieties by industry i from country c in year t. All independent variables are expressed in natural logs. All regressions include country, industry, and year fixed effects.

Column (2) exhibits the coefficients using a fixed effects negative binomial model. We can also observe that new imported varieties are positive and strongly significant. Compared to the baseline specification in Table 3.8, the size of the effect is significantly larger in the same year than with a lag.

In column (3), we present the results at the extensive margin using a linear fixed effects model. We can also see that the impact of new imported varieties is positive and strongly statistically significant in the same year. The magnitude of this contemporaneous effect is slightly larger than the effect with a lag presented in Table 3.9.

Finally, column (4) displays the results at the intensive margin also using a linear fixed effects model. Just as before, new imported varieties have a positive and strong effect on the value of the export share. Interestingly, the contemporaneous effect is also significantly larger than in the baseline specification with a lag exhibited in Table 3.10.

In conclusion, we can notice the contemporaneous effects of new imported varieties on new exported varieties are significantly larger for the number of varieties and for the export share at the intensive margin. A plausible explanation may be the role of new intermediate inputs in processing trade, which we will study in Chapter 4.

3.7.13 Lag Length Increase

To avoid potential reverse causality, we increase the lag length of the independent variables by two and three lags. In columns (1)-(4) of Table 3.23, we report the results having all the independent variables with two lags. We can observe that our results hold on the probability of exporting new varieties, on the number of new exported varieties, and on the export share at the extensive margin, but not on the export share at the intensive margin.

These results are similar to those reported in columns (5)-(8), where we increase the lag length to three periods on the independent variables. Nonetheless, we can now observe that the key covariate became weakly statistically significant in column (7), while remained insignificant in column (8).

Finally, it is worth mentioning that the number of observations and varieties are different in each column because the logit and negative binomial models combined with fixed effects drop time-invariant observations. Moreover, introducing more lags translates into dropping observations from the sample.

Table 3.23: Lag Length Increase

VARIABLES	(1) AME Prob_X_NEW	(2) X_NEW	(3) ExtM_X_NEW	(4) IntM_X_NEW	(5) AME Prob_X_NEW	(6) X_NEW	(7) ExtM_X_NEW	(8) IntM_X_NEW
ln(New Imported Varieties)_cit-2	0.0086***	0.0286***	0.0440***	0.0481				
	(0.0020)	(0.0076)	(0.0157)	(0.0299)				
ln(GDP in PPP)_ct-2	-0.0158*	0.0826***	-0.1736***	-0.0449				
	(0.0489)	(0.0069)	(0.0479)	(0.0705)				
ln(Starting a Business)_ct-2	0.0502***	0.2254***	0.2228***	0.3452***				
	(0.0378)	(0.0273)	(0.0426)	(0.0667)				
ln(New Imported Varieties)_cit-3					0.0085***	0.0464***	0.0317*	0.0165
					(0.0029)	(0.0081)	(0.0167)	(0.0318)
ln(GDP in PPP)_ct-3					0.0018	0.1234***	-0.1592***	-0.0749
					(0.0548)	(0.0079)	(0.0507)	(0.0749)
ln(Starting a Business)_ct-3					0.0305***	0.1211***	0.1639***	0.2067***
					(0.0387)	(0.0278)	(0.0457)	(0.0741)
Constant		-2.6293***	4.7058***	0.9137		-2.7437***	4.6266***	2.2852
		(0.2059)	(1.2283)	(1.8154)		(0.2260)	(1.3102)	(1.9368)
Observations	478,010	479,020	742,400	742,400	410,337	411,435	668,160	668,160
Prob Wald Chi2	0.000	0.000			0.000	0.000		
Number of varieties	47,801	47,902	74,240	74,240	45,593	45,715	74,240	74,240
R-squared			0.006	0.004			0.005	0.002
Country FE	YES	YES	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table reports the results of the estimation sample after increasing the lag length of the independent variables to two and three lags. Columns (1)-(4) exhibit the results after increasing the lag length of the independent variables by two, while columns (5)-(8) increase the lag length by three periods. The dependent variable in columns (1) and (5) corresponds to the probability of exporting new varieties by industry i to country c in time t; this column reports the average marginal effects using a logit model with fixed effects. The dependent variable in columns (2) and (6) is the number of new exported varieties by industry i to country c in time t; this column reports the coefficients using a negative binomial with fixed effects approach. The dependent variable in columns (3) and (7) is the export share of the number of new varieties by industry i to country c in time t (i.e., the extensive margin). On the other hand, the dependent variable in columns (4) and (8) is the export share of the value (in U.S. dollars) of new varieties by industry i to country c in time t (i.e., the intensive margin). Columns (3)-(4) and (7)-(8) are estimated by linear regressions with fixed effects. All independent variables are expressed in natural logs and lagged by two and three years, respectively. All regressions include country, industry, and year fixed effects.

3.7.14 Alternative Fixed Effects Combinations

Industry and Country Time Trends

To tackle a potential omitted variable bias, we include a combination of industry and country time trend fixed effects. Table 3.24 exhibits the results of including industry-year and country-year fixed effects. It is worth mentioning that our control variables dropped from the model because country-year fixed effects capture the variation of GDP and score of starting a business.

In column (1), we use a linear probability model with fixed effects to examine the impact of importing new varieties on the probability of exporting new varieties.[35] We can observe that new imported varieties positively and strongly impact the probability of exporting new varieties.

Table 3.24: Industry and Country Time Trends

VARIABLES	(1) Prob_X_NEW	(2) X_NEW	(3) ExtM_X_NEW	(4) IntM_X_NEW
ln(New Imported Varieties)_cit-1	0.0483***	0.0767***	0.0543***	0.1239***
	(0.0015)	(0.0076)	(0.0130)	(0.0253)
Constant	0.1381***	-0.7371***	1.1491***	1.1515***
	(0.0005)	(0.0041)	(0.0055)	(0.0089)
Observations	890,868	834,922	890,868	890,868
R-squared	0.185		0.044	0.023
Prob Wald Chi2		0.000		
Industry-Year FE	YES	YES	YES	YES
Country-Year FE	YES	YES	YES	YES

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

Notes: This table includes industry and country time trends as fixed effects. The dependent variable in column (1) corresponds to the probability of exporting new varieties by industry i to country c in time t; this column is estimated by a linear probability model with fixed effects. The dependent variable in column (2) is the number of new exported varieties by industry i to country c in time t; this column reports the coefficients using a PPML model with fixed effects. The dependent variable in column (3) is the export share of the number of new varieties by industry i to country c in time t (i.e., the extensive margin). On the other hand, the dependent variable in column (4) is the export share of the value (in U.S. dollars) of new varieties by industry i to country c in time t (i.e., the intensive margin). Both columns (3) and (4) are estimated by linear regressions with fixed effects. The independent variable is expressed in natural logs and lagged by one year. All regressions include industry-year and country-year fixed effects.

In column (2), we now use a Poisson Pseudo-Maximum Likelihood model with fixed effects. We can also notice that the new imported varieties have a positive and strong effect on the number of exports of new varieties. Compared to the baseline results presented in Table 3.8, the size of the coefficients is slightly smaller now. Column (3) exhibits the results at the extensive margin employing a linear fixed effects model. We can see that the main explanatory variable remains positive and strongly significant. Furthermore, the size of the coefficient is consistent with the baseline results in Table 3.9. Finally, column (4) shows the results at the intensive margin using a linear fixed effects model. We can also notice our main explanatory variable remains positive and strongly statistically significant. In terms of the magnitude of the coefficient, this is significantly larger than the baseline specification exhibited in Table 3.10.

[35] To absorb multiple levels of fixed effects, we use two available Stata commands for this subsection: *reghdfe*, which is useful in running linear regressions, and *ppmlhdfe*, which is helpful in running Poisson Pseudo-Likelihood regressions. Nevertheless, there is not an available command yet to run logit regressions absorbing multiple levels of fixed effects; thus, we use a linear probability model instead.

Sector Time Trends

Furthermore, we also include sector-year fixed effects to tackle a potential omitted variable bias. Table 3.25 shows the results of including sector-year and variety fixed effects.[36] In column (1), we display the results using a linear probability model with fixed effects. We can observe that new imported varieties have a positive and strong effect on the probability of exporting new varieties. Compared to previous Table 3.24 that uses the same methodology, the size of the coefficient is cut by half when we include sector-year and variety fixed effects in our specification.

In column (2), we use a PPML estimation with fixed effects. These results suggest that importing new varieties has a positive and strong effect on the number of new exported varieties. Compared to the baseline regression in Table 3.8, the magnitude of the coefficient of the main explanatory variable is now smaller. Moving on to the extensive and intensive margins, we employ linear fixed effects models. In column (3), we present the results at the extensive margin. Here, we can notice that the coefficient of new imported varieties is positive and strongly statistically significant. Furthermore, the magnitude of the coefficient is consistent with the baseline specification in Table 3.9. Column (4) displays the results at the intensive margin. We can also notice that the main explanatory variable is positive and strongly significant. Also, the magnitude of the coefficient is comparable in size to the baseline results presented in Table 3.10.

Table 3.25: Sector Time Trends

VARIABLES	(1) Prob_X_NEW	(2) X_NEW	(3) ExtM_X_NEW	(4) IntM_X_NEW
ln(New Imported Varieties)_cit-1	0.0237***	0.0547***	0.0634***	0.0786***
	(0.0014)	(0.0073)	(0.0137)	(0.0266)
ln(GDP in PPP)_ct-1	0.0105***	0.0584	-0.0358	0.0003
	(0.0028)	(0.0399)	(0.0426)	(0.0648)
ln(Starting a Business)_ct-1	0.0380***	0.3597***	0.2027***	0.3459***
	(0.0026)	(0.0319)	(0.0350)	(0.0547)
Constant	-0.2938***	-3.6955***	1.2091	-0.3283
	(0.0713)	(1.0239)	(1.0890)	(1.6628)
Observations	890,880	607,492	890,880	890,880
R-squared	0.254		0.098	0.094
Prob Wald Chi2		0.000		
Variety FE	YES	YES	YES	YES
Sector-Year FE	YES	YES	YES	YES

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

Notes: This table includes variety and sector time trends as fixed effects. The dependent variable in column (1) corresponds to the probability of exporting new varieties by industry i to country c in time t; this column is estimated by a linear probability model with fixed effects. The dependent variable in column (2) is the number of new exported varieties by industry i to country c in time t; this column reports the coefficients using a negative binomial with fixed effects approach. The dependent variable in column (3) is the export share of the number of new varieties by industry i to country c in time t (i.e., the extensive margin). On the other hand, the dependent variable in column (4) is the export share (in U.S. dollars) of new varieties by industry i to country c in time t (i.e., the intensive margin). Both columns (3) and (4) are estimated by linear regressions with fixed effects. All independent variables are expressed in natural logs and lagged by one year. All regressions include variety and sector-year fixed effects.

[36] Variety fixed effects correspond to country-industry fixed effects.

3.7.15 Input-Output Linkages

We are now interested in analyzing input-output linkages across sectors because manufacturing goods require not only inputs from the same sector but also from other sectors. Therefore, we are interested in examining how new imported varieties in upwards sectors impact new exported varieties in downwards sectors. To perform this analysis, we employ the World Input-Output Database (WIOD), which comprises sector-level data using the ISIC classification at the 2-digits level. To be able to match these ISIC sectors to Harmonized System (HS) codes, we use a correspondence table. It is worth mentioning that we only include manufacturing sectors.

In this exercise, we define the dependent variable as the log value (in U.S. dollars) of exports by sector k to country c in time t.[37] On the other hand, we define the main explanatory variable as the log value (in U.S. dollars) of imports of new varieties by sector k from country c in time t. This import value of new varieties is calculated similar to Javorcik (2004): we multiply the import value (in U.S. dollars) of new varieties times the share of inputs acquired by other sectors in total inputs sourced by sector k using input-output tables.

Table 3.26 examines these input-output linkages across sectors. In column (1), we start by only including the main explanatory variable (i.e., new imported varieties) with a full set of country, sector, and year fixed effects. From this column, we can observe the indirect effect of new imported varieties on exports of new varieties is positive and strongly statistically significant. The following columns incorporate the control variables one by one. In column (2), we can notice that after incorporating GDP as the control variable, our main explanatory variable remains positive and strongly statistically significant. In column (3), we now include the score of starting a business as a control variable. We can see that our main explanatory variable also remains positive and strongly significant.

Table 3.26: Input-Output Linkages: Indirect Effect

VARIABLES	(1) ln(X_NEW_USD)	(2) ln(X_NEW_USD)	(3) ln(X_NEW_USD)	(4) ln(X_NEW_USD)
ln(New Imported Varieties IND)_ckt	0.0268***	0.0260***	0.0260***	0.0256***
	(0.0021)	(0.0021)	(0.0021)	(0.0021)
ln(GDP in PPP)_ct		0.8157***		0.6823***
		(0.0427)		(0.0430)
ln(Starting a Business)_ct			0.5528***	0.4307***
			(0.0341)	(0.0344)
Constant	9.9312***	-10.9447***	7.6536***	-9.3056***
	(0.0143)	(1.0946)	(0.1419)	(1.0844)
Observations	487,278	487,278	487,209	487,209
R-squared	0.038	0.039	0.039	0.040
Number of varieties	64,170	64,170	64,124	64,124
Country FE	YES	YES	YES	YES
Sector FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

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

Notes: This table analyzes across sectors input-output linkages. This table is estimated using linear regressions with fixed effects. The dependent variable stands for the log value (in U.S. dollars) of exports by sector k to country c in time t. The main explanatory variable corresponds to the log value of imports of new varieties by sector k from country c in time t. All regressions include country, sector, and year fixed effects.

[37] We define sectors as sections of the Harmonized System classification.

In column (4), we now include the main explanatory variable along with all the control variables and the full set of country, sector, and year fixed effects. We can conclude that importing new varieties from different sectors also has a positive and strong significant effect on exports of new varieties. Our results suggest that a 1% increase in new imported varieties is associated with an increase of new exported varieties by about 0.03%. However, it is worth noting that the magnitude of this effect is smaller compared to imports of new varieties from the same sector.[38]

Finally, it is worth recognizing that this input-output exercise has some limitations. First, input-output matrices are reported at the sector level (i.e., ISIC 2-digits) instead of at the industry level, which represents a more granular level. Furthermore, the World Input-Output Database is only available for 43 countries, including 28 EU countries and 15 major countries. Thus, there is no available data for most developing countries representing the main traders of new varieties with Mexico. In terms of data accuracy, this input-output exercise loses accuracy when we merge sectors using a HS-ISIC correspondence table. Despite these limitations, this input-output analysis still represents an interesting exercise that allows a better understanding of the magnitude of the effects within and between sectors.

3.7.16 Marginal Effects by Sector

Furthermore, we examine the heterogeneity across sectors to identify those with stronger effects on exports of new varieties. To perform this analysis, we include an interaction term between the main explanatory variable (i.e., the log number of new imported varieties in the previous year) and sectors.[39] Regarding the methodology, we employ average marginal effects in a panel data fixed effects logit model.

Table 3.27 reports the average marginal effects by industry. In column (1), we only include the interaction term with the full set of country, sector, and year fixed effects. Column (2) includes only GDP as a control variable, and column (3) includes only the score of starting a business as the control variable. Column (4) incorporates the interaction term, all the control variables, and the full set of country, sector, and year fixed effects. Our results suggest that most of the impact can be explained by the transport equipment sector, followed by the machinery and electrical equipment sector. Therefore, these two sectors predominantly impact the probability of exporting new varieties. From a policy perspective, trade policies need to be focused on these two sectors as these have the potential to boost exports of new varieties.

In case business chambers perceive that negotiating free trade agreements with Asian countries could potentially damage sensitive sectors of the economy, preferential trade agreements focused on specific sectors could be the solution. Thus, Mexican firms in the transport equipment, machinery, and electrical equipment sectors could benefit from a higher integration in

[38] We also include the results of the direct effect of input-output linkages at the sector level in the Appendix section. Compared to our baseline specifications, we now aggregate data at the sector level instead of at the industry level.

[39] Just as before, we define sectors as sections of the Harmonized System. It is worth mentioning that we only include manufacturing sectors.

global value chains.

Table 3.27: Marginal Effects by Sector

VARIABLES	(1) AME Prob_X_NEW	(2) AME Prob_X_NEW	(3) AME Prob_X_NEW	(4) AME Prob_X_NEW
Sector = 5, Mineral Products	0.0179***	0.0178***	0.0178***	0.0178***
	(0.0053)	(0.0053)	(0.0053)	(0.0053)
Sector = 6, Chemicals and Allied Industries	0.0918***	0.0917***	0.0917***	0.0916***
	(0.0021)	(0.0021)	(0.0021)	(0.0021)
Sector = 7, Plastics and Rubbers	0.1370***	0.1369***	0.1368***	0.1368***
	(0.0038)	(0.0038)	(0.0038)	(0.0038)
Sector = 8, Raw Hides, Skins, Leather, and Furs	0.0987***	0.0987***	0.0987***	0.0987***
	(0.0079)	(0.0079)	(0.0079)	(0.0079)
Sector = 9, Wood and Articles of Wood	0.0102*	0.0103*	0.0102*	0.0102*
	(0.0057)	(0.0057)	(0.0057)	(0.0057)
Sector = 10, Pulp of Wood	0.0932***	0.0931***	0.0931***	0.0930***
	(0.0051)	(0.0051)	(0.0051)	(0.0051)
Sector = 11, Textiles	0.0619***	0.0619***	0.0620***	0.0620***
	(0.0023)	(0.0023)	(0.0023)	(0.0023)
Sector = 12, Footwear and Headgear	0.0992***	0.0992***	0.0993***	0.0993***
	(0.0078)	(0.0078)	(0.0078)	(0.0078)
Sector = 13, Stone and Glass	0.0727***	0.0726***	0.0726***	0.0726***
	(0.0047)	(0.0047)	(0.0047)	(0.0047)
Sector = 14, Precious Stones and Metals	0.0130	0.0129	0.0129	0.0129
	(0.0098)	(0.0098)	(0.0098)	(0.0098)
Sector = 15, Base Metals	0.1146***	0.1145***	0.1145***	0.1144***
	(0.0023)	(0.0023)	(0.0023)	(0.0023)
Sector = 16, Machinery and Electrical Equipment	0.1462***	0.1461***	0.1461***	0.1460***
	(0.0019)	(0.0019)	(0.0019)	(0.0019)
Sector = 17, Transport Equipment	0.1665***	0.1665***	0.1664***	0.1664***
	(0.0043)	(0.0043)	(0.0043)	(0.0043)
Sector = 18, Optical and Medical Instruments	0.1196***	0.1195***	0.1194***	0.1194***
	(0.0038)	(0.0038)	(0.0038)	(0.0038)
Sector = 19, Arms and Ammunition	-0.0179	-0.0181	-0.0179	-0.0180
	(0.0180)	(0.0180)	(0.0180)	(0.0180)
Sector = 20, Miscellaneous Manufactured Articles	0.1180***	0.1181***	0.1181***	0.1181***
	(0.0049)	(0.0049)	(0.0049)	(0.0049)
Sector = 21, Works of Art and Antiques	0.0224*	0.0224*	0.0223*	0.0223*
	(0.0132)	(0.0132)	(0.0132)	(0.0132)
Observations	890,880	890,880	890,880	890,880
Control Variables	NO	YES	YES	YES
Country FE	YES	YES	YES	YES
Sector FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table reports the average marginal effects by sector using a logit model. The main explanatory variable corresponds to an interaction term between the log number of new imported varieties in the previous year and sector. Control variables are incorporated stepwise; thus, column (1) starts by excluding all control variables; column (2) only includes GDP as the control variable; column (3) only includes score of starting a business as the control variable; column (4) incorporates all control variables. All independent variables included are expressed in natural logs and lagged by one year. All regressions include country, sector, and year fixed effects.

3.7.17 Two-Stage Regressions

Finally, we run two-stage regressions with the main objective of discarding potential endogeneity arising from omitted variables and reverse causality. According to Bas & Strauss-Kahn (2014) and Feng et al. (2016), MFN import tariffs constitute a good instrument as this variable has an impact on imported varieties, but does not have an impact on exported varieties. However, using MFN tariffs can be problematic because within an industry, export tariffs may be correlated to import tariffs. Thus, we offer an alternative instrument consisting of applied import tariffs. In other words, we use country-specific tariffs that account for preferential and MFN tariffs.

Table 3.28: Two-Stage Regressions

Second Stage IV Regressions	(1) Second Stage Prob_X_NEW AME	(2) Second Stage X_NEW	(3) Second Stage ExtM_X_NEW	(4) Second Stage IntM_X_NEW
ln(New Imported Varieties)_cit-1	0.1035***	72.2932***	36.9954*	60.1395*
	(0.4622)	(5.0588)	(19.9679)	(31.9811)
ln(GDP in PPP)_ct-1	0.9697***	6.1318***	3.3959*	5.5798*
	(0.2024)	(0.4823)	(1.9085)	(3.0566)
ln(Starting a Business)_ct-1	0.3210***	2.4194***	1.3621**	2.2252**
	(0.0673)	(0.1670)	(0.6564)	(1.0514)
Residuals	-9.9114***	-72.1594***		
	(2.0878)	(5.0591)		
Constant	-28.1963***			
	(5.8327)			
Observations	888,450	605,302	888,450	888,450
Country FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Prob Wald Chi2	0.000	0.000		
Underidentification stat.			4.219	4.219
Prob underident. stat.			0.040	0.040
Weak identification stat.			4.219	4.219
Endogeneity F-test			18.081	21.491
Prob endogeneity test			0.000	0.000
First Stage IV Regressions	First Stage l_IM_NEW_L1	First Stage l_IM_NEW_L1	First Stage l_IM_NEW_L1	First Stage l_IM_NEW_L1
ln(Applied Import Tariffs)_cit-1	-0.0009**	-0.0009**	-0.0009**	-0.0009**
	(0.0004)	(0.0004)	(0.0004)	(0.0004)
ln(GDP in PPP)_ct-1	-0.0970***	-0.0970***	-0.0970***	-0.0970***
	(0.0025)	(0.0025)	(0.0025)	(0.0025)
ln(Starting a Business)_ct-1	-0.0331***	-0.0331***	-0.0331***	-0.0331***
	(0.0025)	(0.0025)	(0.0025)	(0.0025)
Constant	2.8016***	2.8016***	2.8016***	2.8016***
	(0.0613)	(0.0613)	(0.0613)	(0.0613)

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

Notes: This table reports the two-stage regressions. The instrumental variable employed is the log of applied import tariffs in the previous year. The top panel shows the second stage IV regressions, while the bottom panel exhibits the first stage regressions. The dependent variable on the top panel corresponds to the different specifications of exports of new varieties. Column (1) is estimated using an IV Probit model, where the dependent variable is the probability of exporting new varieties; this column reports the average marginal effects. Column (2) uses a Poisson model, where the dependent variable stands for the number of new exported varieties; this column reports the coefficients. Columns (3)-(4) are estimated using a 2SLS model, where the dependent variable corresponds to the export share of new varieties at the extensive and intensive margin, respectively.

Table 3.28 reports the two-stage regressions using applied import tariffs as the instrumental variable. The top panel exhibits the second stage of the IV regressions, while the bottom panel displays the first stage regressions. We include a set of tests comprised of an underidentification statistic, a weak identification statistic, and an endogeneity F-test. Our estimation sample is slightly reduced because we do not possess information of applied tariffs for some industries.[40] All the independent variables are expressed in their log form and lagged by one year.

From the first stage regressions, we find a negative and statistically significant effect of applied import tariffs on new imported varieties. After the instrumentation, we can notice that overall, new imported varieties have a positive and statistically significant effect on exports of new varieties. In column (1), we report the average marginal effects using an IV Probit model; our results suggest that a 1% increase in new imported varieties is associated with an increase of 0.10 percentage points on the probability of exporting new varieties.

In column (2), we employ a Poisson model and report the coefficients; the results suggest that a 1% increase in new imported varieties is associated with an increase in the number of new exported varieties by about 0.72. Columns (3) and (4) are estimated using Two-Stage Least Squares (2SLS); column (3) suggests that a 1% increase in new imported varieties from a country leads to an increase of roughly 0.37 percentage points in the export share of new varieties at the extensive margin to that same country. Likewise, column (4) suggests that a 1% increase in new imported varieties from a country leads to an increase of 0.6 percentage points in the export share of new varieties at the intensive margin to that same country.

We now report a series of tests depending on each methodology. We start by an underidentification test, which tests whether the number of instrumental variables is less compared to the endogenous variables; we confirm that we do not face any underidentification issue as we have one instrument and one endogenous variable (i.e., we have a just-identified model). Then, we use a weak identification test to examine the explanatory power of the instruments for the endogenous variable; this F-statistic is not larger than the critical values; thus, our instrument is weak and do not have a good explanatory power for our endogenous variable.

Finally, we employ an endogeneity test and we conclude that the endogenous regressor is in fact endogenous. As a general conclusion, we need to take these results with caution as the two-stage approach is problematic.[41] In other words, our estimates may remain biased and we would need to look for better instruments.

[40]The WTO does not report MFN and applied tariffs for the following industries: "Inorganic or organic compounds of mercury, whether or not chemically defined, excluding amalgams" (HS 2852), "Phosphides, whether or not chemically defined, excluding ferrophosphorus; other inorganic compounds (including distilled or conductivity water and water of similar purity); liquid air (whether or not rare gases have been removed); compressed air; amalgams, other than amalgams of precious metals" (HS 2853), "Biodiesel and mixtures thereof, not containing or containing less than 70% by weight of petroleum oils or oils obtained from bituminous minerals" (HS 3826), "Machines and apparatus of a kind used solely or principally for the manufacture of semiconductor boules or wafers, semiconductor devices, electronic integrated circuits or flat panel displays; machines and apparatus specified in Note 9 (C) to this Chapter; parts and accessories" (HS 8486), "Machinery parts, not containing electrical connectors, insulators, coils, contacts or other electrical features, not specified or included elsewhere in this Chapter" (HS 8487), "Vacuum cleaners" (HS 8508), and "Sanitary towels (pads) and tampons, napkins and napkin liners for babies and similar articles, of any material" (HS 9619).

[41]We also present the results of the IV regressions using MFN import tariffs as the instrumental variable in the Appendix section.

3.8 Conclusions

This study contributes to the existing literature on the relationship between new imported varieties and exports by revealing a degree of trade complementarities between imports and exports of new varieties at the country level. In other words, by exploiting the bilateral trade component of our database, we examine to what extent importing new varieties from a source country increases exports of new varieties to that same country. We start our analysis by performing a decomposition exercise of the annual growth of traded varieties between new, withdrawn, and continuing varieties. This decomposition exercise aims to reveal the importance of new varieties for trade.

Next, we employ a three-fold empirical strategy on an estimation sample of 74,240 new traded varieties belonging to the manufacturing sector over the period 2005-2016. First, we use a fixed effects logit model to measure the influence of new imported varieties on the probability to export new varieties. Then, we make use of a fixed effects negative binomial model to measure the impact of new imported varieties on the number of new exported varieties. Finally, we employ linear fixed effects regressions to measure the impact of new imported varieties on the export shares of new varieties at the extensive margin (i.e., number of new exported varieties) and at the intensive margin (i.e., value of new exported varieties).

Our results suggest that imports of new varieties have a positive and statistically strong effect on the probability of exporting new varieties. Furthermore, imports of new varieties also have a positive and statistically strong impact on the number of new exported varieties. Finally, imports of new varieties from a source country have a positive and strong significant effect, albeit small, on the export share of new varieties to that same country at both the extensive and intensive margins. In other words, importing new varieties from a source country tends to increase the exports of new varieties to that same country. Our results hold after controlling for GDP and the score of starting a business, and after including the full set of country, industry, and year fixed effects.

We also perform a series of robustness checks that involves incorporating trade gravity variables, excluding source and destination countries, excluding Mexico's main trading partner, using alternative control variables, employing an alternative methodology to the negative binomial model, dealing with zero-value observations in the dependent variable, using a log-log model, examining contemporaneous effects, increasing the lag length of the independent variables, using different combinations of fixed effects, exploring input-output linkages, examining the marginal effects by sector, and using different sub-samples. These different sub-samples include restricting the estimation sample to Mexico's 50 major trade partners, splitting the sample into country sub-samples based on income groups, including only the top trading partners of new varieties, and incorporating only the top industries trading new varieties. Finally, we also make use of two-stage regressions to discard any potential endogeneity.

An interesting finding is that we detect some degree of trade complementarities between imports and exports of new varieties at the country level. In other words, the results suggest that there is a magnifying trade effect on the source country, instead of propagating homogeneously across countries. A plausible explanation for this phenomenon is the knowledge about the country mechanism; thus, Mexican firms tend to adapt their products to demand features (i.e., tastes and preferences) and to regulations of their export markets.

From a policy perspective, this complementarity trade effect is stronger for Spain, Colombia, and China. These three partner countries constitute both the main source and destination countries for new varieties traded with Mexico. In the case of Spain and Colombia, we could also explain this trade complementarity effect by the knowledge about the country mechanism (e.g., common language), rather than by transportation costs. In the case of China, however, the Asian country constitutes a large and dynamic market for Mexican exports.

On the other hand, we can observe that transportation costs and free trade agreements play an important role in the decision-making process of Mexican manufacturing firms that export new varieties. Nonetheless, this does not hold when it comes to importing new varieties. Despite the tariffs associated with importing products from countries that do not benefit from a free trade agreement, Mexican manufacturing firms consider it is still more profitable to import new varieties from these Asian countries.

Due to the above-mentioned reasons, expanding the free trade network to Asian countries could represent an interesting opportunity for Mexican firms that would benefit from lower tariffs or free trade for these new imported varieties. Nonetheless, these policy implications should be taken cautiously as some of these sectors, such as the textile and footwear industries, are quite sensitive and a free trade agreement could negatively impact the domestic production of these goods. Therefore, we highlight the importance of making a consultation process with the relevant business chambers in Mexico to evaluate the benefits of negotiating free trade agreements with Asian countries.

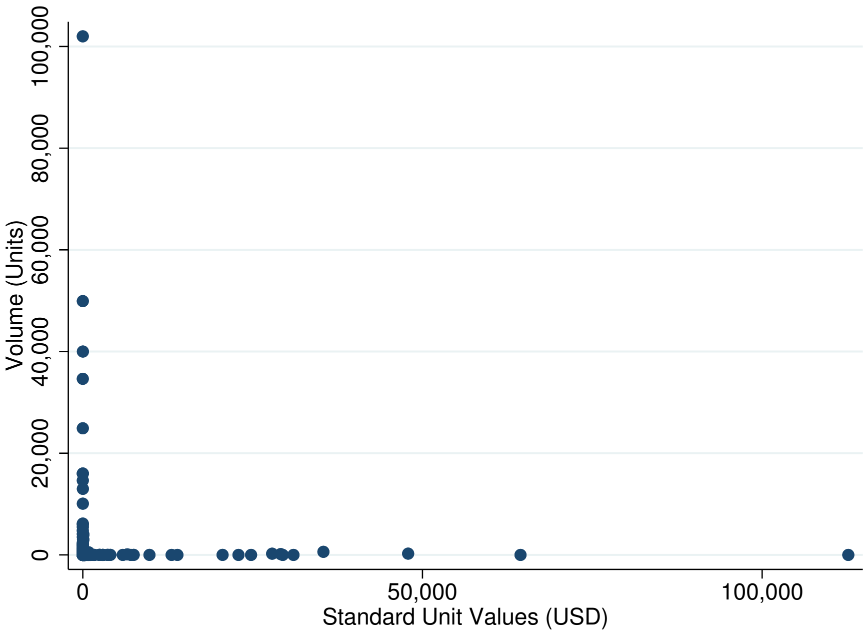
In case business chambers in Mexico perceive that negotiating free trade agreements with Asian countries could have a negative impact on sensitive sectors of the economy, preferential trade agreements focused on specific sectors could be the solution. Thus, Mexican firms in the transport equipment, machinery, and electrical equipment sectors could benefit from a higher integration in global value chains.

Finally, this chapter presents some limitations as the level of disaggregation of the data is not at the firm level. The reason for not using such a granular level is because firm data is confidential and was unavailable for this study. Nevertheless, the data employed in this chapter is at the product level, which is also very granular. In addition to this limitation, Mexican trade datasets do not differentiate between intermediate inputs, capital goods, and final goods. Making a differentiation between intermediate and capital goods constitutes an interesting area for further research as these have implications on global value chains. In the next chapter, we disentangle the effects of new imported intermediate inputs and new imported capital goods on exports of new varieties.

3.9 Appendix

3.9.1 Zero-Value Observations in Original Datasets

Figure 3.4: Zero-Value Observations in Original Datasets



Notes: This scatterplot reports the zero-value observations contained in the Mexican trade datasets. Even though these products report zero trade value, we matched their HS codes with UN Comtrade Standard Unit Values expressed in U.S. dollars. We now provide two examples to illustrate zero-value observations in the trade datasets. As a first example, the maximum data value in the horizontal axis corresponds to "Turbo-jets, turbo-propellers and other gas turbines" (i.e., HS 8411.21.01) imported from Germany in 2013; the reported value is 0 U.S. dollars, while the standard unit value is 112,685 U.S. dollars; furthermore, only one unit of this product was imported. A second example is the maximum data value in the vertical axis, which corresponds to "Ferrous waste and scrap; re-melting scrap ingots of iron or steel" (i.e., HS 7204.49.99) imported from Czech Republic in 2005; the reported value is 0 U.S. dollars, while the standard unit value is 0.25 U.S. dollars, and the volume imported corresponds to 102,003 kgs.

3.9.2 Additional Regressions

Table 3.29: Foreign Direct Investment

VARIABLES	(1) Prob_X_NEW: AME	(2) X_NEW	(3) ExtM_X_NEW	(4) IntM_X_NEW
ln(New Imported Varieties)_cit-1	0.0043***	0.0587***	0.0432**	0.0713*
	(0.0005)	(0.0073)	(0.0202)	(0.0381)
ln(GDP in PPP)_ct-1	-0.0056**	-0.0279***	-0.3317**	-0.3769*
	(0.0580)	(0.0094)	(0.1346)	(0.2044)
ln(Starting a Business)_ct-1	0.0064***	0.0168	0.2300*	0.6549***
	(0.0527)	(0.0342)	(0.1237)	(0.1773)
ln(FDI Inflows)_ct-1	-0.0001	-0.0003	-0.0004	0.0032
	(0.0009)	(0.0007)	(0.0017)	(0.0025)
Constant		1.5048***	9.6025***	9.0087*
		(0.2820)	(3.3950)	(5.2231)
Observations	367,440	368,208	480,084	480,084
Prob Wald Chi2	0.000	0.000		
R-squared			0.043	0.026
Number of varieties	30,620	30,684	40,007	40,007
Country FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table includes FDI inflows as a control variable. The dependent variable in column (1) corresponds to the probability of exporting new varieties by industry i to country c in time t; this column reports the average marginal effects using a logit model with fixed effects. The dependent variable in column (2) is the number of new exported varieties by industry i to country c in time t; this column reports the coefficients using a negative binomial with fixed effects approach. The dependent variable in column (3) is the export share of the number of new varieties by industry i to country c in time t (i.e., the extensive margin). On the other hand, the dependent variable in column (4) is the export share value (in U.S. dollars) of new varieties by industry i to country c in time t (i.e., the intensive margin). Both columns (3) and (4) are estimated by linear regressions with fixed effects. All independent variables are expressed in natural logs and lagged by one year. All regressions include country, industry, and year fixed effects.

Table 3.30: Input-Output Linkages: Direct Effect

VARIABLES	(1) ln(X_NEW_USD)	(2) ln(X_NEW_USD)	(3) ln(X_NEW_USD)	(4) ln(X_NEW_USD)
ln(New Imported Varieties DIR)_ckt	0.0531***	0.0524***	0.0526***	0.0523***
	(0.0085)	(0.0085)	(0.0085)	(0.0085)
ln(GDP in PPP)_ct		0.7940***		0.6658***
		(0.2044)		(0.2053)
ln(Starting a Business)_ct			0.5332***	0.4141**
			(0.1634)	(0.1649)
Constant	9.6409***	-10.6764**	7.4443***	-9.1024*
	(0.0857)	(5.2324)	(0.6812)	(5.1832)
Observations	21,186	21,186	21,183	21,183
R-squared	0.041	0.042	0.042	0.042
Number of varieties	2,790	2,790	2,788	2,788
Country FE	YES	YES	YES	YES
Sector FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

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

Notes: This table reports new traded varieties imported and exported within the same sector. This table is estimated using linear regressions with fixed effects. The dependent variable stands for the log value of exports (in U.S. dollars) by sector k to country c in time t. The main explanatory variable corresponds to the log value of imports of new varieties by sector k from country c in time t. All regressions include country, sector, and year fixed effects.

Table 3.31: Two-Stage Regressions using an Alternative Instrument

Second Stage IV Regressions	(1) Second Stage Prob_X_NEW AME	(2) Second Stage X_NEW	(3) Second Stage ExtM_X_NEW	(4) Second Stage IntM_X_NEW
ln(New Imported Varieties)_cit-1	0.3259***	6.2006***	0.7498**	1.0927**
	(0.0110)	(0.1468)	(0.3239)	(0.4831)
ln(GDP in PPP)_ct-1	0.01721***	-0.0515*	-0.0636	-0.0561
	(0.0022)	(0.0268)	(0.0529)	(0.0789)
ln(Starting a Business)_ct-1	0.0038***	0.2441***	0.1846***	0.3070***
	(0.0059)	(0.0224)	(0.0451)	(0.0672)
Residuals	-0.2032***	-6.0995***		
	(0.0497)	(0.1474)		
Constant	-3.3816***			
	(0.0708)			
Observations	888,450	605,302	888,450	888,450
Prob Wald Chi2	0.000	0.000		
Adj. R2			-0.093	-0.093
Country FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Underidentification stat.			3,038.712	3,038.712
Prob underident. stat.			0.000	0.000
Weak identification stat.			3,050.084	3,050.084
Prob F-test Excluded Instr.			0.000	0.000
Endogeneity test			4.507	4.410
Prob endogeneity test			0.034	0.036
First Stage IV Regressions	First Stage l_IM_NEW_L1	First Stage l_IM_NEW_L1	First Stage l_IM_NEW_L1	First Stage l_IM_NEW_L1
ln(MFN Import Tariffs)_it-1	0.0320***	0.0320***	0.0320***	0.0320***
	(0.0006)	(0.0006)	(0.0006)	(0.0006)
ln(GDP in PPP)_ct-1	-0.0397***	-0.0397***	-0.0397***	-0.0397***
	(0.0026)	(0.0026)	(0.0026)	(0.0026)
ln(Starting a Business)_ct-1	-0.0025	-0.0025	-0.0025	-0.0025
	(0.0025)	(0.0025)	(0.0025)	(0.0025)
Constant	1.1233***	1.1233***	1.1233***	1.1233***
	(0.0645)	(0.0645)	(0.0645)	(0.0645)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table reports the two-stage regressions using MFN import tariffs as an alternative instrument. Therefore, the instrumental variable employed is the log of MFN import tariffs. The top panel shows the second stage IV regressions, while the bottom panel exhibits the first stage regressions. The dependent variable on the top panel corresponds to the different specifications of exports of new varieties. Column (1) is estimated using an IV Probit model, where the dependent variable is the probability of exporting new varieties; this column reports the average marginal effects. Column (2) uses a Poisson model, where the dependent variable stands for the number of new exported varieties; this column reports the coefficients. Columns (3)-(4) are estimated using a 2SLS model, where the dependent variable corresponds to the export share of new varieties at the extensive and intensive margin, respectively.

3.9.3 Negative Binomial Model

The negative binomial model is a count data model employed in the case where the data presents overdispersion. By overdispersion, we refer to the situation where the variance is larger than the mean. As a recap, the Poisson model assumes that the variance of the dependent variable y_i equals its mean. As acknowledge in Greene (2003), Cameron & Trivedi (1990) offer an alternative methodology to the Poisson model that relaxes this assumption. Thus, the authors propose the following hypothesis test:

$H_0 : Var[y_i] = E[y_i]$
$H_1 : Var[y_i] = E[y_i] + \alpha g(E[y_i])$

The procedure is done by regressing:

$$z_i = \frac{(y_i - \hat{\lambda}_i)^2 - y_i}{\hat{\lambda}_i\sqrt{2}}, \tag{3.11}$$

where $\hat{\lambda}_i$ stands for the predicted value from the regression with or without the constant term. This procedure is followed by a t-test to conclude whether the coefficient is significantly different from zero.

3.9.4 Poisson Pseudo Maximum Likelihood Estimator

The Poisson Pseudo Maximum Likelihood (PPML) estimator is originally proposed by Gourieroux et al. (1984) who define this as a Poisson model with specification errors. Thus, the basic Poisson model is defined as:

$$Pr(y_i = j \mid x_i) = \frac{exp(-\lambda)\lambda^j}{j!}, \tag{3.12}$$

where $j = 0, 1, 2, ...,$ and λ stands for:

$$\lambda = exp(x_i'\beta) = exp(\beta_0 + \beta_1 x_{1i} + ... + \beta_k x_{ki}). \tag{3.13}$$

The vector of parameters of interest can be estimated by maximizing the log-likelihood function:

$$lnL(\beta) = \sum_{i=1}^{n}[-exp(x_i'\beta) + (x_i'\beta)y_i - ln(y_i!)]. \tag{3.14}$$

Gourieroux et al. (1984) propose to introduce a disturbance term in the parameter of the Poisson distribution as follows:

$$\lambda_i^* = exp(x_i'\beta + \epsilon_i), \tag{3.15}$$

where ϵ_i stands for a specification error (e.g., due to omitted variables).

Furthermore, Gourieroux et al. (1984) define the PPML estimator by using the following equation:

$$\sum_{i=1}^{n} x_i[-exp(x_i'\beta) + y_i] = 0. \tag{3.16}$$

Later on, Santos Silva & Tenreyro (2006) emphasize the importance of the PPML methodology to estimate log-linear models when heteroskedasticity is present. Furthermore, these last authors highlight the relevance of the PPML estimator in a trade context, such as in the trade gravity equation. The authors also claim the importance of this estimator when dealing with trade datasets that contain a significant amount of zero-value observations; this excess of zeros can be explained by the fact that not all countries trade with each other.

The empirical trade literature often deals with these zero-value observations by using two approaches. One approach consists of truncating the sample; this means to suppress the zero-value observations from the sample as in Kleinert et al. (2015). Another approach is to transform the variables by adding one unit and then, taking the natural logarithm as in Calderón et al. (2007). In contrast, Santos Silva & Tenreyro (2006) propose the PPML estimator as an alternative approach to deal with an excess of zero-value observations in the dependent variable of log-linear models.

Finally, Santos Silva & Tenreyro (2006, 2010) explain that to use the PPML estimator, the conditional mean needs to be specified as: $E[y_i \mid x_i] = exp(x_i'\beta)$.

3.9.5 Harmonized System (HS) Description

SECTION / CHAPTER	DESCRIPTION
SECTION I	**LIVE ANIMALS; ANIMAL PRODUCTS**
1	Live animals.
2	Meat and edible meat offal.
3	Fish and crustaceans, mollusks and other aquatic invertebrates.
4	Dairy produce; birds' eggs; natural honey; edible products of animal origin, not elsewhere specified or included.
5	Products of animal origin, not elsewhere specified or included.
SECTION II	**VEGETABLE PRODUCTS**
6	Live trees and other plants; bulbs, roots and the like; cut flowers and ornamental foliage.
7	Edible vegetables and certain roots and tubers.
8	Edible fruit and nuts; peel of citrus fruit or melons.
9	Coffee, tea, maté and spices.
10	Cereals.
11	Products of the milling industry; malt; starches; inulin; wheat gluten.
12	Oil seeds and oleaginous fruits; miscellaneous grains, seeds and fruit; industrial or medicinal plants; straw and fodder.
13	Lac; gums, resins and other vegetable saps and extracts.
14	Vegetable plaiting materials; vegetable products not elsewhere specified or included.
SECTION III	**ANIMAL OR VEGETABLE FATS AND OILS AND THEIR CLEAVAGE PRODUCTS; PREPARED EDIBLE FATS; ANIMAL OR VEGETABLE WAXES**
15	Animal or vegetable fats and oils and their cleavage products; prepared edible fats; animal or vegetable waxes.
SECTION IV	**PREPARED FOODSTUFFS; BEVERAGES, SPIRITS AND VINEGAR; TOBACCO AND MANUFACTURED TOBACCO SUBSTITUTES**
16	Preparations of meat, of fish or of crustaceans, molluscs or other aquatic invertebrates.
17	Sugars and sugar confectionery.
18	Cocoa and cocoa preparations.
19	Preparations of cereals, flour, starch or milk; pastrycooks' products.
20	Preparations of vegetables, fruit, nuts or other parts of plants.
21	Miscellaneous edible preparations.
22	Beverages, spirits and vinegar.
23	Residues and waste from the food industries; prepared animal fodder.
24	Tobacco and manufactured tobacco substitutes.
SECTION V	**MINERAL PRODUCTS**
25	Salt; sulphur; earths and stone; plastering materials, lime and cement.
26	Ores, slag and ash.
27	Mineral fuels, mineral oils and products of their distillation; bituminous substances; mineral waxes.
SECTION VI	**PRODUCTS OF THE CHEMICAL OR ALLIED INDUSTRIES**
28	Inorganic chemicals; organic or inorganic compounds of precious metals, of rare-earth metals, of radioactive elements or of isotopes.
29	Organic chemicals.
30	Pharmaceutical products.
31	Fertilisers.
32	Tanning or dyeing extracts; tannins and their derivatives; dyes, pigments and other colouring matter; paints and varnishes; putty and other mastics; inks.
33	Essential oils and resinoids; perfumery, cosmetic or toilet preparations.
34	Soap, organic surface-active agents, washing preparations, lubricating preparations, artificial waxes, prepared waxes, polishing or scouring preparations, candles and similar articles, modelling pastes, "dental waxes" and dental preparations with a basis of plaster.
35	Albuminoidal substances; modified starches; glues; enzymes.
36	Explosives; pyrotechnic products; matches; pyrophoric alloys; certain combustible preparations.
37	Photographic or cinematographic goods.
38	Miscellaneous chemical products.
SECTION VII	**PLASTICS AND ARTICLES THEREOF; RUBBER AND ARTICLES THEREOF**
39	Plastics and articles thereof.
40	Rubber and articles thereof.

SECTION / CHAPTER	DESCRIPTION
SECTION VIII	**RAW HIDES AND SKINS, LEATHER, FURSKINS AND ARTICLES THEREOF; SADDLERY AND HARNESS; TRAVEL GOODS, HANDBAGS AND SIMILAR CONTAINERS; ARTICLES OF ANIMAL GUT (OTHER THAN SILK-WORM GUT)**
41	Raw hides and skins (other than furskins) and leather.
42	Articles of leather; saddlery and harness; travel goods, handbags and similar containers; articles of animal gut (other than silk-worm gut).
43	Furskins and artificial fur; manufactures thereof.
SECTION IX	**WOOD AND ARTICLES OF WOOD; WOOD CHARCOAL; CORK AND ARTICLES OF CORK; MANUFACTURES OF STRAW, OF ESPARTO OR OF OTHER PLAITING MATERIALS; BASKETWARE AND WICKERWORK**
44	Wood and articles of wood; wood charcoal.
45	Cork and articles of cork.
46	Manufactures of straw, of esparto or of other plaiting materials; basketware and wickerwork.
SECTION X	**PULP OF WOOD OR OF OTHER FIBROUS CELLULOSIC MATERIAL; RECOVERED (WASTE AND SCRAP) PAPER OR PAPERBOARD; PAPER AND PAPERBOARD AND ARTICLES THEREOF**
47	Pulp of wood or of other fibrous cellulosic material; recovered (waste and scrap) paper or paperboard.
48	Paper and paperboard; articles of paper pulp, of paper or of paperboard.
49	Printed books, newspapers, pictures and other products of the printing industry; manuscripts, typescripts and plans.
SECTION XI	**TEXTILES AND TEXTILE ARTICLES**
50	Silk.
51	Wool, fine or coarse animal hair; horsehair yarn and woven fabric.
52	Cotton.
53	Other vegetable textile fibres; paper yarn and woven fabrics of paper yarn.
54	Man-made filaments.
55	Man-made staple fibres.
56	Wadding, felt and nonwovens; special yarns; twine, cordage, ropes and cables and articles thereof.
57	Carpets and other textile floor coverings.
58	Special woven fabrics; tufted textile fabrics; lace; tapestries; trimmings; embroidery.
59	Impregnated, coated, covered or laminated textile fabrics; textile articles of a kind suitable for industrial use.
60	Knitted or crocheted fabrics.
61	Articles of apparel and clothing accessories, knitted or crocheted.
62	Articles of apparel and clothing accessories, not knitted or crocheted.
63	Other made up textile articles; sets; worn clothing and worn textile articles; rags.
SECTION XII	**FOOTWEAR, HEADGEAR, UMBRELLAS, SUN UMBRELLAS, WALKING-STICKS, SEAT-STICKS, WHIPS, RIDING-CROPS AND PARTS THEREOF; PREPARED FEATHERS AND ARTICLES MADE THEREWITH; ARTIFICIAL FLOWERS; ARTICLES OF HUMAN HAIR**
64	Footwear, gaiters and the like; parts of such articles.
65	Headgear and parts thereof.
66	Umbrellas, sun umbrellas, walking-sticks, seat-sticks, whips, riding-crops and parts thereof.
67	Prepared feathers and down and articles made of feathers or of down; artificial flowers; articles of human hair.
SECTION XIII	**ARTICLES OF STONE, PLASTER, CEMENT, ASBESTOS, MICA OR SIMILAR MATERIALS; CERAMIC PRODUCTS; GLASS AND GLASSWARE**
68	Articles of stone, plaster, cement, asbestos, mica or similar materials.
69	Ceramic products.
70	Glass and glassware.
SECTION XIV	**NATURAL OR CULTURED PEARLS, PRECIOUS OR SEMI-PRECIOUS STONES, PRECIOUS METALS, METALS CLAD WITH PRECIOUS METAL AND ARTICLES THEREOF; IMITATION JEWELLERY; COIN**
71	Natural or cultured pearls, precious or semi-precious stones, precious metals, metals clad with precious metal and articles thereof; imitation jewellery; coin.

SECTION / CHAPTER	DESCRIPTION
SECTION XV	**BASE METALS AND ARTICLES OF BASE METAL**
72	Iron and steel.
73	Articles of iron or steel.
74	Copper and articles thereof.
75	Nickel and articles thereof.
76	Aluminium and articles thereof.
77	(Reserved for possible future use in the Harmonized System)
78	Lead and articles thereof.
79	Zinc and articles thereof.
80	Tin and articles thereof.
81	Other base metals; cermets; articles thereof.
82	Tools, implements, cutlery, spoons and forks, of base metal; parts thereof of base metal.
83	Miscellaneous articles of base metal.
SECTION XVI	**MACHINERY AND MECHANICAL APPLIANCES; ELECTRICAL EQUIPMENT; PARTS THEREOF; SOUND RECORDERS AND REPRODUCERS, TELEVISION IMAGE AND SOUND RECORDERS AND REPRODUCERS, AND PARTS AND ACCESSORIES OF SUCH ARTICLES**
84	Nuclear reactors, boilers, machinery and mechanical appliances; parts thereof.
85	Electrical machinery and equipment and parts thereof; sound recorders and reproducers, television image and sound recorders and reproducers, and parts and accessories of such articles.
SECTION XVII	**VEHICLES, AIRCRAFT, VESSELS AND ASSOCIATED TRANSPORT EQUIPMENT**
86	Railway or tramway locomotives, rolling-stock and parts thereof; railway or tramway track fixtures and fittings and parts thereof; mechanical (including electro-mechanical) traffic signalling equipment of all kinds.
87	Vehicles other than railway or tramway rolling-stock, and parts and accessories thereof.
88	Aircraft, spacecraft, and parts thereof.
89	Ships, boats and floating structures.
SECTION XVIII	**OPTICAL, PHOTOGRAPHIC, CINEMATOGRAPHIC, MEASURING, CHECKING, PRECISION, MEDICAL OR SURGICAL INSTRUMENTS AND APPARATUS; CLOCKS AND WATCHES; MUSICAL INSTRUMENTS; PARTS AND ACCESSORIES THEREOF**
90	Optical, photographic, cinematographic, measuring, checking, precision, medical or surgical instruments and apparatus; parts and accessories thereof.
91	Clocks and watches and parts thereof.
92	Musical instruments; parts and accessories of such articles.
SECTION XIX	**ARMS AND AMMUNITION; PARTS AND ACCESSORIES THEREOF**
93	Arms and ammunition; parts and accessories thereof.
SECTION XX	**MISCELLANEOUS MANUFACTURED ARTICLES**
94	Furniture; bedding, mattresses, mattress supports, cushions and similar stuffed furnishings; lamps and lighting fittings, not elsewhere specified or included; illuminated signs, illuminated name-plates and the like; prefabricated buildings.
95	Toys, games and sports requisites; parts and accessories thereof.
96	Miscellaneous manufactured articles.
SECTION XXI	**WORKS OF ART, COLLECTORS' PIECES AND ANTIQUES**
97	Works of art, collectors' pieces and antiques.

3.9.6 ISIC Manufacturing Sectors

SECTOR	DESCRIPTION
C10-C12	Manufacture of food products, beverages and tobacco products.
C13-C15	Manufacture of textiles, wearing apparel and leather products.
C16	Manufacture of wood and of products of wood and cork, except furniture.
C17	Manufacture of paper and paper products.
C18	Printing and reproduction of recorded media.
C19	Manufacture of coke and refined petroleum products.
C20	Manufacture of chemicals and chemical products.
C21	Manufacture of basic pharmaceutical products and pharmaceutical preparations.
C22	Manufacture of rubber and plastic products.
C23	Manufacture of other non-metallic mineral products.
C24	Manufacture of basic metals.
C25	Manufacture of fabricated metal products, except machinery and equipment.
C26	Manufacture of computer, electronic and optical products.
C27	Manufacture of electrical equipment.
C28	Manufacture of machinery and equipment n.e.c.
C29	Manufacture of motor vehicles, trailers and semi-trailers.
C30	Manufacture of other transport equipment.
C31-C32	Manufacture of furniture; other manufacturing.
C33	Repair and installation of machinery and equipment.

Chapter 4

New Imported Inputs, New Export Varieties: Capital Matters

4.1 Introduction

An important strand of the trade literature focuses on the role of trade on technology transfers from developed to developing countries (Coe et al. 1997, Eaton & Kortum 2001). Similarly, the literature also studies the effects of imports on exports (Bas & Strauss-Kahn 2014, Castellani & Fassio 2019, Damijan et al. 2014, Feng et al. 2016). Furthermore, another strand stresses the importance of imported intermediate inputs on exports but largely neglects the importance of imported capital goods on exports. Also, as far as we know, only Castellani & Fassio (2019) explore the relationship between new imported inputs and new exports, but leave aside the impact of new imported capital goods on new exports.

We consider it worth studying this last relationship as capital goods may be closely related to new and more sophisticated exports. In this chapter, we explore the mechanism behind the relationship between new imported capital goods and new exported varieties: technology embedded in imported products. Furthermore, we also examine the mechanism behind new imported intermediate inputs and new exported varieties: trade processing.

The main contribution of this chapter is to disentangle the effects between new imported intermediate inputs and new imported capital goods on exports of new varieties. Furthermore the chapter sheds light on the role of new imported capital goods, which has been relatively neglected in the literature. Our study also provides empirical evidence of the relationship between new imported varieties disaggregated by end-use and new exported varieties from the perspective of an emerging economy. Finally, Mexico is an emerging economy that heavily relies on imports of capital goods and surprisingly, the importance of this type of imports has not been explored in the literature.

We employ a four-fold empirical strategy on our estimation sample, which is composed of 68,727 new traded varieties belonging to the manufacturing sector over the period 2005-2016. We use a fixed effects negative binomial model to evaluate the impact of new imported intermediate inputs and capital goods on the number of new exported varieties. Then, we employ a fixed effects logit model to study the probability of exporting new varieties based on the number of new imported intermediate inputs and capital goods. Next, we use a linear fixed effects model to analyze the effects of importing new intermediate inputs and capital goods on the export value of new varieties. Finally, we employ a log-first difference estimator with fixed effects to examine the net change in the number of new imported intermediate inputs and capital goods on the net change in the number of new exported varieties.

Our findings suggest that new imported intermediate inputs have positive and strong statistical effects on exports of new varieties measured as number of varieties, probability, export value, and net change. Most importantly, our results also reveal the pivotal role of new imported capital goods on exports of new varieties measured as probabilities, export value, and net change. These results hold after controlling for market size and ease of doing business, and after including a full set of country, industry, and year fixed effects.

Moreover, we also conduct a series of robustness checks, which consist of including traditional trade gravity variables, excluding main trade partners, dividing our sample based on income groups, drawing a sub-sample based on top industries trading new varieties, using other trade variables as controls, employing an alternative methodology to the fixed effects negative binomial model, dealing with zero-value observations in the dependent variable, using a log-log model, looking at contemporaneous effects, increasing the lag length of the independent variables, including different combinations of fixed effects, analyzing input-output linkages across sectors, examining the marginal effects by industry, and running two-stage regressions to allow for potential endogeneity.

Although our results are statistically significant, these may not have a very large economic impact. Thus, it may be interesting to revisit this study considering state heterogeneity as part of future research. In this thesis, we did not perform the analysis at such a granular level because the Mexican authorities do not report imports at the state level due to confidentiality issues.

The rest of the chapter is organized as follows: Section 4.2 reviews the existing literature. Section 4.3 describes the data. Section 4.4 provides the descriptive statistics. Section 4.5 explains the methodology. Section 4.6 shows the results. Section 4.7 presents the robustness analysis. Section 4.8 concludes.

4.2 Literature Review

We begin this section by introducing the importance of trade on technology transfers between countries covered in subsection 4.2.1. In fact, the literature has documented that developing countries can benefit from R&D activities carried out in developed countries through imports. Then, we review the literature on the trade complementarity effects of imports and exports in subsection 4.2.2. Next, we explore the role of imported intermediate inputs and imported capital goods on exports in subsection 4.2.3. Finally, we discuss the self-selection and learning effects in subsection 4.2.4.

4.2.1 Technology Transfers

A strand of the trade literature highlights the key role of trade for technology transfers from developed economies to developing economies. This strand documents how a significant portion of R&D activities is concentrated in a few developed countries (Coe et al. 1997, Eaton & Kortum 2001). Thus, developing countries can benefit from R&D activities carried out in developed countries through imports. Other related topics to technology transfers are the self-selection and learning effects, which we discuss in subsection 4.2.4.

Coe et al. (1997) examine to what extent developing countries benefit from R&D activities performed in developed countries using cross-country data for 77 developing countries over the period 1971-1990. The findings suggest that developing countries benefit from R&D activities performed in developed countries through imports of intermediate inputs and capital goods that embody technology. The results also show a positive relationship between total factor productivity (TFP) in developing economies and R&D activities in developed economies. Moreover, the results suggest a positive relationship between TFP in developing economies and imports of capital goods from developed economies. An interesting feature commented by the authors is that productivity in Latin American countries tends to be more influenced by R&D activities carried out in the United States.

Acharya & Keller (2009) analyze the relationship between international technology transfers and R&D spillover effects using cross-country data for 17 countries from 1973 to 2002. The results imply that technology transfers from abroad have a stronger impact on productivity than domestic R&D activities. Consistent with Coe et al. (1997), the authors claim that geographically closer countries tend to benefit more from these technology transfers from abroad. In a similar vein, Coe et al. (2009) also study the impact of both domestic and foreign R&D activities on total factor productivity using a set of 24 OECD countries over the period 1971-2004. Their results indicate that both domestic and foreign R&D activities have positive effects on TFP. An interesting feature is that the authors also evaluate the impact of institutions on domestic and foreign R&D spillover effects. The authors claim that high-quality institutions have a positive impact on the magnitude of R&D spillovers. To be specific, the authors emphasize the importance of the following institutions: high-quality tertiary education systems, ease of doing business, strong patent protection, and legal systems based on English and German law.

4.2.2 The Link between Imports and Exports

Another strand of the trade literature focuses on the relationship between imports and exports, which was already reviewed in Chapter 3, except for Damijan et al. (2013). Aristei et al. (2013) examine the two-way relationship between exports and imports. Their results suggest that the relationship holds only in one direction only; this direction begins with imports having a positive effect on the probability of exporting. Damijan et al. (2013) document that roughly 70% of firms enrolled in the export activity engages in what they call "pass-on-trade"

(POT), which refers to firms importing and exporting the same products.

Lo Turco & Maggioni (2013) claim that imports from developing countries have a positive impact on the probability of exporting for Italian manufacturing firms. Bas & Strauss-Kahn (2014) also show that imported intermediate inputs positively impact the number of exported varieties. Feng et al. (2016) suggest that importing more intermediate inputs leads to an increase in the export value of firms. Consistent with these authors, Xu & Mao (2018) conclude that importing intermediate inputs can boost firms' export quality. Castellani & Fassio (2019) reveal that importing new inputs has a positive effect on exporting new products. Navas et al. (2020) explain that importing from large markets that are geographically closer may also boost exports and the export value of firms.

As a summary, the empirical trade literature focuses on the impact of imported intermediate inputs on exports (Aristei et al. 2013, Bas & Strauss-Kahn 2014, Castellani & Fassio 2019, Feng et al. 2016, Lo Turco & Maggioni 2013, Navas et al. 2020, Xu & Mao 2018), but somehow neglects the potential effects of imported capital goods with the exception of Damijan et al. (2014). In this chapter, we take a further step by focusing on new traded varieties, which is something that was not accounted by Damijan et al. (2014). Thus, an important contribution of this chapter is to shed light on the importance of importing new capital goods on exports of new varieties for a developing country.

4.2.3 The Role of Imported Intermediates and Capital Goods

This subsection considers the role of imported intermediate inputs and imported capital goods on exports. We start this subsection by exploring the extensive literature on the relationship between imported intermediates and exports. Then, we focus on the few available studies related to the importance of capital goods on exports. Finally, we finish this subsection by discussing the limited literature examining the effects of imports of both intermediate and capital goods on exports and productivity.

The Role of Imported Intermediate Inputs

A vast literature focuses on the impact of imported intermediate inputs on exports. Bas & Strauss-Kahn (2014) focus on the impact of imported intermediate inputs on firm productivity. Nonetheless, the authors also examined the impact of imported intermediate inputs on exported varieties using firm-level data for France over the period 1996-2005. The methodology employed consists of a Two-Stage Least Squares (2SLS) estimator where the instrumental variable corresponds to input tariffs from non-EU countries. The endogenous variable is the number of imported inputs in France, and the dependent variable is the number of exported varieties to EU countries as a bloc. The specification also includes controls for firm size, firm TFP, and firm and year fixed effects. The findings suggest that imported intermediate inputs positively impact the number of exported varieties.

An important feature that we analyze in this chapter, which was not considered in Bas & Strauss-Kahn (2014), is the focus on new varieties. Moreover, these authors did not consider the effects of imported capital goods, which we now consider in this chapter. Furthermore, the results presented by Bas & Strauss-Kahn (2014) are aligned to the strand on technology transfers (Acharya & Keller 2009, Coe et al. 1997, 2009); these authors claim that the import activity allows firms to gain access to inputs embedded with higher quality or with more sophisticated technology. Another feature that is not considered in Bas & Strauss-Kahn (2014) is the bilateral nature of trade activities; in other words, the authors do not differentiate between individual countries but rather consider the EU as a bloc. This chapter arguably offers a more comprehensive analysis where we account for country heterogeneity by focusing on bilateral trade of new varieties.

Feng et al. (2016) examine the impact of imported intermediate inputs on export value using Chinese manufacturing firm-level data during the period 2002-2006. The authors used 2SLS regressions to account for potential endogeneity of firms' decision to import inputs. The instrumental variables employed are import tariffs of inputs, import real exchange rates, and a variable that indicates whether a firm was already processing imports. The endogenous variable corresponds to imported intermediate inputs by firms, and the dependent variable is the export value of firms. The empirical specification also includes export tariffs and export real exchange rates as controls, as well as firm and year fixed effects. The results show that Chinese firms importing more intermediate inputs tend to experience greater export growth. Moreover, firms importing intermediate inputs with higher quality and embedded technology tend to export improved varieties. This last finding is consistent with the literature on technology transfers (Acharya & Keller 2009, Coe et al. 1997, 2009). A drawback in Feng et al. (2016) is the short five-year time span employed compared to this chapter, which considers twelve years. Moreover, the authors do not acknowledge the impact of imported capital goods on exports. Moreover, we take a further step by focusing on new varieties in this chapter.

Another paper using Chinese firm data to investigate the relationship between imported intermediate inputs and exports is Xu & Mao (2018). The methodology employed by the authors consists of linear regressions with fixed effects, where the main explanatory variable is imported intermediate inputs and the dependent variable is firms' export quality. The regression equation includes controls for firm size, average wage, firm profit, credit constraint, government subsidy, exchange rates, and firm ownership; this specification also includes firm and year fixed effects. The findings report that imported intermediate inputs with embedded quality have a positive effect on firms' export quality. These findings suggest that imports of intermediate inputs boost the quality of exports. A pitfall in Xu & Mao (2018) is also the short time span of the dataset consisting of eight years, compared to the twelve-year span in this chapter. Furthermore, we aim to examine this same relationship between imported intermediate inputs and exports; additionally, we study the relationship between imported capital goods and exports. Finally, the scope of our analysis is different as we focus on new varieties only, instead of using all the varieties.

One of the closest papers to this chapter would be Castellani & Fassio (2019), who analyze the linkage between new imported inputs and new exported products using data on Swedish manufacturing firms over the period 2001-2012. The methodology employed consists of a negative binomial model. In this specification, the main explanatory variable corresponds to new imported inputs, and the dependent variable is new exported products. Moreover, the baseline regression equation includes controls for the number of employees, investments, productivity, a dummy variable for having at least one patent in the firm, and dummies for firm ownership (i.e., Swedish group, Swedish MNE, and foreign MNE), as well as firm and year fixed effects. The results suggest that an increase in the number of new imported products leads to an increase in the number of new exported products.

An important contribution of Castellani & Fassio (2019) is that it sheds light on the relationship between new imported varieties and new exported varieties using an extensive dataset of the whole population of manufacturing firms in Sweden. Although Castellani & Fassio (2019) have extensively studied the determinants of new exported products and the probability to start exporting, the authors did not consider the impact on the export value of new varieties and on the net change in the number of new varieties; these two features are now examined in this chapter.

Another closely related paper presented in this chapter is Navas et al. (2020), who examine the indirect role of market size and geographical proximity via imports on exports at the extensive and intensive margins. The methodology employed comprises 2SLS and Poisson regressions relying on data of Italian manufacturing firms from 2000 to 2006. The authors use two instrumental variables; the first instrument corresponds to GDP weighted by a firm's import share of each country, and the second instrument is the log of total imports from other EU countries weighted by the relative importance of a product in firms' total imports. The endogenous variables are the firm's productivity and the TFP-enhancing effect of imported intermediate inputs. The analysis presents two sets of regressions; the first accounts for exports at the extensive margin, where the dependent variable is the export status of a firm denoted with a dummy variable. The second type of regression reports the exports at the intensive margin, where the dependent variable is defined as total exports of a firm to a specific country. All regressions incorporate controls for GDP, distance, trade openness, remoteness, and market costs; these regressions also include firm and year-area fixed effects. The findings suggest that the market size and distance have significant effects on firms' decision to import. Furthermore, both market size and distance have an indirect effect on firms' exports through imports.

The Role of Imported Capital Goods

The trade literature also covers a few studies on the importance of imported capital goods. Eaton & Kortum (2001) develop a model of trade in capital goods, which accounts for technological change embedded in new capital goods. In the paper, the authors acknowledge the importance of this type of imports on productivity due to the beneficial effects of technology transfers. The authors explain that technological innovations are mainly concentrated in a few developed countries. Thus, these developed countries can transfer technological advances through exports of capital goods targeted to developing countries.

Koren & Csillag (2011) explore the impact of importing capital goods on the wage gap within machine operators using firm-level data for Hungary over the period 1994-2004. The authors use pooled cross-section regressions and linear regressions with fixed effects. In both specifications, the dependent variable is the monthly earnings of workers, and the main explanatory variables are the usage of imported machines specific to occupation and exposure to firm-level imports. These regressions also include firm controls (e.g., firm employment and ownership) and individual controls (e.g., gender, educational attainment, and age). The panel approach includes year fixed effects. Consistent with Eaton & Kortum (2001), the authors agree that capital goods are predominantly produced by developed countries. Due to this reason, developing countries heavily rely on imports of capital goods from developed countries. The results indicate that imports of capital goods have a strong and positive effect on wages for machine operators. A plausible explanation for these wage differences is that operators of sophisticated imported machines need to be better skilled compared to those operating domestic machines.

Moreover, Yasar (2013) examines the importance of firms' absorptive capacity of imported machinery using Chinese firm-level data for the year 2003. The methodology consists of a cross-section regression, where the dependent variable is value-added. The explanatory variables are labor, capital input, a dummy variable for imported machinery, the share of high skilled labor, and an interaction term between imported machinery and share of high skilled labor; this interaction term constitutes the main variable of interest. The specification also includes controls, such as firm age, whether the firm's products have an ISO 9000 certification, and capacity utilization, as well as firm size, industry, and region dummies. The paper suggests that importing machinery has a positive effect on firm productivity. Furthermore, firms that have a high capacity of absorption of new technologies benefit more from importing capital goods from developed countries. This last finding complements Eaton & Kortum (2001) and Koren & Csillag (2011), who explain that developing countries tend to import capital goods produced in developed countries, but did not specify which type of firms benefit the most. Finally, this study by Yasar (2013) presents a drawback, which is the fact that the usage of a longitudinal dataset would have constituted a richer analysis.

Although a few papers in the literature highlight the importance of capital goods, these do not attempt to study the impact of new imported capital goods on exports of new varieties, which is now explored in this chapter.

Imported Intermediate Inputs versus Imported Capital Goods

Besides the previous studies analyzing the importance of imported intermediate inputs and capital goods, by separate, we can also find a couple of studies that compare both types of imported goods on exports and productivity. Damijan et al. (2014) examine the impact of churning in imported varieties of intermediate inputs and capital goods on the firm export scope and productivity using Slovenian firm-level data over the period 1994-2008.[1] The authors use a log-first difference estimator with fixed effects for three different specifications. The dependent

[1] Churning refers to the process of adding and withdrawing goods in the product mix.

variables for each specification correspond to the net changes in the log number of exported capital goods, intermediate inputs, and final goods, respectively; these variables are expressed in logs and as first differences. On the other hand, the independent variables are the net changes in the log number of imported capital goods, intermediate inputs, and final goods, as well as import tariffs and import unit value; these explanatory variables are all expressed in logs and as first differences in all the three regression equations. Moreover, these regression equations control for firm size, inward and outward FDI, and include industry, firm, and year fixed effects.

The findings in Damijan et al. (2014) suggest that adjusting the product mix of imported intermediate inputs and capital goods has a strong impact on firms' export scope and on productivity. Although the authors explore the net changes in the number of imported capital goods and intermediate inputs on the net change in the number of exported goods, they did not attempt to study these relationships from the perspective of new varieties, which is now explored in this chapter. A drawback in Damijan et al. (2014) is the level of aggregation of the dataset, which is at the industry level; in this chapter, we use a more granular disaggregation level, which is at the product level (i.e., HS 8-digits). This fine disaggregation level allows the identification of new varieties over time, which is the innovative feature in this chapter.

An interesting feature is that previous studies in the trade literature focus on the impact of tariff reductions on firms' productivity (Amiti & Konings 2007, Goldberg et al. 2009, Topalova & Khandelwal 2011); however, Damijan et al. (2014) suggest that the adjustment of the product mix of imported varieties has a stronger effect on productivity than trade liberalization. In this chapter, we exclusively focus on trade of new varieties using Mexico as a case study. It is worth mentioning that just as we observed in Chapter 3 and aligned to Goldberg et al. (2010*b*), who focus on India, it seems that Mexican firms tend to add products to the product mix while keeping obsolete products instead of withdrawing these products.

Moreover, Caselli (2018) studies the relationship between import activity and productivity using Mexican plant-level data for the period 1994-2003. The author employed different estimators, such as OLS and linear fixed effects models. The dependent variable is defined as productivity, and the independent variables are dummy variables for plants exporting in the previous year, plants importing intermediate inputs in the previous year, and plants importing capital goods in the previous year. The results show that more productive plants tend to become importers of capital goods rather than importers of intermediate inputs. Furthermore, these plants importing capital goods tend to experience higher productivity levels. Also, the results suggest that imported capital goods are more likely to embody technological advances than imported intermediate inputs. In this chapter, we focus instead on the impact of new imported intermediate and capital goods on exports of new varieties using Mexican product-level data.

4.2.4 Self-Selection and Learning Effects

This chapter is somehow related to the discussion on self-selection and learning effects in the trade literature. The following studies presented in this subsection aim to explain the reasons for positive and strong links between productivity and trade activities. Thus, we review an interesting discussion between the self-selection into imports, learning by importing, and learning by exporting mechanisms. We now proceed to discuss the evidence surrounding this strand of the literature.

Self-Selection into Imports

A strand of the trade literature focuses on the decision made by firms to self-select into the import activity. Castellani et al. (2010) use Italian manufacturing firm-level data to examine firm heterogeneity and trade activities. The authors document that only a few firms are responsible for a large share of international transactions; these firms operate within a range of sectors and trade with several countries. In general terms, these few firms engaging in both import and export activities exhibit better performance compared to domestic firms. Furthermore, firms that only import display better performance than those firms only exporting. The results suggest that firms tend to self-select into imports.

These results are similar to Kugler & Verhoogen (2009), where the authors analyze plants that self-select into imports using Colombian manufacturing plant-level data over the period 1982-1988. The results indicate that plants exhibiting higher productivity tend to self-select into imports. Furthermore, more productive plants tend to import inputs embedded with superior quality from abroad. In the case of Colombia, the authors explain that manufacturing plants are more inclined to import inputs of higher quality rather than to purchase inputs from the domestic market.

In the same vein, Kasahara & Lapham (2013) develop a trade model with heterogeneous firms, where producers of final goods must choose to import and export intermediate inputs. Then, the authors use an empirical approach to study the impact of the export and import status on plants' performance using Chilean manufacturing plant-level data over the period 1990-1996. The findings suggest that trade in intermediate inputs and final goods has a positive impact on aggregate productivity. Moreover, trade policies restricting imports of intermediate inputs have a negative effect on exports of final goods due to the trade complementarity effect between imports and exports.

Learning Effects: Learning by Importing or Learning by Exporting?

Turning now to the learning effects, the trade literature divides these effects into two categories: learning by importing and learning by exporting. The learning by importing mechanism refers to firms' ability to improve their productivity after starting to import inputs from abroad. By contrast, the learning by exporting mechanism refers to firms' capacity to increase their productivity after entering the export market. We now proceed to discuss both trade mechanisms.

On the one hand, there is empirical evidence showing that firms tend to learn by importing. Damijan & Kostevc (2015) explore the sequence of firm's participation in trade activities and their learning effects using Spanish manufacturing firm-level data over the period 1990-2008. The findings show that small firms and technologically advanced firms learn by importing, allowing these firms to innovate first and export later on. These findings imply that the sequencing starts from importing goods, followed by making innovations, which promotes exporting goods, leading to further innovations. Similarly, Fernández & Gavilanes (2017) analyze the link between imports and productivity using Ecuadorian manufacturing, wholesales, and retailers data during the period 2009-2012. The authors claim that imports have the potential to generate technology diffusion if two conditions are met. First, firms should be capable of absorbing new technologies from abroad; as the results suggest, imports of developed countries have a positive impact on labor productivity in the manufacturing sector. Second, the national innovation system must be in a mature phase to fully benefit from learning opportunities offered by importing advanced technologies embedded in goods from abroad.

On the other hand, there is also evidence showing that firms can learn by exporting. Clerides et al. (1998) examine the relationship between export activity and productivity using plant-level data for Colombia, Mexico, and Morocco.[2] The findings suggest that relatively more efficient firms tend to become exporters. Moreover, the authors claim that this strong relationship may be explained by self-selection into exporting. Similarly, De Loecker (2013) analyzes how learning by exporting could lead to productivity gains using Slovenian manufacturing firm-level data for the period 1994-2000. The results suggest that exports can lead to productivity gains; however, these effects can differ across manufacturers.

Self-Selection into Importing versus Learning by Importing

Thus far, we have discussed the self-selection and learning effects separately; however, a few studies analyze both effects by exploring the relationship between imports and productivity. Vogel & Wagner (2010) examine this causal link using German manufacturing firm-level data over the period 2001-2005. The findings suggest that more productive firms tend to self-select into imports. However, the authors claim that importing does not have an effect on productivity due to learning by importing. Likewise, Caselli (2018) studies this relationship using Mexican plant-level data for the period 1994-2003. The results indicate that plants tend to self-select into import activities, and also, plants learn by importing. Similarly, Zhou et al. (2020) examine this link between imports and productivity using Chinese manufacturing firm-level data for the period 2000-2005. The findings show that firms with import activity display higher productivity levels compared to non-trading firms. Moreover, the authors argue that this strong positive relationship between imports and productivity is caused by the learning by importing mechanism, rather than the self-selection into importing mechanism.

These findings provide mixed evidence on the self-selection and learning effects literature. First, both Vogel & Wagner (2010) and Caselli (2018) converge on the idea that more productive firms tend to self-select into the import activity, which is aligned to Castellani et al. (2010),

[2] Clerides et al. (1998) present a limitation, which is that the time span and sample size for Mexican firms is not large enough. In the case of Colombian and Moroccan data these are available for almost all plants over the 1981-1991 and 1984-1990 periods. In contrast, Mexico displays information for only 2,800 of the largest firms over the 1986-1990 period. Due to this short 5-year time span and considering that the empirical specification employs variables with three lags, the sample size significantly reduced. Thus, the authors did not employ Mexican data for the empirical approach.

Kasahara & Lapham (2013) and Kugler & Verhoogen (2009). In contrast, Vogel & Wagner (2010) and Caselli (2018) differ on the learning by importing mechanism as the results suggest that this learning by importing mechanism is important for Mexican firms, while it seems irrelevant for German firms. In this second point, Caselli (2018) and Zhou et al. (2020) both agree on the presence of self-selection into importing mechanism. Nonetheless, these studies differ on the fact that Caselli (2018) highlights the importance of both learning by importing and self-selection into importing, while Zhou et al. (2020) only claim the relevance of the learning by importing mechanism.

4.3 Data

Similar to Chapter 3, the database employed consists of a compilation of official datasets. We start with bilateral trade data, which was retrieved from the Mexican Ministry of Economy. These annual datasets on exports and imports are reported at the tariff line-level (HS 8-digits) and are aggregated at the country-level over the period 2003-2016. These datasets contain bilateral trade information for more than 12,000 products per year for over 230 countries and territories. Furthermore, the datasets cover the entire universe of goods (i.e., agricultural, extractive, and manufacturing goods); however, we only focus on manufacturing goods.[3] Moreover, the datasets contain information on the value of the merchandise expressed in U.S. dollars, volume in units, and source and destination countries of the traded merchandise. It is worth mentioning that these trade datasets are no longer updated; thus, the most recent available year in the datasets is 2016.

As previously mentioned, the focus of this study is on new varieties.[4] Therefore, as part of the methodology, we identify new, continuing, and withdrawn varieties. To make a distinction between these categories, one of the criteria to identify new varieties was to rely on the evolution of HS codes over time, as proposed by Colantone & Crinò (2014). Thus, we use two concordance tables sourced from the Integrated Foreign Trade Information System (SIICEX). These concordance tables correspond to the TIGIE 2002-2007 and TIGIE 2007-2012, which capture the added, modified, and suppressed product codes during the analyzed period. Since we are now interested in disentangling the effects of imports of new intermediate inputs and new capital goods, we make use of the UN Broad Economic Categories (BEC), as suggested by the literature (Arkolakis et al. 2008, Damijan et al. 2014, Dean et al. 2011, Feng et al. 2016, Koopman et al. 2012). In particular, we use the correspondence table HS 6-digits (2012) - BEC Rev.5 to distinguish goods by end-use (i.e., intermediate, capital, and consumption goods).

[3]This chapter only considers manufacturing goods, due to the nature of the interest variables; our main explanatory variables are new imported intermediate inputs and new imported capital goods; thus, agricultural and extractive goods do not enter into either of these categories; therefore, we leave these observations out of the estimation sample.

[4]We define a variety as a product-country combination, which is a standard definition in the trade literature (Arkolakis et al. 2008, Broda & Weinstein 2006, Colantone & Crinò 2014, Goldberg et al. 2009, 2010*a*,*b*). Furthermore, a product is defined as the tariff line identified by the Harmonized System (HS) nomenclature at 8-digits.

The control variables included in this study are the GDP of partner countries and score of starting a business. We employ GDP at purchasing power parity (PPP) in constant international dollars, which was retrieved from the World Bank Development Indicators. On the other hand, the World Bank Doing Business dataset includes three important trade-related variables considered in Navas et al. (2020); these are the number of documents to import, cost to import in U.S. dollars per container deflated, and time required to import in days. Nonetheless, a limitation of these last trade-related variables is the short time span covering the period 2006-2015. Due to this reason, we decided to use the score of starting a business as a control variable instead because this last variable has a longer time span covering the period 2004-2020. We use this score as a proxy for the ease of doing business and trading with partner countries.[5] Additionally, we also considered the above-mentioned trade-related variables as controls in the Robustness Analysis section.

Moreover, we use trade gravity variables as part of the robustness checks. These gravity variables are associated with latitude and longitude, official language, landlocked-status, continent, and colonizer, which were sourced from the Geo CEPII Database (Calderón et al. 2007, Imbs 2004, Navas et al. 2020). In the case of bilateral distance, we employ the great-circle distance formula to calculate the distance between Mexico City and each partner country's capital city (Jansen & Stokman 2014, Navas et al. 2020). We also include border, which is a dummy variable that equals one if a country shares a border with Mexico; and zero otherwise (Calderón et al. 2007, Clark & Van Wincoop 2001, Imbs 2004, Jansen & Stokman 2014, Kleinert et al. 2015). Also, we include a free trade agreement (FTA) variable, which is a dummy variable that equals one if a country has a free trade agreement with Mexico; and zero otherwise (Jansen & Stokman 2014). This variable is constructed relying on the Organization of American States (OAS) list of trade agreements.

We also use input-output tables from the World Input-Output Database (WIOD) sourced by the University of Groningen. This database includes information for 28 EU countries and 15 major countries over the 2000-2014 period. This input-output data is reported at the sector level under the International Standard Industrial Classification (ISIC) at a 2-digits level. To match ISIC manufacturing sectors to HS trade data, we use the OECD Correspondence Table HS-ISIC.

We also employ data on applied import tariffs to construct the instrumental variables. This import tariff data reports the average of ad valorem duties, and was retrieved from the World Trade Organization (WTO). Since we aim to disentangle the effects of new imported intermediate inputs and new imported capital goods, we recreate two variables: applied tariffs for intermediate imports and applied tariffs for capital imports. Finally, we sourced country code labels from the International Organization for Standardization (ISO) and from the Mexican Ministry of Economy; this last type of country code corresponds to internal codes used by

[5]The score of starting a business is defined by the World Bank Doing Business dataset as the simple average of the following scores: procedures, time, and cost to start and operate a business, as well as minimum capital requirements. Each of these individual indicators are measured on a scale from 0 to 100, where 0 represents the worst regulatory performance and 100 the best regulatory performance.

Mexican authorities.

4.4 Descriptive Statistics

4.4.1 Decomposition Exercise

The basis of our analysis relies on identifying new, continuing, and withdrawn products from the universe of manufacturing goods traded by Mexico during the period 2003-2016. As in Chapter 3, we define a variety as a product-country combination, which is a standard definition in the trade literature (Arkolakis et al. 2008, Broda & Weinstein 2006, Colantone & Crinò 2014, Goldberg et al. 2009, 2010*a*,*b*).[6]

In line with Colantone & Crinò (2014), we use two criteria to define a new variety. First, when a product is traded (i.e., exported or imported) with a country for the first time. Second, by tracing the evolution of product codes over time using concordance tables provided by Mexican authorities.

To be more specific, we identify a product as new under the next three circumstances. First, the tariff line is introduced to the Harmonized System in time t and does not have any previous code corresponding to it. Second, the tariff line is introduced to the Harmonized System in time t and has one or more previous codes corresponding to it, but none was traded with a particular country in any previous year. Third, the tariff line is not new to the Harmonized System but has not been traded with a particular country in any previous year. Following these criteria, traded varieties can be counted as new only once.

A major difference compared to the previous chapter is that now we focus on disentangling the effects of imports of new intermediate inputs and new capital goods on exports of new varieties. To do so, first we differentiate between intermediate, capital, and consumption goods in a similar manner as Arkolakis et al. (2008), Damijan et al. (2014), Dean et al. (2011), Feng et al. (2016), Koopman et al. (2012); these authors use the UN Broad Economic Categories (BEC).

It is important to point out that the BEC code classification could have two end-uses (e.g., a product could be classified as a capital good, but also as an intermediate input). To avoid consistency issues, we restrict our sample to unique categories. In other words, goods are clearly classified as either capital, consumption, or intermediate inputs (i.e., categories are mutually exclusive). These unique BEC categories represent roughly 90% of the original sample.

Table 4.1 identifies the number of traded manufacturing varieties by Mexico into new, continuing, and withdrawn. We identify a total of 227,005 new exported varieties, which account for 17.8% of the total number of exported varieties; from this number of new exported varieties, roughly 7,610 corresponds to new products (i.e., new HS-8 digit codes).

On the other hand, we ended up with a total of 1,734,200 imported varieties once we restricted our sample to unique BEC categories. From this total, about 63% corresponds to

[6] In this chapter, products are defined as 8-digit tariff lines of the Harmonized System.

Table 4.1: Identification of Traded Varieties

		Total Varieties	**New Varieties**	**Withdrawn Varieties**	**Continuing Varieties**
Exported Varieties	(Freq.)	1,275,607	227,005	5,019	1,043,583
Exported Varieties	*(%)*	*100.0*	*17.8*	*0.4*	*81.8*
Imported Intermediate Inputs	(Freq.)	1,090,021	142,527	703	946,791
Imported Intermediate Inputs	*(%)*	*100.0*	*13.1*	*0.1*	*86.9*
Imported Capital Goods	(Freq.)	366,627	48,786	229	317,612
Imported Capital Goods	*(%)*	*100.0*	*13.3*	*0.1*	*86.6*
Imported Consumption Goods	(Freq.)	277,552	38,322	429	238,801
Imported Consumption Goods	*(%)*	*100.0*	*13.8*	*0.2*	*86.0*

Notes: This table identifies the universe of traded manufacturing goods by Mexico into new, continuing, and withdrawn varieties. Figures displayed in odd rows represent the number of exported varieties, imported intermediate inputs, imported capital goods, and imported consumption goods, respectively. On the other hand, figures in italics display these traded varieties as percentages.

imports of intermediate inputs, 21% to imports of capital goods, and 16% to imports of consumption goods. Thus, we can observe that the largest share of imported varieties corresponds to intermediate inputs, which are associated with processing trade.

We now examine the breakdown of imported varieties by end-use (i.e., intermediate inputs, capital goods, and consumption goods) and by product category (i.e., new, continuing, and withdrawn varieties). First, we identify a total number of 142,527 new imported intermediate inputs; this figure represents 13.1% of the total number of imported intermediate inputs; from this number of new imported intermediate inputs, around 6,030 corresponds to new products (i.e., new HS 8-digit codes).

Then, we identify a total number of 48,786 new imported capital goods; this figure accounts for 13.3% of the total number of imported capital goods; from this number of new imported capital goods, approximately 2,145 corresponds to new products. Finally, we identify a total number of 38,322 new imported consumption goods; this figure represents 13.8% of the total number of imported consumption goods; from this number of new imported consumption goods, around 2,865 corresponds to new products.

To conclude this subsection, it is important to mention that our sample suffered from two reductions. The first reduction was due to the process of matching products to unique BEC categories. The second reduction is associated with the scope of our analysis; since we are exclusively interested in the impact of new imported intermediate inputs and new imported capital goods, we leave consumption goods out of our sample. These two reductions explain why our estimation sample is slightly smaller compared to the previous chapter. Finally, we aggregate these new intermediate inputs and new capital goods at the industry level (i.e., HS 4-digits).

4.4.2 Variable Description

Dependent Variable

The dependent variable consists of new exported varieties to country c by industry i in year t. We employ four different measures of the dependent variable. Some of these measures vary from Chapter 3.

The first measure corresponds to the number of new exported varieties to country c by industry i in year t. We employ this measure of the dependent variable in our fixed effects negative binomial model. Our measure in levels of the dependent variable is defined as:

$$X\_NEW_{cit} = \sum_k X\_PROD^k_{cit}, \tag{4.1}$$

where X_NEW_{cit} corresponds to the sum of new varieties k (in HS 8-digits) of exported products X_PROD to country c belonging to industry i (in HS 4-digits) in year t.

The second measure represents the probability of exporting new varieties to country c by industry i in year t; this dependent variable constitutes the extensive margin. In this chapter, we define the extensive margin as in Navas et al. (2020), where the dependent variable is the export status expressed as a dummy variable. We use this dependent variable in our fixed effects logit model. We define this probability as:

$$Prob\_X\_NEW_{cit} > 0, \tag{4.2}$$

where $Prob\_X\_NEW_{cit}$ is a dummy variable that equals 1 if at least one new variety is exported to country c by industry i in time t; and zero, otherwise.

The third measure of the dependent variable stands for the export value (in U.S. dollars) of new varieties to country c by industry i in time t; this dependent variable represents the intensive margin. In this chapter, we also define this intensive margin as in Navas et al. (2020). We use this measure of the dependent variable in our linear fixed effects models. We define our intensive margin as follows:

$$IntMarg\_V_{cit} = ln(X\_NEW^{USD}_{cit}), \tag{4.3}$$

where $IntMarg\_V_{cit}$ represents the log value (in U.S. dollars) of exports of new varieties to country c by industry i in time t.

The fourth measure denotes the net change in the log number of new exported varieties to country c by industry i in time t. We employ this measure of the dependent variable on a log-first difference model with fixed effects. We define this variable as:

$$\Delta\ ln(X\_NEW_{cit}) = ln(X\_NEW_{cit}) - ln(X\_NEW_{cit-1}), \tag{4.4}$$

where $\Delta\ ln(X\_NEW_{cit})$ stands for the net change in the log number of new exported varieties to country c by industry i in time t.

Main Explanatory Variables

In this chapter, we define two main explanatory variables for the negative binomial, logit, and linear models combined with fixed effects. The first main explanatory variable corresponds to the log number of new imported intermediate inputs from country c by industry i in the previous year. This variable is defined as follows:

$$ln(IM\_NEW^{int}_{cit-1}) = ln(\sum_k IM\_PROD^{int\ k}_{cit-1}), \tag{4.5}$$

where $IM\_NEW^{int}_{cit-1}$ corresponds to the sum of new varieties k (in HS 8-digits) of imported intermediate inputs IM_PROD^{int} from country c belonging to industry i (in HS 4-digits) in the previous year $t-1$.

The second main explanatory variable represents the log number of new imported capital goods from country c by industry i in the previous year. This variable is defined as:

$$ln(IM\_NEW^{cap}_{cit-1}) = ln(\sum_k IM\_PROD^{cap\ k}_{ckit-1}), \tag{4.6}$$

where $IM\_NEW^{cap}_{cit-1}$ corresponds to the sum of new varieties k (in HS 8-digits) of imported capital goods IM_PROD^{cap} from country c belonging to industry i (in HS 4-digits) in the previous year $t-1$.

Furthermore, we define two more main explanatory variables for the log-first difference model with fixed effects. On the one hand, we have the net change in the log number of new imported intermediate inputs. We define this explanatory variable as:

$$\Delta\ ln(IM\_NEW^{int}_{cit}) = ln(IM\_NEW^{int}_{cit}) - ln(IM\_NEW^{int}_{cit-1}), \tag{4.7}$$

where $\Delta\ ln(IM\_NEW^{int}_{cit})$ stands for the net change in the log number of new imported intermediate inputs from country c by industry i in time t.

On the other hand, we have the net change in the log number of new imported capital goods. We define this other explanatory variable as follows:

$$\Delta\ ln(IM\_NEW^{cap}_{cit}) = ln(IM\_NEW^{cap}_{cit}) - ln(IM\_NEW^{cap}_{cit-1}), \tag{4.8}$$

where $\Delta\ ln(IM\_NEW^{cap}_{cit})$ stands for the net change in the log number of new imported capital goods from country c by industry i in time t.

Like in the previous chapter, to avoid the log of zero, we add one unit to the main explanatory variables before taking the natural logarithm.

Control Variables

In terms of the control variables, we include the log of GDP at PPP in constant international dollars: ln(GDP in PPP)$_{ct-1}$. We also use the log of the score of starting a business: ln(Starting a Business)$_{ct-1}$. This score proxies the ease of doing business with partner countries. All control variables are expressed in natural logs and lagged by one year to tackle potential reverse causality.[7]

As previously mentioned in Chapter 3, despite FDI plays an important role in processing trade, we did not include FDI inflows as a control variable in our baseline specifications because we only have data for 47 major investment countries. Just as before, our focus is on new varieties, and one of the criteria to be considered a new variety is that the product is traded with a country for the first time. Therefore, there is no FDI data for an important amount of new source and destination countries. Nevertheless, we include FDI inflows as a control variable in the Appendix section. Nonetheless, FDI has an insignificant effect on exports of new varieties.

4.4.3 Summary Statistics

Table 4.2 reports the summary statistics of our estimation sample. This estimation sample comprises 68,727 new varieties over the period 2005-2016. This number of new varieties is the result of product-country combinations belonging to manufacturing industries. In other words, these varieties are new products traded with specific partner countries. We aggregate these new products (i.e., HS 8-digits) into industries (i.e., HS 4-digits), which is standard in the trade literature.

We employ four different measures of the dependent variable. These dependent variables are the number of new exported varieties (X_NEW); the probability of exporting new varieties ($Prob\_X\_NEW$), which represents the extensive margin; the log of the export value in U.S. dollars ($IntMarg\_V$), which represents the intensive margin; and the net change in the log number of new exported varieties (X_NEW_D1). From this table, we can observe a maximum number of 46 new exported varieties (i.e., product-country combination) within an industry. We can also notice that the probability of exporting a new variety is 11.05% for Mexico.[8]

In this study, we now make a distinction between the number of new imported intermediate inputs and new imported capital goods; these two explanatory variables account for 44 and 34 new varieties, respectively. We also use another set of independent variables, representing

[7] As mentioned before, we add one unit to the independent variables before taking the natural logarithm.

[8] This percentage slightly differs from Chapter 3 since we now reduced the estimation sample to intermediate and capital goods, leaving consumption goods out of the sample.

Table 4.2: Summary Statistics

Variables	Labels	(1) N	(2) Mean	(3) Std.Dev.	(4) Min	(5) Max
Number of New Exported Varieties	X_NEW	824,724	0.1576	0.5836	0	46
Probability of Exporting New Varieties	$Prob\_X\_NEW$	824,724	0.1105	0.3135	0	1
ln(Export Value in USD)	$IntMarg\_V$	824,724	0.8171	2.5437	0	22.4346
Net Change in (log) No. of New Exported Varieties	X_NEW_D1	824,724	-0.0017	0.3611	-3.3673	3.8501
No. New Imported Intermediate Inputs	$IM\_NEW^{int}_{cit}$	824,724	0.0933	0.4552	0	44
No. New Imported Capital Goods	$IM\_NEW^{cap}_{cit}$	824,724	0.0300	0.2574	0	34
ln(No. New Imported Intermediate Inputs)_cit-1	$ln(IM\_NEW^{int})_{cit-1}$	824,724	0.0642	0.2353	0	3.9512
ln(No. New Imported Capital Goods)_cit-1	$ln(IM\_NEW^{cap})_{cit-1}$	824,724	0.0213	0.1405	0	3.5553
Net Change in (log) No. New Imported Intermediates	$\Delta ln(IM\_NEW^{int})_{cit}$	824,724	-0.0089	0.2803	-3.8501	3.6636
Net Change in (log) No. New Imported Capital Goods	$\Delta ln(IM\_NEW^{cap})_{cit}$	824,724	-0.0035	0.1618	-3.2189	3.5553
ln(GDP in PPP)_ct-1	$ln(GDP\ in\ PPP)_{ct-1}$	824,724	25.9776	2.0641	18.9133	30.5546
ln(Starting a Business)_ct-1	$ln(Starting\ a\ Business)_{ct-1}$	824,724	4.2828	0.2883	0.7885	4.6052
ln(No. Documents to Import)_ct-1	$ln(Docs2Import)_{ct-1}$	677,104	1.7780	0.4058	0.6931	3.0445
ln(Import Costs)_ct-1	$ln(Cost2Import)_{ct-1}$	677,104	7.2625	0.5889	5.9092	9.8975
ln(Time to Import)_ct-1	$ln(Time2Import_{ct-1}$	677,104	2.7653	0.6423	1.3863	4.7622
ln(Distance)	$ln(Distance)_c$	821,160	8.9501	0.6948	6.9680	9.7742
Border	$Border_c$	821,160	0.0332	0.1790	0	1
Landlocked	$Landlocked_c$	821,160	0.1216	0.3269	0	1
Continent	$Continent_c$	821,160	0.3145	0.4643	0	1
Language	$Language_c$	821,160	0.2049	0.4036	0	1
Colonizer	$Colonizer_c$	821,160	0.2007	0.4005	0	1
Free Trade Agreement	FTA_{ct-1}	824,724	0.3873	0.4871	0	1
ln(Applied Interm Import Tariffs)_it-1	$ln(Applied^{int})_{it-1}$	593,824	1.1446	1.1170	0	3.6288
ln(Applied Capital Import Tariffs)_it-1	$ln(Applied^{cap})_{it-1}$	168,809	1.3344	1.1225	0	3.8712

Notes: The estimation sample is conformed by 68,727 new varieties over 12 years.

the first difference of the log number of new imported intermediate inputs and new imported capital goods.

In terms of our control variables, we employ GDP at PPP in constant international dollars and score of starting a business. It is worth mentioning that the mean score of starting a business is 76, where Guinea-Bissau obtained the lowest score of 2.2 over the period 2006-2009 and New Zealand obtained the highest score of 100 during the period 2009-2016. All the independent variables are expressed in their log form and lagged by one year to avoid potential reverse causality. Also, we add one unit to all our independent variables before taking the natural logarithm to avoid the log of zero, which is undefined.[9]

Furthermore, we use trade-related variables as part of our robustness checks; these include number of documents to import, costs to import, and time to import. The mean number of documents required to import goods is 6 documents; France only requests 2 documents while Rwanda requests 21 documents. Also, the mean import costs are 1,756 U.S. dollars per container; on the one hand, Singapore is the most efficient country with import costs of 368 U.S. dollars per container; on the other hand, Uzbekistan is the less efficient destination with import costs of 19,881 U.S. dollars per container.[10] Finally, the mean time to import is 20 days; Singapore is again the most efficient country requiring only 4 days to import; again Uzbekistan

[9]Adding one unit and taking the natural logarithm is an approach often used in the empirical trade literature to deal with zero-value observations (Calderón et al. 2007). Nonetheless, we also use a Poisson Pseudo Maximum Likelihood (PPML) estimator as an alternative approach to deal with zero-value observations as part of the Robustness Analysis section.

[10]According to the World Bank Doing Business dataset, Uzbekistan had the highest costs to import with 19,881 U.S. dollars per container in 2006, but these import costs have declined over time. In 2015, this country had import costs of approximately 6,500 U.S. dollars per container.

is the less efficient country requiring 117 days to import.

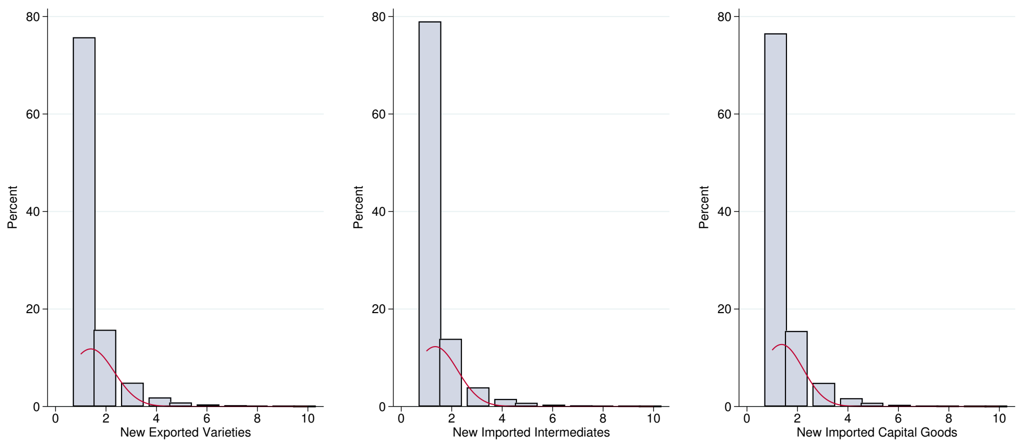
We also include trade gravity variables, which consist of bilateral distance, shared border, landlocked-status, common continent, language, and colonizer, as well as free trade agreement status. Finally, we also use two instrumental variables as part of our study: applied import tariffs for intermediate inputs and applied import tariffs for capital goods; these instruments are incorporated in our two-stage regressions at the end of the Robustness Analysis section.

4.4.4 Industry Distribution of New Traded Varieties

Consistent with Table 4.2, the telecommunications industry (i.e., 8517) registered the highest frequency of 46 new varieties exported to the United States in 2007. From the imports side, the autoparts industry (i.e., 8707) recorded the highest frequency of 44 new intermediate inputs imported from India in 2007. Furthermore, the telecommunications industry (i.e., 8517) registered again the highest frequency of 34 new capital goods imported from China in 2007 during the examined period.[11]

Figure 4.1 displays the distribution of industries trading new varieties over the examined period 2005-2016. We truncated the scale of the horizontal axis to the ten new varieties with the largest concentration of industries. On the one hand, we can observe that about 75% of industries exported only one new variety. On the other hand, we can notice that roughly 79% of industries imported only one new intermediate input, while 76% of industries imported only one new capital good.

Figure 4.1: Industry Distribution of New Traded Varieties



Notes: The left figure exhibits the industry distribution of new exported varieties. The middle figure displays the industry distribution of new imported intermediate inputs. The right figure presents the industry distribution of new imported capital goods.

Overall, we can observe that these three distributions are right-skewed, resembling binomial

[11] Industry 8517 corresponds to "telephone sets, including telephones for cellular networks or for other wireless networks; other apparatus for the transmission or reception of voice, images or other data, including apparatus for communication in a wired or wireless network". On the other hand, industry 8708 refers to "parts and accessories of the motor vehicles of headings 87.01 to 87.05; these motor vehicles comprise tractors; vehicles for the transport of ten or more persons; station wagons and racing cars; and vehicles for the transport of goods, lorries, fire fighting vehicles, concrete-mixer lorries, among others".

distributions, with most of the observations concentrated in the first four new varieties. The shape of our distributions are consistent to the distribution of new exported and imported products for Sweden (Castellani & Fassio 2019); nonetheless, Swedish firms trade a larger amount of new varieties, and consequently, these distributions are more spread out.

4.4.5 New Traded Varieties by Sector

Table 4.3 exhibits the number of new traded varieties aggregated at the sector-level of the HS classification for the year 2016.[12] What stands out in the table is that a large amount of imports are concentrated in new intermediate inputs rather than in new capital goods. A plausible explanation is that Mexico is intensively involved in global value chains (GVCs).[13] More precisely, the country tends to focus on the last stages of the production chain (e.g., in assembly activities). An exception here is new varieties in the machinery and electrical sector, where the imports of new capital goods surpass new intermediate inputs. Again, this can be explained by Mexico's participation in global value chains, where machinery is needed.

Table 4.3: Number of New Traded Varieties by Sector

Group Sector	New Exported Varieties	New Imported Intermediate Inputs	New Imported Capital Goods
25–27 Mineral Products	111	131	0
28–38 Chemicals and Allied Industries	813	839	0
39–40 Plastics and Rubbers	462	384	0
41–43 Raw Hides, Skins, Leather, and Furs	70	44	0
44–49 Wood and Wood Products	292	303	0
50–63 Textiles	1,900	842	7
64–67 Footwear and Headgear	887	13	0
68–71 Stone and Glass	252	251	0
72–83 Metals	1,009	752	83
84–85 Machinery and Electrical	1,756	811	952
86–89 Transportation	296	178	42
90–97 Miscellaneous	610	144	238
Total	8,458	4,692	1,322

Notes: This table displays the number of new traded varieties by sector. The reference year is 2016.

4.4.6 Source and Destination Countries of New Varieties

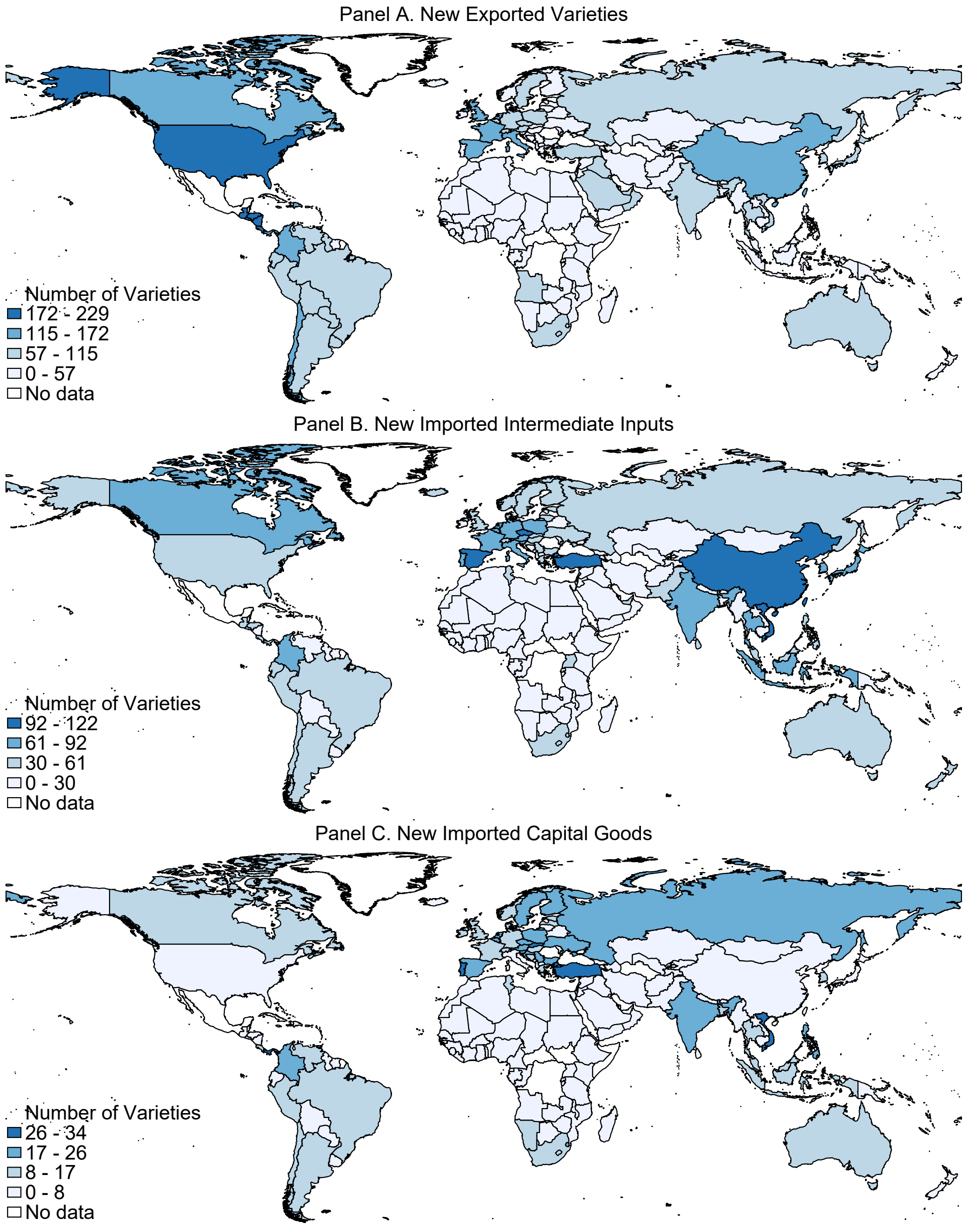
In Figure 4.2, we present a spatial visualization of Mexico's main trade partners for new manufacturing varieties in 2016. The color range displayed in these maps is determined by utilizing an equal interval classification. The aim of these maps is to clearly show the top destination countries for new exported varieties, as well as the main source countries of new imported intermediate inputs and new imported capital goods.

In Panel A, we can observe that the main destination countries of new exported varieties are the United States (229 varieties), Nicaragua (226 varieties), Costa Rica (205 varieties), El Salvador (202 varieties), and Honduras (199 varieties). Spain (170 varieties), Canada (170 varieties), and China (140 varieties) occupy the 8th, 9th, and 15th positions, respectively. We can infer from this map that most of the new varieties are exported to countries located in

[12] In this chapter, we define sectors as sections of the Harmonized System classification.

[13] The OECD defines global value chains as the fragmentation of the production chain into different stages across different countries.

Figure 4.2: Source and Destination Countries of New Traded Varieties



Notes: Panel A exhibits the frequency of Mexican exports of new manufacturing goods to the different destination countries. Panel B shows the frequency of Mexican imports of new manufacturing intermediate inputs from source countries. Panel C presents the frequency of Mexican imports of new manufacturing capital goods from source countries. The reference year for these maps is 2016.

the same continent; this implies that geographical distance matters, as well as having a free trade agreement (i.e., the United States-Mexico-Canada Agreement and the Central America Agreement).[14]

In Panel B, we now look at the main source countries of new imported intermediate inputs; these are Turkey (122 varieties), Czech Republic (101 varieties), China (96 varieties), Vietnam (92 varieties), and Spain (92 varieties). Something that may appear striking is that the United States (59 varieties) is not among the main source countries of new imported intermediate inputs. Nonetheless, this could be explained by an increasing participation of developing

[14]The member states of the Central American Agreement are Costa Rica, El Salvador, Guatemala, Honduras, Mexico, and Nicaragua.

countries in the different stages involving global value chains (OCDE 2013).

Finally, Panel C shows the main source countries of new imported capital goods; these are Portugal (34 varieties), Turkey (33 varieties), Panama (32 varieties), Slovak Republic (30 varieties), and Vietnam (28 varieties). Surprisingly, Spain (22 varieties), the USA (22 varieties), and China (5 varieties) are not among the top source countries of new imported capital goods. Again, a plausible explanation could be the increasing participation of developing countries in GVCs.

4.5 Methodology

The methodology in this study consists of four different empirical approaches: fixed effects negative binomial model, fixed effects logit model, linear fixed effect model, and log-first difference estimator with fixed effects. All these approaches include fixed effects to exploit the characteristics of our panel data covering the period 2005-2016. This section presents in detail each of these empirical approaches.

4.5.1 Fixed Effects Negative Binomial Model

First, we start by examining the impact of importing new intermediate inputs and new capital goods on the number of new varieties exported by Mexico. The best empirical approach to examine this relationship is by using a negative binomial model with fixed effects. The reason for selecting this approach is because the dependent variable is a count variable exhibiting overdispersion around the mean; this means that the variance of the dependent variable is larger than its mean.[15] In other words, the dependent variable follows a discrete distribution where the sample is concentrated on just a few values (see Figure 4.1). Thus, our first baseline regression is defined as follows:

$$X\_NEW_{cit} = \alpha + \beta_1 \, ln(IM\_NEW^{int}_{cit-1}) \; + \beta_2 \, ln(IM\_NEW^{cap}_{cit-1}) + \gamma X_{ct-1} + \nu_c + \nu_i + \nu_t + \varepsilon_{cit}, \tag{4.9}$$

where X_NEW_cit denotes the number of new exported varieties by industry i to country c in year t. Compared to the previous empirical chapter, where the regression equation was inspired in Castellani & Fassio (2019), we now distinguish between new imported intermediate inputs and new imported capital goods. Thus, Eq.(4.9) includes two main explanatory variables: $IM\_NEW^{int}_{cit-1}$ representing the number of new imported intermediate inputs by industry i from country c in year $t-1$, and $IM\_NEW^{cap}_{cit-1}$ denoting the number of new imported capital goods by industry i from country c in the previous year $t-1$. We also include control variables in X_{ct-1}; these controls are GDP at PPP international dollars of partner countries (proxies

[15] An alternative methodology would be a Poisson model with fixed effects, which assumes that the mean of the distribution equals the variance. We include this alternative methodology in the Robustness Analysis section. Nonetheless, the negative binomial model seems to be a better empirical approach for our sample as it relaxes this equality assumption. A drawback for both negative binomial and Poisson models is that observations for time-invariant variables are dropped when we include fixed effects.

the market size) and score of starting a business (proxies the ease of doing business). All explanatory variables are expressed in natural logs and lagged by one year to avoid potential endogeneity. Finally, we also include a set of country (ν_c), industry (ν_i), and year (ν_t) fixed effects.[16]

Regarding our control variables, GDP is considered standard in the trade literature (see, for example, Jansen & Stokman (2014) and Navas et al. (2020)). On the other hand, variables extracted from the World Bank Doing Business dataset are also standard. For example, Navas et al. (2020) employ three measurements to proxy market costs, which include the number of documents to import, costs to import in U.S. dollars per container deflated, and time to import in days. A limitation of these three measures is the short time span of the available data encompassing the period 2006-2015. Thus, we use instead the score of starting a business, which has a longer time span running from 2004 to 2020. Nonetheless, we also use the three trade-related variables mentioned above (i.e., number of documents to import, costs to import, and time to import) in the Robustness Analysis section.

4.5.2 Fixed Effects Logit Model

The next step is to measure the impact of new imported varieties on new export varieties at the extensive and intensive margins. We start by focusing on the extensive margin. In the previous empirical chapter, we examined the relationship between imports of new varieties on the probability of exporting new varieties using a logit model with fixed effects; this specification is similar to one presented in the Robustness Analysis section in Castellani & Fassio (2019).

Having considered the relationship between new imported varieties and the probability of exporting new varieties, we are now interested in disentangling the effects of importing new intermediate inputs and new capital goods on the probability of exporting new varieties. To do so, we use a logit model with fixed effects as the dependent variable is a binary variable. Therefore, our next baseline specification for the extensive margin is presented as follows:

$$Prob(X\_NEW_{cit} > 0) = \beta_1\, ln(IM\_NEW^{int}_{cit-1}) + \beta_2\, ln(IM\_NEW^{cap}_{cit-1}) + \gamma X_{ct-1} + \nu_c + \nu_i + \nu_t + \varepsilon_{cit}, \quad (4.10)$$

Where $Prob(X\_NEW_{cit} > 0)$ is our dependent variable that represents the probability of exporting new varieties by industry i to country c in year t (i.e., the extensive margin). This dummy variable equals one if at least one new variety is exported to a country, and zero otherwise. We have now a set of two main explanatory variables in Eq.(4.10): $IM\_NEW^{int}_{cit-1}$ representing the number of new imported intermediate inputs by industry i from country c in the previous year $t-1$, and $IM\_NEW^{cap}_{cit-1}$ denoting the number of new imported capital goods by industry i from country c in the previous year $t-1$. We include a vector of control variables, X_{ct-1}, that includes GDP and score of starting a business. We also add a set of country (ν_c),

[16] We declare the dataset to be a panel by establishing country-industry paired observations as the cross-section observations, and year as the time series observations. Thus, including both country and industry fixed effects in our regression equations is equivalent to including variety fixed effects.

industry (ν_i), and year (ν_t) fixed effects to this specification. All independent variables are expressed in their logarithmic form and lagged by one year.

4.5.3 Linear Fixed Effects Model

Building on from the idea of disentangling between the extensive and intensive margins, we now proceed to analyze the intensive margin. We define the intensive margin as the export value of new varieties, which is in line with Navas et al. (2020). We now examine the relationship of new imported intermediate inputs and new imported capital goods on the export value of new varieties. We use a linear regression model with fixed effects as the dependent variable is a continuous variable. The baseline regression for the intensive margin is the following:

$$ln(IntMarg\_V_{cit}) = \alpha + \beta_1\, ln(IM\_NEW^{int}_{cit-1}) + \beta_2\, ln(IM\_NEW^{cap}_{cit-1}) + \gamma X_{ct-1} + \nu_c + \nu_i + \nu_t + \varepsilon_{cit}, \tag{4.11}$$

where $IntMarg\_V_{cit}$ corresponds to the dependent variable that is defined as the log value (in U.S. dollars) of exports of new varieties by industry i to country c at time t.[17]

Similar to the extensive margin specification, we have two main explanatory variables in Eq.(4.11). On the one hand, we have $IM\_NEW^{int}_{cit-1}$ corresponding to the number of new imported intermediate inputs by industry i from country c in the previous year $t-1$. On the other hand, we have $IM\_NEW^{cap}_{cit-1}$ denoting the number of new imported capital goods by industry i from country c in the previous year $t-1$. We also include two control variables in X_{ct-1}: GDP and score of starting a business. Moreover, we include a set of country (ν_c), industry (ν_i), and year (ν_t) fixed effects. Once again, all the right-hand side variables are expressed in their logarithmic form and lagged by one year.

4.5.4 Log-First Difference Estimator

Finally, we explore the impact of the net change in the log number of new imported intermediate inputs and new imported capital goods on the net change in the log number of new exported varieties. To examine this relationship, we use a log-first difference estimator with fixed effects along the lines of Damijan et al. (2014). This last baseline regression is defined as follows:

$$\Delta\, ln(X\_NEW_{cit}) = \alpha + \beta_1 \Delta ln(IM\_NEW^{int}_{cit}) + \beta_2 \Delta ln(IM\_NEW^{cap}_{cit}) + \delta X_{ct} + \nu_c + \nu_i + \nu_t + \varepsilon_{cit}, \tag{4.12}$$

Where $\Delta ln(X\_NEW_{cit})$ stands for the net change in the log number of new exported varieties by industry i to country c at time t. Eq.(4.12) contains two main explanatory variables also expressed in logs and as first differences. First, we have $\Delta ln(IM\_NEW^{int}_{cit})$, which represents the net change in the log number of new imported intermediate inputs by industry i

[17] It is worth mentioning that the difference between the dependent variable in Eq.(4.9) and Eq.(4.11). On the one hand, X_NEW_cit in Eq.(4.9) is a count variable representing the number of new exported varieties. On the other hand, $IntMarg\_V_{cit}$ in Eq.(4.11) is a continuous variable representing the export value (in U.S. dollars) of new varieties.

from country c in time t. This equation also contains $\Delta ln(IM\_NEW^{cap}_{cit})$, which corresponds to the net change in the log number of new imported capital goods by industry i from country c in time t. The regression specification also incorporates controls for GDP and the score of starting a business. A full set of country (ν_c), industry (ν_i), and year (ν_t) fixed effects are also included. All variables given in X_{ct-1} are expressed in their log form.

4.6 Results

The section below presents the results from our four empirical approaches. We start by employing a negative binomial model with fixed effects to estimate the impact of importing new intermediate inputs and new capital goods on the number of new exported varieties. Next, we evaluate the effects of new intermediate and new capital imports at the extensive and intensive margins. The extensive margin is defined as the probability of exporting new varieties, which is a binary dependent variable; thus, the selected empirical approach consists of a logit model with fixed effects. On the other hand, the intensive margin is defined as the export value of new varieties, where we use a linear regression model with fixed effects. Finally, we use a log-first difference estimator with fixed effects to examine the impact of the net change in new imported intermediates and new capital goods on the net change of new exported varieties. In summary, this section reveals the importance of new imports of both intermediate inputs and capital goods on exports of new varieties.

4.6.1 Number of New Exported Varieties

We begin by estimating the effects of new imported intermediate inputs and new imported capital goods on the number of new exported varieties. For this first baseline specification, we use a negative binomial model with fixed effects as this is the most appropriate methodology for a count dependent variable with overdispersion around the mean; this was previously discussed in the Methodology section. The estimation sample is constituted by 68,727 new varieties over the period 2005-2016. All the right-hand side variables are expressed in their logarithmic form and lagged by one year.

Table 4.4 reports the coefficients after estimating Eq.(4.9) using a fixed effects negative binomial model. In columns (1)-(2), we start by estimating negative binomial regressions using a cross-section approach.[18] The idea behind this exercise is to include an overdispersion parameter α. If this dispersion parameter equals zero, then a Poisson model would be a better approach. On the other hand, if this dispersion parameter is significantly greater than zero, this means that the data presents overdispersion. We can observe that *alpha* is larger than zero, which confirms that our data presents overdispersion; thus, a negative binomial model constitutes a more appropriate approach for our specification.

[18] The number of varieties is not reported in columns (1)-(2) since the unit of observation in the cross-section specifications is not industry-country paired observations (i.e., varieties) as in the panel specifications.

Table 4.4: Number of New Exported Varieties

VARIABLES	(1) Cross-sec X_NEW	(2) Cross-sec X_NEW	(3) Panel X_NEW	(4) Panel X_NEW	(5) Panel X_NEW	(6) Panel X_NEW	(7) Panel X_NEW	(8) Panel X_NEW
ln(No. New Imported Intermediate Inputs)_cit-1	0.5104***		0.0616***		0.0621***	0.0592***	0.0591***	0.0571***
	(0.0124)		(0.0137)		(0.0137)	(0.0137)	(0.0137)	(0.0137)
ln(No. New Imported Capital Goods)_cit-1		0.7242***		0.0228	0.0256	0.0230	0.0220	0.0202
		(0.0213)		(0.0204)	(0.0205)	(0.0205)	(0.0204)	(0.0205)
ln(GDP in PPP)_ct-1						0.0171***		0.0122**
						(0.0056)		(0.0056)
ln(Starting a Business)_ct-1							0.2800***	0.2749***
							(0.0266)	(0.0267)
Constant	-1.8910***	-1.8702***	-0.4743***	-0.4666***	-0.4762***	-0.9212***	-1.6295***	-1.9254***
	(0.0038)	(0.0038)	(0.0157)	(0.0157)	(0.0158)	(0.1450)	(0.1108)	(0.1750)
Observations	824,724	824,724	543,084	543,084	543,084	543,084	543,084	543,084
Number of varieties			45,257	45,257	45,257	45,257	45,257	45,257
Alpha	4.851	4.893						
Prob Wald Chi2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Country FE	NO	NO	YES	YES	YES	YES	YES	YES
Industry FE	NO	NO	YES	YES	YES	YES	YES	YES
Year FE	NO	NO	YES	YES	YES	YES	YES	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table constitutes the baseline to examine the number of new exported varieties by industry i to country c in time t. The table reports the coefficients of Eq.(4.9) employing a fixed effects negative binomial model. The estimation sample is conformed by 68,727 new varieties over the period 2005-2016; however, some observations were dropped from the sample due to the nature of the methodology employed. All independent variables are expressed in natural logs and lagged by one year. All panel regressions include country, industry, and year fixed effects.

In the following columns, we continue to use a negative binomial model now exploiting the panel structure of our database; thus, we include a full set of country, industry, and year fixed effects.[19] All regressions report the coefficients. In column (3), we only include one of our main explanatory variables (i.e., number of new imported intermediate inputs) and the full set of country, industry, and year fixed effects. We can notice the number of new imported intermediate input has a positive and strong statistical effect on the number of new exported varieties. In column (4), we now include only the other main explanatory variable, along with the full set of country, industry, and year fixed effects. We can notice that the number of new imported capital goods is insignificant in this specification.

In column (5), we introduce both main explanatory variables, and we can observe that our results remain similar to the coefficients presented individually in the previous columns. In the following columns, we introduce the control variables one by one. In column (6), we add GDP in our specification; we can notice that the magnitude and direction of the coefficients of our two main explanatory variables do not change substantially. In column (7), we now introduce the score of starting a business as our unique control variable; we can see that the coefficients of our main explanatory variables look similar to column (5).

Finally, we present our baseline specification in column (8), which includes both main explanatory variables, the control variables, and the full set of country, industry, and year fixed effects. We can observe that our results do not change dramatically once all the right-hand side variables are included. In other words, the variable for the number of new imported intermediate inputs remains positive and strongly significant. By contrast, the number of new imported capital goods remains statistically insignificant. The control variables (i.e., GDP and

[19] The overdispersion parameter α is not displayed for the panel specifications because Stata does not possess an option to include this parameter using a fixed effects negative binomial command.

score of starting a business) are both positive and significant. Our findings suggest that a 1% increase in new imported intermediate inputs is associated with an increase in the number of new exported varieties by about 0.0006, ceteris paribus.

As part of the tests reported in Table 4.4, we include the Wald Chi-square test, which is used to examine whether the explanatory variables included in the model are significantly improving the fit of the model. The null hypothesis is that the coefficients for the explanatory variables are simultaneously equal to zero. The table reports the p-values associated with the Wald test and we can conclude that we can reject the null hypothesis; thus, we can interpret that including these variables significantly improves the fit of the model.

4.6.2 Extensive Margin: Probability to Export New Varieties

Turning now to the extensive margin, we want to measure the probability of exporting new varieties, $Prob(X\_NEW_{cit} > 0)$, as a function of importing new intermediate inputs and new capital goods. Therefore, we use a logit model with fixed effects to estimate Eq.(4.10). The dependent variable, $Prob(X\_NEW_{cit} > 0)$, is a dummy variable that equals one if at least one new variety is exported by industry i to country c in year t; and zero, otherwise. The estimation sample is composed of 68,727 new varieties over the period 2005-2016. All the right-hand side variables are expressed in logs and lagged by one year. Also, all the specifications include a full set of country, industry, and year fixed effects.[20] As part of the tests, we also include the p-values of the Wald Chi-square test; we can interpret that including these explanatory variables significantly improve the fit of the model.

Table 4.5: Extensive Margin: Probability to Export New Varieties

VARIABLES	(1) Prob_X_NEW	(2) Prob_X_NEW	(3) Prob_X_NEW	(4) Prob_X_NEW	(5) Prob_X_NEW	(6) Prob_X_NEW
ln(No. New Imported Intermediate Inputs)_cit-1	0.0139***		0.0141***	0.0001***	0.0093***	0.0002***
	(0.0045)		(0.0045)	(0.0001)	(0.0030)	(0.0001)
ln(No. New Imported Capital Goods)_cit-1		0.0134**	0.0139**	0.0001**	0.0089**	0.0002**
		(0.0067)	(0.0270)	(0.0270)	(0.0270)	(0.0270)
ln(GDP in PPP)_ct-1				0.0003***		0.0007***
				(0.0436)		(0.0450)
ln(Starting a Business)_ct-1					0.0553***	0.0012***
					(0.0346)	(0.0357)
Observations	542,964	542,964	542,964	542,964	542,964	542,964
Number of varieties	45,247	45,247	45,247	45,247	45,247	45,247
Prob Wald Chi2	0.000	0.000	0.000	0.000	0.000	0.000
Country FE	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table constitutes the baseline to analyze the probability of exporting new varieties by industry i to country c in time t (i.e., the extensive margin). The table reports the average marginal effects of Eq.(4.10) using a fixed effects logit model. The estimation sample is conformed by 68,727 new varieties over the period 2005-2016; however, some observations were dropped from the sample due to the nature of the methodology employed. Thus, the number of observations differs from the Summary Statistics table because the likelihood function is only identified from switchers (i.e., 0 to 1, or 1 to 0); therefore, observations that remain always 1 or always 0 do not contribute to the likelihood function (i.e., log 1 = 0 and log 0 is undefined). All independent variables are expressed in natural logs and lagged by one year. All regressions include country, industry, and year fixed effects.

[20] The number of observations differs to that of Table 4.4 due to the inclusion of fixed effects in a logit model, where the log likelihood function is only identified through changes in the dependent variable (i.e., switchers from 0 to 1, or 1 to 0).

Table 4.5 reports the average marginal effects of the logit model with fixed effects. Column (1) constitutes the starting point, where we only include the number of new imported intermediate inputs as one of our main explanatory variables, along with the full set of country, industry, and year fixed effects. The results suggest that new imported intermediate inputs have a positive and strong statistical effect on the probability of exporting new varieties.

In column (2), we only include the number of new imported capital goods, together with the full set of country, industry, and year fixed effects. We can notice that this variable of interest is positive and statistically significant in this specification. Compared to Table 4.4, this result implies that imports of new capital goods play an important role on the probability of exporting new varieties.

In column (3), we now include both explanatory variables; we can observe that our results are similar to those in the previous columns; nonetheless, it is worth noting an increase in the significance level of importing new capital goods. Like before, the next columns introduce the control variables stepwise. In column (4), we include GDP in our specification, which does not dramatically alter our results. In column (5), we add the score of starting a business as another control; we can observe that the size of our main explanatory variables remains similar to the previous columns.

Our baseline specification is presented in column (6); this specification includes both explanatory variables, controls, and a full set of country, industry, and year fixed effects. We can notice that our results hold throughout these specifications. Both new imported intermediate inputs and new capital goods display positive and statistically significant outcomes. Moreover, GDP and score of starting a business are also positive and statistically significant.

The results suggest that a 10% increase in new imported intermediate inputs increases the probability of exporting new varieties by 0.002 percentage points, ceteris paribus. We find a similar effect in terms of the magnitude when considering new imported capital goods. From our estimation sample, about 11% of exports were recorded as new varieties. Therefore, our results are statistically significant, albeit not economically meaningful.[21] Despite this non economically significant outcome at the country level, it may be interesting to perform the analysis at the state level due to heterogeneity shown in Chapter 2.[22]

It is also worth mentioning that the effects of these explanatory variables fall substantially when we include the control variables (i.e., GDP and score of starting a business). These results support Aristei et al. (2013), who found that imports have a positive impact on the probability of exporting. Our results are also aligned to Lo Turco & Maggioni (2013), who suggest that importing from low income countries has a strong and positive impact on the probability to start exporting. This chapter complements these results as we now disentangle imports into intermediate inputs and capital goods.

[21]For example, if there is a 10% increase in new imported intermediate inputs, this give us 0.112. Thus, 100*((0.112-0.11)/0.11) = 1.82%. This suggests there is a 1.82% increase in exports of new varieties in our sample.

[22]We could not perform this additional analysis at the state level because the National Institute of Statistics and Geography (INEGI) does not report imports at the state level due to confidentiality issues.

4.6.3 Intensive Margin: Export Value of New Varieties

We now proceed to investigate the intensive margin, which aim is to analyze the impact of importing new intermediate inputs and new capital goods on the export value of new varieties. To perform this analysis, we employ a linear fixed effects model to estimate Eq.(4.11). As a recap, the estimation sample is comprised of 68,727 new varieties over the period 2005-2016. All the right-hand side variables are expressed in logs and lagged by one year to avoid potential endogeneity. Moreover, all the specifications include a full set of country, industry, and year fixed effects.

Table 4.6: Intensive Margin: Export Value of New Varieties

VARIABLES	(1) IntMarg_V	(2) IntMarg_V	(3) IntMarg_V	(4) IntMarg_V	(5) IntMarg_V	(6) IntMarg_V
ln(No. New Imported Intermediate Inputs)_cit-1	0.0707***		0.0716***	0.0711***	0.0704***	0.0702***
	(0.0145)		(0.0145)	(0.0145)	(0.0145)	(0.0145)
ln(No. New Imported Capital Goods)_cit-1		0.1010***	0.1029***	0.1020***	0.1013***	0.1009***
		(0.0316)	(0.0316)	(0.0316)	(0.0316)	(0.0316)
ln(GDP in PPP)_ct-1				0.1236***		0.0612***
				(0.0228)		(0.0230)
ln(Starting a Business)_ct-1					0.2605***	0.2487***
					(0.0208)	(0.0212)
Constant	0.8412***	0.8444***	0.8378***	-2.3484***	-0.2358***	-1.7651***
	(0.0089)	(0.0088)	(0.0089)	(0.5872)	(0.0860)	(0.5829)
Observations	824,724	824,724	824,724	824,724	824,724	824,724
Number of varieties	68,727	68,727	68,727	68,727	68,727	68,727
R-squared	0.006	0.006	0.006	0.006	0.006	0.006
Country FE	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

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

Notes: This table constitutes the baseline to examine the log of value of exports (in U.S. dollars) of new varieties by industry i to country c at time t (i.e., the intensive margin); this measure of the intensive margin is consistent to Navas et al. (2020). This table estimates Eq.(4.11) using a linear fixed effects model. The estimation sample is conformed by 68,727 new varieties over the period 2005-2016. All independent variables are expressed in natural logs and lagged by one year. All regressions include country, industry, and year fixed effects.

Table 4.6 provides the results of these linear regressions with fixed effects. Column (1) only includes the number of new imported intermediate inputs, along with a full set of country, industry, and year fixed effects. We can observe that the coefficient of interest is positive and strongly significant. This result is consistent with Feng et al. (2016), who also found a positive relationship between firm's imported intermediate inputs and firm's export value. In column (2), we include only the number of new imported capital goods, as well as the whole set of country, industry, and year fixed effects. We can also notice that this coefficient is positive and strongly significant, which is in line with Table 4.5.

In column (3), we now introduce both variables of interest and we can observe that our results do not dramatically change to those displayed in the previous columns. Then, we introduce the control variables one by one. In column (4), we can notice that the magnitude and significance level of our two main explanatory variables hold after including GDP. In column (5), we now add the score of starting a business as our only control; we can also notice that our coefficients of interest hold.

Column (6) represents our baseline specification, where both explanatory variables and control variables are included, as well as the full set of country, industry, and year fixed effects. We can notice that both imports of new intermediate inputs and new capital goods, as well as GDP and score of starting a business, report positive and statistically strong coefficients. These results suggest that a 1% increase in new imported intermediate inputs may increase the export value of new varieties by about 0.001. Also, a 1% increase in new imported capital goods may lead to an increase in the export value of new varieties of about 0.001. An important contribution in this chapter is that we shed light on the role of new imported capital goods, which have been somehow neglected in the trade literature.

4.6.4 Net Change in the Number of New Exported Varieties

What follows is an examination of the impact of the net change in new imported intermediate inputs and in new imported capital goods on the net change in new exported varieties. We use a log-first difference estimator with fixed effects. The estimation sample is formed by 68,727 new varieties over the period 2005-2016. The dependent variable and main explanatory variables are expressed as first differences of the logged variables. Also, all the regressions include a full set of country, industry, and year fixed effects. In this specification, the coefficients can be interpreted as elasticities.

Table 4.7 shows the results of estimating Eq.(4.12) using a log-first difference estimator with fixed effects. In column (1), we only include the net change in the number of new imported intermediate inputs with the full set of country, industry, and year fixed effects. We can notice that the coefficient is positive and strongly significant. In column (2), we evaluate the net change in the number of new imported capital goods with the full set of country, industry, and year fixed effects. Similarly, the coefficient is positive and strongly significant.

In column (3), we can observe that the net changes in both the number of new imported intermediate inputs and new imported capital goods have positive and strong significant effects on the net change of the number of new exported varieties. These coefficients remain stable compared to the previous columns, where we separately explored the effects of these two main explanatory variables. Next, we introduce our controls in a stepwise manner. In column (4), we add GDP in our specification, and we can observe that our coefficients remain stable. In column (5), we include the score of starting a business as our sole control variable, and we can also notice that our coefficients of interest hold.

In column (6), we present our baseline results for this specification that includes all the explanatory variables and control variables, as well as the full set of country, industry, and year fixed effects. We can observe that both explanatory variables have positive and statistically significant effects on the net change of the number of new exported varieties. Our results suggest that a 1% increase in the net change of new imported intermediate inputs can lead to an increase in the net change of the number of new exported varieties by about 0.16%. Also, a 1% increase in the net change of new imported capital goods can lead to an increase in the net change of the number of new exported varieties by about 0.23%.

Table 4.7: Net Change in the Number of New Exported Varieties

VARIABLES	(1) X_NEW_D1	(2) X_NEW_D1	(3) X_NEW_D1	(4) X_NEW_D1	(5) X_NEW_D1	(6) X_NEW_D1
Δln(No. New Imported Intermediates)_cit	0.1528***		0.1548***	0.1548***	0.1548***	0.1548***
	(0.0026)		(0.0026)	(0.0026)	(0.0026)	(0.0026)
Δln(No. New Imported Capital Goods)_cit		0.2277***	0.2317***	0.2317***	0.2317***	0.2317***
		(0.0054)	(0.0054)	(0.0054)	(0.0054)	(0.0054)
ln(GDP in PPP)_ct-1				-0.0110***		-0.0108***
				(0.0021)		(0.0022)
ln(Starting a Business)_ct-1					-0.0031	-0.0011
					(0.0020)	(0.0021)
Constant	-0.0192***	-0.0212***	-0.0163***	0.2683***	-0.0034	0.2658***
	(0.0014)	(0.0014)	(0.0014)	(0.0547)	(0.0084)	(0.0547)
Observations	824,724	824,724	824,724	824,724	824,724	824,724
Number of varieties	68,727	68,727	68,727	68,727	68,727	68,727
R-squared	0.017	0.013	0.028	0.028	0.028	0.028
Country FE	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

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

Notes: This table constitutes the baseline to explore the net change in the log number of new exported varieties by industry i to country c in time t. The table reports the coefficients of Eq.(4.12) employing a log-first difference estimator with fixed effects. The estimation sample is conformed by 68,727 new varieties over the period 2005-2016. The main explanatory variables are expressed as log-first differences, while control variables are expressed in natural logs and lagged by one year. All regressions include country, industry, and year fixed effects.

An important conclusion to highlight is that these results imply that new imported capital goods have a slightly larger effect compared to new imported intermediate inputs on the net change in new exported varieties. These findings are in line with Damijan et al. (2014) for Slovenia, who provide evidence that the net change of imported intermediate inputs and capital goods have positive and significant effects on the net change of exports; although, the authors show that imported intermediate inputs have a larger effect compared to imported capital goods for Slovenia. In this chapter, we now take one step further, and we extend the analysis by focusing on new varieties only. In contrast to Damijan et al. (2014), we provide empirical evidence that the effects of the net change of new imported capital goods are larger than the net change of new imported intermediate inputs on the net change of exports of new varieties.

4.7 Robustness Analysis

In the section that follows, we present a series of robustness checks of the different empirical approaches employed in this chapter. The structure of most of the tables under this section is as follows: the first column exhibits the coefficients using a fixed effects negative binomial model; the second column reports the average marginal effects employing a fixed effects logit model; the third column displays the coefficients using a linear fixed effects model; and the fourth column shows the results applying a log-first difference estimator with fixed effects.

Moving on to the robustness checks, we begin this section by including trade gravity variables to our empirical specifications. Next, we exclude the United States and Spain from our estimation sample. Then, we divide our estimation sample into two country sub-samples based on income groups (i.e., high-income OECD countries and low- and middle-income countries). Then, we focus on the top industries trading new varieties from our estimation sample. After that, we also use other trade-related control variables in our specifications. Later, we employ an alternative methodology to the fixed effects negative binomial model. Next, we use an alternative approach to deal with zero-value observations in the dependent variable. Then, we try a log-log model. Then, we also look at the contemporaneous effects of importing new varieties. After that, we increase the lag length of the independent variables. This is followed by using different combinations of fixed effects. Later, we analyze input-output linkages across sectors. After that, we examine the marginal effects by industry. Finally, we use two-stage regressions to tackle potential endogeneity.

4.7.1 Trade Gravity Variables

The first robustness check consists of including standard trade gravity variables to our baseline specifications. These gravity variables comprise bilateral distance, shared border, free trade agreement status, landlocked-status, common continent, common language, and common colonizer.[23] These last two gravity variables invoke shared historical and cultural linkages.

All the independent variables are expressed in their natural logarithmic form and lagged by one year. As these gravity variables only vary across countries and are time invariant, we cannot include country fixed effects; otherwise, these observations would drop from our estimation sample. Despite this limitation, we include industry and year fixed effects in all the specifications. It is worth commenting that the number of observations differs in the first two columns; the reason is that negative binomial and logit models combined with fixed effects drop observations for time-invariant variables. On top of this, the logit model drops observations that do not contribute to the likelihood function.[24]

Table 4.8 provides the results after including trade gravity variables in our models. Under these specifications, we can now observe that our two main explanatory variables (i.e., new imported intermediate inputs and capital goods) are positive and statistically significant through all four specifications. It is worth noticing that the magnitude of the coefficients are now larger in all the specifications compared to the baseline regressions presented in Tables 4.4-4.6.

Moving on to the trade gravity variables, we can infer that partner countries located in the same continent, as well as countries where Spanish constitutes the official language, have both positive and strong effects on exporting new varieties (i.e., on number of varieties, probability, and export value). A plausible explanation conditional to distance is that being in the same

[23]Bilateral distance is computed as the geographical log distance between Mexico City and each capital city using the great-circle distance formula; this variable is a standard proxy for transportation costs. Border refers to sharing a border; in this case, Mexico shares borders with the United States, Guatemala, and Belize. Free trade agreement status means that both countries are part of a bilateral or multilateral free trade agreement; this dummy variable equals one if the partner country has a free trade agreement with Mexico; and zero, otherwise. Continent refers to the geographical location of the partner country being in the Americas. Language means that Spanish is the partner country's official language. Finally, colonizer refers to being a former colony of Spain.

[24]The likelihood function is only identified from switchers (i.e., 0 to 1, or 1 to 0); therefore, observations that remain always 1 or always 0 do not contribute to this likelihood function (i.e., $\log 1 = 0$ and $\log 0$ is undefined).

Table 4.8: Trade Gravity Variables

VARIABLES	(1) X_NEW	(2) AME Prob_X_NEW	(3) IntMarg_V	(4) X_NEW_D1
ln(No. New Imported Intermediate Inputs)_cit-1	0.0729***	0.0102***	0.1840**	
	(0.0123)	(0.0149)	(0.0752)	
ln(No. New Imported Capital Goods)_cit-1	0.1118***	0.0165***	0.3537***	
	(0.0183)	(0.0229)	(0.0882)	
Δln(No. New Imported Intermediates)_cit				0.1603***
				(0.0014)
Δln(No. New Imported Capital Goods)_cit				0.2402***
				(0.0024)
ln(GDP in PPP)_ct-1	0.1865***	0.0177***	0.1494***	-0.0010***
	(0.0021)	(0.0023)	(0.0057)	(0.0002)
ln(Starting a Business)_ct-1	0.2642***	0.0255***	0.1995***	-0.0018
	(0.0144)	(0.0160)	(0.0163)	(0.0016)
ln(Distance)	-0.4652***	-0.0448***	-0.3345***	0.0004
	(0.0099)	(0.0115)	(0.0150)	(0.0013)
Border	-0.4186***	-0.4355***	-0.2205***	-0.0054**
	(0.0177)	(0.0211)	(0.0260)	(0.0026)
Free Trade Agreement	-0.0554***	-0.0058***	-0.0810***	-0.0014
	(0.0079)	(0.0089)	(0.0081)	(0.0009)
Landlocked	-0.0450***	-0.0047***	-0.0393***	0.0016
	(0.0123)	(0.0134)	(0.0077)	(0.0013)
Continent	0.3407***	0.0356***	0.3028***	-0.0027
	(0.0135)	(0.0158)	(0.0196)	(0.0019)
Language	0.3580***	0.0326***	0.3210***	-0.0030
	(0.0145)	(0.0170)	(0.0201)	(0.0020)
Colonizer	-0.1453***	-0.0143***	-0.0678***	-0.0009
	(0.0145)	(0.0170)	(0.0153)	(0.0020)
Constant	-2.7083***	-4.9071***	-0.9639***	0.0145
	(0.1076)	(0.1954)	(0.1717)	(0.0139)
Observations	820,740	820,740	821,160	821,160
Prob Wald Chi2	0.000	0.000		
R-squared			0.035	0.030
Country FE	NO	NO	NO	NO
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table includes the standard trade gravity variables as additional controls. The dependent variable in column (1) corresponds to the number of new exported varieties by industry i to country c in time t; this column reports the coefficients using a negative binomial with fixed effects approach. The dependent variable in column (2) is the probability of exporting new varieties by industry i to country c in time t (i.e., the extensive margin); this column reports the average marginal effects using a logit model with fixed effects. The dependent variable in column (3) is the log value of exports (in U.S. dollars) of new varieties by industry i to country c at time t (i.e., the intensive margin); this column reports the coefficients employing a linear regression with fixed effects. The dependent variable in column (4) is the net change in the log number of new exported varieties by industry i to country c in time t; this column is calculated employing a log-first difference estimator with fixed effects. All time-variant independent variables are expressed in natural logs. All regressions include industry and year fixed effects.

continent may reduce transportation costs (e.g., timing or transportation mode). Also, having a common language facilitates trade, which also includes trade in new varieties; this explanation is consistent with the trade literature (Calderón et al. 2007, Imbs 2004) and with our results in Chapter 3.

Whereas geographical distance, shared border, free trade agreement status, landlocked-status, and common colonizer have negative and statistically strong effects on exporting new varieties to those countries (i.e., on number of varieties, probability, and export value). First, we can interpret that more distant partner countries incur in higher transportation costs; thus, these higher transportation costs may hinder exports to those countries. Second, a partner country that is landlocked may also be associated with higher transportation costs; therefore, these costs may also disincentivize exports to those countries. Thus far, negative coefficients of distance (Calderón et al. 2007, Imbs 2004, Navas et al. 2020) and landlocked status (Calderón et al. 2007) are consistent with the trade literature.

Interestingly, we expect positive coefficients for sharing a border, having a free trade agreement, and having a common colonizer. Nevertheless, we can observe in Section 4.4 that source countries of new varieties are concentrated in Europe and Asia, which do not share a border; some of these countries may not even have free trade agreements with Mexico, nor have share a common heritage; an exception is Spain, which has a free trade agreement through the EU and also shares historical ties with Mexico. From a trade policy perspective, we have identified that some of the main source countries of new imported varieties are concentrated in Asia. This means that new imported intermediate inputs and capital goods sourced from these countries pay tariffs, which translates into higher costs of these goods. Therefore, we could interpret this situation as a motivation for Mexican policy makers to negotiate trade agreements with source countries located in Asia.

4.7.2 Exclusion of Main Partner Countries

Now, we proceed to analyze our four different specifications once we take the United States and Spain out of our estimation sample, separately. The reason why we drop the United States from the sample is because the U.S. represents Mexico's main trade partner. In a similar exercise, we also exclude Spain from the sample since this Iberian country has strong historical and cultural linkages with Mexico. Thus, we want to confirm that our results hold even if we exclude these countries. All regressions contain the full set of independent variables, as well as industry, country, and year fixed effects. The inconsistency of the number of observations across models can be attributed to the fact that negative binomial and logit models combined with fixed effects drop observations for time-invariant variables; in addition, the logit model drops observations for previously discussed non-switchers.[25]

Table 4.9 displays the results for the estimation sample excluding the United States and Spain, separately. The first four columns report our results after excluding the United States from the sample, while the last four columns exhibit the results after excluding Spain from the sample. As we can observe, the coefficients of the number of new imported intermediate inputs remain positive and strongly significant throughout all our specifications (i.e., on number of varieties, probability, export value, and first differences). On the other hand, the coefficients of the number of new imported capital goods are also positive and statistically significant under most of our models (i.e., on the probability, export value, and first differences).

These results are consistent with the baseline results presented in the Results section. Nonetheless, we proceed to compare the magnitude of the coefficients under these specifications in contrast to the baseline results. On the one hand, the magnitude of the coefficients in columns (1) and (5) of Table 4.9 are comparable in size to the baseline results in Table 4.4 under the fixed effects negative binomial model. Also, the size of the coefficients in columns

[25] As previously mentioned, the non-switchers refer to observations always 1 or always 0, which do not contribute to the likelihood function (i.e., log 1 = 0 and log 0 is undefined).

Table 4.9: Exclusion of Main Partner Countries

VARIABLES	(1) Excl.USA X_NEW	(2) AME Excl.USA Prob_X_NEW	(3) Excl.USA IntMarg_V	(4) Excl.USA X_NEW_D1	(5) Excl.ESP X_NEW	(6) AME Excl.ESP Prob_X_NEW	(7) Excl.ESP IntMarg_V	(8) Excl.ESP X_NEW_D1
ln(No. New Imported Intermediates)_cit-1	0.0523***	0.0005***	0.0676***		0.0574***	0.0007***	0.0710***	
	(0.0139)	(0.0002)	(0.0144)		(0.0139)	(0.0002)	(0.0145)	
ln(No. New Imported Capital Goods)_cit-1	0.0142	0.0005*	0.0958***		0.0281	0.0008**	0.1054***	
	(0.0206)	(0.0272)	(0.0314)		(0.0207)	(0.0274)	(0.0318)	
Δln(No. New Imported Intermediates)_cit				0.1495***				0.1526***
				(0.0025)				(0.0026)
Δln(No. New Imported Capital Goods)_cit				0.2275***				0.2273***
				(0.0053)				(0.0054)
ln(GDP in PPP)_ct-1	0.0346***	0.0013***	0.0435*	-0.0104***	0.0123**	0.0016***	0.0336	-0.0101***
	(0.0058)	(0.0451)	(0.0230)	(0.0022)	(0.0056)	(0.0456)	(0.0230)	(0.0022)
ln(Starting a Business)_ct-1	0.2764***	0.0025***	0.2288***	-0.0004	0.2776***	0.0042***	0.2560***	-0.0014
	(0.0269)	(0.0358)	(0.0212)	(0.0021)	(0.0268)	(0.0359)	(0.0212)	(0.0021)
Constant	-2.4906***		-1.2307**	0.2544***	-1.9403***		-1.0882*	0.2497***
	(0.1813)		(0.5799)	(0.0544)	(0.1759)		(0.5812)	(0.0544)
Observations	535,320	535,212	815,604	815,604	534,648	534,540	813,540	813,540
Number of varieties	44,610	44,601	67,967	67,967	44,554	44,545	67,795	67,795
Prob Wald Chi2	0.000	0.000			0.000	0.000		
R-squared			0.006	0.026			0.006	0.027
Country FE	YES	YES	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table reports the results after excluding the United States and Spain from the estimation sample. Columns (1)-(4) report the results after excluding only the United States from the sample, while columns (5)-(8) show the results after excluding only Spain from the sample. The dependent variable in columns (1) and (5) corresponds to the number of new exported varieties by industry i to country c in time t; these columns report the coefficients using a negative binomial with fixed effects approach. The dependent variable in columns (2) and (6) is the probability of exporting new varieties by industry i to country c in time t (i.e., the extensive margin); these columns report the average marginal effects using a logit model with fixed effects. The dependent variable in columns (3) and (7) is the log value of exports (in U.S. dollars) of new varieties by industry i to country c at time t (i.e., the intensive margin); these columns report the coefficients employing linear regressions with fixed effects. The dependent variable in columns (4) and (8) is the net change in the log number of new exported varieties by industry i to country c in time t; these columns are calculated employing a log-first difference estimator with fixed effects. All independent variables are expressed in natural logs. All regressions include country, industry, and year fixed effects.

(3) and (7) are comparable to those presented in Table 4.6 under the linear fixed effects model. Finally, the magnitude of the coefficients in columns (4) and (8) are comparable to the baseline results in Table 4.7 under the log-first difference estimator with fixed effects. On the other hand, the coefficients in columns (2) and (6) of Table 4.9 are larger compared to Table 4.5 under the fixed effects logit model. Overall, these results are in line with our baseline results presented in the Results section.

4.7.3 Income Profile

Similar to Lo Turco & Maggioni (2013), we also divide our estimation sample into two categories: high-income OECD countries and low and middle-income countries.[26] We are now interested in exploring how the income profile of partner countries shape the effects of new imported varieties (i.e., intermediate inputs and capital goods) on new export varieties. As previously mentioned, all regressions include the full set of independent variables and industry,

[26]High-income OECD countries are defined by the OECD as members with a GNI per capita income above 12,236 U.S. dollars in 2016; therefore, the OECD classifies the following countries under this category: Australia, Austria, Belgium, Canada, Chile, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, South Korea, Latvia, Lithuania, Luxembourg, the Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, United Kingdom, and the United States.

country, and year fixed effects. As previously explained, the discrepancy in the sample size across columns can be explained by the fact that negative binomial and logit models combined with fixed effects drop observations for time-invariant variables; also, the logit model drops observations that do not contribute to the likelihood function.

Table 4.10: Income Profile

VARIABLES	(1) High Income X_NEW	(2) AME High Income Prob_X_NEW	(3) High Income IntMarg_V	(4) High Income X_NEW_D1	(5) Low_Middle X_NEW	(6) AME Low_Middle Prob_X_NEW	(7) Low_Middle IntMarg_V	(8) Low_Middle X_NEW_D1
ln(No. New Imported Intermediates)_cit-1	0.0293	0.0000	0.0405*		0.0777***	0.0285***	0.0929***	
	(0.0195)	(0.0000)	(0.0218)		(0.0193)	(0.0062)	(0.0193)	
ln(No. New Imported Capital Goods)_cit-1	0.0098	0.0000	0.1142**		0.0241	0.0168*	0.0845**	
	(0.0281)	(0.0374)	(0.0478)		(0.0299)	(0.0392)	(0.0418)	
Δln(No. New Imported Intermediates)_cit				0.1662***				0.1452***
				(0.0038)				(0.0035)
Δln(No. New Imported Capital Goods)_cit				0.2549***				0.2111***
				(0.0081)				(0.0071)
ln(GDP in PPP)_ct-1	0.0285*	0.0000***	0.0829	-0.0326***	0.0121*	-0.0112	0.0168	-0.0080***
	(0.0156)	(0.1332)	(0.0826)	(0.0084)	(0.0070)	(0.0540)	(0.0253)	(0.0023)
ln(Starting a Business)_ct-1	0.1832**	0.0000***	0.3456***	-0.0117	0.2643***	0.0710***	0.2269***	0.0001
	(0.0892)	(0.1114)	(0.0714)	(0.0076)	(0.0297)	(0.0382)	(0.0225)	(0.0022)
Constant	-1.9482***		-2.9035	0.9122***	-1.8919***		-0.4893	0.1859***
	(0.5803)		(2.2719)	(0.2295)	(0.2192)		(0.6299)	(0.0563)
Observations	177,228	177,180	284,472	284,472	365,856	365,784	540,252	540,252
Number of varieties	14,769	14,765	23,706	23,706	30,488	30,482	45,021	45,021
Prob Wald Chi2	0.000	0.000			0.000	0.000		
R-squared			0.008	0.044			0.005	0.020
Country FE	YES	YES	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table reports the results for country sub-samples based on income levels. Columns (1)-(4) report the results for high-income OECD countries, while columns (5)-(8) show the results for low- and middle-income countries. The dependent variable in columns (1) and (5) corresponds to the number of new exported varieties by industry i to country c in time t; these columns report the coefficients using a negative binomial with fixed effects approach. The dependent variable in columns (2) and (6) is the probability of exporting new varieties by industry i to country c in time t (i.e., the extensive margin); these columns report the average marginal effects using a logit model with fixed effects. The dependent variable in columns (3) and (7) is the log value of exports (in U.S. dollars) of new varieties by industry i to country c at time t (i.e., the intensive margin); these columns report the coefficients employing a linear fixed effects model. The dependent variable in columns (4) and (8) is the net change in the log number of new exported varieties by industry i to country c in time t; these columns are calculated employing a log-first difference estimator with fixed effects. All independent variables are expressed in natural logs. All regressions include country, industry, and year fixed effects.

Table 4.10 shows the results for both income sub-samples. The first set of four columns corresponds to high-income OECD countries. We can observe that the coefficients of new imported intermediate inputs and new imported capital goods are statistically insignificant on the number of varieties and on the probability to export new varieties. These main explanatory variables are weakly significant on export value but strongly significant on first differences for high-income OECD countries.

The second set of four columns conforming the sub-sample on low and middle-income countries is quite revealing. We can observe that the impact of new imported intermediate inputs is positive and statistically significant throughout all specifications (i.e., on the number of varieties, probability, export value, and first differences). On the other hand, the impact of new imported capital goods is also positive and statistically significant in most of the specifications (i.e., on the probability, export value, and fist differences).

Compared to Chapter 3, these results are quite striking. Now that we disaggregate our data by end-use, we can detect a different pattern. We now observe that importing new intermediate inputs from developed countries has no effect on exports of new varieties; this differs from the previous chapter, where we found that at the aggregate level, importing new varieties from a developed countries have an impact on exports of new varieties. However, we need to remember that there is an increasing trend of developing countries engaging in global value chains. Therefore, our results are consistent in revealing that Mexico trades more new intermediate inputs with other developing countries than with developed countries. These results are aligned to the processing trade mechanism. Furthermore, this interpretation is in line with the maps of the source and destination countries for new varieties exhibited in the Descriptive Statistics section.

Similar to Lo Turco & Maggioni (2013), our findings suggest that importing from low-income countries has a larger effect on exports compared to importing from high-income countries. Unlike these authors, this chapter makes several contributions to the existing trade literature. Firstly, we make a distinction between new imports of intermediate inputs and new imports of capital goods. Secondly, we focus our analysis not only on the probability of exporting new varieties, but we extend it to the number of exported varieties, export value, and net change of new exported varieties. Finally, we complement Lo Turco & Maggioni (2013) findings as we provide empirical evidence from the perspective of a developing country, and we show bilateral trade patterns with other countries based on their income profile.

4.7.4 Main Industries Trading New Varieties

Another interesting feature is to extract a sub-sample of the top industries trading new varieties from our estimation sample. This sub-sample includes a total of 609 industries belonging to the chemicals and allied industries sectors (i.e., Chapters 28 to 38 of the Harmonized System), textiles (i.e., Chapters 50 to 63 of the Harmonized System), metals (i.e., Chapters 72 to 83 of the Harmonized System), and machinery and electrical sector (i.e., Chapters 84 and 85 of the Harmonized System) during the period from 2005 to 2016.

Once again, all the independent variables are expressed in logs and lagged by one year. Moreover, we include the full set of industry, country, and year fixed effects. The differences in the sample size across columns may be explained by the fact that negative binomial and logit regressions combined with fixed effects drop observations for time-invariant variables; also, logit regressions discard observations that do not contribute to the likelihood function.

Table 4.11 reveals the estimation of this sub-sample of top industries trading new varieties. We can observe that the coefficients of new imported intermediate inputs remain positive and strongly significant, while coefficients of new imported capital goods remain positive and statistically significant in most of our specifications (i.e., on the probability, export value, and first differences); these findings are consistent with our regression tables shown in the Results section. Moreover, the size of the coefficients of our explanatory variables (i.e., new imports of intermediates and capital goods) is magnified for this sub-sample of top industries. A plausible explanation for this increase is that we are now including industries with the highest concentration of new traded varieties; thus, the effects are larger.

Table 4.11: Main Industries Trading New Varieties

VARIABLES	(1) X_NEW	(2) AME Prob_X_NEW	(3) IntMarg_V	(4) X_NEW_D1
ln(No. New Imported Intermediate Inputs)_cit-1	0.0656***	0.0005***	0.0787***	
	(0.0166)	(0.0002)	(0.0182)	
ln(No. New Imported Capital Goods)_cit-1	0.0358	0.0006**	0.1214***	
	(0.0225)	(0.0003)	(0.0374)	
Δln(No. New Imported Intermediates)_cit				0.1617***
				(0.0031)
Δln(No. New Imported Capital Goods)_cit				0.2502***
				(0.0065)
ln(GDP in PPP)_ct-1	0.0036	0.0011**	0.0412	-0.0096***
	(0.0066)	(0.0005)	(0.0290)	(0.0028)
ln(Starting a Business)_ct-1	0.2724***	0.0027***	0.2564***	-0.0025
	(0.0325)	(0.0004)	(0.0274)	(0.0027)
Constant	-1.7802***		-1.2450*	0.2426***
	(0.2094)		(0.7359)	(0.0697)
Observations	339,912	339,876	515,904	515,904
Number of varieties	28,326	28,323	42,992	42,992
Prob Wald Chi2	0.000	0.000		
R-squared			0.006	0.033
Country FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table is based on a sub-sample of the main trading industries of new varieties; these industries belong to the chemicals and allied industries, textiles, metals, and to the machinery and electrical sectors. This sample is composed by 42,992 new varieties over the period 2005-2016. The dependent variable in column (1) corresponds to the number of new exported varieties by industry i to country c in time t; this column reports the coefficients using a negative binomial with fixed effects approach. The dependent variable in column (2) is the probability of exporting new varieties by industry i to country c in time t (i.e., the extensive margin); this column reports the average marginal effects using a logit model with fixed effects. The dependent variable in column (3) is the log value of exports (in U.S. dollars) of new varieties by industry i to country c at time t (i.e., the intensive margin); this column reports the coefficients employing a linear regression with fixed effects. The dependent variable in column (4) is the net change in the log number of new exported varieties by industry i to country c in time t; this column is calculated employing a log-first difference estimator with fixed effects. All independent variables are expressed in natural logs. All regressions include country, industry, and year fixed effects.

4.7.5 Alternative Control Variables

We now proceed to include other trade-related control variables extracted from the World Bank Doing Business dataset. These control variables are the number of documents to import, costs to import in U.S. dollars per container deflated, and time to import in days. The reason why we did not include these trade-related control variables in our baseline specifications is due to the short time span of these variables (i.e., the data is available for the period 2006-2015).

All regressions include logged and lagged independent variables and a full set of industry, country, and year fixed effects. The sample sizes may vary from one column to another due to the previously discussed drawback of the negative binomial and logit models; these two models combined with fixed effects result in dropping observations for time-invariant variables; also, the logit model drops observations for non-switchers that, therefore, do not contribute to the likelihood function. In addition, these sample sizes are smaller compared to the baseline results as the available time span for these new control variables is shorter.

Table 4.12 reports the regressions using these different control variables instead of the score of starting a business. In columns (1)-(4), we use the log number of documents to import as our control variable, along with logged GDP.[27] We can observe that new imported intermediate varieties remain positive and strongly significant (i.e., on export value and first differences). On the other hand, the coefficients for the number of documents to import are negative and statistically significant. A likely explanation is that requiring more paperwork at Customs may hinder the number of new varieties imported, which could potentially be used to produce new varieties for the export market.

Table 4.12: Alternative Control Variables

VARIABLES	(1) X_NEW	(2) AME Prob_NEW	(3) IntMar_V	(4) X_NEW_D1	(5) X_NEW	(6) AME Prob_NEW	(7) IntMar_V	(8) X_NEW_D1	(9) X_NEW	(10) AME Prob_NEW	(11) IntMar_V	(12) X_NEW_D1
$ln(IM\_NEW^{int})_{cit-1}$	0.0465***	0.0001	0.0586***		0.0423***	0.0021	0.0568***		0.0470***	0.0001	0.0593***	
	(0.0157)	(0.0001)	(0.0163)		(0.0157)	(0.0015)	(0.0163)		(0.0157)	(0.0001)	(0.0163)	
$ln(IM\_NEW^{cap})_{cit-1}$	0.0027	0.0001	0.0838**		-0.0006	0.0018	0.0821**		0.0038	0.0001	0.0849**	
	(0.0233)	(0.0309)	(0.0363)		(0.0233)	(0.0309)	(0.0363)		(0.0233)	(0.0309)	(0.0363)	
$\Delta ln(IM\_NEW^{int})_{cit}$				0.1520***				0.1520***				0.1520***
				(0.0028)				(0.0028)				(0.0028)
$\Delta ln(IM\_NEW^{cap})_{cit}$				0.2295***				0.2295***				0.2295***
				(0.0059)				(0.0059)				(0.0059)
ln(GDP in PPP)_ct-1	-0.0039	0.0005***	0.0867***	-0.0071***	-0.0069	0.0132***	0.0453*	-0.0084***	0.0025	0.0003***	0.1071***	-0.0043*
	(0.0064)	(0.0557)	(0.0273)	(0.0025)	(0.0064)	(0.0559)	(0.0273)	(0.0025)	(0.0064)	(0.0565)	(0.0274)	(0.0024)
ln(Import Docs)_ct-1	-0.1394***	-0.0002**	-0.0643**	-0.0113***								
	(0.0225)	(0.0325)	(0.0277)	(0.0033)								
ln(Import Costs)_ct-1					-0.1743***	-0.0209***	-0.2184***	-0.0078***				
					(0.0176)	(0.0280)	(0.0212)	(0.0022)				
ln(Import Time)_ct-1									0.0096	0.0001	0.1146***	0.0156***
									(0.0165)	(0.0272)	(0.0250)	(0.0028)
Constant	-0.1018		-1.2210*	0.2279***	1.0179***		1.3503*	0.2968***	-0.5517***		-2.2087***	0.0871
	(0.1777)		(0.7106)	(0.0644)	(0.2260)		(0.7469)	(0.0699)	(0.1848)		(0.7180)	(0.0649)
Observations	411,905	411,730	677,104	677,104	411,905	411,730	677,104	677,104	411,905	411,730	677,104	677,104
Number of varieties	41,639	41,620	68,727	68,727	41,639	41,620	68,727	68,727	41,639	41,620	68,727	68,727
Prob Wald Chi2	0.000	0.000			0.000	0.000			0.000	0.000		
R-squared			0.007	0.027			0.007	0.027			0.007	0.027
Country FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table reports the results of the estimation sample using different trade-related control variables instead of the score of starting a business. Columns (1)-(4) report the results using the log number of documents to import as a control variable. Columns (5)-(8) exhibit the results using the log costs to import in U.S. dollars per container deflated as a control variable. Columns (9)-(12) display the results using the log time to import in days as a control variable. The dependent variable in columns (1), (5), and (9) corresponds to the number of new exported varieties by industry i to country c in time t; these columns report the coefficients using a negative binomial with fixed effects approach. The dependent variable in columns (2), (6), and (10) is the probability of exporting new varieties by industry i to country c in time t (i.e., the extensive margin); these columns report the average marginal effects using a logit model with fixed effects. The dependent variable in columns (3), (7), and (11) is the log value of exports (in U.S. dollars) of new varieties by industry i to country c at time t (i.e., the intensive margin); these columns report the coefficients employing a linear fixed effects model. The dependent variable in columns (4), (8), and (12) is the net change in the log number of new exported varieties by industry i to country c in time t; these columns are calculated employing a log-first difference estimator with fixed effects. All independent variables are expressed in natural logs. All regressions include country, industry, and year fixed effects.

In columns (5)-(8), we now include the log costs to import and the log of GDP as our control variables.[28] Our results are consistent with the previous set of four columns, where new imported intermediates remain positive and statistically significant for the same two specifications (i.e., on export value and first differences). The coefficients of our import cost variable are negative and strongly significant through these four specifications. A plausible explanation is that import costs may also disincentivize technology transfers of new varieties.

[27] The World Bank defines the number of documents to import as the total amount of documents per import shipment required by law or by relevant agencies, including government ministries, customs authorities, port authorities, and other related agencies.

[28] The World Bank defines import costs as those related to importing a 20-foot container of goods by sea transport through the following stages: document preparation, customs clearance and inspections, inland transport and handling, and port and terminal handling. These costs are measured in U.S. dollars per container deflated.

In the last set of columns (9)-(12), we employ the log of days to import and the log of GDP as control variables.[29] Similarly, the coefficients of new imported intermediate inputs are positive and statistically significant for the same specifications (i.e., on export value and first differences). Contrary to expectations, days to import exhibit positive coefficients when these are statistically significant. A plausible explanation could be that source countries of new imported varieties are located in other continents (e.g., in Asia), translating into more days required to import goods from these remote locations.

These results with different sets of control variables are similar to our baseline results in Section 4.6 for the fixed effects negative binomial model, linear fixed effects model, and log-first difference specifications; however, the magnitude of the coefficients of our main explanatory variables (i.e., new imported intermediate inputs and new imported capital goods) are slightly smaller compared to our baseline results presented in Tables 4.4, 4.6, and 4.7. On the other hand, these results do not hold for the baseline specification using a fixed effects logit model in Table 4.5; our main explanatory variables are now insignificant when we include these different sets of trade-related variables as controls.

4.7.6 Fixed Effects Poisson Model

In this subsection, we use a Poisson model with fixed effects as an alternative methodology to the negative binomial model with fixed effects to estimate Eq.(4.9). As a recap, we are interested in examining the relationship between new imported intermediate inputs and capital goods on the number of new exported varieties. All the right-hand side variables are expressed in their logarithmic form and lagged by one year. The specifications also include a full set of country, industry, and year fixed effects. It is worth mentioning that a drawback of the Poisson model combined with fixed effects is that the model drops observations for time-invariant variables.

Table 4.13 exhibits the results after running the Poisson model with fixed effects specification. We can observe that the coefficients of the number of new imported intermediate inputs remain positive and strongly significant, while the number of new imported capital goods remain positive, albeit insignificant. These results are consistent with those presented in Table 4.4 for the fixed effects negative binomial approach; nonetheless, the magnitude of the coefficients under the Poisson model are slightly larger compared to the negative binomial model.

On a final note, we need to keep in mind that the Poisson model assumes that the mean of the distribution equals the variance. Thus, these results need to be taken with caution and may only be useful as a robustness check. In conclusion, the negative binomial model seems to be a more suitable empirical approach for our estimation sample as it relaxes the assumption that the mean of the distribution equals the variance.

[29]The World Bank Doing Business dataset states that the time to import refers to the calendar days associated to import a 20-foot container of goods by sea transport through the following stages: document preparation, customs clearance and inspections, inland transport and handling, and port and terminal handling.

Table 4.13: Fixed Effects Poisson Model

VARIABLES	(1) X_NEW	(2) X_NEW	(3) X_NEW	(4) X_NEW
ln(No. New Imported Intermediate Inputs)_cit-1	0.0615***	0.0605***	0.0609***	0.0605***
	(0.0183)	(0.0182)	(0.0182)	(0.0182)
ln(No. New Imported Capital Goods)_cit-1	0.0290	0.0282	0.0271	0.0268
	(0.0243)	(0.0243)	(0.0243)	(0.0243)
ln(GDP in PPP)_ct-1		0.2020***		0.0931*
		(0.0478)		(0.0491)
ln(Starting a Business)_ct-1			0.3536***	0.3342***
			(0.0374)	(0.0385)
Observations	543,084	543,084	543,084	543,084
Number of varieties	45,257	45,257	45,257	45,257
Prob Wald Chi2	0.000	0.000	0.000	0.000
Country FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

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

Notes: This table reports the results of estimating Eq.(4.9) employing a Poisson model with fixed effects. This approach displays the alternative results to Table 4.4 using a negative binomial model with fixed effects. The estimation sample is conformed by 68,727 new varieties over the period 2005-2016; however, some observations were dropped from the sample due to the nature of the methodology employed. The dependent variable is the number of new exported varieties by industry i to country c in time t. All independent variables are expressed in natural logs and lagged by one year. All regressions include country, industry, and year fixed effects.

4.7.7 Zero-Value Observations

The aim of this subsection is to deal with the issue of zero-value observations in the dependent variable. As mentioned before, this issue is common in trade datasets, as some countries may not trade with all the possible partner countries. The issue with zero-value observations is that these may lead to distorted estimates for semi-log models. Thus, we employ a Poisson Pseudo-Maximum-Likelihood (PPML) estimator proposed by Santos Silva & Tenreyro (2006), which was also used by Navas et al. (2020), to deal with zero-value observations. This technique is useful for linear and count data models.

In Table 4.14, we report the coefficients of Eqs.(4.9) and (4.11) using a PPML approach with fixed effects. Columns (1)-(4) estimate Eq.(4.9), where the dependent variable is a count variable representing the number of new exported varieties. We can observe that the coefficients of the main explanatory variables (i.e., new imported intermediate inputs and new imported capital goods) remain positive and significant throughout the different specifications; also, the magnitude of the coefficients of new imported intermediate inputs are now three times larger compared to Table 4.4, where we used a fixed effects negative binomial model. Interestingly, the magnitude of the coefficient of new imported capital goods is now statistically significant compared to Table 4.4.

On the other hand, columns (5)-(8) estimate Eq.(4.11), where the dependent variable is the log value of exports (in U.S. dollars) of new varieties (i.e., the intensive margin). We observe that our main explanatory variables remain positive and statistically significant. The magnitude of our coefficients of interest are now slightly larger compared to Table 4.6, where we used a linear fixed effects model.

Table 4.14: Zero-Value Observations

VARIABLES	(1) X_NEW	(2) X_NEW	(3) X_NEW	(4) X_NEW	(5) IntMarg_V	(6) IntMarg_V	(7) IntMarg_V	(8) IntMarg_V
ln(No. New Imported Intermediate Inputs)_cit-1	0.1553***	0.1547***	0.1546***	0.1544***	0.0901*	0.0893*	0.0893*	0.0888*
	(0.0557)	(0.0540)	(0.0540)	(0.0540)	(0.0475)	(0.0469)	(0.0468)	(0.0469)
ln(No. New Imported Capital Goods)_cit-1	0.1274**	0.1268**	0.1261**	0.1259**	0.1093***	0.1083***	0.1078***	0.1074***
	(0.0501)	(0.0495)	(0.0494)	(0.0493)	(0.0386)	(0.0381)	(0.0381)	(0.0381)
ln(GDP in PPP)_ct-1		0.1978		0.0898		0.2795		0.1852
		(0.2101)		(0.2223)		(0.1776)		(0.1836)
ln(Starting a Business)_ct-1			0.3507**	0.3320*			0.3298**	0.2921**
			(0.1770)	(0.1881)			(0.1411)	(0.1453)
Constant	-1.3193***	-6.4974	-2.8182***	-5.0890	0.1593***	-7.1638	-1.2498**	-5.9394
	(0.0039)	(5.4991)	(0.7570)	(5.5518)	(0.0025)	(4.6528)	(0.6031)	(4.6395)
Observations	824,304	824,304	824,304	824,304	824,304	824,304	824,304	824,304
Prob Wald Chi2	0.005	0.010	0.003	0.008	0.006	0.006	0.003	0.003
Country FE	YES	YES	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES

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

Notes: This table reports the coefficients of Eqs.(4.9) and (4.11) employing a Poisson Pseudo Maximum Likelihood (PPML) model with fixed effects; the aim is to deal with an excess of zero-value observations in the depent variable. Columns (1)-(4) estimate Eq.(4.9), where the dependent variable is a count variable representing the number of new exported varieties. Columns (5)-(8) estimate Eq.(4.11), where the dependent variable is the log value of exports (in U.S. dollars) of new varieties (i.e., the intensive margin). All independent variables are expressed in natural logs and lagged by one year. All regressions include country, industry, and year fixed effects.

4.7.8 Log-Log Model

We also try a log-log model to study the impact of the log number of new imported intermediate inputs and new imported capital goods on the log number of new exported varieties. Instead of running a fixed effects negative binomial model, we now run a linear fixed effects model. Table 4.15 exhibits the results of this log-log model. In column (1), we exclusively include both main explanatory variables along with the full set of country, industry, and year fixed effects. We can notice that both, new imported intermediate inputs and new imported capital goods, have positive and strong effects on exports of new varieties. Nevertheless, we can notice the size of the coefficient of new imported capital goods is slightly larger than the coefficient of new imported intermediate inputs.

The following columns incorporate the control variables in a stepwise manner. Thus, column (2) includes GDP as a control variable. Here, we can notice that the coefficients of the main explanatory variables remain positive and strongly significant. In column (3), we now include the score of starting a business. We can also notice that the coefficients of both main explanatory variables remain stable.

Finally, column (4) incorporates both main explanatory variables, all the control variables, and the full set of country, industry, and year fixed effects. We can observe that both imports of new intermediate inputs and new capital goods remain positive and strongly statistically significant. Our results suggest that a 1% increase in new imported intermediate inputs is associated with an increase of the number of new exported varieties by about 0.009%, ceteris paribus. Furthermore, our findings also suggest that a 1% increase in new imported capital goods is associated with an increase of the number of new exported varieties by about 0.01%, ceteris paribus. Compared to the baseline results presented in Table 4.4, we can notice the

Table 4.15: Log-Log Model

VARIABLES	(1) l_X_NEW	(2) l_X_NEW	(3) l_X_NEW	(4) l_X_NEW
ln(No. New Imported Intermediate Inputs)_cit-1	0.0087***	0.0086***	0.0085***	0.0085***
	(0.0018)	(0.0018)	(0.0018)	(0.0018)
ln(No. New Imported Capital Goods)_cit-1	0.0118***	0.0117***	0.0116***	0.0116***
	(0.0037)	(0.0037)	(0.0037)	(0.0037)
ln(GDP in PPP)_ct-1		0.0143***		0.0077***
		(0.0024)		(0.0024)
ln(Starting a Business)_ct-1			0.0280***	0.0265***
			(0.0022)	(0.0022)
Constant	0.0935***	-0.2761***	-0.0217**	-0.2141***
	(0.0010)	(0.0612)	(0.0090)	(0.0606)
Observations	824,724	824,724	824,724	824,724
R-squared	0.007	0.007	0.007	0.007
Number of varieties	68,727	68,727	68,727	68,727
Country FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

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

Notes: This table estimates a log-log model using a linear fixed effects approach. In this specification, our dependent variable is the log number of new exported varieties by industry i to country c in time t. All independent variables are expressed in natural logs and lagged by one year. All panel regressions include country, industry, and year fixed effects.

impact of new imported capital goods is now strongly statistically significant under this log-log specification.

4.7.9 Contemporaneous Effects

We are now interested in examining the contemporaneous effects of new imported intermediate inputs and new imported capital goods on exports of new varieties. To perform this analysis, we define all the right-hand side variables in time t instead of in time $t-1$. Therefore, our main explanatory variables are now the log number of new imported intermediate inputs by industry i from country c in year t, and new imported capital goods by industry i from country c in year t.

Table 4.16 shows the results of this contemporaneous effects associated with imports of new intermediate inputs and new capital goods on exports of new varieties. In column (1), we exhibit the coefficients using a fixed effects negative binomial model. We can observe that both, new imported intermediate inputs and new imported capital goods, are positive and strongly significant. Compared to the baseline specification in Table 4.4, the size of the coefficient of new imported intermediate inputs is significantly larger in the same year than with a lag. Interesting, the coefficient of new imported capital goods is now strongly statistically significant in the same year compared to the specification with a lag.

Column (2) reports the average marginal effects using a fixed effects logit model. We can notice that both, new imported intermediate inputs and new imported capital goods, are also positive and strongly statistically significant. Furthermore, the size of the effects are larger compared to our baseline specification with a lag presented in Table 4.5.

Column (3) displays the results using a linear fixed effects model. Like before, new imported intermediate inputs and new imported capital goods have positive and strong effects on the

Table 4.16: Contemporaneous Effects

	(1)	(2)	(3)
		AME	
VARIABLES	X_NEW	Prob_X_NEW	IntMarg_V
ln(No. New Imported Intermediate Inputs)_cit	1.0143***	0.0011***	1.2826***
	(0.0113)	(0.0001)	(0.0213)
ln(No. New Imported Capital Goods)_cit	1.1178***	0.0012***	1.9782***
	(0.0173)	(0.0304)	(0.0453)
ln(GDP in PPP)_cit	0.0136**	0.0002***	0.1039***
	(0.0058)	(0.0466)	(0.0221)
ln(Starting a Business)_cit	0.2985***	0.0002***	0.2310***
	(0.0290)	(0.0393)	(0.0228)
Constant	-2.1367***		-2.9102***
	(0.1845)		(0.5619)
Observations	542,544	542,424	824,724
Prob Wald Chi2	0.000	0.000	
R-squared			0.025
Number of varieties	45,212	45,202	68,727
Country FE	YES	YES	YES
Industry FE	YES	YES	YES
Year FE	YES	YES	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table explores the contemporaneous effects of importing new intermediate inputs and new capital goods on exports of new varieties. The dependent variable in column (1) corresponds to the number of new exported varieties by industry i to country c in time t; this column reports the coefficients using a fixed effects negative binomial model. The dependent variable in column (2) is the probability of exporting new varieties by industry i to country c in time t (i.e., the extensive margin); this column reports the average marginal effects using a fixed effects logit model. The dependent variable in column (3) is the log value of exports (in U.S. dollars) of new varieties by industry i to country c at time t (i.e., the intensive margin); this column reports the coefficients employing a linear fixed effects model. All independent variables are expressed in natural logs. All regressions include country, industry, and year fixed effects.

value (in U.S. dollars) of new exported varieties. Again, the contemporaneous effects are also significantly larger than in the baseline specification with a lag exhibited in Table 4.6.

In conclusion, we can notice that the magnitude of the impact of new imported intermediate inputs on exports of new varieties is stronger in the same year than with a lag. A plausible explanation for a larger contemporaneous effect of imports of new intermediate inputs may be processing trade. This is in line with global value chains, where production is fragmented into different stages across countries. On the other hand, it is also worth noticing that the impact of new imported capital goods on new exported varieties is also significant in the same year.

4.7.10 Lag Length Increase

This subsection addresses potential reserve causality. To tackle this potential issue, we increase the lag length of the independent variables by two lags. All regressions include the full set of independent variables, as well as industry, country, and year fixed effects. The inconsistency on the sample size across columns can be explained by the combination of fixed effects in our negative binomial and logit models; also, the logit model drops observations that do not contribute to the likelihood function; moreover, introducing lags also translates into dropped observations from our sample.

Table 4.17: Lag Length Increase

VARIABLES	(1) X_NEW	(2) AME Prob_X_NEW	(3) IntMarg_V
ln(No. New Imported Intermediate Goods)_cit-2	0.0454***	0.0088*	0.0396***
	(0.0144)	(0.0045)	(0.0150)
ln(No. New Imported Capital Goods)_cit-2	-0.0029	0.0047	0.0423
	(0.0215)	(0.0282)	(0.0323)
ln(GDP in PPP)_ct-2	0.0045	-0.0149	-0.0139
	(0.0059)	(0.0494)	(0.0252)
ln(Starting a Business)_ct-2	0.2217***	0.0674***	0.2202***
	(0.0277)	(0.0370)	(0.0231)
Constant	-1.6839***		0.1573
	(0.1839)		(0.6378)
Observations	479,919	479,798	755,997
Number of variety	43,629	43,618	68,727
Prob Wald Chi2	0.000	0.000	
R-squared			0.007
Country FE	YES	YES	YES
Industry FE	YES	YES	YES
Year FE	YES	YES	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table reports the results of the estimation sample after increasing the lag length of the independent variables by two lags. The dependent variable in column (1) corresponds to the number of new exported varieties by industry i to country c in time t; this column reports the coefficients using a negative binomial with fixed effects approach. The dependent variable in column (2) is the probability of exporting new varieties by industry i to country c in time t (i.e., the extensive margin); this column reports the average marginal effects using a logit model with fixed effects. The dependent variable in column (3) is the log value of exports (in U.S. dollars) of new varieties by industry i to country c at time t (i.e., the intensive margin); this column reports the coefficients employing a linear regression with fixed effects. All independent variables are expressed in natural logs. All regressions include country, industry, and year fixed effects.

Table 4.17 displays the results after increasing the number of lags of our independent variables by two periods. We can observe new imported intermediate inputs remains positive and statistically significant with two lags through most of our specifications (i.e., on number of varieties, probability, and export value).

4.7.11 Alternative Fixed Effects Combinations

Industry and Country Time Trends

To tackle a potential omitted variable bias, we include a combination of industry and country time trend fixed effects. Table 4.18 exhibits the results of including industry-year and country-year fixed effects. It is worth mentioning that our control variables dropped from the model because country-year fixed effects capture the variation of GDP and score of starting a business.

In column (1), we use a Poisson Pseudo-Maximum Likelihood model with fixed effects. We can notice that the new imported intermediate inputs and new imported capital goods have positive and strong effects on the number of new exported varieties. Compared to the baseline results presented in Table 4.4, the size of the coefficient of new imported intermediate inputs is larger now. Furthermore, the effect of new imported capital goods is now strongly statistically significant.

Table 4.18: Industry and Country Time Trends

VARIABLES	(1) X_NEW	(2) Prob_X_NEW	(3) IntMarg_V	(4) l_X_NEW_D1
ln(No. New Imported Intermediate Inputs)_cit-1	0.1147***	0.0075***	0.1068***	
	(0.0182)	(0.0021)	(0.0193)	
ln(No. New Imported Capital Goods)_cit-1	0.0856***	0.0205***	0.2566***	
	(0.0230)	(0.0037)	(0.0355)	
Δln(No. New Imported Intermediates)_cit-1				0.1590***
				(0.0026)
Δln(No. New Imported Capital Goods)_cit-1				0.2368***
				(0.0052)
Constant	-1.1209***	0.1096***	0.8049***	0.0002
	(0.0044)	(0.0004)	(0.0032)	(0.0002)
Observations	763,859	824,700	824,700	824,700
Prob Wald Chi2	0.000			
R-squared		0.120	0.121	0.068
Industry-Year FE	YES	YES	YES	YES
Country-Year FE	YES	YES	YES	YES

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

Notes: This table includes industry and country time trends as fixed effects. The dependent variable in column (1) corresponds to the number of new exported varieties by industry i to country c in time t; this column reports the coefficients using a fixed effects negative binomial model. The dependent variable in column (2) is the probability of exporting new varieties by industry i to country c in time t (i.e., the extensive margin); this column reports the average marginal effects using a fixed effects logit model. The dependent variable in column (3) is the log value of exports (in U.S. dollars) of new varieties by industry i to country c at time t (i.e., the intensive margin); this column reports the coefficients employing a linear fixed effects model. The dependent variable in column (4) is the net change in the log number of new exported varieties by industry i to country c in time t; this column is calculated employing a log-first difference estimator with fixed effects. All regressions include industry-year and country-year fixed effects.

Column (2), shows the results after employing a linear probability model with fixed effects to examine the impact of new imported intermediate inputs and new imported capital goods on the probability of exporting new varieties.[30] We can observe that new imported imported intermediate inputs are positive and strongly statistically significant. Interestingly, the impact of new imported capital goods is also positive and strongly significant on the probability of exporting new varieties. Furthermore, the magnitude of the effects is larger for new imported capital goods than for new imported intermediate inputs. Column (3) exhibits the results of our linear fixed effects model. We can also see that both main explanatory variables remain positive and strongly significant. Furthermore, the magnitude of the coefficients is considerably larger than the baseline results presented in Table 4.6. In column (4), we use a log-first difference estimator with fixed effects. Here, we can observe that both net changes of new imported intermediate inputs and new imported capital goods are also positive and strongly statistically significant. These results are in line with our baseline specifications presented in Table 4.7.

[30] To absorb multiple levels of fixed effects, we use two available Stata commands for this subsection: *reghdfe*, which is useful in running linear regressions, and *ppmlhdfe*, which is helpful in running Poisson Pseudo-Likelihood regressions. Nevertheless, there is not an available command yet to run logit regressions absorbing multiple levels of fixed effects; thus, we use a linear probability model instead.

Sector Time Trends

Moreover, we also include sector-year fixed effects to tackle a potential omitted variable bias. Table 4.19 exhibits the results of including sector-year and variety fixed effects.[31] In column (1), we employ a PPML estimator with fixed effects. These results suggest that importing new intermediate inputs has a positive and strong significant effect on the number of new exported varieties. In contrast, importing new capital goods does not have an effect on exports of new varieties. These results are in line with our baseline specification presented in Table 4.4.

Table 4.19: Sector Time Trends

VARIABLES	(1) X_NEW	(2) Prob_X_NEW	(3) IntMarg_V	(4) l_X_NEW_D1
ln(No. New Imported Intermediate Inputs)_cit-1	0.0694***	0.0053***	0.0602***	
	(0.0190)	(0.0017)	(0.0145)	
ln(No. New Imported Capital Goods)_cit-1	0.0057	0.0089***	0.0853***	
	(0.0258)	(0.0034)	(0.0317)	
Δln(No. New Imported Intermediates)_cit-1				0.1553***
				(0.0026)
Δln(No. New Imported Capital Goods)_cit-1				0.2323***
				(0.0054)
ln(GDP in PPP)_ct-1	0.1508***	0.0105***	0.0813***	-0.0100***
	(0.0488)	(0.0028)	(0.0234)	(0.0021)
ln(Starting a Business)_ct-1	0.3361***	0.0260***	0.2401***	-0.0009
	(0.0374)	(0.0026)	(0.0212)	(0.0020)
Constant	-6.3273***	-0.2747***	-2.3276***	0.2638***
	(1.2503)	(0.0716)	(0.5973)	(0.0548)
Observations	542,204	824,724	824,724	824,724
Prob Wald Chi2	0.000			
R-squared		0.176	0.184	0.064
Variety FE	YES	YES	YES	YES
Sector-Year FE	YES	YES	YES	YES

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

Notes: This table includes variety and sector time trends as fixed effects. The dependent variable in column (1) corresponds to the number of new exported varieties by industry i to country c in time t; this column reports the coefficients using a fixed effects negative binomial model. The dependent variable in column (2) is the probability of exporting new varieties by industry i to country c in time t (i.e., the extensive margin); this column reports the average marginal effects using a fixed effects logit model. The dependent variable in column (3) is the log value of exports (in U.S. dollars) of new varieties by industry i to country c at time t (i.e., the intensive margin); this column reports the coefficients employing a linear fixed effects model. The dependent variable in column (4) is the net change in the log number of new exported varieties by industry i to country c in time t; this column is calculated employing a log-first difference estimator with fixed effects. All regressions include variety and sector-year fixed effects.

Column (2) displays the results using a linear probability model with fixed effects. We can observe that both, new imported intermediate inputs and new imported capital goods, have positive and strong effects on the probability of exporting new varieties. Compared to the previous Table 4.18 that uses the same methodology, the size of the coefficients are smaller when we use sector-year and variety fixed effects in our specification.

Column (3) exhibits the results after employing a linear fixed effects model. We can also observe that both, new imported intermediate inputs and new imported capital goods, remain positive and strongly significant. Furthermore, the magnitude of the coefficients are slightly smaller than the baseline results in Table 4.6. In column (4), we employ a log-first difference estimator with fixed effects. Here, we can also notice that both net changes of new imported intermediate inputs and of new imported capital goods are also positive and strongly statisti-

[31]Similar to the previous chapter, variety fixed effects correspond to country-industry fixed effects.

cally significant. These results are in line with our baseline specifications presented in Table 4.7.

4.7.12 Input-Output Linkages

We are now interested in studying the input-output linkages across sectors because manufacturing goods require not only inputs from the same sector but also from other sectors. Therefore, we are interested in examining how new imported intermediate inputs in upwards sectors impact new exported varieties in downwards sectors. To perform this analysis, we employ the World Input-Output Database (WIOD), which reports data at the sector level using the ISIC classification at 2-digits. We then match ISIC sectors to Harmonized System (HS) codes using a correspondence table. It is worth mentioning that we only include manufacturing sectors in this analysis.

In this exercise, the dependent variable is defined as the log value (in U.S. dollars) of exports by sector k to country c in time t.[32] On the other hand, we define two main explanatory variables. First, we have the log value (in U.S. dollars) of new imported intermediate inputs by sector k from country c in time t. Then, we have the log value (in U.S. dollars) of new imported capital goods by sector k from country c in time t. These two main explanatory variables are calculated in line with Javorcik (2004). Thus, we multiply the value (in U.S. dollars) of new imports (i.e., intermediate inputs and capital goods by separate) times the share of inputs acquired by other sectors in total inputs sourced by sector k using input-output tables.

Table 4.20 examines these input-output linkages across sectors. In column (1), we start by only including new imported intermediate inputs with a full set of country, sector, and year fixed effects. From this column, we can observe that the indirect effect of new imported intermediate inputs on exports of new varieties is positive and strongly statistically significant. In column (2), we only include new imported capital goods with the full set of country, sector, and year fixed effects. We can also notice that the indirect effect of new imported capital goods on exports of new varieties is positive and strongly significant. Column (3) includes both main explanatory variables. We can observe that both explanatory variables remain positive and statistically significant. The following columns incorporate the control variables one by one.

Finally, column (6) includes both main explanatory variables, all the control variables, and the full set of country, sector, and year fixed effects. We can conclude that both, imports of new intermediate inputs and imports of new capital goods, have positive and strong statistical effects on exports of new varieties. Furthermore, the magnitude of the coefficients are comparable.

Our results suggest that a 1% increase in new imported intermediate inputs leads to an increase of new exported varieties by about 0.02%, ceteris paribus. Likewise, our findings suggest that 1% increase in new imported capital goods leads to an increase of new exported varieties by about 0.03%, ceteris paribus. It is worth noting that the magnitude of these effects

[32] We define sectors as sections of the Harmonized System classification.

Table 4.20: Input-Output Linkages: Indirect Effect

VARIABLES	(1) l(X_NEW_USD)	(2) l(X_NEW_USD)	(3) l(X_NEW_USD)	(4) l(X_NEW_USD)	(5) l(X_NEW_USD)	(6) l(X_NEW_USD)
ln(New Imp Interm IND)_ckt	0.0264***		0.0250***	0.0243***	0.0241***	0.0237***
	(0.0022)		(0.0022)	(0.0022)	(0.0022)	(0.0022)
ln(New Imp Capital IND)_ckt		0.0336***	0.0305***	0.0289***	0.0295***	0.0284***
		(0.0035)	(0.0035)	(0.0035)	(0.0035)	(0.0035)
ln(GDP in PPP)_ct				0.7844***		0.6494***
				(0.0425)		(0.0428)
ln(Starting a Business)_ct					0.5565***	0.4408***
					(0.0331)	(0.0335)
Constant	9.9956***	10.0363***	9.9717***	-10.1024***	7.6791***	-8.4636***
	(0.0136)	(0.0125)	(0.0139)	(1.0874)	(0.1380)	(1.0791)
Observations	487,094	487,094	487,094	487,094	487,025	487,025
R-squared	0.039	0.038	0.039	0.040	0.039	0.040
Number of varieties	64,239	64,239	64,239	64,239	64,193	64,193
Country FE	YES	YES	YES	YES	YES	YES
Sector FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

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

Notes: This table analyzes across sectors input-output linkages. This table is estimated using a linear fixed effects model. The dependent variable stands for the log value of exports (in U.S. dollars) by sector k to country c in time t. The main explanatory variables correspond to the log value of new imported intermediate inputs by sector k from country c in time t and the log value of new imported capital goods by sector k from country c in time t. All regressions include country, sector, and year fixed effects.

are smaller compared to new imports from the same sector.[33]

To conclude this subsection, this input-output exercise has some limitations. First, input-output tables are reported at the sector level (i.e., ISIC 2-digits) instead of at the industry level; thus, this exercise is performed at a more aggregate level compared to the baseline specifications. Moreover, the World Input-Output Database is only available for 43 countries, including 28 EU countries and 15 major countries. Thus, there is no available data for most developing economies representing Mexico's main traders of new varieties. In terms of data accuracy, this input-output exercise loses accuracy when sectors are merged using an HS-ISIC correspondence table. Although this input-output exercise exhibits some limitations, it is worth exploring this exercise, as it provides an overview of the magnitude of the effects within and between sectors.

4.7.13 Marginal Effects by Sector

Furthermore, we examine the heterogeneity across sectors. The goal now is to identify those sectors with stronger effects on exports of new varieties. To perform this analysis, we run two separate sets of regressions where we interact the main explanatory variables with sectors.[34] Our results are reported as average marginal effects resulting from a fixed effects logit model.

[33] We also include the results of the direct effect of input-output linkages at the sector level in the Appendix section. Compared to our baseline specifications, we now aggregate data at the sector level instead of at the industry level.

[34] Like before, we define sectors as sections of the Harmonized System. It is worth mentioning that we only include manufacturing sectors in this analysis.

Marginal Effects of Intermediate Inputs by Sector

We start our analysis by interacting the log number of new imported intermediate inputs and sectors. Table 4.21 reports the average marginal effects of intermediate inputs by industry. In column (1), we include only the interaction term with the full set of country, sector, and year fixed effects. Column (2) incorporates only GDP as a control variable, and column (3) includes only the score of starting a business as the control variable. Column (4) includes the interaction term, all the control variables, and the full set of country, sector, and year fixed effects. Our results suggest that an important size of the impact can be explained by the transport equipment sector. Thus, we can infer that new imported intermediate inputs by the transport equipment sector has a significant impact on the probability of exporting new varieties.

Table 4.21: Marginal Effects of Intermediate Inputs by Sector

VARIABLES	(1) AME Prob_X_NEW	(2) AME Prob_X_NEW	(3) AME Prob_X_NEW	(4) AME Prob_X_NEW
Sector = 5, Mineral Products	-0.0388***	-0.0389***	-0.0389***	-0.0390***
	(0.0083)	(0.0083)	(0.0083)	(0.0083)
Sector = 6, Chemicals and Allied Industries	0.0368***	0.0367***	0.0367***	0.0366***
	(0.0025)	(0.0025)	(0.0025)	(0.0025)
Sector = 7, Plastics and Rubbers	0.0360***	0.0359***	0.0359***	0.0358***
	(0.0048)	(0.0048)	(0.0048)	(0.0048)
Sector = 8, Raw Hides, Skins, Leather, and Furs	-0.0530***	-0.0530***	-0.0530***	-0.0530***
	(0.0180)	(0.0180)	(0.0180)	(0.0180)
Sector = 9, Wood and Articles of Wood	-0.0514***	-0.0513***	-0.0514***	-0.0513***
	(0.0104)	(0.0104)	(0.0104)	(0.0104)
Sector = 10, Pulp of Wood	-0.0023	-0.0025	-0.0025	-0.0026
	(0.0081)	(0.0081)	(0.0081)	(0.0081)
Sector = 11, Textiles	-0.0419***	-0.0419***	-0.0419***	-0.0419***
	(0.0057)	(0.0057)	(0.0057)	(0.0057)
Sector = 12, Footwear and Headgear	-0.1996***	-0.1994***	-0.1997***	-0.1996***
	(0.0506)	(0.0506)	(0.0506)	(0.0506)
Sector = 13, Stone and Glass	-0.0052	-0.0053	-0.0053	-0.0053
	(0.0064)	(0.0064)	(0.0064)	(0.0064)
Sector = 14, Precious Stones and Metals	-0.1264***	-0.1265***	-0.1265***	-0.1266***
	(0.0364)	(0.0364)	(0.0364)	(0.0364)
Sector = 15, Base Metals	0.0304***	0.0303***	0.0303***	0.0303***
	(0.0031)	(0.0031)	(0.0031)	(0.0031)
Sector = 16, Machinery and Electrical Equipment	0.0595***	0.0594***	0.0595***	0.0594***
	(0.0034)	(0.0034)	(0.0034)	(0.0034)
Sector = 17, Transport Equipment	0.1382***	0.1382***	0.1382***	0.1381***
	(0.0056)	(0.0056)	(0.0056)	(0.0056)
Sector = 18, Optical and Medical Instruments	-0.0481***	-0.0481***	-0.0482***	-0.0482***
	(0.0134)	(0.0134)	(0.0134)	(0.0134)
Sector = 20, Miscellaneous Manufactured Articles	-0.0230	-0.0230	-0.0230	-0.0230
	(0.0157)	(0.0157)	(0.0157)	(0.0157)
Observations	824,711	824,711	824,711	824,711
Control Variables	NO	YES	YES	YES
Country FE	YES	YES	YES	YES
Sector FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table reports the average marginal effects by sector of intermediate inputs using a logit model. The main explanatory variable corresponds to an interaction term between the log number of new imported intermediate inputs in the previous year and sector. Control variables are incorporated stepwise; thus, column (1) starts by excluding all control variables; column (2) only includes GDP as a control variable; column (3) only includes score of starting a business as a control variable; column (4) incorporates all control variables. All independent variables are expressed in natural logs and lagged by one year. All regressions include country, sector, and year fixed effects.

Marginal Effects of Capital Goods by Sector

Furthermore, we also interact the log number of new imported capital goods and sectors. Table 4.22 presents the average marginal effects of capital goods by industry. In column (1), we include only the interaction term with the full set of country, sector, and year fixed effects. Column (2) incorporates only GDP as a control variable, and column (3) includes only the score of starting a business as the control variable. Column (4) includes the interaction term, all the control variables, and the full set of country, sector, and year fixed effects. Our results suggest that an important size of the effect can be explained by the base metals, optical and medical instruments, and machinery and electrical equipment sectors. We can infer that new imported capital goods of these three sectors have significant effects on the probability of exporting new varieties.

Table 4.22: Marginal Effects of Capital Goods by Sector

VARIABLES	(1) AME Prob_X_NEW	(2) AME Prob_X_NEW	(3) AME Prob_X_NEW	(4) AME Prob_X_NEW
Sector = 11, Textiles	-0.0512	-0.0515	-0.0511	-0.0514
	(0.0324)	(0.0324)	(0.0324)	(0.0324)
Sector = 15, Base Metals	0.0895***	0.0894***	0.0894***	0.0894***
	(0.0070)	(0.0070)	(0.0070)	(0.0070)
Sector = 16, Machinery and Electrical Equipment	0.0508***	0.0506***	0.0506***	0.0505***
	(0.0029)	(0.0029)	(0.0029)	(0.0029)
Sector = 17, Transport Equipment	-0.0541***	-0.0543***	-0.0544***	-0.0545***
	(0.0167)	(0.0167)	(0.0167)	(0.0167)
Sector = 18, Optical and Medical Instruments	0.0639***	0.0638***	0.0638***	0.0637***
	(0.0050)	(0.0050)	(0.0050)	(0.0050)
Sector = 19, Arms and Ammunition	-0.1209	-0.1212	-0.1211	-0.1213
	(0.0786)	(0.0787)	(0.0786)	(0.0787)
Sector = 20, Miscellaneous Manufactured Articles	-0.0837***	-0.0841***	-0.0839***	-0.0841***
	(0.0259)	(0.0259)	(0.0260)	(0.0260)
Observations	824,721	824,721	824,721	824,721
Control Variables	NO	YES	YES	YES
Country FE	YES	YES	YES	YES
Sector FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: This table reports the average marginal effects by sector of capital goods using a logit model. The main explanatory variable corresponds to an interaction term between the log number of new imported capital goods in the previous year and sector. Control variables are incorporated stepwise; thus, column (1) starts by excluding all control variables; column (2) only includes GDP as a control variable; column (3) only includes score of starting a business as a control variable; column (4) incorporates all control variables. All independent variables are expressed in natural logs and lagged by one year. All regressions include country, sector, and year fixed effects.

This exercise is important because from a policy perspective, trade policies could be tailored to specific sectors depending on the type of goods Mexican firms are interested in importing. This chapter allows to identify specific sectors that could represent the core of future preferential trade agreements. In this way, Mexican firms belonging to these strong sectors could benefit from a higher integration in global value chains.

4.7.14 Two-Stage Regressions

A potential problem with the estimates so far is that of endogeneity arising from potential omitted variables and reverse causality. Therefore, we use applied import tariffs of intermediate inputs and of capital goods as instrumental variables, and employ these instruments in two-stage regressions. The usage of MFN import tariffs as an instrumental variable is standard in the empirical trade literature (Bas & Strauss-Kahn 2014, Feng et al. 2016).

The idea behind MFN import tariffs of intermediate inputs and of capital goods as instruments is that these will not impact exports directly, but will have an indirect impact through imports of new intermediate inputs and capital goods, respectively. Nevertheless, using MFN tariffs can be problematic because within an industry, export tariffs may be correlated to import tariffs. To tackle this potential problem, we employ applied import tariffs instead.[35] Furthermore, these applied import tariffs are country-specific and account for preferential and MFN tariffs.

We use different methodologies for our two-stage regressions: Poisson model with fixed effects, Two-Stage Least Squares (2SLS) regressions with fixed effects, and a log-first difference estimator in two stages. Our estimation sample is now reduced due to a lack of information on applied tariffs for some industries.[36] All the independent variables are expressed in logs, and regressions include a full set of country, industry, and year fixed effects.

Table 4.26 provides the results of our two-stage regressions using applied import tariffs of intermediate inputs and of capital goods as our instrumental variables. The top panel shows the second stage of the IV regressions, while the bottom panel displays the first stage of the regressions. Columns (1)-(2) exhibit the regression models where the endogenous variable is the number of new imported intermediate inputs. Furthermore, column (3) displays the model where the endogenous variable is the net change in the log number of new imported intermediate inputs. On the other hand, columns (4)-(5) present the regression models where the endogenous variable is the number of new imported capital goods. Moreover, column (6) shows the model where the endogenous variable is the net change in the log number of new imported capital goods.

From the first stage regressions, we can observe there is a positive and significant relationship between applied import tariffs of intermediate inputs and new imported intermediate inputs

[35] Nevertheless, we also display the results of the IV regressions using MFN import tariffs of intermediate inputs and capital goods as instrumental variables in the Appendix section.

[36] The WTO does not report MFN and applied tariffs for the following industries: "Inorganic or organic compounds of mercury, whether or not chemically defined, excluding amalgams" (HS 2852), "Phosphides, whether or not chemically defined, excluding ferrophosphorus; other inorganic compounds (including distilled or conductivity water and water of similar purity); liquid air (whether or not rare gases have been removed); compressed air; amalgams, other than amalgams of precious metals" (HS 2853), "Biodiesel and mixtures thereof, not containing or containing less than 70% by weight of petroleum oils or oils obtained from bituminous minerals" (HS 3826), "Machines and apparatus of a kind used solely or principally for the manufacture of semiconductor boules or wafers, semiconductor devices, electronic integrated circuits or flat panel displays; machines and apparatus specified in Note 9 (C) to this Chapter; parts and accessories" (HS 8486), "Machinery parts, not containing electrical connectors, insulators, coils, contacts or other electrical features, not specified or included elsewhere in this Chapter" (HS 8487), "Vacuum cleaners" (HS 8508), and "Sanitary towels (pads) and tampons, napkins and napkin liners for babies and similar articles, of any material" (HS 9619).

Table 4.23: Two-Stage Regressions

Second Stage IV Regressions	(1) Second Stage X_NEW	(2) Second Stage IntMarg_V	(3) Second Stage X_NEW_D1	(4) Second Stage X_NEW	(5) Second Stage IntMarg_V	(6) Second Stage X_NEW_D1
ln(No. New Imp Interm Inputs)_cit-1	7.4279***	4.6296***				
	(0.7002)	(0.6430)				
Δln(No. New Imp Intermediates)_cit-1			0.4650***			
			(0.1531)			
ln(No. New Imp Capital Goods)_cit-1				2.1082**	1.4949	
				(1.0234)	(1.1787)	
Δln(No. New Imp Capital Goods)_cit-1						0.4807*
						(0.2673)
Observations	349,284	593,824	593,824	108,748	168,809	168,809
Number of varieties	29,361	50,030	50,030	9,194	14,290	14,290
Country FE	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Prob Wald Chi2	0.000			0.000		
Adj. R2		-0.348	-0.168		-0.113	-0.111
Underidentification stat.		261.032	48.050		75.685	17.488
Prob underident. stat.		0.000	0.000		0.000	0.000
Weak identification stat.		261.157	48.054		75.722	17.490
Endogeneity F-test		61.144	4.524		1.454	1.018
Prob endogeneity test		0.000	0.033		0.228	0.313
First Stage IV Regressions	First Stage IM_NEW_int	First Stage IM_NEW_int	First Stage IM_NEW_int	First Stage IM_NEW_cap	First Stage IM_NEW_cap	First Stage IM_NEW_cap
ln(Interm Applied Import Tariffs)_cit-1	0.0063***	0.0063***	-0.0037***			
	(0.0004)	(0.0004)	(0.0005)			
ln(Capital Applied Import Tariffs)_cit-1				0.0078***	0.0078***	-0.0051***
				(0.0009)	(0.0009)	(0.0012)
Constant	0.0817***	0.0817***	-0.0087***	0.0926***	0.0926***	-0.0113***
	(0.0005)	(0.0005)	(0.0008)	(0.0014)	(0.0014)	(0.0018)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table reports the two-stage regressions. The instrumental variables employed are the log of applied import tariffs of intermediate inputs in time $t-1$ and the log of applied import tariffs of capital goods in time $t-1$. The top panel shows the second stage IV regressions, while the bottom panel exhibits the first stage regressions. Columns (1) and (4) are estimated using an IV Poisson model, where the dependent variable is the number of new exported varieties; these columns report the coefficients. Columns (2) and (5) are estimated using a 2SLS model, where the dependent variable is the export value (in U.S. dollars) of new varieties. Finally, columns (3) and (6) are estimated using a log-first difference estimator in two stages, where the dependent variable is the net change in the log number of new exported varieties.

throughout the different models, except for the first difference of the endogenous variable, which displays a negative and significant relationship. On the other hand, we can also notice a positive and strong relationship between applied import tariffs of capital goods and new imported capital goods, except for the first difference of the endogenous variable, which exhibits a negative and significant relationship. A plausible explanation for our positive effect could be that nowadays, most of the manufacturing products have very low applied tariffs.

Columns (1) and (4) report the results using an IV Poisson model. After the instrumentation, we can observe that both, new imports of intermediate inputs and of new capital goods, have positive and statistically significant effects on the number of new exported varieties. Our results suggest that a 1% increase in new imported intermediate inputs is associated with an increase in the number of new exported varieties by about 0.07. Also, a 1% increase in the number of new imported capital goods leads to an increase in the number of new exported varieties by about 0.02. We can now observe that both effects are larger in comparison to the model where exogeneity was assumed presented in Table 4.4; more importantly, the effects of new imported capital goods are now statistically significant.

Columns (2) and (5) are estimated employing 2SLS model. After the instrumentation, we can notice that imports of new intermediate inputs are positive and strongly significant, while imports of new capital goods are insignificant. These results suggest that a 1% increase in the number of new imported intermediate inputs leads to an increase in the export value of new varieties by about 0.05. In comparison to the model where exogeneity was assumed exhibited in Table 4.6, the magnitude of the effect of new imported intermediate inputs is now larger.

Columns (3) and (6) display the results after using a log-first difference estimator in two stages. In the second stage of the regressions, we can observe that both endogenous variables (i.e., new imports of intermediate inputs and new capital goods) are positive and statistically significant. Our findings suggest that a 1% increase in the net change number of new imported intermediate inputs leads to an increase in the net change number of new exported varieties by about 0.5%. Also, a 1% increase in the net change number of new imported capital goods leads to an increase in the net change number of new exported varieties by about 0.5%. Compared to the model where we assumed exogeneity presented in Table 4.7, the effects are now larger. As a final remark, these two-stage regressions have identified that our findings are robust in terms of statistical significance and direction of impact, and confirm the importance of importing new intermediate inputs, as well as the relevance of importing new capital goods, on exports of new varieties.

We also report a series of tests depending on each methodology. We begin with an underidentification test to examine whether the number of instruments is less compared to the number of endogenous variables; we confirm that we do not face any underidentification issue as we only use one instrument for each endogeneous variable. Next, we use a weak identification test; the F-statistic is larger than the critical values; thus, our instruments are not weak and have a good explanatory power for our endogeneous variables. Finally, we also report an endogeneity test. On the one hand, we conclude that the endogenous regressor is in fact endogenous in columns (2) and (3). On the other hand, we conclude that the endogenous regressor can be treated as exogenous in columns (5) and (6).

4.8 Conclusions

The empirical trade literature has widely focused on the impact of importing intermediate inputs on exports, but has somehow neglected the importance of importing capital goods on exports. The main contribution of this chapter is to disentangle the effects between new imported intermediate inputs and new imported capital goods on exports of new varieties. Thus, we show that new imported capital goods play an important role on exporting new varieties. This chapter also contributes to the existing literature in two other ways. First, the chapter focuses exclusively on new varieties, which can uncover technological transfers from abroad through imports of new varieties. Second, we explore these trade relationships between imports and exports of new varieties from the perspective of an emerging economy that intensively trades with both developed and developing countries.

We employ four different empirical approaches on our estimation sample, which is composed of 68,727 new traded varieties belonging to the manufacturing sector over the period 2005-2016. We start by using a fixed effects negative binomial model to evaluate the importance of new imported intermediate inputs and new imported capital goods on the number of new exported varieties. Then, we employ a fixed effects logit model to examine the effects of importing new intermediate inputs and new capital goods on the probability of exporting new varieties. Next, we use a linear fixed effects model to analyze the impact of importing new intermediate inputs and new capital goods on the export value of new varieties. Finally, we use a log-first difference estimator to study the impact of the net change in new imported intermediate inputs and in new capital goods on the net change in the number of new exported varieties.

Our results suggest that new imported intermediate inputs have positive and strong effects on exports of new varieties measured in number of varieties, probability, export value, and net changes. Furthermore, these findings also reveal the importance of new imported capital goods on new exported varieties measured as the probability, export value, and net changes. Our results hold even after controlling for market size and ease of doing business, and after incorporating the full set of country, industry, and year fixed effects.

We perform a series of robustness checks that consists of including trade gravity variables, excluding main partner countries from our sample, dividing our sample into two sub-samples based on income groups, including only top industries trading new varieties, using other trade variables as controls, employing an alternative methodology to the fixed effects negative binomial model, dealing with zero-value observations in the dependent variable, using a log-log model, looking at contemporaneous effects, increasing the lag length of the independent variables, including different combinations of fixed effects, analyzing input-output linkages across sectors, and examining the marginal effects by industry. We also use two-stage regressions, where our instruments are applied import tariffs of intermediate inputs and of capital goods, to discard potential endogeneity.

Our findings also suggest that importing new intermediate inputs from low- and middle-income countries has a larger effect on exports of new varieties compared to importing from high-income countries. A plausible explanation is that Mexican firms tend to adapt their products according to the tastes, preferences, and regulations of their export markets. A rationale could be that these firms tend to export new varieties to foreign markets that are similar to the domestic market. In other words, exporting to Latin American countries with comparable levels of development and a common language may be viewed as more appealing to Mexican firms. For example, Mexican firms do not need to incur in additional packaging costs in terms of labelling to export to Latin American countries, as these countries have a common language.

From a trade policy perspective, we have identified that some of the main source countries of new imported intermediate inputs and of new capital goods are located in Asia. As Mexico does not possess free trade agreements with these Asian nations, imports from these countries pay tariffs; these tariffs translate into higher costs for firms importing intermediate inputs and capital goods from these countries. Therefore, negotiating free trade agreements with strategic trade partners located in Asia could represent a motivation for Mexican policymakers to promote further integration with this strategic region and consolidate Mexico's role in global value chains.

Nonetheless, as acknowledged in the previous chapter, this policy recommendation should be taken with caution as some Mexican industries could be negatively impacted as a result of a trade liberalization process with Asian countries. Therefore, it is of vital importance to follow an appropriate consultation process with the relevant business chambers in Mexico to evaluate the advantages and disadvantages of negotiating trade agreements with countries located in Asia. Despite this, it may be interesting evaluating the possibility of negotiating preferential trade agreements with Asian nations focused on identified high-performance sectors of new imported intermediate inputs and of new imported capital goods.

Finally, our analysis faces two main limitations. First, trade datasets at the product level are no longer updated; thus, our analysis covers until 2016, which constitutes the most recent available year. Second, other relevant trade-related variables that could be used as controls (i.e., number of documents to import, import costs, and time to import) have a short time span covering the period 2006-2015. Due to the nature of the analysis, the first two years of the time series are dropped from the sample (i.e., due to the lag of the independent variables by one year and to the first difference of the log-first difference estimator approach); thus, we end up with a sample including eight years instead of twelve. Nonetheless, we recognize the importance of these trade-related variables; thus, we include these variables as part of the Robustness Analysis.

4.9 Appendix

4.9.1 Additional Regressions

Table 4.24: Foreign Direct Investment

VARIABLES	(1) X_NEW	(2) AME Prob_X_NEW	(3) IntMarg_V
ln(No. New Imported Intermediate Inputs)_cit-1	0.0054	-0.0021	0.0381*
	(0.0151)	(0.0049)	(0.0195)
ln(No. New Imported Capital Goods)_cit-1	-0.0127	0.0047	0.1057**
	(0.0221)	(0.0299)	(0.0430)
ln(GDP in PPP)_ct-1	-0.0382***	0.0028	-0.1464***
	(0.0078)	(0.0652)	(0.0553)
ln(Starting a Business)_ct-1	-0.0188	0.0056	0.0784
	(0.0399)	(0.0586)	(0.0508)
ln(FDI Inflows)_ct-1	0.0005	0.0001	-0.0005
	(0.0008)	(0.0010)	(0.0007)
Constant	0.8291***		4.6589***
	(0.2568)		(1.4075)
Observations	323,880	323,772	462,828
Prob Wald Chi2	0.000	0.000	
R-squared			0.012
Number of varieties	26,990	26,981	38,569
Country FE	YES	YES	YES
Industry FE	YES	YES	YES
Year FE	YES	YES	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table includes FDI as a control variable. The dependent variable in column (1) corresponds to the number of new exported varieties by industry i to country c in time t; this column reports the coefficients using a fixed effects negative binomial model. The dependent variable in column (2) is the probability of exporting new varieties by industry i to country c in time t (i.e., the extensive margin); this column reports the average marginal effects using a fixed effects logit model. The dependent variable in column (3) is the log value of exports (in U.S. dollars) of new varieties by industry i to country c at time t (i.e., the intensive margin); this column reports the coefficients employing a linear fixed effects model. All regressions include country, industry, and year fixed effects.

Table 4.25: Input-Output Linkages: Direct Effect

VARIABLES	(1) l(X_NEW_USD)	(2) l(X_NEW_USD)	(3) l(X_NEW_USD)	(4) l(X_NEW_USD)	(5) l(X_NEW_USD)	(6) l(X_NEW_USD)
ln(New Imp Interm DIR)_ckt	0.0304***		0.0287***	0.0283***	0.0282***	0.0280***
	(0.0083)		(0.0082)	(0.0082)	(0.0082)	(0.0082)
ln(New Imp Capital DIR)_ckt		0.0465***	0.0445***	0.0431***	0.0438***	0.0428***
		(0.0098)	(0.0097)	(0.0097)	(0.0097)	(0.0097)
ln(GDP in PPP)_ct				0.7580***		0.6270***
				(0.2036)		(0.2051)
ln(Starting a Business)_ct					0.5398***	0.4283***
					(0.1594)	(0.1611)
Constant	9.8952***	9.9692***	9.8113***	-9.5862*	7.5883***	-7.9960
	(0.0752)	(0.0621)	(0.0782)	(5.2107)	(0.6643)	(5.1698)
Observations	21,178	21,178	21,178	21,178	21,175	21,175
R-squared	0.039	0.039	0.040	0.041	0.041	0.042
Number of varieties	2,793	2,793	2,793	2,793	2,791	2,791
Country FE	YES	YES	YES	YES	YES	YES
Sector FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

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

Notes: This table reports new traded varieties within the same sector. This table is estimated using a linear fixed effects model. The dependent variable stands for the log value of exports (in U.S. dollars) by sector k to country c in time t. The main explanatory variables correspond to the log value of new imported intermediate inputs by sector k from country c in time t and the log value of new imported capital goods by sector k from country c in time t. All regressions include country, sector, and year fixed effects.

Table 4.26: Two-Stage Regressions using Alternative Instruments

Second Stage IV Regressions	(1) Second Stage X_NEW	(2) Second Stage X_NEW	(3) Second Stage Prob_X_NEW AME	(4) Second Stage Prob_X_NEW AME	(5) Second Stage IntMarg_V	(6) Second Stage IntMarg_V	(7) Second Stage X_NEW_D1	(8) Second Stage X_NEW_D1
ln(No. New Imp Interm Inputs)_cit-1	11.4933***		0.3054***		6.6302***			
	(0.3914)		(0.0241)		(0.3767)			
ln(No. New Imp Capital Goods)_cit-1		6.3342***		0.3577***		4.0911***		
		(0.4173)		(0.0241)		(0.4560)		
Δln(No. New Imp Intermediates)_cit							0.4886***	
							(0.1881)	
Δln(No. New Imp Capital Goods)_cit								0.3764**
								(0.1909)
ln(GDP in PPP)_ct-1	-0.1286***	-0.3026***	0.0068***	0.0108***	-0.0523	-0.0271	-0.0055	0.0009
	(0.0400)	(0.0585)	(0.0026)	(0.0027)	(0.0335)	(0.0572)	(0.0047)	(0.0085)
ln(Starting a Business)_ct-1	0.2375***	0.2905***	-0.0110***	-0.0030	0.1219***	0.2028***	-0.0025	-0.0065
	(0.0337)	(0.0539)	(0.0079)	(0.0137)	(0.0296)	(0.0543)	(0.0038)	(0.0074)
Residuals	-11.4102***	-6.2729***	-0.2652***	-0.3045***				
	(0.3922)	(0.4187)	(0.1428)	(0.1231)				
Constant			-0.8926***	-0.4800***				
			(0.0636)	(0.1561)				
Observations	350,448	108,554	593,824	168,809	593,824	168,809	593,824	168,809
Number of varieties	29,461	9,173			50,030	14,290	50,030	14,290
Prob Wald Chi2	0.000	0.000	0.000	0.000				
Adj. R2					-0.620	-0.258	-0.183	-0.072
Country FE	YES	YES	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Underidentification stat.					922.584	576.922	32.430	32.816
Prob underident. stat.					0.000	0.000	0.000	0.000
Weak identification stat.					924.147	579.073	32.432	32.822
Prob F-test Excluded Instr.					0.000	0.000	0.000	0.000
Endogeneity test					448.195	88.177	3.498	0.692
Prob endogeneity test					0.000	0.000	0.061	0.406

First Stage IV Regressions	First Stage IM_NEW_int	First Stage IM_NEW_cap	First Stage IM_NEW_int	First Stage IM_NEW_cap	First Stage IM_NEW_int	First Stage IM_NEW_cap	First Stage IM_NEW_int_D1	First Stage IM_NEW_cap_D1
ln(MFN Interm Import Tariffs)_it-1	0.0152***		0.0152***		0.0152***		-0.0039***	
	(0.0005)		(0.0005)		(0.0005)		(0.0007)	
ln(MFN Capital Import Tariffs)_it-1		0.0331***		0.0331***		0.0331***		-0.0108***
		(0.0014)		(0.0014)		(0.0014)		(0.0019)
ln(GDP in PPP)_ct-1	-0.0162***	-0.0140***	-0.0162***	-0.0140***	-0.0162***	-0.0140***	0.0078**	0.0111
	(0.0027)	(0.0050)	(0.0027)	(0.0050)	(0.0027)	(0.0050)	(0.0037)	(0.0068)
ln(Starting a Business)_ct-1	0.0026	0.0033	0.0026	0.0033	0.0026	0.0033	-0.0013	-0.0027
	(0.0026)	(0.0050)	(0.0026)	(0.0050)	(0.0026)	(0.0050)	(0.0036)	(0.0068)
Constant	0.4750***	0.3920***	0.4750***	0.3920***	0.4750***	0.3920***	-0.2039**	-0.2740
	(0.0665)	(0.1231)	(0.0665)	(0.1231)	(0.0665)	(0.1231)	(0.0918)	(0.1687)

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: This table reports the two-stage regressions using MFN import tariffs as alternative instruments. Thus, the instrumental variables employed are the log of MFN intermediate import tariffs and the log of MFN capital import tariffs. The top panel shows the second stage IV regressions, while the bottom panel exhibits the first stage regressions. Columns (1)-(2) are estimated using an IV Poisson model, where the dependent variable is the number of new exported varieties; these columns report the coefficients. Columns (3)-(4) are estimated employing an IV Probit model, where the dependent variable is the probability of exporting new varieties; these columns report the average marginal effects. Columns (5)-(6) are estimated using a 2SLS model, where the dependent variable is the export value of new varieties. Finally, columns (7)-(8) are estimated using a log-first difference estimator in two stages, where the dependent variable is the net change in the log number of new exported varieties.

Chapter 5

Conclusion

5.1 Overall Aim of the Thesis

The overall aim of this thesis is to make different contributions to the existing trade literature on global value chains from the perspective of an emerging economy. Thus, we aim to study different mechanisms including FDI, trade processing, and technology embedded in imports. The thesis encompasses three empirical chapters on trade topics using advanced econometric techniques. These trade-related topics are especially interesting for Mexican policymakers in the context of the country's increasing participation and integration in global value chains, as well as for future negotiations of free and preferential trade agreements.

5.2 Summary of Findings and Policy Implications

This section summarizes the findings of each of the empirical chapters conforming this thesis. Furthermore, we also discuss the policy implications, as well as the limitations of these empirical chapters, and areas for future research.

5.2.1 FDI, Trade, and Business Cycle Comovements

The first empirical chapter examines the impact of FDI and bilateral trade on business cycle comovements from the perspective of an emerging economy. Although the trade literature on the determinants of business cycle comovements has acknowledged the role of FDI (Hsu et al. 2011, Jansen & Stokman 2014), this determinant has not been explored for developing countries. Thus, the aim of this first empirical chapter is to provide evidence that FDI also plays an important role on the transmission of business cycles for an emerging economy.

To address this relationship between FDI and business cycle comovements, we use a linear fixed effects model on our estimation sample of paired observations between 47 major partner countries and all 32 Mexican states over different time periods. We also make use of the novel Hamilton regression filter to detrend GDP time series employed on the construction of the dependent variable.

The findings of the first empirical chapter suggest that FDI inflows have a positive and statistically significant effect on business cycle comovements for an emerging economy.

From a policy perspective, we observe that Spain is mainly driving the impact of FDI on business cycle comovements with Mexican states. Thus, a policy recommendation would be to continue diversifying markets and to attract other investment partners so that in case of an economic downturn in Spain, the transmission of the business cycles smooths as Mexican states can rely on other partner countries.

Furthermore, we also observe state heterogeneity among Mexican states. On the one hand, we can notice northern states perform better in terms of GDP per capita, trade, and FDI inflows. On the other hand, we can observe the opposite on southern states with a low performance in all of these economic indicators, except for two states that highly rely on the oil and gas industry. Based on this state heterogeneity, it may be worth evaluating the creation of industry clusters in the southern region of the county to promote their economic growth. Moreover, it may be worth attracting FDI inflows from emerging economies to these southern states. By doing this, it may be possible to integrate southern states to global value chains. Thus, these southern states could benefit from FDI and trade spillover effects.

5.2.2 New Traded Varieties and Source Countries: Evidence of Trade Complementarities

The second empirical chapter focuses on the relationship between new imported varieties and exports of new varieties by revealing a degree of trade complementarities between countries. Even though the existing literature focuses on the impact of imported goods on exports, the bilateral trade component has been somehow neglected. This second empirical chapter aims to reveal trade complementarities at the country level. In other words, we show how importing new varieties from a source country impacts exports of new varieties to that same country.

To address this relationship between imports and exports of new varieties, we first use a decomposition exercise of the annual growth of traded varieties to identify the relevance of new varieties. Then, we use three different approaches consisting of logit, negative binomial, and linear models combined with fixed effects. The estimation sample in our analysis is composed of 74,240 new traded varieties in the manufacturing sector over the time lapse from 2005 to 2016.

The findings of the second empirical chapter suggest that importing new varieties has a positive and strong effect on the exports of new varieties. We also conclude that there is a degree of trade complementarity between imports and exports of new varieties at the country level. In other words, our results suggest that the trade effect magnifies on the source country instead of propagating homogeneously across countries.

From a policy perspective, this complementarity trade effect is stronger for Spain, Colombia, and China. These three partner countries constitute both main source and destination countries for new traded varieties with Mexico. In the case of Spain and Colombia, we could explain this trade complementarity effect by the knowledge about the country mechanism (e.g., common language) rather than by transportation costs. Thus, we could infer that Mexican firms tend to adapt their products to the taste, preferences, and regulations of their export markets.

Moreover, transportation costs and free trade agreements play an important role in the decision-making process of Mexican firms that export new varieties. Nonetheless, this does not hold when it comes to importing new varieties. Despite the tariffs associated with importing products from countries that do not benefit from free trade agreements, Mexican manufacturing firms consider it is still more profitable to import new varieties from Asian countries. Thus, expanding the free trade network to Asian countries could represent an interesting opportunity for Mexican firms to benefit from lower tariffs or even free trade on imports of new varieties.

5.2.3 New Imported Inputs, New Export Varieties: Capital Matters

The third empirical chapter analyzes the impact of new imported intermediate inputs and new imported capital goods on exports of new varieties. Although the empirical literature has widely studied the importance of importing intermediate inputs on exports (Aristei et al. 2013, Castellani & Fassio 2019, Feng et al. 2016, Lo Turco & Maggioni 2013, Navas et al. 2020), the role of importing capital goods have not been widely explored, except for Damijan et al. (2014). Therefore, this last empirical chapter aims to show the importance of importing new capital goods on exports of new varieties for an emerging economy.

To address the impact of new imported intermediate inputs and new imported capital goods on exports of new varieties, we first use a decomposition exercise to identify new traded varieties similar to the second empirical chapter. Then, we match these new varieties with UN Broad Economic Categories (BEC) to disentangle intermediate, capital, and consumption goods. Next, we use four different approaches consisting of negative binomial, logit, linear, and log-first difference regressions combined with fixed effects on our estimation sample; this estimation sample consists of 68,727 new traded varieties belonging to the manufacturing sector over the period between 2005 and 2016.

The findings of the third empirical chapter confirm a positive relationship between new imported intermediate inputs and new exported varieties. An interesting result is that we also unveil the importance of new imported capital goods on exports of new varieties. From a trade policy perspective, we have identified that some of the main source countries of new imported intermediate inputs and capital goods are located in Asia. We have to consider that imported goods from these Asian countries translate into higher costs due to tariffs. Therefore, negotiating free or preferential trade agreements with strategic trade partners located in Asia could represent a motivation for Mexican policymakers to promote further integration with this strategic region and consolidate Mexico's role in global value chains.

Nonetheless, these policy recommendations should be taken cautiously as some of these sectors, such as the textile and footwear industries, are quite sensitive, and a free trade agreement could negatively impact the domestic production of these goods. Therefore, it is worth highlighting the importance of making the appropriate consultation process with the relevant business chambers in Mexico to evaluate the benefits of negotiating free trade agreements with

Asian countries. Otherwise, it could be worth negotiating preferential agreements focused on specific sectors depending on the type of imported goods (i.e., intermediate inputs or capital goods). As a result, Mexico could strengthen its participation in global value chains.

5.3 Limitations and Future Research

The three empirical chapters presented in this thesis make interesting contributions to the empirical trade literature related to business cycle comovements, trade complementarities, and learning by importing. Nevertheless, each of these chapters present limitations that we further discuss in this section.

The first empirical chapter encounters data limitations in terms of the limited sample size of investing countries in Mexican states. Despite this small number of countries, the sample roughly covers 98% of FDI inflows received in Mexican states. Furthermore, Mexican authorities do not report data on imports at the state level due to confidentiality issues. Future research could be focused on revisiting this study having larger time series and having data at a more granular level; however, it is worth mentioning that trade data at the firm level is confidential for Mexico.

The second empirical chapter also presents some data limitations as the disaggregation level of trade datasets is not at the firm level. As mentioned above, the reason for not using such a granular level is because of confidentiality issues related to firm data. Despite this drawback, we employ product-level data, which is the finest level available. Another pitfall is that Mexican trade datasets do not clearly distinguish between intermediate, capital, and final goods. Understanding the role of new imported intermediate inputs and new imported capital goods represents an interesting research area, which is examined in the third empirical chapter.

The last empirical chapter faces two main limitations. First, Mexican trade datasets at the product level are no longer updated since 2016; thus, our analysis covers until that last available year. Second, trade-related variables sourced from the World Bank Doing Business dataset have a short time span covering the period from 2006 to 2015. Due to the methodology employed to identify new varieties and the nature of some of the estimators, we end up with significantly fewer observations. Despite these drawbacks, we incorporate these trade-related variables as part of the Robustness Analysis. Finally, it may be interesting to revisit this last study considering state heterogeneity. However, as mentioned before, there is no available data on imports at the state level.

References

Acharya, R. C. & Keller, W. (2009), 'Technology transfer through imports', *Canadian Journal of Economics* **42**(4), 1411–1448.

Alviarez, V., Cravino, J. & Levchenko, A. A. (2017), 'The growth of multinational firms in the Great Recession', *Journal of Monetary Economics* **85**, 50–64.

Amiti, M. & Konings, J. (2007), 'Trade liberalization, intermediate inputs, and productivity: Evidence from Indonesia', *American Economic Review* **97**(5), 1611–1638.

Aristei, D., Castellani, D. & Franco, C. (2013), 'Firms' exporting and importing activities: Is there a two-way relationship?', *Review of World Economics* **149**(1), 55–84.

Arkolakis, C., Demidova, S., Klenow, P. J. & Rodríguez-Clare, A. (2008), 'Endogenous variety and the gains from trade', *American Economic Review* **98**(2), 444–450.

Backus, D. K., Kehoe, P. J. & Kehoe, T. J. (1992), 'In search of scale effects in trade and growth', *Journal of Economic Theory* **58**(2), 377–409.

Bahar, D., Hausmann, R. & Hidalgo, C. A. (2014), 'Neighbors and the evolution of the comparative advantage of nations: Evidence of international knowledge diffusion?', *Journal of International Economics* **92**(1), 111–123.

Bas, M. & Strauss-Kahn, V. (2014), 'Does importing more inputs raise exports? Firm-level evidence from France', *Review of World Economics* **150**(2), 241–275.

Bergin, P. R., Feenstra, R. C. & Hanson, G. H. (2009), 'Offshoring and volatility: Evidence from Mexico's maquiladora industry', *American Economic Review* **99**(4), 1664–1671.

Bernard, A. B., Redding, S. J. & Schott, P. K. (2010), 'Multiple-product firms and product switching', *American Economic Review* **100**(1), 70–97.

Boehm, C. E., Flaaen, A. & Pandalai-Nayar, N. (2019), 'Input Linkages and the Transmission of Shocks: Firm-Level Evidence from the 2011 Tohoku Earthquake', *Review of Economics and Statistics* **101**(1), 60–75.

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

Broda, C. & Weinstein, D. E. (2010), 'Product creation and destruction: Evidence and price implications', *American Economic Review* **100**(3), 691–723.

Burstein, A., Kurz, C. & Tesar, L. (2008), 'Trade, production sharing, and the international transmission of business cycles', *Journal of Monetary Economics* **55**(4), 775–795.

Calderón, C., Chong, A. & Stein, E. (2007), 'Trade intensity and business cycle synchronization: Are developing countries any different?', *Journal of International Economics* **71**(1), 2–21.

Cameron, A. C. & Trivedi, P. K. (1990), 'Regression-based tests for overdispersion in the Poisson model', *Journal of Econometrics* **46**(3), 347–364.

Canova, F. (1998), 'Detrending and business cycle facts', *Journal of Monetary Economics* **41**, 475–512.

Caselli, F., Koren, M., Lisicky, M. & Tenreyro, S. (2020), 'Diversification Through Trade', *Quarterly Journal of Economics* **135**(1), 449–502.

Caselli, M. (2018), 'Do all imports matter for productivity? Intermediate inputs vs capital goods', *Economia Politica* **35**(2), 285–311.

Castellani, D. & Fassio, C. (2019), 'From new imported inputs to new exported products. Firm-level evidence from Sweden', *Research Policy* **48**(1), 322–338.

Castellani, D., Serti, F. & Tomasi, C. (2010), 'Firms in international trade: Importers' and exporters' heterogeneity in Italian manufacturing industry', *World Economy* **33**(3), 424–457.

Clark, T. E. & Van Wincoop, E. (2001), 'Borders and business cycles', *Journal of International Economics* **55**(1), 59–85.

Clerides, S. K., Lach, S. & Tybout, J. R. (1998), 'Is learning by exporting important? Micro-dynamic evidence from Colombia, Mexico, and Morocco', *Quarterly Journal of Economics* **113**(3), 903–947.

Coe, D. T., Helpman, E. & Hoffmaister, A. W. (1997), 'North-south R&D spillovers', *Economic Journal* **107**(440), 134–149.

Coe, D. T., Helpman, E. & Hoffmaister, A. W. (2009), 'International R&D spillovers and institutions', *European Economic Review* **53**(7), 723–741.

Colantone, I. & Crinò, R. (2014), 'New imported inputs, new domestic products', *Journal of International Economics* **92**(1), 147–165.

Corbae, D., Ouliaris, S. & Phillips, P. C. (2002), 'Band spectral regression with trending data', *Econometrica* **70**(3), 1067–1109.

Cravino, J. & Levchenko, A. A. (2017), 'Multinational firms and international business cycle transmission', *Quarterly Journal of Economics* **132**(2), 921–962.

Damijan, J. P., Konings, J. & Polanec, S. (2013), 'Pass-on trade: Why do firms simultaneously engage in two-way trade in the same varieties?', *Review of World Economics* **149**(1), 85–111.

Damijan, J. P., Konings, J. & Polanec, S. (2014), 'Import Churning and Export Performance of Multi-product Firms', *World Economy* **37**(11), 1483–1506.

Damijan, J. P. & Kostevc, Č. (2015), 'Learning from Trade through Innovation', *Oxford Bulletin of Economics and Statistics* **77**(3), 408–436.

De Hoyos, R. E. & Iacovone, L. (2013), 'Economic Performance under NAFTA: A Firm-Level Analysis of the Trade-productivity Linkages', *World Development* **44**(1), 180–193.

De Loecker, J. (2013), 'Detecting learning by exporting', *American Economic Journal: Microeconomics* **5**(3), 1–21.

Dean, J. M., Fung, K. C. & Wang, Z. (2011), 'Measuring vertical specialization: The case of China', *Review of International Economics* **19**(4), 609–625.

Di Giovanni, J. & Levchenko, A. A. (2010), 'Putting the parts together: Trade, vertical linkages, and business cycle comovement', *American Economic Journal: Macroeconomics* **2**(2), 95–124.

Di Giovanni, J. & Levchenko, A. A. (2012), 'Country Size, International Trade, and Aggregate Fluctuations in Granular Economies', *Journal of Political Economy* **120**(6), 1083–1132.

Di Giovanni, J., Levchenko, A. A. & Mejean, I. (2017), 'Large firms and international business cycle comovement', *American Economic Review* **107**(5), 598–602.

Eaton, J. & Kortum, S. (2001), 'Trade in capital goods', *European Economic Review* **45**(7), 1195–1235.

Feng, L., Li, Z. & Swenson, D. L. (2016), 'The connection between imported intermediate inputs and exports: Evidence from Chinese firms', *Journal of International Economics* **101**, 86–101.

Fernández, J. & Gavilanes, J. C. (2017), 'Learning-by-importing in emerging innovation systems: evidence from Ecuador', *Journal of International Trade and Economic Development* **26**(1), 45–64.

Frankel, J. A. & Rose, A. K. (1998), 'The Endogenity of the Optimum Currency Area Criteria', *Economic Journal* **108**(449), 1009–1025.

Gabaix, X. (2011), 'The Granular Origins of Aggregate Fluctuations', *Econometrica* **79**(3), 733–772.

Garetto, S., Oldenski, L. & Ramondo, N. (2016), 'The Dynamics of Multinational Activity: Evidence from U.S. Firms', *Society for Economic Dynamics, 2016 Meeting Papers 1431* .

Goldberg, P. K., Khandelwal, A. K., Pavcnik, N. & Topalova, P. (2010*a*), 'Imported intermediate inputs and domestic product growth: Evidence from India', *Quarterly Journal of Economics* **125**(4), 1727–1767.

Goldberg, P. K., Khandelwal, A. K., Pavcnik, N. & Topalova, P. (2010*b*), 'Multiproduct firms and product turnover in the developing world: Evidence from India', *Review of Economics and Statistics* **92**(4), 1042–1049.

Goldberg, P., Khandelwal, A., Pavcnik, N. & Topalova, P. (2009), 'Trade Liberalization and New Imported Inputs', *American Economic Review* **99**(2), 494–500.

Gopinath, G. & Neiman, B. (2014), 'Trade adjustment and productivity in large crises', *American Economic Review* **104**(3), 793–831.

Gourieroux, C., Monfort, A. & Trognon, A. (1984), 'Pseudo Maximum Likelihood Methods: Applications to Poisson Models', *Econometrica* **52**(3), 701–720.

Greene, W. W. H. . (2003), *Econometric Analysis 5th Edition*.

Guarascio, D. & Pianta, M. (2017), 'The gains from technology: new products, exports and profits', *Economics of Innovation and New Technology* **26**(8), 779–804.

Hamilton, J. (2017), 'Why you should never use the Hodrick-Prescott filter', *Review of Economics and Statistics* **100**(5), 831–843.

Hanson, G. H. (1997), 'The Effects of Offshore Assembly on Industry Location: Evidence from U.S. Border Cities', *NBER Working Paper No. 5400* pp. 1–35.

Hausmann, R. & Hidalgo, C. A. (2011), 'The network structure of economic output', *Journal of Economic Growth* **16**(4), 309–342.

Helpman, E. (1984), 'A Simple Theory of International Trade with Multinational Corporations', *Journal of Political Economy* **92**(3), 451–471.

Helpman, E., Melitz, M. J. & Yeaple, S. R. (2004), 'Export versus FDI with heterogeneous firms', *American Economic Review* **94**(1), 300–316.

Hsu, C. C., Wu, J. Y. & Yau, R. (2011), 'Foreign direct investment and business cycle comovements: The panel data evidence', *Journal of Macroeconomics* **33**(4), 770–783.

Imbs, J. (2004), 'Trade, Finance, Specialization, and Synchronization', *Review of Economics and Statistics* **86**(3), 723–734.

Isogawa, D., Nishikawa, K. & Ohashi, H. (2012), 'New-to-Market Product Innovation and Firm Performance: Evidence from a firm-level innovation survey in Japan', *RIETI Discussion Paper Series* (12-E-077).

Jansen, W. J. & Stokman, A. C. J. (2014), 'International business cycle co-movement: the role of FDI', *Applied Economics* **46**(4), 383–393.

Javorcik, B. S. (2004), 'Does foreign direct investment increase the productivity of domestic firms? in search of spillovers through backward linkages', *American Economic Review* **94**(3), 605–627.

Kandilov, I. & Leblebicioğlu, A. (2012), 'Trade liberalization and investment: Firm-level evidence from Mexico', *World Bank Economic Review* **26**(2), 320–349.

Kasahara, H. & Lapham, B. (2013), 'Productivity and the decision to import and export: Theory and evidence', *Journal of International Economics* **89**(2), 297–316.

Kleinert, J., Martin, J. & Toubal, F. (2015), 'The few leading the many: Foreign affiliates and business cycle comovement', *American Economic Journal: Macroeconomics* **7**(4), 134–159.

Koopman, R., Wang, Z. & Wei, S. J. (2012), 'Estimating domestic content in exports when processing trade is pervasive', *Journal of Development Economics* **99**(1), 178–189.

Koren, M. & Csillag, M. (2011), 'Machines and machinists: Capital-Skill Complementarity from an International Trade Perspective', *CEPR Discussion Paper No. 8317* .

Kugler, M. & Verhoogen, E. (2009), 'Plants and imported inputs: New facts and an interpretation', *American Economic Review* **99**(2), 501–507.

Kurz, C. & Senses, M. Z. (2016), 'Importing, exporting, and firm-level employment volatility', *Journal of International Economics* **98**, 160–175.

Lawless, M. (2010), 'Deconstructing gravity: trade costs and extensive and intensive margins', *Canadian Journal of Economics/Revue canadienne d'économique* **43**(4), 1149–1172.

Lo Turco, A. & Maggioni, D. (2013), 'On the Role of Imports in Enhancing Manufacturing Exports', *World Economy* **36**(1), 93–120.

Navas, A., Serti, F. & Tomasi, C. (2020), 'The role of the gravity forces on firms' trade', *Canadian Journal of Economics* **53**(3), 1059–1097.

OCDE (2013), 'Interconnected Economies: Benefiting from Global Value Chains', *OECD Publishing* pp. 1–269.
URL: *http://dx.doi.org/10.1787/9789264189560-en*

OECD (2013), 'Global value chains: Mexico', *OECD Publishing* pp. 1–4.
URL: *http://www.oecd.org/sti/ind/GVCs - MEXICO.pdf*

OECD (2017), 'The Size and Sectoral Distribution of State-Owned Enterprises', *OECD Publishing* pp. 1–79.
URL: *https://doi.org/10.1787/9789264280663-en*

Ouliaris, S., Tan, S. M. & Chew, J. (2014), 'Exchange Rate Pass-Through Over the Business Cycle in Singapore', *IMF Working Papers* **11**(141).

Poncet, S. & Mayneris, F. (2013), 'French firms penetrating Asian markets: Role of export spillovers', *Journal of Economic Integration* **28**(2), 354–374.

Rivera-Batiz, L. A. & Romer, P. M. (1991), 'Economic Integration and Endogenous Growth', *The Quarterly Journal of Economics* **106**(2), 531–555.

Sandqvist, A. P. (2017), 'Dynamics of sectoral business cycle comovement', *Applied Economics* **49**(47), 4742–4759.

Santos Silva, J. M. & Tenreyro, S. (2010), 'On the existence of the maximum likelihood estimates in Poisson regression', *Economics Letters* **107**(2), 310–312.

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

Saviotti, P. P. & Frenken, K. (2008), 'Export variety and the economic performance of countries', *Journal of Evolutionary Economics* **18**(2), 201–218.

Schüler, Y. (2018), 'On the cyclical properties of Hamilton's regression filter', *Deutsche Bundesbank Discussion Paper No. 03/2018* .

StataCorp. (2017), 'Stata: Release 15', *Statistical Software. College Station, TX: StataCorp LLC.* .

Topalova, P. & Khandelwal, A. (2011), 'Trade liberalization and firm productivity: The case of India', *Review of Economics and Statistics* **93**(3), 995–1009.

Vogel, A. & Wagner, J. (2010), 'Higher productivity in importing German manufacturing firms: Self-selection, learning from importing, or both?', *Review of World Economics* **145**(4), 641–665.

Xiang, C. (2005), 'New goods and the relative demand for skilled labor', *Review of Economics and Statistics* **87**(2), 285–298.

Xu, J. & Mao, Q. (2018), 'On the relationship between intermediate input imports and export quality in China', *Economics of Transition* **26**(3), 429–467.

Yasar, M. (2013), 'Imported capital input, absorptive capacity, and firm performance: Evidence from firm-level data', *Economic Inquiry* **51**(1), 88–100.

Zhou, S., Wang, Y. & Yang, M. (2020), 'Importing and Firm Productivity in China: The Self-Selection Effect or the Learning Effect?', *Emerging Markets Finance and Trade* **56**(7), 1462–1473.

Zlate, A. (2016), 'Offshore production and business cycle dynamics with heterogeneous firms', *Journal of International Economics* **100**, 34–49.