On the variety of Mexico's export goods

INSTITUTE OF DEVELOPING ECONOMIES

IDE Discussion Papers are preliminary materials circulated
to stimulate discussions and critical comments

IDE DISCUSSION PAPER No. 510

On the Variety of Mexico's Export Goods

Tadashi ITO *

March 2015

Abstract

This paper examines the evolution of the variety of Mexico's export goods using disaggregated trade data. Both the econometric estimation analyses using the raw data and the one using an improved version of Feenstra and Kee's (2004, 2007) methodology proposed in this paper show that NAFTA membership does not enhance the variety of Mexico's export goods. This finding contrasts with NAFTA's positive association with the increase in export variety found in the literature.

Keywords: Mexico, NAFTA, Export goods variety
JEL classification: F14, F15

---

* Director, Technological Innovation and Economic Growth Studies Group, Inter-Disciplinary Studies Centre, IDE (tadashi_itoATide.go.jp) – please change AT to @.

The Institute of Developing Economies (IDE) is a semigovernmental, nonpartisan, nonprofit research institute, founded in 1958. The Institute merged with the Japan External Trade Organization (JETRO) on July 1, 1998. The Institute conducts basic and comprehensive studies on economic and related affairs in all developing countries and regions, including Asia, the Middle East, Africa, Latin America, Oceania, and Eastern Europe.

INSTITUTE OF DEVELOPING ECONOMIES (IDE), JETRO
3-2-2, WAKABA, MIHAMA-KU, CHIBA-SHI
CHIBA 261-8545, JAPAN

On the Variety of Mexico's Export Goods

Tadashi Ito◆

March 2015

Abstract

This paper examines the evolution of the variety of Mexico's export goods using disaggregated trade data. Both the econometric estimation analyses using the raw data and the one using an improved version of Feenstra and Kee's (2004, 2007) methodology proposed in this paper show that NAFTA membership does not enhance the variety of Mexico's export goods. This finding contrasts with NAFTA's positive association with the increase in export variety found in the literature.

Keywords: NAFTA, Export goods variety

JEL Classification: F14, F15

◆Institute of Developing Economies, Japan, E-mail: Tadashi_Ito@ide.go.jp. I would like to thank seminar participants at the Graduate Institute, in Geneva and Keio University in Tokyo for their valuable comments. This paper is forthcoming at *Estudios Económicos.*

Introduction

Since the 1960s, following the Prebisch–Singer hypothesis of terms of trade deterioration, policymakers in developing countries have been aiming to increase the variety of goods that they export. Several Asian countries such as South Korea and Taiwan are examples of remarkable export-led growth. Some economists have found an association between the degree of export diversification and economic growth. For example, Sachs and Warner (1995) suggest a correlation between low export diversification and slow growth.[1] However, despite the desire of policymakers to increase the variety of exported goods, until recently trade economists had not paid much attention to this issue, mainly because classical trade theories focus on the advantages of *specialization* into products in which countries have comparative advantages rather than *diversification* or increasing the variety of export goods.[2]

However, the theoretical contribution by Melitz (2003) triggered a rapid increase in research on export diversification. His model is referred to in the literature as the heterogeneous firms trade model. It shows that a reduction in trade barriers allows an initially non-exporting firm to become an exporter. Melitz (2003) also incorporated firm heterogeneity into Krugman's (1980) model of trade under increasing returns and monopolistic competition. Firms differ in their levels of productivity and thus their marginal costs, while they also incur sales-related costs in both domestic and export markets. Such sales-related costs are assumed to be higher for export markets than for domestic markets.

---

[1] More recently, using disaggregated panel data, Imbs and Wacziarg (2003); Klinger and Lederman (2004); and Cadot, Carrère,and Strauss-Kahn (2007) showed that poor countries tend to have lower degrees of export diversification.

[2] There is no single established definition on the terminology of "export diversification" and "export variety" in the literature. This paper uses "export diversification" and "export variety" interchangeably for the same meaning. However, diversification (variety) has two dimensions: diversification (variety) of *destinations* and diversification (variety) of *goods*. The analysis in Section 2 deals with export diversification (variety) of destinations *and* goods, while the analysis in Section 3 deals with export diversification of goods only because of the use of US import data, as in Feenstra and Kee (2007).

Firms that have sufficiently high productivity levels can thus absorb export-related costs and can be active in both domestic and export markets. By contrast, firms that have low levels of productivity cannot be active even in domestic markets, while those that have medium levels serve only the domestic market. The borders among these three types of firms are determined by two cut-off productivity conditions.

This paper focuses on the case of Mexico, which has undergone a series of large-scale trade liberalization activities culminating in the formation of NAFTA. Despite the apparent positive effect of NAFTA on the Mexican economy shortly after its formation (Lederman, Maloney and Serven, 2004, offered a generally positive assessment of NAFTA's effect on the Mexican economy), economic performance over the past decade in Mexico has remained "lacklustre, trailing that of many other developing nations" (Hanson, 2010). It is thus worth revisiting NAFTA's impact on the Mexican economy. Specifically, this paper focuses on the change in the variety of Mexico's export goods.

The contributions of this paper are threefold. First, it shows the trends in the goods exported to major US trade partners. Second, it uses disaggregated trade data to find that Mexico's unilateral trade liberalization has increased the possibility of any particular good being exported. Third, it points out a potential problem when selecting a comparison country for the computation of the variety index (Feenstra and Kee, 2007) and thus proposes a variety index with a multi-county base. It then carries out an econometric analysis using this improved version and shows that NAFTA is not associated with an increase in the variety index, which is at odds with the findings presented in the literature.

One of the early papers related to the topic of this paper was Helpman, Melitz and Rubinstein (2008), which developed a theory-based gravity model of trade that predicts *positive as well as zero trade flows* across pairs of countries. Using total trade value data for

158 countries in 1986, this paper showed that among 24,806 possible bilateral export relationships, only 11,146 pairs have non-zero exports. More importantly, it found that the usual gravity equation variables also affect the *probability* of the occurrence of a bilateral export relationship.

While Helpman, Melitz and Rubinstein (2008) studied the export relationships between country pairs, others have examined the extensive margins or trade value generated by the sale of new goods. Kehoe and Ruhl (2009), for example, analyzed the relation between trade liberalization and extensive margins and find an overall positive association. This paper studied NAFTA, but did not conduct an econometric analysis to explain its effects. Moreover, its analysis of the Mexico–US trade agreement used data from 1989, five years before the formation of NAFTA. Baier, Bergstran and Feng (2011) employed a panel econometric methodology to examine the effects of economic integration agreements on aggregate trade flows using a gravity equation model and finds economically and statistically significant effects on both intensive and extensive margins. Feenstra and Kee (2007) proposed an index for export variety and found a positive association between NAFTA membership and an increase in the export variety index.

Our paper departs from the literature in three aspects. First, unlike the studies mentioned above, we use raw disaggregated data to identify increases in the number of product categories exported. Second, we acknowledge a potential problem in the selection of a comparison base (Feenstra and Kee, 2007) and propose a revised version of the index. Finally, econometric analysis we present shows that NAFTA is not associated with an increase in the variety index.

Section 1 describes the evolution of zeros. Section 2 presents the econometric analyses using raw disaggregated data. The variety index is discussed in Section 3. The final section concludes.

1. The Evolution of Zeros

Recent studies of the exports of new goods show a very large number of zeros in the world trade matrix. According to Baldwin and Harrigan (2011, page 72), "the United States imported in nearly 17,000 different 10-digit HS categories from 228 countries, for a total of over 3.8 million potential trade flows. Over 90% of these potential trade flows are zeros." Hummels and Klenow (2005) further showed that "60 percent of the greater export of larger economies in their sample of 126 countries is due to the increase of the number of exported products."

Mexico was the third largest import partner for the US in the period 1989–2006. For Mexico, the US is by far its largest trade partner both for imports and for exports. The share of the US in Mexico's exports has steadily increased and reached 89% in 2006. We analyze Mexico's exports using US import data since they are available at the most disaggregated level of classification, namely 10-digit HS codes. The data come from Feenstra and Kee (2004) and cover the years 1989 to 2006.[3] The 10-digit HS codes during this period cover, in total, 24,818 items.[4]

3 US trade data at the 10-digit level, which the US Census Bureau makes publicly available, are only usable from 1992, whereas The Center for International Data at UC Davis makes them freely available from 1989 to 2006 at http://cid.econ.ucdavis.edu/data/sasstata/usiss.html. It also makes available US trade data with a US trade statistics code at the seven-digit level for 1972–1988.

4 This number differs from the "nearly 17,000" of Baldwin and Harrigan (2008) since they count those categories that registered a positive import value from at least one country in a single year. Here, 24,818 is the number of 10-digit HS categories that had imports from at least one country during the period 1989–2006.

Table 1 shows the evolution of the number of zeros in US imports from its top 20 import partners plus Colombia and Chile, from 1989 to 2006. The order of the countries in the table represents the ranking of exports into the US over this period, except Colombia (ranked 29) and Chile (ranked 40).[5] Mexico ranks number three following Canada and Japan. There were 19,105 zeros in 1989, which decreased to 15,993 in 2006. The last row in Table 1 shows the reduction in zeros during the investigated period. The number for Mexico (-3112) is second only to China, a huge country that is rapidly and dramatically increasing its exports across the world. In this sense, China can be considered to be an exceptional case. Mexico had more zeros, or non-exported items, than Korea and Taiwan in 1989. In 2006, the opposite was true. In other words, Mexico has surpassed Korea and Taiwan in terms of the variety of its exports to the US, and is close to the level of France.

======= Table 1 =======

However, it is important to seek data from well before 1986 in order to assess the evolution in the number of exported goods during Mexico's unilateral trade liberalization from 1986 and during the formation of NAFTA from 1994. Although the 10-digit HS data are only available from 1989, US trade data (with a trade statistics seven-digit code) are available from 1972 to 1988. Table 2 shows the evolution of zeros for 1972–1988. The decrease in the number of zeros for Mexico is lower than for the other major US import partners during this period. China registers the largest decrease in the number of zeros, but this decrease is less impressive than the decrease it obtained during the 1989–2006 period.

======= Table 2 =======

5 We included Colombia and Chile as good candidates for comparison with Mexico since both are Latin American countries and their economic sizes are similar. The total number of countries analyzed was limited by the technical limitations of the software used, which was STATA MP Quadcore 64 bit with a 16 GB RAM computer.

Since the absolute level of the number of zeros in the initial year of each of the time series (i.e., 1989 and 1972, respectively) varies by country, we indexed the number of zeros by taking 1989 as the base year for the 1989–2006 period and 1988 as the base year for the 1972–1988 period in order to permit comparison. The result is shown in Figure 1.[6] Imports from China underwent the largest decrease in this zero index, while Canada also displayed a steady and rapid decrease in the index. The decrease in the index of Mexico from the 1970s to the first half of the 1980s was modest compared with the other major US trade partners, but it accelerated from around 1985 to 1997 and retained this higher rate of decrease thereafter. Although the US HS 10-digit trade data at from the Center for International Data at UC Davis are available only up to 2006, the number of zeros between 2006 and 2012 may give some useful information, especially because of the crisis in 2008-2009 and subsequent recovery. Thus Figure 2 shows the zero index (using 1991 as the base year) of US imports from its major import partners at HS 6-digit level from 1991 to 2012. There is a clear rise in the number of zeros during the crisis and some recovery after the crisis. This may indicate that business cycles have a stronger effect on trade flows than do trade agreements such as NAFTA.[7] To see the potential effects of economic downturn during the crisis from the side of Mexico's export, Figure 3 shows the zero index of Mexico's exports at HS 6-digit to its major markets.[8] As in the previous case, the zero index rose in the crisis period, especially for the developed countries, notably the US, which were bitterly hit by the crisis.

======= Figure 1, Figure 2, and Figure 3 =======

6 This shows only the index change for the top 10 US trade partners for simplicity.

7 To avoid clutter but still make possible a sensible comparison, only some of the major US import partner countries are chosen.

8 Mexico's export data at HS 6-digit are available only from 1990. To avoid clutter but still make possible a sensible comparison, only some of the major export destination countries within its 20 largest partners are chosen.

2. ECONOMETRIC ANALYSIS USING RAW TRADE DATA

This section performs an econometric analysis using raw trade data to examine whether the probability of a particular product being exported is associated with Mexico's periods of trade liberalization. As noted in the previous section, it is imperative to use trade data from a sufficiently long period of time before 1986 in order to appropriately capture the effects of the two major series of Mexico's trade liberalization: the unilateral trade liberalization from 1986 and inception of NAFTA from 1994. It is logically straightforward to assume that NAFTA may have had a positive impact on Mexico's export variety because the US eliminated its import tariffs on Mexican goods. On the other hand, Mexico's unilateral trade liberalization was an initiative on the side of Mexico to reduce its import tariffs, and thus did not directly work to reduce its trade partners' import tariffs. However, this initiative may have increased Mexico's export diversification through two channels. The first channel is due to better access to affordable inputs, which may eventually have led to an increase of export variety. Indeed, the Mexican government's first action in its series of unilateral trade liberalization was the elimination of license requirements, official import prices, and quantitative restrictions, in order to improve Mexican products competitiveness.[9] The other is Mexico's accession to the General Agreement on Tariffs and Trade (GATT) in 1987, which itself was made possible by Mexico's unilateral trade liberalization in 1986. We use five-digit SITC trade data, which are the only data available with consistent product codes for a sufficiently long period.[10] We use Mexico's export data on the 50 largest export

9 A Mexican refrigerator manufacturer had once opposed to the formation of NAFTA, but eventually became one of the largest refrigerator suppliers in the world by procuring high quality compressors from the US. This anecdote is a well-known example of positive effects of imported intermediate inputs on productivity. (Amiti and Konings (2007).

10 HS data (e.g., at the 10-digit level for the US) are only available from 1989, while US trade data with a US trade statistics code at the seven-digit level for 1972–1988 have different code systems, which precludes connecting the two data sets. Detailed explanation on trade data codes is in the appendix.

destinations at the five-digit SITC level for the longest possible date range, namely 1962–2010. The following equation is then estimated using a Probit model:

$$\Pr(y=1|x)=\int_{-\infty}^{\beta' x}\phi(v)dv=\Phi(\beta' x)$$

where y takes 1 when the dependent variable (i.e., the trade value) takes a positive number, and 0 otherwise. x is the vector of explanatory variables, namely the GDP of destination countries; the distance between Mexico and destination countries; the NAFTA dummy, which takes 1 if the destination country is the US and the years are on or after 1994, and 0 otherwise; the Mexico unilateral liberalization dummy, which takes 1 if the years are on or after 1986; the common language dummy; and the dummies for years, destinations, and two-digit SITC codes.[11] β' is the vector of parameters for these variables. $\phi(v)$ is a standard normal density function.

The summary statistics are in Table 3 and the estimation results are in Table 4. The large number of observations -more than 2 million- comes from 48 years times 50 partner countries times approximately 1000 SITC codes. The first column only includes the NAFTA dummy, while the second column includes both the NAFTA dummy and Mexico's unilateral liberalization dummy. The variable of interest, NAFTA, shows negative and statistically highly significant coefficients, -0.147. Contrary to the sign of NAFTA, Mexico's unilateral liberalization dummy shows a statistically significant positive coefficient with a relatively large magnitude, 0.781. Namely, this estimation result suggests that NAFTA is *negatively* associated with an increase in the probability of a particular product being exported while Mexico’s unilateral liberalization is positively associated.

---

11 Including a more disaggregated SITC dummy, such as four-digit codes, exceeded the capacity of the software/computer at hand.

======= Table 3 and Table 4 =======

However, these results might have occurred because of the small change in the number of zeros after 2000, as shown in Table 1 and Figure 1. Thus, the same estimation was carried out for 1972–2001 as a robustness check (i.e., to make it comparable with the 1972–2001 study period of Feenstra and Kee, 2007). The results in Table 5 still show the statistically significant negative coefficient estimate for the NAFTA dummy and the statistically significant positive coefficient estimate for Mexico's unilateral liberalization dummy, although the magnitude is much attenuated in the latter. The smaller coefficient estimate for Mexico's unilateral liberalization dummy during the shorter period of study (1972–2001) seems reasonable because the number of zero trade after 2002 declined only slightly, as is shown above, reducing the relative positive impact of the unilateral liberalization.

======= Table 5 =======

Another issue which should be considered is that the US has always been the main destination of the Mexican exports. Therefore, Mexico's unilateral liberalization might have coincided with tariff reduction of the US on Mexican products. Unfortunately, the US tariff data are available only from 1989, which hinders an econometric analysis. However, the duty amounts collected at the US custom office are available. Thus, as measures of the US average tariffs, we have computed the ratio of duties collected to dutiable imports and the ratio of duties collected to total imports.[12] Table 6 shows the US average tariffs across all import partners. The average tariffs are almost constant in the 1980s to the beginning of the 1990s. Table 7 shows the ratio of duties collected to dutiable imports, while Table 8 shows the ratio of duties collected to total imports by the US top 10 import partners. In both cases,

---

[12] This method of estimating average tariffs is also used in Baldwin (2010).

the tariffs on Mexican products in the 1980s to the beginning of the 1990s are almost constant. Given this evidence, we can discard the possibility that the US tariff reduction on Mexican products was the real cause of the positive impact of Mexico's unilateral trade liberalization documented above. These results contrast with previous findings of NAFTA's positive association with diversification or an increase in the variety of Mexico's exports.

======= Table 6, Table 7 and Table 8 =======

Because this effect might be different across sectors, the same estimation was carried out for the machinery sector, which typically has the largest trade values.[13] The estimation results in Table 9 (for products in Machinery sector only) show the statistically significant negative coefficient for NAFTA dummy, -0.192, which is similar in its magnitude with the case in Table 4 (for products in all industries), -0.147, and also shows the statistically significant positive coefficient for Mexico's unilateral liberalization, 0.818, which is close to 0.781 in Table 4. Table 10 (for 1972–2001) shows the coefficient estimates for NAFTA and Mexico's unilateral liberalization with the expected signs and the smaller magnitude of Mexico's unilateral liberalization, being consistent with the case of the products in all industries (Table 4 and Table 5). The other control variables have the expected signs, except for the distance variable in Table 9, which is probably caused by the rapid decrease of zero trade (higher incidence of a product being exported) with China and Chile, which are distant from Mexico, as can be seen in Figure 3.

======= Table 9 and Table 10 =======

---

[13] I appreciate the suggestions of Fukunari Kimura on this point.

Panel unit root tests

When estimating a gravity model using panel data, there is a potential problem caused by nonstationarity (Quah, 1994). Zwinkels and Beugelsdijk (2010) point out the lack of treatment for nonstationarity in the gravity model literature and argue that ignoring nonstationarity can lead to overestimated coefficients. Thus, we performed panel unit root tests on our data. Among the several tests available for panel unit root tests, we used Im et al. (2003) for reasons discussed in the appendix.

Although we planned to perform the test for the whole panel, the technical limitations of the statistical software[14] caused us to divide the panel data according to two-digit HS code instead. For the import value variable (i.e., the dependent variable), the null hypothesis of all the series having unit roots was rejected at the 0.1% level for each of the 99 sub-panels except one (HS66), for which the null was still rejected at the 1% level. We did not need to divide the GDP data into sub-groups because they vary only by year and by country and thus did not exceed the capacity of the software. The null hypothesis of unit root was rejected at the 1% significance level. Given these results, our data were shown to have no nonstationarity issues.

3. VARIETY INDEX

Although a simple count of products is intuitive, this approach suffers from a lack of underlying theories. A theory-based methodology for measuring trade variety was proposed by Feenstra and Kee (2007), which draws on Feenstra (1994), which has been widely

14 Since we have a three-dimensional panel (i.e., time, country, and product), we need to generate a panel ID variable, which is a combination of country and product. Since the number of combinations of country and product is huge, it exceeds the storage capacity of even a 16 GB RAM computer.

employed by other researchers, including Hummels and Klenow (2005) and Broda and Weinstein (2006). This section introduces Feenstra and Kee's variety index, but points out a potential problem in the selection of a comparison country. Our estimation results using an improved index show that NAFTA membership is not positively associated with the variety of Mexico's exports, which is at odds with the result found in Feenstra and Kee (2007).

Feenstra and Kee's variety index

Consider the set of exports from countries a and c. They differ but have some product varieties in common. We denote this common set by $I \equiv \left(I_t^a \cap I_t^c\right) \neq \varnothing$. Feenstra and Kee (2004, 2007) show that the variety index of country c compared with that of the base country a at time t , $\psi_{a,t}^{c,t}$ can be computed as

$$\psi_{a,t}^{c,t} = \frac{\lambda_t^a\left(I\right)}{\lambda_t^c\left(I\right)} \text{ , where } \lambda_t^c\left(I\right) \equiv \frac{\sum_{i\in I} p_{it}^c q_{it}^c}{\sum_{i\in I_{it}^c} p_{it}^c q_{it}^c} \quad \text{and} \quad \lambda_t^a\left(I\right) \equiv \frac{\sum_{i\in I} p_{it}^a q_{it}^a}{\sum_{i\in I_{it}^a} p_{it}^a q_{it}^a}$$

When the values of the products exported only by the base country a , $\sum_{i\in I_{it}^a} p_{it}^a q_{it}^a$, is high, holding the others constant, the variety index is low. When the values of the products exported only by the country c , $\sum_{i\in I_{it}^c} p_{it}^c q_{it}^c$, is high, holding the others constant, the variety index is high. This is the variety index at a certain point in time by country. In addition to this cross-country aspect, the variety index changes over time. In other words, we have two dimensions: one of the cross-country and the other of the time-series.

By using US trade data at the 10-digit HS code level for 1989–2001 and US trade statistics for 1972–1988, Feenstra and Kee (2007) compute Mexico's export variety for 1972–2001 based on worldwide exports from all countries to the US averaged over time as the comparison base. They compute the variety index only for Mexico in seven industry groups and run regressions using the NAFTA dummy.

We argue that the variety indexes should be computed not only for Mexico but also for other countries in order to assess whether the NAFTA dummy shows any association with the variety index, because this dummy captures the effects specific to Mexico and the years after 1994 rather than industry-specific trade policy. Moreover, the index numbers change depending on which base is taken as the reference case. While Feenstra and Kee (2007) use only one base, following the convention in the index number problem, we compute an index of a particular country at a particular time with each country and each year as the base, and then take the Fisher index, which is the geometric mean of these index numbers.[15] Thus, the variety index of country c at time t that we propose is

$$\Psi_t^c = \left(\prod_a \prod_t \psi_{a,t}^{c,t} \right)^{1/at}$$

Another problem of using the worldwide exports from all countries to the US, averaged over time, as the comparison base is that the index is distorted by the export values of large exporters. If Mexico increases the variety of its export products by one product, the variety index increases, but the amount of new products exported by Mexico does not matter. Instead, the total amount across countries and averaged over time enters the computation. Thus, an increase in Mexico's variety index may be caused simply by a substantial increase in China's export volume to the US rather than by an increase specific to Mexico. Put differently, when Mexico increases the number of export goods from n at time t to n+1 at time t+1 and from n+1 at time t+1 to n+2 at time t+2, the changes in the index are different. More importantly, changes are substantially affected by worldwide exports to the US. As before, we propose using the Fisher index to overcome this problem (see the appendix for an illustration). As in Feenstra and Kee (2007), the export variety index is computed for

15 The Fisher index was first proposed as the geometric mean of the Paasche and Laspeyres price indexes.

1972–1988 using US import data at the seven-digit code level and for 1989–2006 using US import data at the 10-digit HS code level.[16]

The computed Fisher index for the largest 10 exporters to the US market is shown in Table 11. It is notable that the very low index number for China in 1972 rose sharply toward 2006. For comparison purposes, the variety index computed by the Feenstra and Kee (2007) methodology is shown in Table 12. The notable difference is that China's index numbers at the beginning of the study period are closer to those of Mexico compared with the Fisher variety index. The difference in the index numbers between Mexico and China is clearly smaller in the original Feenstra and Kee index. The correlation coefficient between Mexico and China in the original Feenstra and Kee index is extremely high (0.9748), while that in the Fisher index is 0.8939. The extremely high correlation coefficient (close to one) using Feenstra and Kee (2007)'s methodology is probably caused by the distortion of the index by the worldwide export value to the US in their methodology.

======= Table 11 and Table 12 =======

The following equation is estimated by the fixed effects panel regressions as in Feenstra and Kee (2007):

$$Variety_{it} = \beta_0 + \beta_1 NAFTA + \tilde{\beta}_2 Year + \tilde{\beta}_3 Country + \varepsilon_{it}$$

where i represents origin (exporter) countries and t represents years. *Year* is a vector of the year dummies, and *Country* is a vector of country dummies. The Fisher export variety indexes are computed for the 50 largest exporters to the US market for the maximum period

16 Owing to the difference in the trade statistics codes (seven-digit vs. 10-digit), these indexes are inconsistent between 1988 and 1989. As in Feenstra and Kee (2007), we re-scaled the earlier indexes so that export variety in 1988 equals that in 1989 for each country.

of 35 years (1972-2006), thus giving 1392 observations,[17] as in Table 13. The estimation results for the period of 1972-2006 are in Table 13. The first column shows the fixed-effects estimator, and the second column the random effects estimator. The Hausman test's null hypothesis that the random effects estimator is consistent is rejected, leading us to take the fixed effects as the appropriate estimator. The NAFTA dummy shows a statistically insignificant coefficient estimate. To address possible heteroskedasticity and/or autocorrelation of errors, the third column shows the estimates with cluster-robust standard errors. Notably, the NAFTA dummy has a statistically significant negative coefficient, -0.144, whereas the Mexico unilateral liberalisation dummy shows a statistically significant positive coefficient, 0.233, which is in line with the econometric analysis using raw trade data in Section 2. The same estimation is presented in Table 14 with the period limited to 2001 to make it comparable with the findings of Feenstra and Kee (2007). The NAFTA dummy in the third column becomes insignificant while the Mexico unilateral liberalisation dummy shows a statistically significant positive coefficient, 0.232, which is very close to 0.233 in Table 13.

======= Table 13 and Table 14 =======

In addition to this benchmark model, we also estimate the model adding the GDP of the origin country and the distance from the origin country to the US. The log is taken for all variables. Table 15 shows the estimation results. The results of the Hausman test lead us to take the fixed effects estimator as the appropriate one. The NAFTA dummy in the fixed-effects (column 1) is statistically insignificant whereas the one in the fixed-effects with cluster-robust standard errors (column 3) shows a statistically significant negative

[17] It is not 1750 (=35 times 50) because the data are not available for the maximum years of 35 years for many countries.

coefficient. The coefficient estimates for the Mexico unilateral liberalization dummy show a statistically positive coefficient both in the column 1 and in the column 3. The last column (column 4) shows the case for the least-square dummy-variables, in order to obtain the coefficient estimate for the distance variable. The case for the years 1972-2001 is in Table 16. The results are qualitatively the same, with a somewhat smaller magnitude for the NAFTA dummy. All the findings in Table 13 to Table 16 are thus at odds with those of Feenstra and Kee (2007).

======= Table 15 and Table 16 =======

CONCLUDING REMARKS

Despite the optimistic views expressed about NAFTA's effects on the Mexican economy at the time of the agreement and the positive assessment by studies carried out since the mid-2000s, Mexico lags behind many other middle-income countries in terms of its economic performance. This paper studied the evolution of the variety of Mexico's export goods using disaggregated trade data. Both a regression using the raw data, and another one using an improved version of Feenstra and Kee's (2004, 2007) methodology, proposed in this paper, show that NAFTA membership does not enhance the variety of Mexico's export goods. This finding contrasts with the literature, which shows a positive association between NAFTA and export variety. The paper, on the other hand, finds that Mexico's unilateral trade liberalization had a positive impact on the variety of Mexico's exports.

TABLES AND FIGURES

Table 1: The number of zeros in US imports from the 20 largest import partners plus Colombia and Chile: 1989–2006

Year	Canada	Japan	Mexico	China	Germany	Taiwan	United Kingdom	Korea	France	Italy	Singapore
1989	15179	15748	19105	19198	15600	17859	15884	18954	16601	16809	22163
1990	14931	15673	19124	18918	15481	17926	15848	19061	16597	16745	22199
1991	14756	15516	19043	18634	15401	17947	15733	19320	16585	16643	22185
1992	14678	15721	18994	18272	15467	18049	15731	19410	16637	16676	22167
1993	14461	15726	18761	17842	15357	17932	15463	19229	16384	16376	22083
1994	13907	15370	18084	17254	14876	17690	15033	18903	15895	15823	21928
1995	13484	15377	17176	16925	14743	17608	14839	18737	15705	15594	21910
1996	13260	15211	16478	16555	14531	17410	14475	18725	15439	15249	21943
1997	12729	14826	15980	15793	14020	16949	13992	18283	15000	14714	21611
1998	12990	15072	16245	15578	14283	17141	14175	17853	15140	14924	21670
1999	12922	15065	16161	15160	14280	17103	14382	17516	15116	14941	21627
2000	12847	15076	16117	14621	14104	16950	14272	17474	14959	14611	21575
2001	13010	15292	16250	14506	14201	17077	14503	17489	15026	14764	21765
2002	12724	15180	16105	13769	13991	16898	14443	17143	15041	14460	21699
2003	12844	15157	16203	13379	14054	16856	14546	17246	15040	14525	21670
2004	12920	15126	16095	12837	13987	16724	14460	17069	15053	14457	21703
2005	12912	14986	16009	12126	14078	16741	14640	16850	15126	14424	21643
2006	12959	14892	15993	11691	14017	16746	14672	16938	15096	14270	21463
Change, 1989–2006	−2220	−856	−3112	−7507	−1583	−1113	−1212	−2016	−1505	−2539	−700

Year	Malaysia	Thailand	Venezuela	Hong Kong	Brazil	Saudi Arabia	Philippines	Switzerland	Indonesia	Colombia	Chile
1989	22982	21904	23453	19379	20897	24589	22349	19139	23208	23362	23676
1990	22846	21777	23078	19468	21122	24629	22422	19170	23061	23211	23687
1991	22651	21636	23347	19481	21127	24627	22400	19229	22914	23053	23677
1992	22489	21396	23505	19600	21044	24559	22309	19265	22688	23031	23678
1993	22286	21110	23446	19574	20951	24542	22235	19150	22309	23034	23737
1994	22052	20850	23390	19473	20678	24522	22051	18748	22190	22967	23627
1995	21937	20692	23533	19384	20860	24444	21935	18573	21978	23048	23642
1996	21843	20690	23348	19395	20859	24436	21878	18413	21771	23013	23625
1997	21682	20508	23289	19196	20829	24374	21756	17944	21483	22927	23602
1998	21603	20341	23411	19480	20932	24279	21670	18158	21313	22961	23555
1999	21601	20182	23454	19569	20646	24288	21654	18287	21223	22732	23495
2000	21391	19957	23432	19351	20146	24171	21519	18131	21118	22490	23475
2001	21493	20005	23396	19517	20076	24165	21562	18287	21128	22436	23395
2002	21417	19794	23398	19347	19614	24241	21496	18346	21011	22237	23249
2003	21404	19675	23432	19257	19234	24299	21457	18435	20979	22017	23119
2004	21228	19488	23438	19238	18963	24263	21431	18414	20906	21911	23112
2005	21073	19302	23480	19032	18846	24246	21388	18377	20718	21723	23002
2006	20933	19149	23681	18924	18775	24228	21223	18305	20568	21735	23030
Change, 1989-2006	-2049	-2755	228	-455	-2122	-361	-1126	-834	-2640	-1627	-646

Source: Author’s calculation based on US imports data at HS 10 digit level.

Table 2: The number of zeros in US imports from the 20 largest import partners plus Colombia and Chile: 1972–1988

Year	Canada	Japan	Mexico	China	Germany	Taiwan	United Kingdom	Korea	France	Italy	Singapore
1972	19892	19404	21911	23247	19757	21844	19665	22682	20321	20456	23372
1973	19681	19606	21669	23153	19709	21816	19514	22572	20218	20469	23291
1974	19506	19521	21390	23031	19675	21725	19510	22356	20188	20449	23208
1975	19548	19482	21485	22905	19745	21546	19613	22179	20170	20411	23173
1976	19435	19236	21414	22696	19505	21201	19413	21798	20031	20271	23057
1977	19512	19344	21510	22791	19627	21356	19486	21852	20194	20406	23152
1978	18379	18411	20708	22515	18589	20488	18450	21249	19145	19420	22745
1979	18305	18388	20711	22303	18608	20387	18442	21386	19155	19403	22659
1980	17906	17978	20707	21736	18146	20125	18137	21199	18936	19292	22550
1981	17686	17852	20705	21312	18019	19922	18084	20914	18682	19003	22404
1982	17551	17514	20670	21071	17855	19536	18008	20680	18510	18751	22346
1983	17469	17173	20203	20855	17485	19190	17768	20393	18169	18375	22270
1984	17041	16570	19953	20511	17045	18716	17123	19919	17571	17824	22075
1985	16532	15217	19981	19719	16054	17736	16199	19135	16410	16183	21756
1986	16544	15490	19744	19593	16125	17882	16425	19048	16523	16476	21750
1987	16422	15960	19441	19383	16328	17935	16438	18953	16913	16871	21763
1988	16359	16352	19377	19199	16481	18110	16710	18977	17095	17058	21753
Change, 1972-1988	-3533	-3052	-2534	-4048	-3276	-3734	-2955	-3705	-3226	-3398	-1619

Year	Malaysia	Thailand	Venezuela	Hong Kong	Brazil	Saudi Arabia	Philippines	Switzerland	Indonesia	Colombia	Chile
1972	23637	23501	23631	21631	22831	23915	23231	21450	23765	23203	23818
1973	23584	23396	23608	21633	22547	23923	23092	21474	23738	23132	23842
1974	23580	23317	23548	21626	22534	23915	22984	21383	23686	23061	23798
1975	23519	23315	23620	21505	22589	23914	22979	21483	23635	23088	23748
1976	23434	23168	23635	21199	22593	23921	22827	21322	23593	23004	23632
1977	23453	23139	23710	21384	22640	23916	22840	21471	23579	23178	23667
1978	23271	22936	23669	20653	22202	23906	22451	20789	23473	23069	23581
1979	23189	22880	23657	20669	22131	23884	22438	20793	23377	23175	23570
1980	23178	22798	23603	20516	22121	23870	22403	20574	23396	23140	23600
1981	23097	22655	23597	20315	21818	23827	22255	20415	23269	23187	23615
1982	23117	22511	23570	20133	21678	23800	22224	20297	23213	23218	23589
1983	23003	22459	23407	20003	21364	23776	22210	20008	23225	23137	23454
1984	22841	22191	23302	19562	20558	23735	22001	19624	23038	22987	23370
1985	22521	21794	23302	18285	20210	23674	21580	19124	22561	22844	23331
1986	22548	21859	22964	18577	20471	23636	21728	18994	22732	22840	23101
1987	22462	21759	23007	18784	20699	23728	21748	19028	22629	22717	23036
1988	22443	21632	23026	19069	20623	23752	21779	19232	22558	22636	22888
Change, 1972-1988	-1194	-1869	-605	-2562	-2208	-163	-1452	-2218	-1207	-567	-930

Source: Author's calculation based on US trade statistics imports data at the 7 digit level.

Table 3: Summary statistics 1962-2010

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum	Expected sign for coefficient estimate
Log of export value	2218524	0.533	1.604	0	17.382	Non-applicable
Log of GDP	1956108	24.599	2.298	17.277	30.312	Positive
Log of distance	2218524	8.694	0.812	6.969	9.719	Negative
Common language	2218524	0.367	0.482	0	1	Positive
NAFTA	2218524	0.007	0.084	0	1	Positive/Negative/Neutral
Mexico unilateral liberalization	2218524	0.510	0.500	0	1	Positive/Negative/Neutral

Table 4: Estimation results: Probit using five-digit SITC data for 1962–2010

	(1)	(2)
Log of GDP	0.221***	0.221***
	(41.54)	(41.54)
Log of Distance	-0.926***	-0.828***
	(-49.98)	(-48.77)
Common language	0.695***	0.00514
	(7.96)	(0.22)
NAFTA	-0.147***	-0.147***
	(-10.52)	(-10.52)
Mexico Unilateral Liberalization		0.781***
		(29.78)
Number of observations	1956108	1956108
Pseudo R-squared	0.243	0.243

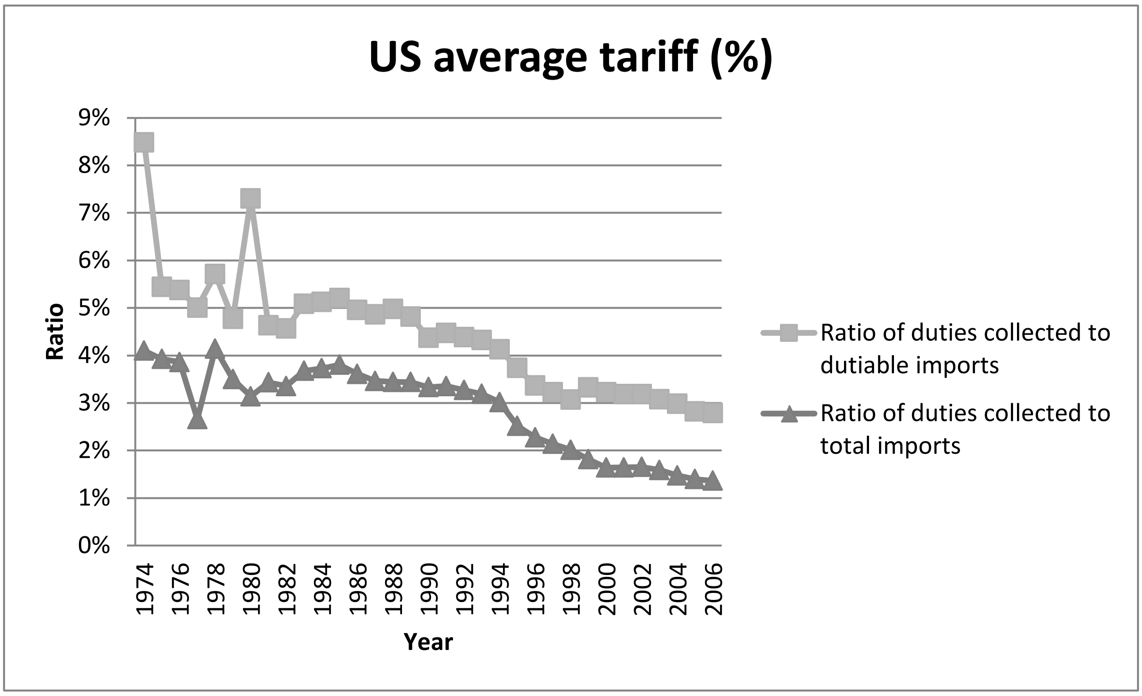
t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Estimation results: Probit using five-digit SITC data for 1972–2001

	(1)	(2)
Log of GDP	0.193***	0.193***
	(26.37)	(26.37)
Log of Distance	-0.799***	-0.783***
	(-38.65)	(-37.10)
Common language	0.609***	0.645***
	(21.56)	(22.93)
NAFTA	-0.216***	-0.216***
	(-10.99)	(-10.99)
Mexico Unilateral Liberalization		0.0985***
		(7.34)
Number of observations	1191960	1191960
Pseudo R-squared	0.230	0.230

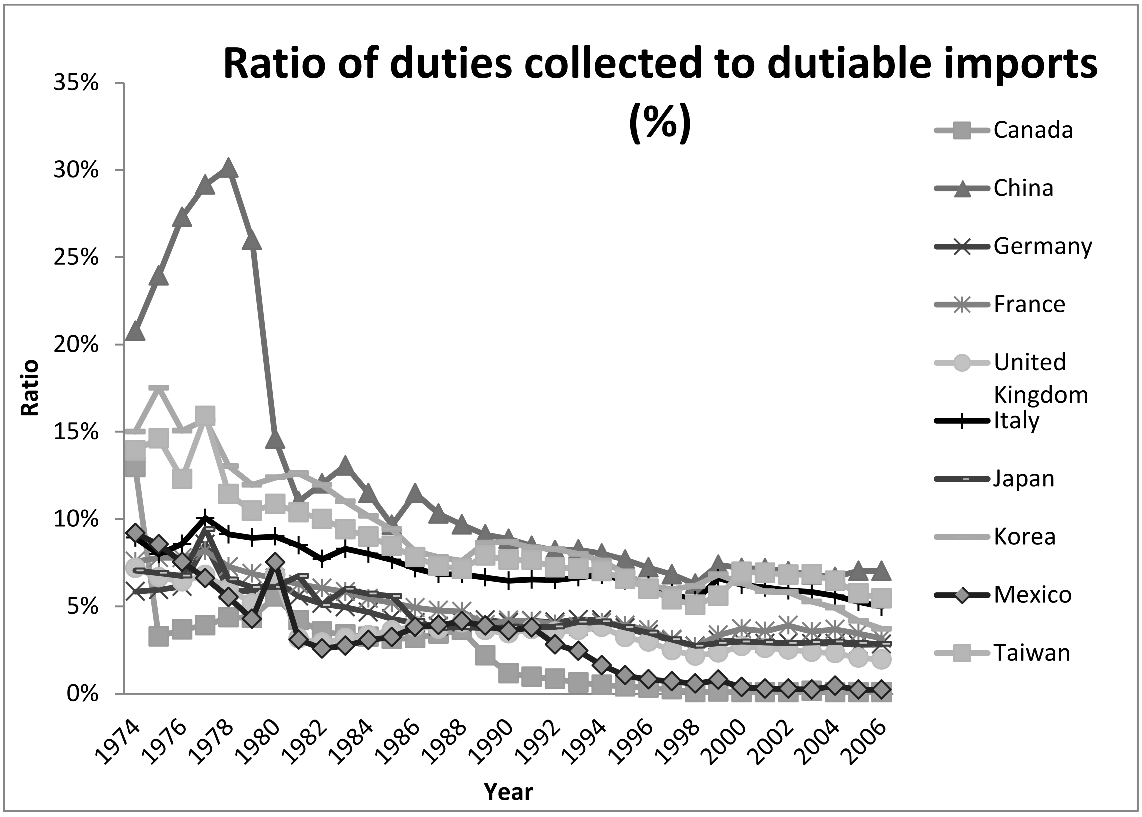
t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: US average tariffs, 1974 to 2006



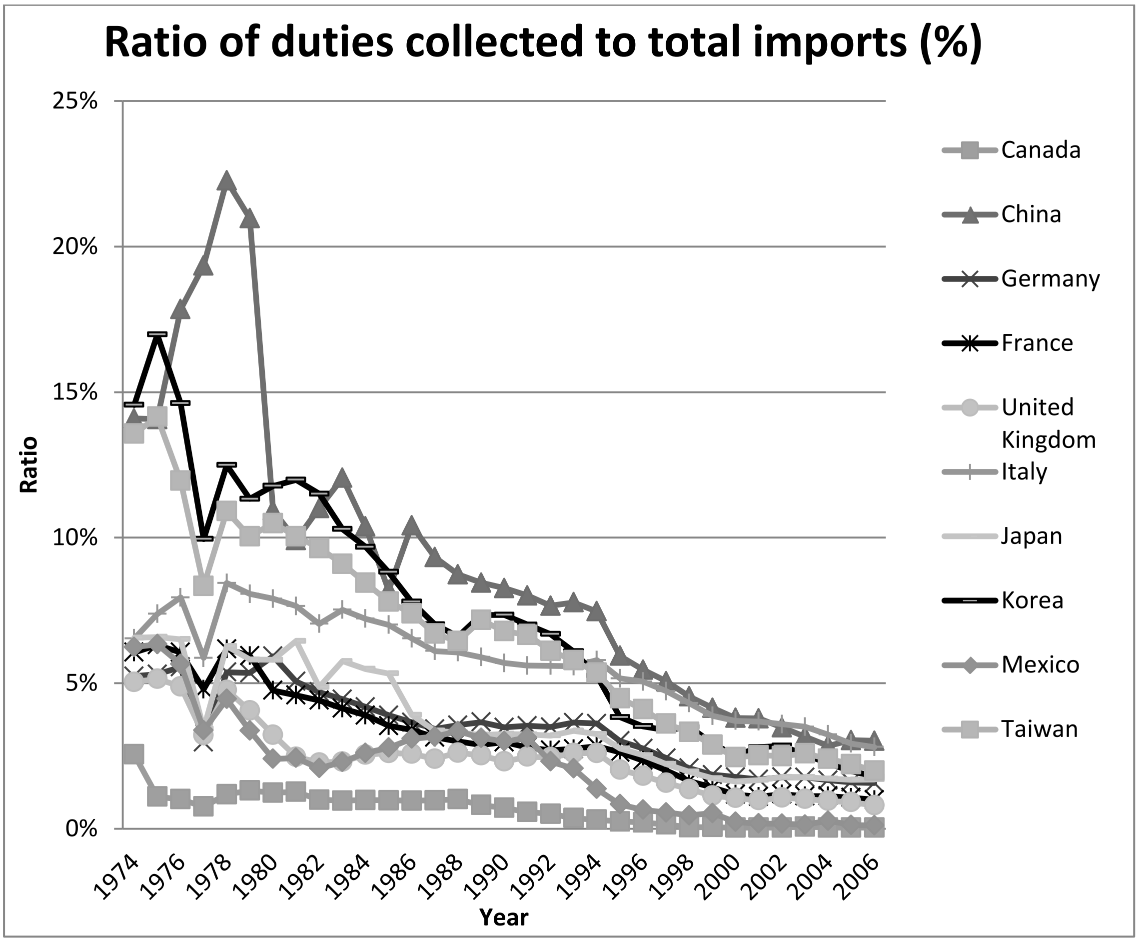
Source: Author's computation from the data at The Center for International Data at UC Davis.

Table 7: US average tariffs, 1974 to 2006, Ratio of duties collected to dutiable imports



Source: Author's computation from the data at The Center for International Data at UC Davis.

Table 8: US average tariffs, 1974-2006, Ratio of duties collected to total imports



Source: Author's computation from the data at The Center for International Data at UC Davis.

Table 9: Estimation results: Probit for the machinery sector using five-digit SITC data for 1972–2010

	(1)	(2)
Log of GDP	0.295***	0.295***
	(21.07)	(21.07)
Log of Distance	2.992**	2.992**
	(2.84)	(2.84)
Common language	5.541***	5.541***
	(3.90)	(3.90)
NAFTA	-0.192***	-0.192***
	(-3.73)	(-3.73)
Mexico Unilateral Liberalization		0.818***
		(19.82)
Number of observations	213817	213817
Pseudo R-squared	0.273	0.273

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 10: Estimation results: Probit for the machinery sector using five-digit SITC data for 1972–2001

	(1)	(2)
Log of GDP	0.242***	0.242***
	(12.79)	(12.79)
Log of Distance	-0.548***	-0.548***
	(-13.45)	(-13.45)
Common language	1.749***	1.061***
	(18.33)	(4.93)
NAFTA	-0.389***	-0.389***
	(-5.24)	(-5.24)
Mexico Unilateral Liberalization		0.346***
		(8.26)
Number of observations	130290	130290
Pseudo R-squared	0.246	0.246

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 11: Export variety index of the top 10 exporters to the US for 1972–2006

Year	Canada	China	Germany	France	United Kingdom	Italy	Japan	Korea	Mexico	Taiwan
1972	1.227	0.049	0.592	0.429	0.744	0.331	0.668	0.248	0.298	0.269
1973	1.236	0.060	0.537	0.423	0.799	0.353	0.544	0.229	0.357	0.408
1974	1.196	0.093	0.664	0.605	0.811	0.350	0.727	0.294	0.337	0.362
1975	1.194	0.129	0.535	0.455	0.785	0.419	0.753	0.277	0.306	0.369
1976	1.284	0.137	0.558	0.543	0.833	0.444	0.677	0.392	0.387	0.398
1977	1.730	0.103	0.764	0.687	1.114	0.642	0.938	0.293	0.698	0.430
1978	3.319	0.178	1.276	1.224	1.903	1.164	1.446	0.562	1.333	0.803
1979	2.905	0.252	1.974	1.191	1.912	1.175	2.102	0.515	1.236	0.747
1980	2.625	0.368	1.997	1.411	1.735	1.183	2.431	0.712	0.851	0.667
1981	2.799	0.486	1.800	1.407	1.905	1.280	2.357	0.907	0.995	0.628
1982	3.389	0.531	2.214	1.475	2.009	1.260	2.466	0.944	0.946	0.857
1983	3.329	0.527	2.161	1.830	1.759	1.341	2.579	1.466	1.106	0.962
1984	3.584	0.607	1.880	1.934	1.924	1.480	2.451	1.075	1.220	1.279
1985	3.700	0.791	2.119	2.191	1.836	1.600	2.745	1.204	1.132	0.890
1986	3.618	1.343	2.633	1.991	2.056	1.569	3.578	1.586	1.367	1.684
1987	3.695	1.443	2.674	1.959	2.065	1.456	2.892	2.073	1.613	1.033
1988	3.870	0.891	2.460	1.841	2.184	1.500	3.845	1.202	1.698	0.988
1989	3.870	0.891	2.460	1.841	2.184	1.500	3.845	1.202	1.698	0.988
1990	3.734	0.854	2.791	1.910	2.301	1.532	3.705	1.073	1.897	0.961
1991	3.632	0.993	2.666	2.049	2.184	1.451	3.316	1.345	1.941	0.775
1992	3.714	1.124	2.766	2.119	2.251	1.477	3.339	1.422	1.896	0.756
1993	3.783	1.184	2.822	2.196	2.422	1.610	3.222	1.484	1.976	0.841
1994	3.396	1.351	2.801	2.157	2.510	1.628	2.990	1.749	1.987	0.969
1995	3.530	1.389	2.963	2.342	2.672	1.371	3.132	2.158	2.206	1.003
1996	3.502	1.554	3.004	2.128	2.551	1.424	2.949	2.135	2.249	0.966
1997	3.557	1.651	3.165	2.173	2.622	1.617	2.862	1.766	2.135	0.958
1998	3.496	1.655	3.209	2.233	2.605	1.641	2.972	1.670	2.093	0.899
1999	3.548	1.760	3.352	2.390	2.625	1.615	2.833	1.719	2.122	1.084
2000	3.341	1.695	3.306	2.439	2.510	1.831	2.773	2.047	1.986	1.010
2001	3.305	1.678	3.325	2.273	2.556	1.725	2.795	1.713	2.081	1.136
2002	3.098	1.463	3.134	2.130	2.419	1.572	2.715	1.826	1.843	0.981
2003	3.015	1.547	3.245	2.268	2.635	1.553	2.613	1.868	1.775	1.013
2004	3.143	1.596	3.195	2.156	2.585	1.573	2.470	1.949	1.835	1.119
2005	3.137	1.521	3.006	2.026	2.671	1.806	2.520	2.195	1.771	1.240
2006	3.053	1.666	2.765	2.087	2.694	1.685	2.860	2.164	1.961	1.424

Table 12: The original Feenstra and Kee (2007) variety index for China and Mexico

Year	China	Mexico
1972	0.017	0.074
1973	0.021	0.072
1974	0.023	0.077
1975	0.028	0.071
1976	0.036	0.102
1977	0.034	0.104
1978	0.062	0.192
1979	0.075	0.196
1980	0.098	0.184
1981	0.123	0.179
1982	0.140	0.217
1983	0.145	0.248
1984	0.179	0.267
1985	0.256	0.303
1986	0.247	0.305
1987	0.276	0.325
1988	0.268	0.324
1989	0.268	0.324
1990	0.268	0.325
1991	0.285	0.344
1992	0.286	0.345
1993	0.307	0.368
1994	0.346	0.394
1995	0.374	0.434
1996	0.392	0.471
1997	0.434	0.489
1998	0.428	0.489
1999	0.454	0.498
2000	0.459	0.494
2001	0.465	0.487
2002	0.497	0.508
2003	0.515	0.526
2004	0.546	0.533
2005	0.548	0.546
2006	0.579	0.558

Table 13: Estimation results – Fisher variety index 1972–2006

	(1) Fixed-effects	(2) Random effects	(3) Fixed effects (cluster-robust standard errors)
NAFTA	-0.144	-0.144	-0.144**
	(-0.84)	(-0.83)	(-2.70)
Mexico unilateral liberalisation	0.233	0.254	0.233**
	(1.38)	(1.50)	(3.26)
Constant	-1.526**	0.152	-1.526**
	(-20.96)	(1.45)	(-16.52)
R-squared	0.625		0.625
Number of observations	1392	1392	1392
Test of joint significance	F(36,1308) = 60.51 Prob > F = 0.0000	Wald chi2(36) = 2126.89 Prob > chi2 = 0.0000	-
Hausman test	chi2(15) = 40.28 Prob > chi2 = 0.0004		

t statistics in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 14: Estimation results – Fisher variety index 1972–2001

	(1) Fixed-effects	(2) Random effects	(3) Fixed effects (cluster-robust standard errors)
NAFTA	-0.0460	-0.0464	-0.0460
	(-0.26)	(-0.25)	(-1.00)
Mexico unilateral liberalisation	0.232	0.250	0.232**
	(1.45)	(1.55)	(3.30)
Constant	-1.431**	0.0313	-1.431**
	(-20.76)	(0.29)	(-23.90)
R-squared	0.657		0.657
Number of observations	1153	1153	1153
Test of joint significance	F(31,1074) = 66.24 Prob > F = 0.0000	Wald chi2(31) = 2001.10 Prob > chi2 = 0.0000	-
Hausman test	chi2(15) = 38.23 Prob > chi2 = 0.0008		

t statistics in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 15: Estimation results – Fisher variety index 1972-2006 with other control variables

	(1) Fixed-effects	(2) Random effects	(3) Fixed effects (cluster-robust standard errors)	(4) Least-square dummy-variables (cluster-robust standard errors)
Log of GDP	0.310**	0.370**	0.310**	0.310**
	(8.51)	(13.59)	(4.54)	(4.46)
Log of distance	.	-0.162	.	-0.250**
	.	(-1.64)	.	(-6.86)
NAFTA	-0.240	-0.252	-0.240**	-0.240**
	(-1.57)	(-1.63)	(-4.89)	(-4.80)
Mexico unilateral liberalisation	0.290+	0.305*	0.290**	0.290**
	(1.92)	(2.02)	(4.15)	(4.08)
Constant	-7.932**	-8.133**	-7.932**	-6.447**
	(-8.20)	(-7.28)	(-4.35)	(-4.24)
R-squared	0.683		0.683	0.871
Number of observations	1371	1371	1371	1371
Test of joint significance	F(37,1286) = 74.91 Prob > F = 0.0000	Wald chi2(38) = 2807.51 Prob > chi2 = 0.0000	-	-
Hausman test	chi2(17) = 40.74 Prob > chi2 = 0.0010			

t statistics in parentheses
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 16: Estimation results – Fisher variety index 1972-2001 with other control variables

	(1) Fixed-effects	(2) Random effects	(3) Fixed effects (cluster-robust standard errors)	(4) Least-square dummy-variables (cluster-robust standard errors)
Log of GDP	0.284**	0.360**	0.284**	0.284**
	(7.60)	(12.64)	(4.26)	(4.17)
Log of distance	.	-0.155	.	-0.243**
	.	(-1.47)	.	(-6.74)
NAFTA	-0.105	-0.113	-0.105*	-0.105*
	(-0.66)	(-0.71)	(-2.35)	(-2.30)
Mexico unilateral liberalisation	0.282*	0.300*	0.282**	0.282**
	(2.01)	(2.12)	(4.04)	(3.96)
Constant	-7.169**	-7.853**	-7.169**	-5.726**
	(-7.36)	(-6.66)	(-4.14)	(-3.77)
R-squared	0.718		0.718	0.894
Number of observations	1133	1133	1133	1133
Test of joint significance	F(32,1053) = 83.83 Prob > F = 0.0000	Wald chi2(33) = 2681.03 Prob > chi2 = 0.0000	-	-
Hausman test	chi2(17) = 49.68 Prob > chi2 = 0.0000			

t statistics in parentheses
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Figure 1: The evolution of the zero index in US imports from the 20 largest import partners plus Colombia and Chile: 1972–2006

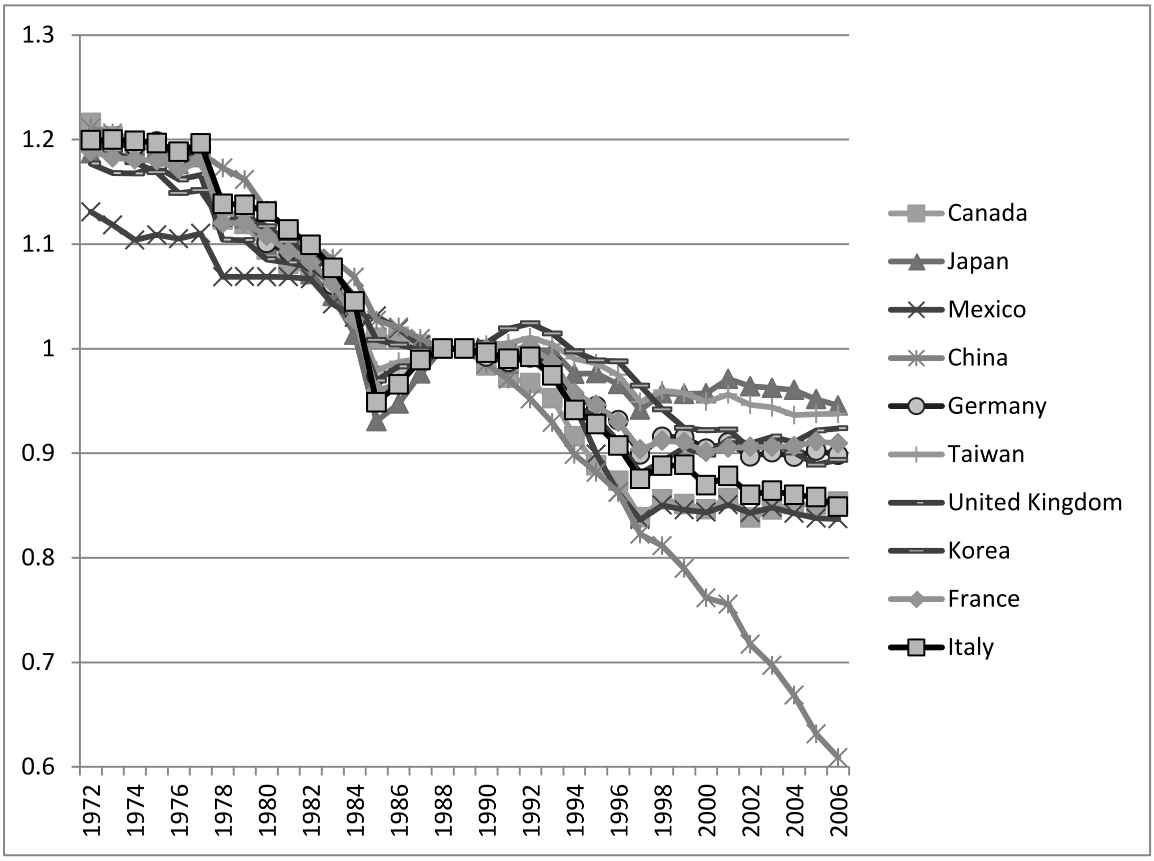


Figure 2: Zero index in US imports from some major import partners at HS 6-digit data 1991-2012

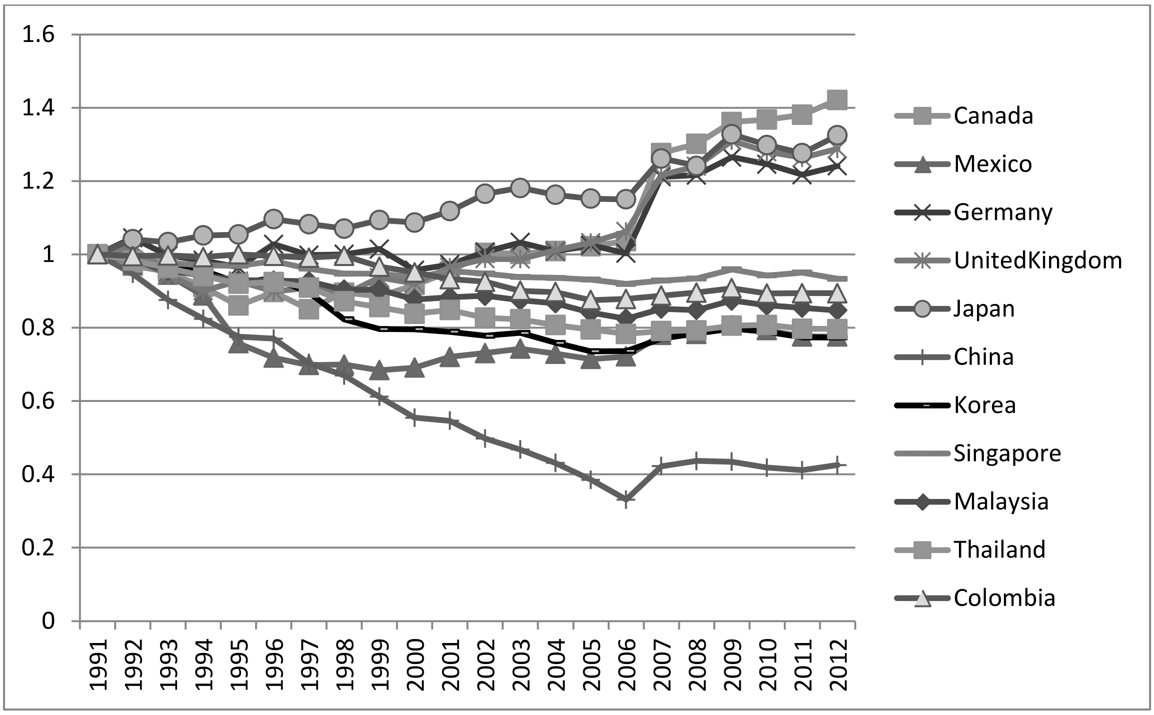
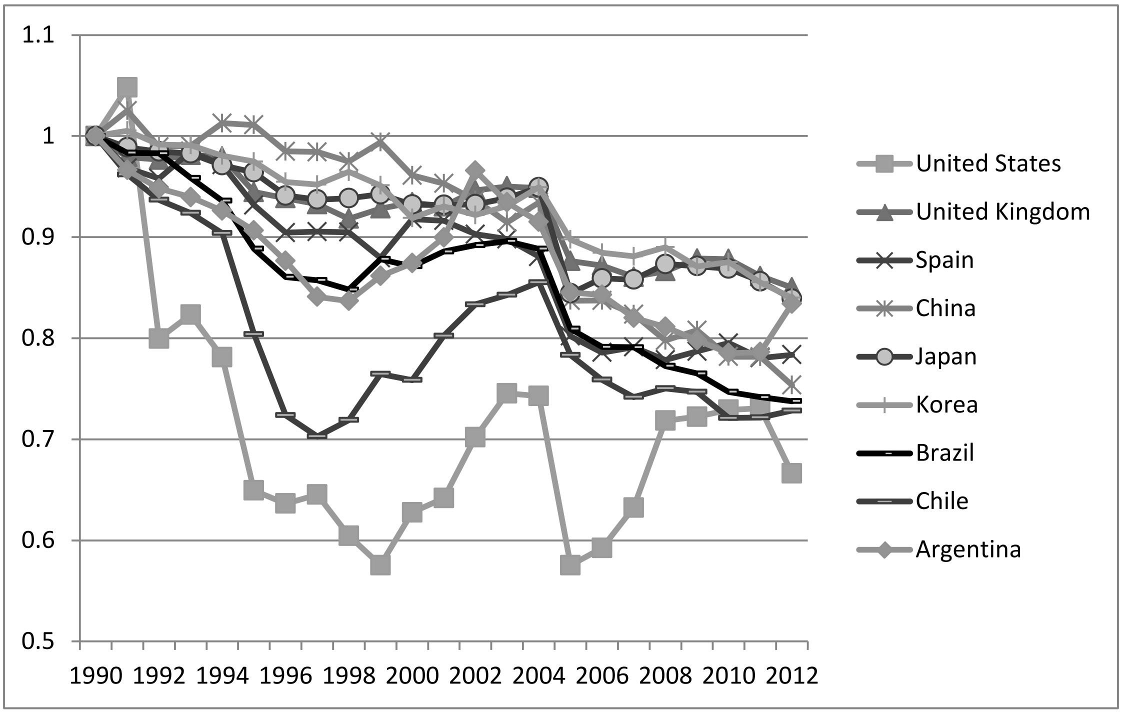


Figure 3: Zero index of Mexico's exports to its major markets, 1990-2012



REFERENCES

Amiti, M. and Konings, J., 2007. Trade liberalization, intermediate inputs, and productivity: Evidence from Indonesia. *American Economic Review*, 97(5): 1611–1638.

Baier, S. L., Bergstran, J. H., and Feng, M., 2011. Economic integration agreements and the margins of international trade. Working paper available at http://www.nd.edu/~jbergstr/Working_Papers/Baier%20Bergstrand%20Feng%20Margins%20and%20EIAs.pdf.

Baldwin, R., 2010. Unilateral tariff liberalization. NBER Working Paper 16600.

Baldwin, R. and Harrigan, J., 2011. Zeros, quality and space: Trade theory and trade evidence. *American Economic Journal*, 3(2), 60–88.

Breitung, J., 2000. The local power of some unit root tests for panel data. In B. H. Baltagi (eds.). *Advances in Econometrics, Volume 15: Nonstationary Panels, Panel Cointegration, and Dynamic Panels*, 161-178. Amsterdam: JAI Press.

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

Cadot, O., Carrère C., and Strauss-Kahn, V., 2007, Export Diversification: What's behind the Hump? CEPR Discussion Paper No. 6590.

Feenstra, R., 1994. New product varieties and the measurement of international prices. *American Economic Review*, 84(1), 157–177.

Feenstra, R. and Kee H., 2004. On the measurement of product variety in trade. *American Economic Review*, 94(2), 145–149.

Feenstra, R. and Kee H., 2007. Trade liberalisation and export variety: A comparison of Mexico and China. *World Economy*, 1(1), 5–21.

Hanson, G., 2010. Why isn't Mexico rich?. NBER Working Paper No. 16470.

Helpman, E., Melitz, M., and Rubinstein, Y., 2008. Estimating trade flows: trading partners and trading volumes. *Quarterly Journal of Economics*, 123(2), 441–487.

Harris, R., and Tzavalis, E., 1999. Inference for unit roots in dynamic panels where the time dimension is fixed. *Journal of Econometrics* 91: 201-226.

Hummels, D. and Klenow, P., 2005. The variety and quality of a nation's exports. *American Economic Review*, 704–723.

Imbs, J. and Wacziarg, R., 2003. Stages of diversification., *American Economic Review*, 9(1), 63–86.

Im, K., Pesaran, M., Shin, Y., 2003. Testing for unit roots in heterogeneous panels., *Journal of Econometrics* 115 (2003) 53-74

Kehoe, T. and Ruhl, K., 2009. How important is the new goods margin in international trade?" Federal Reserve Bank of Minneapolis Research Department Staff Report 324.

Klinger, B. and Lederman, D., 2004. Discovery and development: An empirical exploration of 'new' products", World Bank Policy Research Working Paper Series No. 3450.

Krugman, P., 1980. Scale economies, product differentiation and the pattern of trade *American Economic Review*, 70, 950–959.

Lederman, D., Maloney, W., and Serven, L., 2004. Lessons from NAFTA, World Bank. November 2004.

Levin, A., Lin. C. and Chu C., 2002. Unit root tests in panel data: Asymptotic and finite sample properties. *Journal of Econometrics* 108, 1–24.

Melitz, M., 2003. The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, 71(6), 1695–1725.

Quah, D., 1994. Exploiting cross-section variation for unit root inference in dynamic data. *Economics Letters* 44 (1994) 9-19.

Sachs, J. and Warner, A., 1995. Economic reform and the process of global integration", Brookings Papers on Economic Activity, 1, 1–118.

Zwinkels, R., Beugelsdijk, S., 2010. Gravity equations: Workhorse or Trojan horse in explaining trade and FDI patterns across time and space?" *International Business Review*, 19: 102-115.

Appendix

Appendix to Section 2

Trade data code

Exports and imports are recorded by each custom authority, using each country's own product code. The United States, for example, keeps records of its trade at its own 10 digit code, while Japan uses its own 9 digit system. These codes were internationally harmonized at the six-digit level by the United Nations. Thus, the codes are identical across customs authorities up to six-digit, thus called the Six-digit International Harmonized System code.

Harmonized System code:

The HS code has three tiers: 2-digit, 4-digit, and 6-digit. The broadest category of 2-digit has 96 codes in total, while the most disaggregated category of 6-digit code has approximately 5000 codes. The Harmonized System has been revised every five years and the most recent version is HS 2012. The earliest available year differs across reporting nations, with 1988 being the earliest possible year.

Standard International Trade Classification code:

The United Nations also prepares the Standard International Trade Classification (SITC) code. The trade data using this code is compiled by the United Nations, using the data reported by the member countries using the Harmonized system. This code has 5 tiers: 1-digit, 2-digit, 3-digit, 4-digit, and 5-digit. The broadest category of 1-digit has 10 codes, while the most disaggregated category of 5-digit has approximately 1200 codes. The earliest available year differs across reporting nations. The data are available from 1962, the earliest available year, for some countries. SITC has been revised periodically and the most recent version is SITC revision 4.

Both HS and SITC could be used for trade analyses. However, it is better to use HS code for product level trade analysis, such as impacts of tariff reduction, while SITC is appropriate for analyses of longer time-series because SITC data are available from 1960s, as mentioned above. SITC data is also more appropriate for analyses of trade structures at industry level because, as the United Nations mentions, "The commodity groupings of SITC reflect (a) the materials used in production, (b) the processing stage, (c) market practices and uses of the products, (d) the importance of the commodities in terms of world trade, and (e) technological changes."

Panel unit root tests

Several panel unit root tests have been proposed in the literature, including Levin,, Lin and Chu (2002); Harris and Tzavalis (1999); Breitung (2000); and Im, Pesaran and Shin (2003) (hereinafter IPS). In general, panel unit root tests are based on the following regression:

$$\Delta y_{it} = \alpha_i + \rho_i y_{i,t-1} + \varepsilon_{it}$$

where *i* represents the individual and *t* represents time.

We use IPS for several reasons. First, while the first three tests require a balanced panel, IPS allows the use of an unbalanced panel. Second, while the first three tests have a major limitation of assuming an identical value for ρ ($\rho_i = \rho$), IPS relaxes this assumption and allows for different ρ values.

The null hypothesis of IPS is that all the variables have unit roots, against the alternative in which the fraction of panels that are stationary are non-zero. More specifically,

$$H_1 : \rho_i < 0, i = 1, 2, ..., N_1, \rho_i = 0, i = N_1 + 1, N_1 + 2, ..., N$$

such that

$$\lim_{N \to \infty} \frac{N_1}{N} = \delta, 0 < \delta \le 1$$

Then, the null and alternative hypotheses can be written as

$H_0 : \delta = 0$

$H_1 : \delta > 0$

In order to have a finite number of δ as N goes to infinity, N_1 must be large enough.[1]

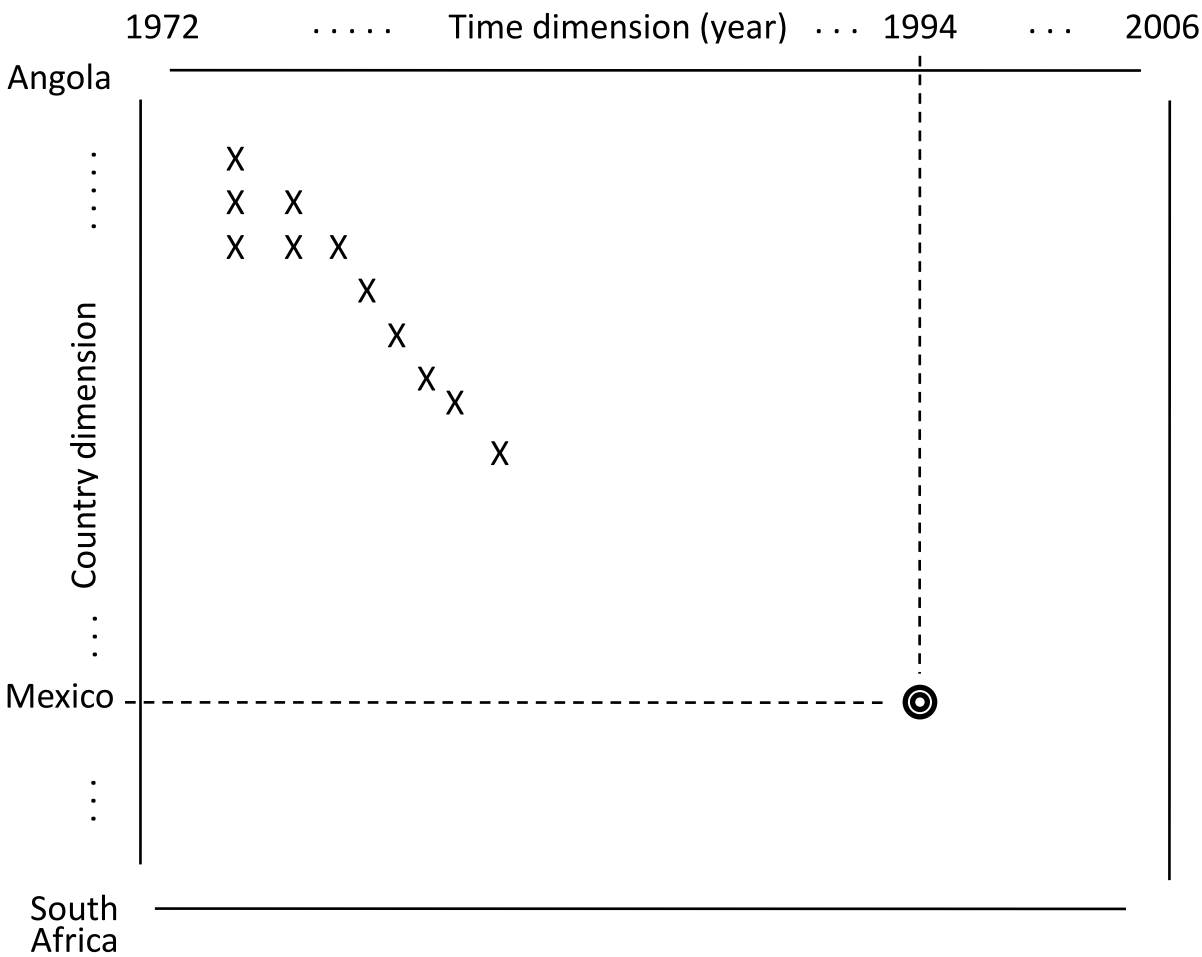
It would be beneficial to have the alternative hypothesis of each series being stationary, but no such test is available.[2]

1 As Hadri (2000) notes, classical hypothesis testing requires strong evidence to the contrary to reject the null hypothesis. Thus, Hadri (2000) proposes a test where the null and alternative are reversed. Although we planned to perform the Hadri test, it requires strongly balanced panels.

2 For homogeneous panels, i.e., $\rho_i = \rho \ \forall i$, Bayoumi and MacDonald (2000) propose the alternative of each series being stationary.

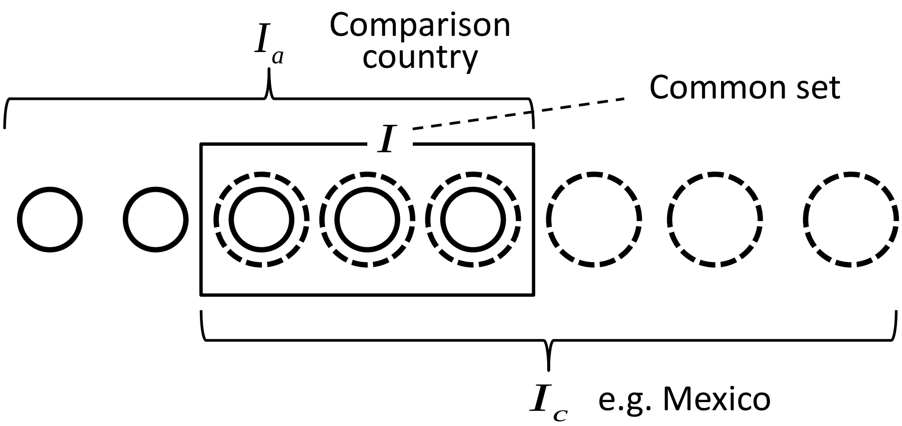
Appendix to Section 3

The essence of Feenstra and Kee's (2007) variety index is best explained in an illustration. As the following figure illustrates, there are hundreds of possible comparison bases (some of which are represented by crosses in the figure) on which to compute the variety index of Mexico for 1994 (represented by the double circle). Feenstra and Kee (2007) choose to take all products with the amount averaged over time as the comparison base. Instead, we propose computing the index for a particular country in a particular year, say Mexico in 1994, using each comparison base and taking the geometric mean in the spirit of the Fisher index.

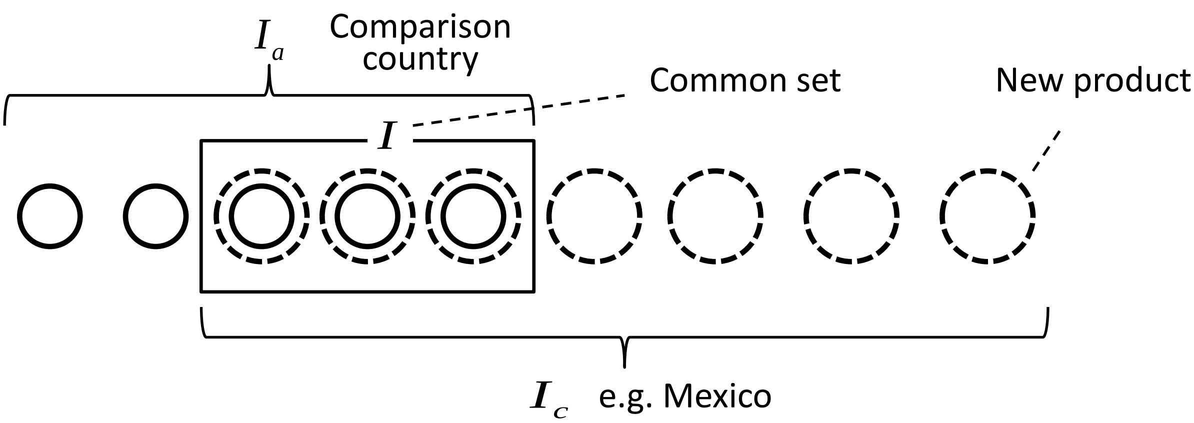


The next illustration shows the problem caused by the selection of Feenstra and Kee (2007). Put simply, consider the variety index of country c at time t and that of comparison country a at time t. Thus, here we ignore changes over time and focus on cross-country differences for the sake of simplicity. The typical case is illustrated in the following figure. As described above, the variety index is computed as

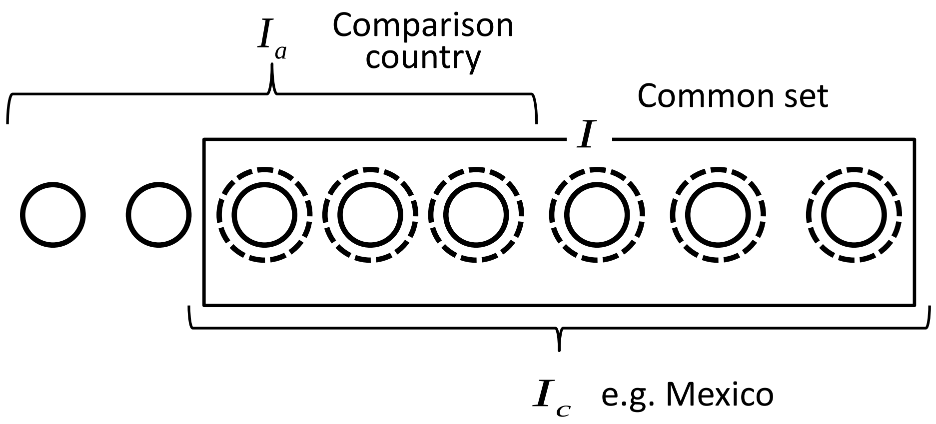
$$\psi_{a,t}^{c,t} \equiv \frac{\lambda_t^a(I)}{\lambda_t^c(I)}, \text{ where } \lambda_t^c(I) \equiv \frac{\sum_{i\in I} p_{it}^c q_{it}^c}{\sum_{i\in I_{it}^c} p_{it}^c q_{it}^c} \text{ and } \lambda_t^a(I) \equiv \frac{\sum_{i\in I} p_{it}^a q_{it}^a}{\sum_{i\in I_{it}^a} p_{it}^a q_{it}^a}$$



When country c, say Mexico, starts to export a new product that is not exported by country a, as in the following figure, $\lambda_t^c(I)$ decreases, whereas there is no change in $\lambda_t^a(I)$ and consequently the variety index ψ increases.



By contrast, the following figure illustrates the case of Feenstra and Kee's (2007) selection of comparison base.



In this case, when Mexico starts to export a new product, that product is also exported by other countries because the comparison base is the products exported by all countries to the US. Therefore, the picture changes to the following. In this case, there is no change in $\lambda_t^c(I)$ despite changes in the variety of country *c*'s exports ($\lambda_t^c(I)$ is always 1), while the denominator of $\lambda_t^a(I)$ changes. Note that only the export values of country a enter the computation of $\lambda_t^a(I)$. In other words, total export value across countries and averaged over time of the new product is the only factor that changes the variety index. The export value of Mexico does not change the index at all.