# Trade, access to varieties and patterns of consumption

Amandine Aubry*

Preliminary work- Please do not cite or quote

February 2013

**Abstract**

This paper empirically test whether product diversity has driven the evolution in U.S. expenditure shares relative to the price effect and the income effect. In particular, I examine whether the growth of international trade amplifies changes in product diversity available to consumers. This channel enables, not only to assess how incremental trade liberalization affects the structure of consumption in a particular economy, but it also makes possible to identify the causal relation. The identification strategy shows that changes in product diversity have a sizable effect on the evolution of consumption patterns, especially relative to the price effect. Moreover, it highlights the negative demand shocks that affect goods sectors exposed to trade. Finally, it identifies the product diversity as a new channel towards which a productive shock in a country can affect its trading partners.

**JEL Classification** : F12 F14 F41 D12

**Keywords** : pattern of consumption, extensive margin, gains from variety, trade liberalization

*IRES - Place Montesquieu, 3, B-1348 Louvain-la-Neuve, Belgium (E-mail: amandine.aubry@uclouvain.be)

# 1 Introduction

Consumer spending has been a key element of the U.S. economic growth. While its composition has changed over time (McCully (2011)), little analysis has been done to study its causes in the presence of differentiated products. One reason is the perception of the consumer choice as an inappropriate area for policy intervention. This approach follows a mistaken logic because even though the consumer's choice results from a sovereign decision, it is constrained by factors such as the range of choices. Bils and Klenow (2001) suggest the variety growth in consumer goods as a key factor in the shift of the spending share. Indeed, if the set of varieties available expands, consumers can better match their consumption to their taste. One can then expect consumers to allocate more resources in sectors subject to relatively larger variety growth.

The starting point of this study is the observation that international trade has drastically grown since the WWII. This growth can be decomposed in terms of volume (intensive margin) and in terms of varieties available to the consumers (extensive margin). Variety expansion is not uniform across sectors. Some of them have been subject to relatively larger cost reductions and others have expanded from the development of comparative advantage in some countries. International trade then provides to consumers a new choice environment which alters their spending. This relation between international trade, product diversity expansion and the shift in consumer spending has raised many concerns. If variety expansion shifts consumer's spending, it then operates a reallocation of resources. This redistribution has been cited but never tested or quantified. Notwithstanding, it plays a major role in the economy of a country. Indeed, change in the consumer's demand affects the production and thereby the employment. Moreover, consumption causes external costs (such as the depletion of resources and its future availability) and the change in the pattern of consumption redistributes those costs and may increase them. Finally, commodities serve as a means of social identity and communication. Their social ends influence the pattern of consumption. The upsurge of spread of global "brand-name" varieties can create social bonds and can be a powerful source of exclusion. All those questions raise concerns in political debates and necessitate a better understanding in order to face the challenges of tomorrow. However, all of them premise that international trade integrates the consumer markets by making more varieties available and that this variety expansion shifts the pattern of spending. While the first assumption has been tested and corroborated, the second one has never been assessed. This analysis fills this gap by assessing how incremental trade liberalization affects the structure of consumption in a particular economy.

Focusing on American trade data over the 1990's and 2000's, I show that changes in the product diversity have a sizable effect on the evolution of the pattern of consumption. To perform this analysis, I first illustrate the variations in the pattern of expenditure as well as the asymmetric growth of varieties imported in the United States between 1993 and 2006. I then provide evidence of a strong positive correlation between the changes in the number of varieties imported and the change in the expenditure share allocated to a particular sector. In particular, a 10 % expansion of variety increases the expenditure share by 9.84 % and the real expenditure share by 10.64 % in a HS 6-digit sector and by 10 % in a HS 4-digit sector controlling for the price effect as well as the income effect. This correlation proves strongly robust across several alternative specifications and econometrical methods. However, this positive relationship between the change in the number of varieties and the change in the expenditure share is compatible with two explanations. The traditional literature in new trade theory assumes that changes in the expenditure drive the variations in the number of varieties. However, a few papers such Bils and Klenow (2001) suggest that the variety growth could be the cause of the evolution of the pattern of consumption. To identify the causal effect of the change in the number of variety on the evolution in the expenditure share, I exploit technological shocks occurring in emerging economies. The development of those economies provides them the opportunity to produce and to export new varieties. Therefore, these shocks would exogenously expand the range of foreign products available to American consumers. This identification strategy speaks in favor of a causal effect of the changes in the number of varieties imported on the expenditure shares. Moreover, it highlights two additional results. First, it shows that demand shocks were negative for the goods sectors exposed to trade. Such shocks seem to reflect a reallocation of expenditure from goods sectors toward service sectors driven by the fast growth in price in the medical care sector. Second, the identification strategy enables to find a new channel towards which a productive shock in a country can affect its trading partners.

My analysis is related to the literature on the relationship between patterns of consumption and product diversity. Most of the previous studies assume that the optimum product diversity is driven by love for variety and the level of expenditure or income per capita. Falkinger and Zweimuller (1996) study the dependence of the degree of product diversity on the average income level and on the size distribution. Foellmi, Hepenstrick, and Zweimuller (2010) and Foellmi and Zweimuller (2004) also analyze how the level and the distribution of income affect the equilibrium mark-up and the optimum product diversity. Murata (2009) takes a step further by not only considering the optimum composition diversity but also the composition of goods. He analyzes theoretically how the composition and the degree of diversity are driven by the technological feasibility and desirability. All those theoretical models assume that income affects the pattern of consumption and from this disruption, they explain how the optimum product diversity differs across countries. However, UNDP (1998) underlines the general change in the pattern of consumption of both developed and developing economies over time. A common factor to every economy is globalization. Over time, trade cost has been reduced and economies have grown to produce more goods in their comparative advantage sectors. Therefore, nowadays, consumers face larger product diversity regardless their income. This phenomenon is investigated in this analysis. Starting from Bils and Klenow (2001), I reverse the relation between the product diversity and the pattern of consumption and analyze how consumers react to this increase in product diversity. To connect globalization to the change in the pattern of consumption through the extensive margin, I build this analysis on the framework initiated by Feenstra (1994) and Broda and Weinstein (2006). These works aim to quantify the aggregate gains from trade along the extensive margin. They were extended to Costa Rica by Arkolakis, Demidova, Klenow, and Rodriguez-Clare (2008) and to a finer decomposition of gains from trade between growth of product diversity and pro-competitive effect in Feenstra and Weinstein (2010). Blonigen and Soderbery (2010) also refine the concept of gains from trade by using a market-based dataset to define good varieties at a more precise level. A large empirical field has emerged to analyze the impact of globalization on the extensive margin. For instance, Debaere and Mostashari (2010) quantify the contribution of tariff to changes in the number of varieties imported. Kehoe and Ruhl (2009) show that changes in the extensive margin are driven by the trade liberalization but provide an alternative definition of the extensive margin (named relative threshold contrary to what they call the zero threshold value).

Those studies only quantify the gains from trade without considering their allocation in the economy and assuming that the level of expenditure drives the optimum product diversity. A new strand of literature analyzes how the growth of the number of intermediate inputs imported in an economy affects the domestic optimum product diversity (Colantone and Crino (2011), Goldberg, Khandelwal, Pavcnik, and Topalova (2010)). I contribute to this literature into two ways. First, I analyze how the gains from trade along the extensive margin are distributed across sectors and how they affect the pattern of consumption. Second, I assume that the optimum product diversity is not only affected by the level of expenditure but also by the development of technological shocks in trading partners as well as trade costs reduction. I then show how the optimum product diversity can be the cause of the expenditure change instead of being its consequence.

By considering changes in the extensive margin due to technological shocks in trading partners, this paper is also connected to studies analyzing how productivity shocks are transmitted across countries through trade. Hsieh and Klenow (2009) show that the misallocation of inputs across firms in emerging economies such China and India is important and thereby leaves room for faster development without requiring innovation. Hsieh and Ossa (2011) highlight the extensive margin as a new channel through which the productivity shocks can be transmitted. They conclude to a small transmission of the positive Chinese productivity shocks especially to the United States. However, they need to assume a constant structure of expenditure to derive their results. di Giovanni, Levchenko, and Zhang (2012) contrast these results by embedding the bilateral relation between China and United States in a global framework and conclude that it is important to compare the Chinese productivity not to an individual country but rather its similarity to the weighted world average productivity. I complete those studies by shedding light on a new channel through which productivity shocks in a trading partner can affect the economy of the importer; the changes in the pattern of consumption.

The paper is organized as follows. Section 2 reports stylized facts and describes the data. A theoretical framework is presented in Section 3. Section 4 provides evidence of a positive correlation between the changes in the extensive margin and the changes in the pattern of consumption as well as the identification strategy to establish the causal relation. Finally, Section 5 concludes the paper by discussing the implications of the results.

# 2 Data Description and Stylized Facts

## 2.1 Data Sources and Definitions

The data come from the UN Comtrade Statistics Database and from United States International Trade Commission (USITC hereafter). The data are reported in the Harmonized System (HS hereafter) classification codes at the 6-digit level in the UN Comtrade Statistics database and at the 10-digit level in the USITC. I primarily focus on the period between 1993 and 2006, but I also check the robustness of my identification strategy with data from 1989. The Harmonized System is an internationally standardized system which theoretically covers all commodities in international trade. If this classification greatly facilitates the comparison of countries in terms of flows of commodities, the recurrent classification changes lead to potential measurement errors. To minimize them, I first focus the analysis on commodities expressed in kilos in the HS classification to make products comparable as possible. Those commodities represent 68.39% of the whole dataset. Second, I organize the sample in order to make different years truly comparable. The HS nomenclature is amended every four to six years. The purpose of these amendments is to bring the HS nomenclature in line with the current international trade patterns, technological progresses and customs practices. In 1996 and in 2002, *structural changes* were implemented.[1] Those changes preclude a comparison of commodities over time since one code might not represent the same product from one year to another year. I disregard the commodities that have been redefined or reclassified at HS 6-digit level. Finally, the imports have been removed from countries having been divided, reunified after 1991 or reclassified in the database. I am left with 2099 HS 6-digit goods and 196 countries.

Another crucial task was to define the concept of "new varieties". I follow the Armington assumption. A variety is defined as a HS 6-digit good produced in a particular country. For instance, dark chocolate is a good while Belgian dark chocolate is one variety. A new variety is a variety that does not have a positive record in that HS 6-digit category previously. This definition of a variety is broad. However, my identification strategy requires information on varieties imported by other countries. This constraint prevents the use of more disaggregated data. Moreover, the use of Feenstra (1994)' s price index enables to capture precise information while using aggregated data.[2]

[1] Those modifications consist in merging, splitting categories or both at the same time (called complex changes).

[2] The analysis has also been carried on at a HS 4-digit level. At this level, I define a new variety as a HS 6-digit/partnername product.

## 2.2 Stylized Facts

This section explores the association between the growth of trade, the expansion of imported varieties and the U.S. pattern of consumption. The goal of this analysis is to first show how trade liberalization has made available a wide variety of products to American consumers. Second, I illustrate the correlation between the growth of imported varieties and the allocation of expenditure across sectors.

Trade flows have been expanding for many decades. The share of imports in U.S. GDP has been rising from 10.94 % in 1993 to 16.83 % in 2006.[3] The augmentation in the U.S. import has been accompanied by a rise in the number of imported varieties.[4] 157,911 varieties were imported either in 1993 or in 2006. 40% of those varieties were imported in 2006 but not in 1993 while 19 % of varieties were not anymore available on the American good market in 2006. In other words, in 2006 more imported varieties were available to American consumers.

Table 1: Changes in the number of varieties by sector between 1993 and 2006

| Variable | Number of HS 6-digit categories | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| Difference in level | 1135 | 5.92 | 5.14 | 1 | 36 |
| Difference in percentage | 1135 | 36.41 | 18.14 | 3.70 | 71.42 |

I disregard the HS 6-digit sectors for which the growth of the number of varieties is either less than 10% of the median growth between 1993 and 2006 or more than 2 times.

Table 1 scrutinizes this surge of varieties per sector. It shows that the number of imported varieties has risen over time by 36.41% on average per HS 6-digit sector. Moreover, it highlights the heterogeneity of the varieties expansion across sector.

Those observations pin down that trade liberalization has changed the environment of consumption by altering the range of products available to consumers. The asymmetric distribution in the number of varieties must affect consumer's decision. Indeed, in the sectors that have been subject to larger variety expansions, consumers can better match their expenditure to their preferences. One can then expect to observe a change in the allocation of expenditure across sectors. Such changes may help to explain the observed change in the composition of consumers' expenditure.[5] Table 12, in appendix, illustrates the evolution of the Personal Consumption Expenditure (hereafter PCE) in the United States for the 1988-2007 period.[6] It indicates large

[3] World Development Indicators

[4] As I mentioned in the previous section, I define a good as a six-digit category and a varieties as the import of a particular good from a specific country.

[5] For further information on the changes in the consumption pattern refer toUNDP (1998) for a worldwide analysis and to McCully (2011) for the U.S. case

[6] PCE is the primary measure of consumer spending on goods and services in the U.S. economy.

variations in the expenditure share spent in each category.

Those observed patterns show that the relation between product diversity and the variations in the pattern of consumption deserves a deeper analysis.

The following two decompositions aim to better understand how product diversity affects the consumption pattern. The total expenditure spent to import in sector g at time t, $X_{gt}$ can be decomposed into an extensive margin and an intensive one:

$$X_{gt} = n_{gt}\bar{X}_{gt} \tag{1}$$

where $n_{gt}$ is the total number of varieties imported in sector g at time t and $\bar{X}_{gt}$ is the average expenditure spent in each imported variety which captures the price effect.

Table 2 reports the regression decomposition for 2006. The coefficients in each row sum to unity. The changes in the number of varieties account for 25.3% of the expenditure variation across sectors in 2006.[7] Columns (3) and (4) report the results when the quantity is considered instead of the expenditure.

Table 2: Decomposition of the expenditure in 2006

| | (1)<br>$\ln(n_{06})$ | (2)<br>$ln(\overline{X_{06}})$<br>(expenditure) | (3)<br>$ln(n_{06})$ | (4)<br>$ln(\overline{Q_{06}})$<br>(quantity) |
|---|---|---|---|---|
| $\ln(X_{2006})$ | 0.253*** | 0.747*** | 0.120*** | 0.880*** |
| | (0.005) | (0.005) | (0.005) | (0.005) |
| Constant | -1.495*** | 1.495*** | 0.940*** | -0.940*** |
| | (0.085) | (0.085) | (0.082) | (0.082) |
| Observations | 3079 | 3079 | 3079 | 3079 |
| $R^2$ | 45.5 | 87.9 | 14.6 | 90.1 |

Standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01
The table reports 2007 OLS decomposition of variation
in the U.S. imports along two margins: the number of trading partners
the number of trading partners and the intensive margin.

The contribution of the intensive and the extensive margins can also be analyzed over time. I follow Bernard and al. (2009)'s methodology to decompose the change in aggregate U.S. trade between period t-1 and t. Let $\triangle x_t$ denotes the change in the total U.S. imports between t-1 and t. $\triangle x_t$ can be decomposed into the increase due to the entry of new variety, the decrease due to the exit of existing importers, and the change due to increases or decreases in trade for the

[7] Bernard, Jensen, Redding, and Schott (2009) find that the contribution of the intensive margin is higher for the import than for the export and they note that such observation may reflect the fact that the concentration of trade amongst importers may be higher than amongst exporters. The large contribution of the intensive margin may then be particular to the United States.

continuing firms.

$$\triangle x_t = \underbrace{\sum_{c \in N} x_{ct} - \sum_{c \in E} x_{ct-1}}_{\text{extensive margins}} + \underbrace{\sum c \in G \triangle x_{ct}}_{\text{intensive margin}} \tag{2}$$

where c is the trading country, N is the set of new trade countries, E is the set of existing trade countries exiting and G is the set of countries continuing to trade. Table 3 decomposes the total U.S. import variation into the contribution of the margins described above from 1992 to 2004. The first 12 columns report annual changes, the next two report 7 years changes (from 1992 to 1998 and from 1998 to 2004) and the last column reports the 12 years change (from 1992 and 2004).

Table 3: Decomposition in the variations of U.S. imports over time

| | 92-93 | 93-94 | 94-95 | 95-96 | 96-97 | 97-98 | 98-99 | 99-00 | 00-01 | 01-02 | 02-03 | 03-04 | 92-98 | 98-04 | 92-04 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Extensive Margins** | | | | | | | | | | | | | | | |
| New varieties (%) | 38.26 | 21.22 | 32.54 | 83.75 | 11 | 14.42 | 14.36 | 13.67 | 21.09 | - | 11.39 | 9.22 | 40.16 | 17.26 | 29 |
| Exit (%) | -30.50 | -8.99 | -6.84 | -11.46 | -6.48 | -7.11 | -13.75 | -5.36 | -46.58 | - | -5.23 | -6.12 | -2.48 | -16.43 | -9 |
| Net Entry (%) | 7.76 | 12.23 | 25.7 | 72.29 | 4.52 | 7.31 | 0.61 | 8.31 | -25.49 | - | 6.16 | 3.1 | 37.68 | 0.83 | 20 |
| **Intensive margin** | 92.24 | 87.77 | 74.30 | 27.71 | 95.44 | 92.68 | 99.4 | 91.67 | -74.51 | - | 93.84 | 96.9 | 62.32 | 99.17 | 80 |
| **Total change in import** | | | | | | | | | | | | | | | |
| percentage | 6.5 | 15 | 18.59 | 15.9 | 10.99 | 7.8 | 4.9 | 12.86 | -3 | 0 | 9.8 | 17.4 | 102 | 47.67 | 149.67 |
| $ billion | 0.88 | 22.6 | 31.1 | 31.7 | 25.4 | 20 | 13.6 | 87.3 | -9.97 | 0 | 31.1 | 60.6 | 140 | 132 | 272 |

Table 3 shows a positive growth in the number of varieties over time except for the period corresponding to the American economic recession (2001-2002) which explains why the net entry is so low between 1998 and 2004. The exceptional growth in 1995-1996 was driven by the imports of vehicle and aircraft accessories mainly from European countries, Mexico, Taiwan and Korea.[8] As Bernard and al. (2009), I find that the short-run changes in the U.S. imports are largely accounted for by the intensive margin while the long-run decomposition highlights a large growth of the extensive margin (29 %). These observations explain why I will concentrate the analysis on long differences instead of annual changes.

This section has focused on the relationship between growth of the number of varieties and the evolution of the U.S. pattern of consumption, leaving aside the sources of this expansion. Several explanations involving a globalization process (coupled with an assumption that goods are differentiated by country) explain this increase in the number of imported varieties. The reduction in trade costs may have made cheaper to source new varieties from different countries.

[8] Such large growth may have been driven by the European demand of aircrafts. Indeed, neither the MFN nor the Mexican tariffs in those sectors have fallen by a large amount but from 1993 to 1997, Europe has deregulated its sky and has seen the emergence of low cost companies using American aircrafts. Outsourcing might be another explanation.

Moreover, Emerging economies such China or India have also started to produce additional varieties. The following table shows the ranking of countries at the beginning of the time frame and at the end.

Table 4: Ranking of countries in terms of number of goods imported by the United States

| | Ranking in year | | | | | | |
|---|---|---|---|---|---|---|---|
| Country | 1992 | 1996 | 2004 | Country | 1992 | 1996 | 2004 |
| Canada | 1 | 1 | 1 | Switzerland | 11 | 11 | 15 |
| Germany | 2 | 2 | 3 | *Korea, Rep.* | 12 | 13 | **10** |
| United Kingdom | 3 | 3 | 4 | Belgium-Luxembourg | 13 | 12 | 14 |
| Japan | 4 | 4 | 6 | Hong Kong, China | 14 | 17 | 19 |
| France | 5 | 5 | 5 | Brazil | 15 | 18 | 16 |
| Italy | 6 | 6 | 7 | Spain | 16 | 14 | 13 |
| Mexico | 7 | 7 | 8 | Sweden | 17 | 16 | 18 |
| *China* | 8 | 8 | **2** | *India* | 18 | 15 | **11** |
| *Taiwan, China* | 9 | 10 | 9 | Austria | 19 | 20 | 21 |
| *Netherlands* | 10 | 9 | 12 | Australia | 20 | 19 | 17 |

Table 5: Top 30 countries in 1996 included.

Countries are ranked by the number of varieties they export to the United States. The first column ranks the countries from the highest to lowest for 1992, and the following columns rank them for subsequent years. Relevant information can be pinned down from table 5. First of all, high-income economies and proximate economies are ranked among the largest exporters. Canada and Mexico are well ranked which may reflect free trade areas and other trade liberalization. The economic growth coupled with trade liberalization also appears to have played an important role in the product diversity growth. China, Korea, India or Indonesia rose dramatically in the ranking. Those countries contributed heavily to the increase in available varieties for the American consumers. China has known a growth of 68% in the varieties it exports to the United States, India has increased the number of exporting varieties by 90% and Turkey by 148%.

Table 5 reveals potential causes of changes in varieties and in the composition of the exporters to the U.S.. The economic growth of Asian countries as well as trade liberalization has changed the number of varieties available to the American consumers. Such growth has not been uniform across sectors. Countries will export more varieties in sectors where they have a comparative advantage. Therefore, the emergence of some countries may explain the heterogeneous changes in the optimum product diversity available in each sector.

## 3 Theoretical Framework

This section presents a theoretical framework for quantifying the effect of changes in the number of varieties available for the consumer on the pattern of expenditure in presence of horizontally differentiated products. This relationship necessitates a functional form for the utility function that endogeneizes the pattern of consumption (i.e. an interdependence between sectors).

Broda and Weinstein (2006) define the preferences of a representative agent by a three level utility function. They specify the upper level utility function as

$$U^{\kappa} = (D^{(\kappa-1)/\kappa} + M^{(\kappa-1)/\kappa})^{\kappa/(\kappa-1)}; \qquad \kappa > 1 \tag{3}$$

where M is the composite import good which will be defined below. D is the domestic good and $\kappa$ is the elasticity of substitution between domestic goods and imports. This functional form creates a separability between imported and domestic goods which allows distinguishing import price index from the domestic one.

The composite imported good is defined as

$$M_t = (\sum_{g \in G} \theta_{gt} M_{gt}^{\frac{\gamma-1}{\gamma}})^{\frac{\gamma}{\gamma-1}}; \qquad \gamma > 1 \tag{4}$$

where $M_{gt}$ is the subutility derived from the consumption of imported good g at time t,$\gamma$ denotes the elasticity of substitution among imported goods, and G is the set of all imported goods.[9] $\theta_{gt}$ is a taste parameter.

$M_{gt}$ is defined by the nonsymmetric CES function represented by

$$M_{gt} = (\sum_{c \in C} d_{gct}^{\frac{1}{\sigma_g}} M_{gct}^{\frac{\sigma_g-1}{\sigma_g}})^{\frac{\sigma_g}{\sigma_g-1}} \sigma_g > 1; \qquad \forall g \in G \tag{5}$$

where $\sigma_g$ is the elasticity of substitution between varieties within a particular industry g. C is the set of all countries. $d_{gc}$ denotes a taste or quality parameter for variety c from good g.[10]

The representative consumer uses a two stage budgeting. In the second stage, for a given expenditure on import allocation $E_{import}$, expenditure share in sector g is

$$s_{gt} = \frac{\theta_{gt} P_{gt}^{1-\gamma}}{\sum_{i \in G} \theta_{it} P_{it}^{1-\gamma}} \tag{6}$$

[9] while it would be interesting to analyze the case $\gamma < 1$ since the latter represents the elasticity of substitution between sectors. However, the consumer surplus derived from additional varieties requires $\gamma > 1$

[10] the country of origin is taken as the demarcation of a variety.

Since I am interested in the change in the expenditure share, I take the ratio of eq. 6 between two periods, t-1 and t:

$$\frac{s_{gt}}{s_{gt-1}} = \frac{\theta_{gt}}{\theta_{gt-1}} \left[ \frac{P^M_{gt}(I_{gt}, \vec{d_{gt}})}{P^M_{gt-1}(I_{gt-1}, \vec{d_{gt-1}})} \right]^{1-\gamma} \left[ \frac{\sum_{i \in G} \theta_{it} P^{1-\gamma}_{it}}{\sum_{i \in G} \theta_{it-1} P^{1-\gamma}_{it-1}} \right] \tag{7}$$

where $P^M_{gt}(I_{gt}, \vec{d_{gt}})$ denotes the price index for import of the single good, $M_{gt}$ and the denominator characterizes the aggregate price index.[11] Following Feenstra (1994) and Broda and Weinstein (2006), these indexes play a pivotal role in this study since they enable to disentangle the price effect from the "variety"' effect. The index theory drawn up on Diewert (1976)'s dual theory considers the representative consumer's problem in terms of cost to reach a certain level of satisfaction. The presence of more varieties enables the representative consumer to better match her tastes with her consumption and then decrease her cost to reach a certain level of satisfaction.

The unit cost function of the utility function defined in (5) is the following:

$$c(p_{gct}, I_{gt}, \vec{d_{gt}}) = (\sum_{c \in I_{gt}} d_{gct} p_{gct}^{1-\sigma_g})^{\frac{1}{1-\sigma_g}} \tag{8}$$

where $I_{gt}$ is a subset of goods available in time t.

Sato (1976)and Vartia (1976) derive a price index for a constant common set of varieties ($I = I_{t-1} \cap I_t$) and a constant taste parameter, $d_{gc}$:

$$\frac{c(p_{gct}, I_{gt}, \vec{d_{gc}})}{c(p_{gct-1}, I_{gt-1}, \vec{d_{gc}})} = P_{SV}(\vec{p_{t-1}}, \vec{p_t}, \vec{m_{t-1}}, \vec{mt}) = \prod_{i \in I} (\frac{p_{gct}}{p_{gct-1}})^{w_{gct}(I)} \tag{9}$$

where

$$w_{gct}(I) \equiv \frac{s_{gct} - s_{gct-1}/ln(s_{gct}) - ln(s_{gct-1})}{\sum_{i \in I} s_{gct} - s_{gct-1}/ln(s_{gct}) - ln(s_{gct-1})}; \tag{10}$$

$$s(I) = \frac{p_{gct} m_{gct}}{\sum_{i \in I} p_{gct} m_{gct}} \tag{11}$$

While the tractability of this price index makes it appealing, it only concerns a common set of goods available between the two periods. Feenstra (1994) modifies this exact price index to incorporate changes in varieties of a single good as long as there is some overlap in the varieties

[11] I remind that a good is characterized by a price index because each good has several varieties.

available between the two periods ($I \neq \varnothing$) and a constant parameter $d_{gc}$:

$$\pi_g^F(I_g) = \frac{c(p_{gct}, I_{gt}, \vec{d_{gc}})}{c(p_{gct-1}, I_{gt-1}, d_{gc})} = P_{SV}(\vec{p_{t-1}}, \vec{p_t}, \vec{m_{t-1}}, \vec{mt})(\frac{\lambda_{gt}(I)}{\lambda_{gt-1}(I)})^{\frac{1}{\sigma_g - 1}} \tag{12}$$

where the weights $w_{gct}(I)$ is defined as in (10) and $\lambda_{gt}(I)$ and $\lambda_{gt-1}(I)$ are defined as the following

$$\lambda_{r,g} = \frac{\sum_{c \in I} p_{gcr} m_{gcr}}{\sum_{c \in I_{gr}} p_{gcr} m_{gcr}}; \qquad r = t-1, t \tag{13}$$

The exact price index with variety changes equals the conventional exact price index multiplied by an additional term, $\left(\frac{\lambda_{gt}(I)}{\lambda_{gt-1}(I)}\right)^{\frac{1}{\sigma_g - 1}}$ which captures the role of new and disappearing varieties. $\lambda_{gr}(I) \leq 1$ can be interpreted as the period r expenditure on varieties in the common set I relative to the period r total expenditure. Therefore $1 - \lambda_{gr}(I)$ can be interpreted as the period r expenditure on new varieties relative to the period r total expenditure. Therefore, when there is a greater number of new varieties in period r, the value of $\lambda_{gr}$ will tend to be lower which leads to a greater fall in the cost of living by an amount that depends of the elasticity of substitution between variety, $\sigma_g$. In other words, more varieties lower the cost of reaching a certain level of satisfaction and this decrease depends how similar is the new variety to the variety already consumed.

Feenstra's price index is limited to an intra-sectoral analysis and assumed a fixed expenditure share for each sector. Broda and Weinstein (2006) aggregate Feenstra's price index to the sector level by taking its geometric mean weighted by the logarithmic mean of the expenditure share allocated across sectors for some overlap in the varieties available between the two periods ($I \neq \varnothing$) and a constant parameter $\theta_g$:

$$\pi^{BW} = \frac{c(\vec{p_{gt}}, I_{gt}, \vec{d_{gct}}, b_{gt})}{c(\vec{p_{gt-1}}, I_{gt-1}, \vec{d_{gt-1}}, b_{gt-1})} = \prod_{g \in G} (P_{SV\,gt}(I_g))^{w_{gt}} (\frac{\lambda_{gt}(I)}{\lambda_{gt-1}(I)})^{\frac{w_{gt}(I)}{\sigma_g - 1}} \tag{14}$$

where $w_{gt}(I) \equiv \frac{s_{gt} - s_{gt-1}/ln(s_{gt}) - ln(s_{gt-1})}{\sum_{i \in G} s_{gt} - s_{gt-1}/ln(s_{gt}) - ln(s_{gt-1})}$

Substituting eq. (12) and eq. (14) into eq. (7), I obtain:

$$\frac{s_{gt}}{s_{gt-1}} = \frac{\theta_{gt}}{\theta_{gt-1}} (\pi_{gt}^F)^{1-\gamma} (\pi_t^{B\&W})^{\gamma - 1} \tag{15}$$

This expression provides a relationship between the extensive margin and the expenditure share devoted to a particular sector g. Other things unchanged, an expansion of the number of varieties available for consumer in a particular sector improves the ability of the consumer to match her expenditure to her preferences and then decreases her cost to reach a certain level of satisfaction. It decreases the price index $\pi_g^F(I_g)$ and increases the expenditure in g by an amount that depends on the elasticity of substitution between sectors and the taste parameter, $\theta_{gt}$ and decreases the expenditure share in other sectors.

Expressed in logarithm and decomposing Feenstra's price index between a price effect and a " variety" effect, the changes in import demand in a particular sector g is the following:

$$\begin{aligned}\triangle\, ln(s_{gt}) &= \triangle ln(\theta_{gt}) + (\gamma-1)ln(\pi_t^{B\&W}) + (1-\gamma)ln\left(P_{SV\,gt}(I_g)\right) \\ &\quad + \frac{(1-\gamma)}{\sigma_g-1}ln\Big(\frac{\lambda_{gt}(I)}{\lambda_{gt-1}(I)}\Big)\end{aligned} \tag{16}$$

If the varieties are symmetric in a standard monopolistic competition model all varieties will be equally priced at $p_{git}$ and consumed in the same quantity. Therefore, the "variety component"' becomes the ratio between the number of varieties of good g consumed in periods t and t-1 and eq. (17) can be rewritten:

$$\begin{aligned}\triangle\, ln(s_{gt}) &= (\gamma-1)ln(\pi_t^{B\&W}) + (1-\gamma)ln\left(P_{SV\,gt}(I_g)\right) \\ &\quad + \frac{(1-\gamma)}{\sigma_g-1}ln\Big(\frac{n_{gt}(I)}{n_{gt-1}(I)}\Big)\end{aligned} \tag{17}$$

The aggregate price index affect every expenditure share in an identical way, it will then be captured by a year fixed effect. The following regression can then be estimated. Notice that $\theta$ captures any preferences shocks other than the one induced by the relative change in the number of varieties. I assume the latter constant since the objective is to quantify how the expansion of varieties reveals better *given* preferences. Moreover, the constant preference over 10 years or 16 years might be reasonable. Finally, the relative preference of a sector is captured by the initial size of the market.

$$\triangle ln(s_{gt}) = \alpha + D_t + \beta_1 ln\left(P_{SV\,gt}(I_g)\right) + \beta_2 ln\Big(\frac{n_{gt}(I)}{n_{gt-1}(I)}\Big) + s_{i0} + \epsilon_g \tag{18}$$

This relation is traditionally analyzed in the other direction. The preferences change the expenditure shares which then affect the optimum diversity. In this study, I reverse the relation to study how the expenditure shares respond to an alteration of the consumer's environment.

# 4 Empirical Analysis

## 4.1 Reduced-form Evidence

This section provides some baseline correlations followed by a number of robustness checks. The identification issues are discussed in subsection 4.3.
The following equation is assessed.[12]

$$ln(\frac{x_{i,2006}}{x_{i,1993}}) = \beta_0 + \beta_1 ln(\frac{n_{i,2006}}{n_{i,1993}}) + \beta_2 ln(\frac{p_{i,2006}}{p_{i,1993}}) + ln(x_{i,1993}) + D_j + \epsilon_{i,t} \quad (19)$$

where i represents the sector. $x_i$ measures either the value of imports in nominal U.S. dollars of product i in year t or the quantity (in kilos and denoted q in the following tables) of imports of this product. Those are transformed in share through the use of the constant $\beta_0$. $n_i$ is the total number of variety imported in a particular sector i. $p_i$ represents the unit value of the import at sector i. The initial value, ln($x_{i,1993}$), captures the importance of the initial size of sector i at time t. The fixed effect, $D_j$, aims to capture the potential effect of the changes in income in the allocation of expenditure in the aggregated industry. A change in income can affect the pattern of consumption. If the preferences are non homothetic, this change is captured in the disturbance which might be correlated with the independent variables. The coefficient of interest, $\beta_1$, captures the elasticity of the expenditure share with respect to the extensive margin. Finally, the standard errors are clustered at the sector level and adjusted for heteroskedasticity.

Table 6 shows the results for eq. (19). Columns (1), (3), (5), (7) and (8) assess the impact of the extensive margin and the price on the CIF trade value while columns (2),(4), (6) estimate the quantity expressed in kilos. Two levels of aggregation are assessed. The first 6 columns show the results when the regressions are carried on at the HS 6-Digit level while the last two columns describe the results at the HS 4-digit level. At the HS 6-digit level, the number of varieties follows the Armington assumption while at the HS 4-digit level, a variety is a country-variety.

The first two columns show a positive correlation between the change in the number of varieties and the change in the expenditure share. My concern is that the results might be upward biased due to the large drop of varieties imported in 2001 and 2002 during the American recession.[13]This drop can bias the results because the price is computed for varieties which have been continuously traded from 1993 to 2006. Given the large drop in 2001 and 2002, the price

[12] As Trefler (2004) describes, the short timeline (1993-2006) does not provide the opportunity to take the first difference annually.

[13]57,000 and 80,000 HS 6-digit partnername varieties disappeared in those years respectively while on average the number of HS 6-digit partername varieties disappearing was around 20,000.

captures few products which are mainly imported from developed economies. Therefore, the importance of the extensive margin might be inflated relative to the intensive margin. This prediction seems to be confirmed in columns (3) and (4). Indeed, those columns show the results when the years 2001 and 2002 are disregarded and for this sample the coefficient of the extensive margin is slightly lower and the coefficient of the price is higher. Those results have also a higher predictive power.

Columns (5) and (6) include a fixed effect capturing time invariant shocks at a HS 4-digit level of aggregation. The potential concern is that an omitted bias is induced by an income effect in presence of non homothetic preferences. The reallocation across sectors could be driven by the changes in the income instead of the changes in the environment. The reallocation of income usually happened at a higher level of aggregation. The estimate of the coefficient of the extensive margin is lower. This result indicates that the extensive margin might have captured a part of the income effect. Those regressions are my favourite since they have the highest predictive power and they capture the change in the environment, the price effect as well as the income effect. Finally, in the last two columns, the regression is estimated at the HS 4-digit level of aggregation. The effect of the extensive margin is similar than at the HS 6-digit level.

For all regressions, one can note the ambiguous impact of the price on the expenditure share. This result can be explained by the fact that, on the one hand, the unit value is computed from the trade value. On the other hand, there is the negative impact of the price on quantities. The constant shows a positive and significant sign which may reflect the positive trend of the United States to import. Finally, one can observe the negative and significant impact of the initial market size on the expenditure. Small sectors then experience larger effects through the extensive margin.

Those results reveal key information. First of all, the positive correlation between the change in the extensive margin and the expenditure share. Secondly, the impact of the extensive margin is stronger than the one of the price (and then than the intensive margin effect).

Table 6: Baseline Specification

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | Δ lnx | Δ lnq | Δ lnx | Δ lnq | Δ lnx | Δ lnq | Δ lnx | Δ lnx |
| Δ lnn | 1.194*** | 1.235*** | 1.093*** | 1.166*** | 0.984*** | 1.064*** | 1.105*** | 1.016*** |
| | (0.109) | (0.120) | (0.073) | (0.081) | (0.112) | (0.128) | (0.100) | (0.116) |
| Δ lnp | -0.003 | -0.108*** | -0.033 | -0.258*** | -0.051 | -0.239*** | 0.143** | 0.149** |
| | (0.034) | (0.040) | (0.034) | (0.043) | (0.036) | (0.041) | (0.070) | (0.070) |
| $\ln(x_{1993})$ | -0.094*** | | -0.120*** | | -0.191*** | | | -0.085*** |
| | (0.019) | | (0.014) | | (0.022) | | | (0.020) |
| $\ln(q_{1993})$ | | -0.091*** | | -0.119*** | | -0.193*** | | |
| | | (0.014) | | (0.012) | | (0.021) | | |
| Constant | 2.189*** | 1.935*** | 2.539*** | 2.302*** | 3.711*** | 3.414*** | 0.659*** | 2.157*** |
| | (0.320) | (0.220) | (0.235) | (0.192) | (0.351) | (0.314) | (0.050) | (0.370) |
| Observations | 1123 | 1123 | 2099 | 2099 | 2099 | 2099 | 662 | 662 |
| $R^2$ | 18.9 | 19.7 | 20.0 | 23.8 | 21.7 | 24.0 | 27.6 | 26.6 |
| sample period | 1993-2006 | 1993-2006 | 1993-2006 | 1993-2006 | 1993-2006 | 1993-2006 | 1993-2006 | 1993-2006 |
| | | | 01 and 02 excl. | 01 and 02 excl. | 01 and 02 excl. | 01 and 02 excl. | 01 and 02 excl. | 01 and 02 excl. |
| sector aggregation | HS6 | HS6 | HS6 | HS6 | HS6 | HS6 | HS4 | HS4 |
| sector FE | No | No | No | No | Yes,HS4 | Yes, HS4 | No | Yes,HS2 |

Standard errors in parentheses.They are clustered by sectors and adjusted for the heteroskedasticity.

All specifications are estimated by OLS. In column (5), (6) and (8), I control for higher level of aggregation sectoral effect.

* $p<0.1$, ** $p<0.05$, *** $p<0.01$

## 4.2 Sensitivity Analysis

Table 7 describes the results of analysis assessing the robustness of the above positive correlation between the number of varieties and the expenditure share. In rows (1) and (2), I regress the logarithm of the expenditure share on the logarithm of the relative number of varieties and the relative price to assess whether the constant in the previous regression generates the shares instead of the variables in level. I obtain exactly the same results than the ones found in columns (5) and (6) in table 6. In rows (3) and (4), I trim the distribution of each variable by disregarding the HS 6-digit sectors for which the value is either less than 10% of the median value between 1993 and 2006 or more than 10 times the median. The results are stronger suggesting that positive correlation is not driven by outliers. Given the change in pattern that happens around 2001/02, I stack the 7-years equivalent first differences for the two periods, 1993 to 1999 and 1999 to 2006. This stacked first difference is similar to a three-periods fixed effects model. The results are sligthly lower but still hold. In rows (7) and (8), I question whether the representative consumer needs some time to adjust her allocation decision. Therefore, I regress the long difference of the expenditure between 1994 and 2005 on the changes in the number of varieties between 1993 and 2006. The coefficients are lower than the ones found in table 6. In other words, the representative consumer adjusts instantaneously her behaviour to the change of her environment.

All the above results apply to all the products imported by the United States while this analysis mostly concerns final goods. The Bureau of Economic Analysis (BEA) classifies goods into six principal "end-use" categories in order to identify the end-use of goods. This classification is very broad and includes few products in the consumption goods categories. For instance, wine and computers are not included in consumer goods. Alternatively, Antras, Chor, Fally, and Hillberry (2012) define a finer measure of upstreamness of products which enable to identify their end-use. This measure captures the average position of an industry output in the value chain. The measure of upstreamness ranges from a minimum of 1 (final goods) to a maximum of 4.65 (Petrochemicals). The weakness of this measure in the context of this analysis is the absence of a clear cutoff between final goods and intermediate goods. In order to define precisely what is a consumption good, I combine both classifications. To use the measure of upstreamess defined by Antras et al. (2012), I use the concordance table of 2002 provided by the BEA to apply the measure to HS classification and I drop all products with a measure of upstreamness above 1.5. I then use the concordance made by Feenstra, Romalis, and Schott (2002) to identify goods classified as consumption goods by the "End-Use Commodity Category" of the BEA. I finally screen the goods selected by those two classifications and I drop goods that are unlikely to be used as final goods.[14] The group of products derived is broader and includes goods such a car, computer or food products that are mainly used as final goods. The coefficients of the extensive margin in rows (9) and (10) are still positive and have a higher value than when all goods are considered. In other words, the changes in the environment seem to affect stronger the final goods.[15]

Finally, I use a nonlinear method of estimation to assess the quality of the log-linearization estimated in table 6. Santos Silva and Tenreyro (2006, 2009) question the validity of the log-linearization of multiplicative models in presence of heteroskedasticity and non-negative values such a trade flow data. The nonlinear transformation of the dependent variable changes the properties in a non-trivial way and in presence of heteroskedastic errors, the transformed errors will generally be correlated with the covariates leading to inconsistent estimators. Santos Silva and Tenreyro (2006, 2009) propose to apply the Pseudo Poisson Maximum Likelihood (PPML) estimator. They apply it on the gravity equation on cross-sectional data. Hausman, Hall, and Griliches (1984) propose fixed-effects poisson procedures which can be applied to panel data. While Hausman et al. (1984) impose strong restrictions on the mean and on the error term, Wooldridge (1999) shows that consistent PPML estimator only requires assumptions on the conditional mean. In rows (11) and (12), I implement the PPML method corrected for the potential heteroskedaticity on eq. (19). The drawback of this method when applied to long difference is that it needs non negative data while the changes in the expenditure share can be negative. Therefore, in rows (11) and (12), sectors which have been shrinking are dropped.
To apply the method on whole the sample, I implement the PPML method on the following regression:

$$x_{i,t} = \beta_0 + \beta_1 ln(n_{i,t}) + \beta_2 ln(p_{i,t}) + D_t + D_i + \epsilon_{i,t} \tag{20}$$

where i represents the sector. $x_i$ measures either the value of imports in nominal U.S. dollars of product i in year t or the quantity (in kilos) of imports of this product. $n_i$ is the total number of variety imported in a particular sector i. $p_i$ captures the price effect. $D_t$ and $D_i$ are year fixed effect and sector fixed effect respectively. The regression assesses the expenditure share by to the use of year fixed effect. The standard errors are clustered at the sector level.

[14] Those goods are defined in appendix.
[15] The regression runned on a sample of final goods as defined by the BEA's classification provides the similar results (1.674 for the expenditure share and 1.825 when the quantities are estimated) but only includes 162 sectors.

The positive correlation still holds but the magnitude is much lower suggesting a potential bias of the log linearization in rows (13) and (14). The importance of the extensive margin relative to the price is also much lower than in columns (5) and (6) of table 6. However, the gap still exists as well as the positive correlation. For sake of clarity, the coefficient of the price is not included in table 7 but the same conclusions can be drawn up as the one derived in table 6; there is an ambiguous effect of the price change on the expenditure share and its effect is weaker than the extensive margin effect.

Table 7: Baseline Specification: Sensitivity Analysis

| | | | Δ ln(n) | Std err. | Obs. | $R^2$ |
|---|---|---|---|---|---|---|
| Share | (1) | Δ ln(x) | 0.984*** | (0.112) | 2099 | 21.7 |
| | (2) | Δ ln(q) | 1.064*** | (0.128) | 2099 | 24.0 |
| Trimmed | (3) | Δ ln(x) | 1.042*** | (0.124) | 1460 | 18.6 |
| | (4) | Δ ln(q) | 1.248*** | (0.135) | 1460 | 20.9 |
| Stacked | (5) | Δln(x) | 0.560*** | (0.067) | 4198 | 14.3 |
| | (6) | Δ ln(q) | 0.614*** | (0.080) | 4198 | 17.3 |
| lag | (7) | Δ ln(x) | 0.824*** | (0.106) | 2099 | 17.4 |
| | (8) | Δ ln(q) | 0.848*** | (0.121) | 2099 | 18.9 |
| Consumption | (9) | Δ ln(x) | 1.239** | (0.479) | 329 | 15.9 |
| | (10) | Δ ln(q) | 1.499*** | (0.547) | 329 | 16.6 |
| Poisson | (11) | Δ x | 0.897*** | (0.217) | 1252 | - |
| | (12) | Δ q | 0.636*** | (0.169) | 1252 | - |
| | (13) | x | 0.612*** | (0.050) | 25188 | - |
| | (14) | q | 0.406*** | (0.064) | 25188 | - |

Standard errors in parentheses.They are clustered by sector and adjusted for heteroskedasticity.
Times dummies are included in regression (13) and (14) in order to compute shares.
* $p<0.1$, ** $p<0.05$, *** $p<0.01$

To summarize, a positive correlation is identified between the change in the extensive margin and the change in the expenditure share allocated to each HS 6-digit or HS 4-digit sectors. This correlation is stronger than the one between the price change and the expenditure share change and it holds when the income effect is controlled. The sensitivity analysis confirms this positive correlation and brings to the light the instantaneous adjustments as well as the stronger impact for consumption goods.

## 4.3 Identification Strategy

The above positive correlation between the changes in the number of varieties and the changes in the expenditure share is compatible with two explanations. The traditional literature in new trade theory assumes that changes in the expenditure drive the variations in the number of varieties. In other words, the representative consumer's taste changes over time. Those changes are captured by the expenditure pattern and then affect the number of varieties consumed. However, a few papers such a Bils and Klenow (2001) suggest that the variety growth could be the cause of the evolution in the pattern consumption. If the set of varieties available expands in a particular sector relative to another one, the representative consumer can better match her tastes to her consumption. One can then expect the representative consumer to allocate more resources in sectors subject to relatively larger variety growth. Therefore, the causal relation is ambiguous and requires a deeper investigation. Moreover, a demand shock can also affect the price leading to a second potential endogenous variable. Therefore, at least two instruments are required to identify the causal effect of the variety expansion relative to the price effect on the expenditure share.

### 4.3.1 Instruments

This subsection discusses the instruments I use to solve for the potential endogeneity of the price and of the changes in the number of varieties imported. I use the variety-specific unit transportation cost for the U.S. to instrument the c.i.f. price as Khandelwal (2010).[16] This instrument varies across countries, industries and years which makes it possible to use with long-difference.

I also need an instrument to identify the causal effect of the change in the number of variety on the change in the expenditure share. In other words, I need to isolate the variations in the extensive margin driven by factors independent of the U.S. expenditure share. The development of emerging countries is an example of factors that disturb the distribution of the number of varieties imported in a particular economy. Indeed, this exogenous variation in the variety expansion can capture technological shocks occurring abroad by enabling third countries to produce and exports new varieties. These shocks would exogenously expand the range of foreign products available to the American consumers. A large literature has been growing on the transmission of the productivity shocks (especially in emerging economies including mostly studies on China)

[16] Those data are sourced from Feenstra et al. (2002). To compute a unit transportation cost at HS 4-digit level, I construct the variable with the same methodology used to compute a unit value at HS 4-digit level.

on the U.S.. As Autor, Dorn, and Hanson (2012) show, growth in low income countries exports over time is driven by China's transition to a market oriented economy and the successive trade reforms which have increased linkages with other countries as well as have decreased the trade costs. As a result, China has moved up the value chain quickly and becomes a major economic player. Moreover, those shocks may grow even over medium run. Indeed, Hsieh and Klenow (2009) show that the misallocation of inputs across firms in emerging economies such China and India is important and then leaves room for faster developments without requiring innovation. This trade flows growth has also been through the extensive margin as highlighted in Table 8. Those results are in line with the observations made by Feenstra (1994), Broda and Weinstein (2006) and Debaere and Mostashari (2010).

Table 8: Country contribution to growth in U.S. varieties (1972/1988 and 1991/2004)

| country | Contribution 72-88 | country | Contribution 91-04 |
|---|---|---|---|
| Taiwan | 5.21 | China | 5.19 |
| Korea | 4.96 | India | 5.09 |
| Mexico | 4.33 | Turkey | 4.27 |
| China | 4.22 | South Africa | 3.62 |
| Canada | 4.19 | Poland | 3.52 |
| Hong Kong | 4.00 | Spain | 3.03 |
| Italy | 3.43 | Indonesia | 3.02 |
| Germany | 3.27 | Thailand | 2.61 |
| France | 3.20 | Korea | 2.38 |
| Japan | 3.03 | Brazil | 2.38 |
| United Kingdom | 2.89 | Bulgaria | 2.36 |
| Brazil | 2.85 | Mexico | 2.26 |
| Israel | 2.70 | NewZealand | 2.09 |
| Thailand | 2.28 | Argentina | 2.03 |
| Singapore | 2.22 | Romania | 2.03 |
| Switzerland | 2.18 | Australia | 1.99 |

The development of those economies either through technology's improvements, successful trade reforms or trade costs reduction has not only been passed on the increase of the number of varieties exported in the U.S. but also in other regions of the world. Indeed, di Giovanni & al. (2012) underline the global feature of the emergence of Chinese exports. Therefore, to identify this supply-driven component of the changes in the extensive margin, I instrument the growth in the number of imported varieties to the U.S. using the contemporaneous composition and growth of the average change in the extensive margin imported by eighteen other developed economies.[17] The average between those countries has been taken to alleviate the potential correlation of demand shocks between those countries. Indeed, some non OECD countries show different economic patterns. Figure 1 shows the evolution of the GDP of the U.S. and of this

[17] Those countries are those that have comparable trade data covering the sample period and exclude Canada due to the similarities it shares with the U.S.. The countries are New Zealand Australia, Greece, Austria, Norway, Portugal, Hungary, Iceland, Ireland, Japan, Korea, Rep., Spain, Sweden, Switzerland, Finland, Netherlands, Denmark, United Kingdom. They are selected according to the World Bank classification of high economies.

average high economy as well as the evolution of their import. One can observe the clear positive trend for the U.S. which is absent for the artifact economy. Another potential concern would be that the positive trend in the U.S. could affect the decision of the other developed countries to import new varieties. Indeed, the U.S. has been subject to an economic growth in the 90's. A part of this growth has been converted into a growth in the import from all countries and then also from those developed economies. Then, those economies may have allocated this incremental income due to additional exports into the purchase of new varieties. In such case, the changes in the number of varieties imported by those developed economies may be correlated with the U.S. expenditure change. Such correlation should be treated by the time effect if all sectors were affected in the same way and by the sector fixed effect. However, such potential shock can also be assessed. Table 9 shows the potential correlation between the long difference of the U.S export (which proxies the U.S. economic growth) before the period covered by the sample with the long difference of the number of varieties imported by this average high economy. This correlation is not significant. Finally, regarding to the literature on the role of the extensive margin and productivity growth (Goldberg et al. (2010),Colantone and Crino (2011)), the economic growth in the U.S. could have boosted the exports to China and then could have been the source of the technological shock. However, U.S. is not an important importer Of China. China imports almost 60 % of its imports from other Asian countries (Ghosh and Rao (2010)). Notwithstanding, the potential correlation between the demand shocks across countries is a source of potential concern and additional measures described below have been taken to alleviate this potential endogeneity.

Figure 1: Evolution of The U.S. relative to the average developed economy

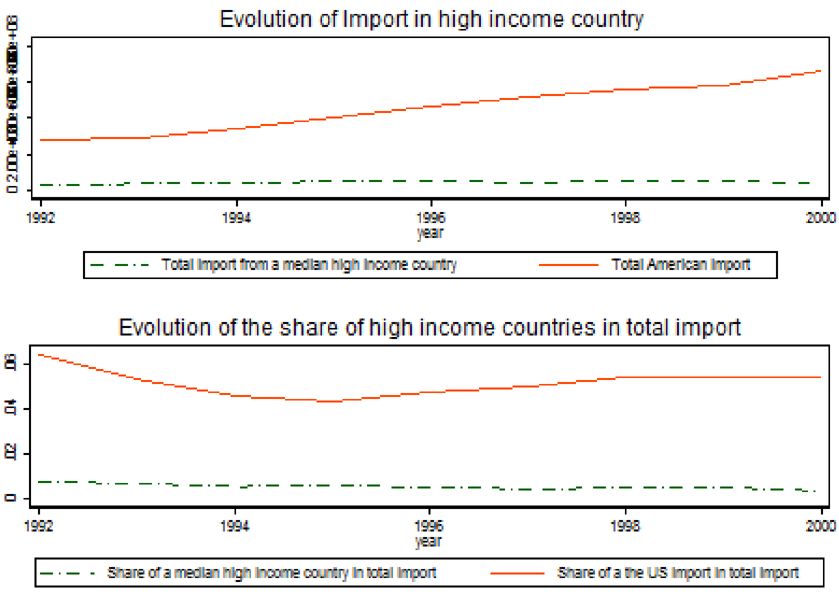


Table 9

| | (1) | (1) |
|---|---|---|
| | $\Delta ln(x)$ | $\Delta ln(x)$ |
| | HS4 level | HS6 level |
| $\Delta ln(n)$ | 0.224 | 0.041 |
| | (0.142) | (0.086) |
| Constant | 0.016 | 0.043*** |
| | (0.022) | (0.015) |
| Observations | 762 | 2808 |
| $R^2$ | 0.005 | 0.000 |

Standard errors in parentheses

* p<0.1, ** p<0.05, *** p<0.01

### 4.3.2 Results

The instrumental variable strategy is inspired by Autor et al. (2012) and Colantone and Crino (2011). They identify the variations in the changes in the number of varieties imported in either a HS 6-digit or HS 4-digit sector due to technological shocks or trade barriers reduction. The underlying assumption to this strategy is the common within industry changes in the extensive margin in the U.S. and in other developed economies in technological shocks as well as falling trade cost.
Figure 2 reveals the high predictive power of the instrument used.

Figure 2: First-Stage predictions

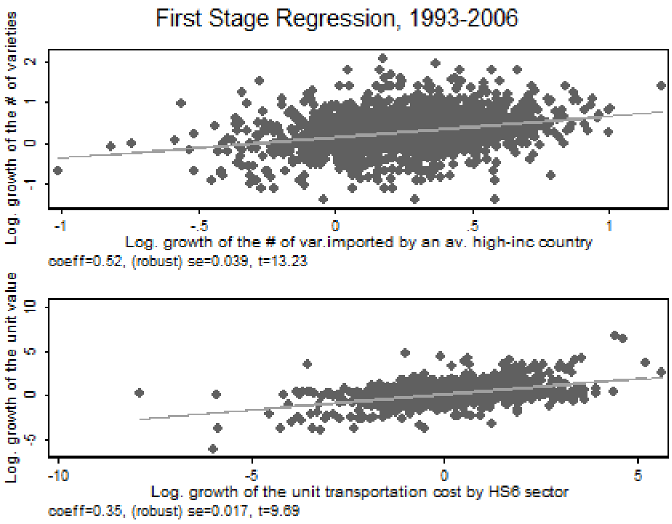


Table 10 estimates eq. (19) at HS 6-digit and HS 4-digit level by Two-Stage Least Squares (2SLS). The errors are clustered at the sector level to account for serial correlation across sectors and also adjusted for the potential heteroskedasticity. In columns (1) and (2), I instrument the number of varieties imported in the U.S. by the number of varieties imported in other developed economies and the price is instrumented by unit product-specific transportation cost. The F-statistic for excluded instruments is well above 10.[18] The coefficient of the extensive margin is higher than the one estimated in the baseline specification. The coefficient of the price is significantly negative and higher than the one derived from the baseline specification. The latter shows that a demand shock potentially upward biases the estimates. The results derived for the extensive margin are unexpected. Indeed, the coefficient is higher than the one derived with

[18]The rule-of-thumb which assesses the predictive power of the instrument.

an OLS. Since the instrument variable strategy aims to purge the variables from any demand variations, it means that the demand shock downward biases the effect. In other words, the sectors analyzed were affected by a negative demand sector taste shock that has changed over time.

Different explanations can clarify the above results. First, I proxy the changes in the pattern of consumption with the changes in the pattern of expenditure spent in the import sector. Such method limits the ability to observe the changes in taste towards domestic goods. Those sectors may have been subject to a reallocation of expenditure from imported varieties to domestic varieties within each sector. Another explanation could be a reallocation of the expenditure from good sectors to services/ non tradable sectors as the evolution of the PCE shows. Indeed, those results are in line with the observations described in section 2.2. Between 1988 and 2007, the expenditure share has grown by 6% to the detriment of the goods sectors. However, the growth of the service sector has been driven by the growth of the price (+3.50%) while the good sector has been subject to a growth of consumption in terms of quantity (+3.28%) while the price has grown at a lower rate than the average (+1.20 %). Therefore, the sectors considered for this analysis (manufacturing and goods expressed in kilos) may have been subject to negative demand shocks and reallocations of their expenditure towards services such medical care as the PCE indicates. Such explanation in line with the observations described in section 2.2 would explain why the value of the estimates by OLS of the extensive margin is lower than the one derived by 2SLS. It would also explain why the difference is stronger when the values are analyzed and is less marked when quantities are considered. However, those conclusions are not at the core of this analysis. The 2SLS seems to be necessary but even though I purge the results from demand variations, the conclusions found in section 4 still hold.

Columns (3) and (4) include a fixed effect at HS 4-digit sector to capture a potential income effect. The F-statistic for excluded instruments is well above 10 for each variable. The coefficient of the extensive margin is positive and significant. It is also higher than ones obtained with an OLS regression. Finally, columns (5) and (6) regress eq. (19) at HS 4-digit level. While the instrument used for the extensive margin passed the F-test, the one capturing the potential endogeneity of the price failed. Such results were expected. Indeed, the price at a higher level of aggregation capture many information especially on a so long difference. Therefore, I regress eq. (20) at HS 4-digit level using 2SLS method. The F-statistic for excluded instruments is above 10 and the coeffient still positive and significant.

Table 10: Identification

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | Δ ln(x) | Δ ln(q) | Δ ln(x) | Δ ln(q) | Δ ln(x) | Δ ln(x) | ln(x) |
| Δ ln(n) | 3.542*** | 3.694*** | 3.586*** | 3.484*** | 2.251*** | 1.838*** | 2.326*** |
| | (0.274) | (0.296) | (0.523) | (0.563) | (0.313) | (0.573) | (0.313) |
| Δ ln(p) | -0.109 | -0.226** | 0.012 | -0.188 | -0.158 | 0.108 | -0.228*** |
| | (0.103) | (0.114) | (0.108) | (0.117) | (0.296) | (0.291) | (0.078) |
| $\ln(x_{1993})$ | -0.010 | | -0.051 | | -0.001 | -0.053 | |
| | (0.022) | | (0.038) | | (0.028) | (0.033) | |
| $\ln(q_{1993})$ | | -0.042** | | -0.089*** | | | |
| | | (0.017) | | (0.033) | | | |
| Constant | 0.087 | 0.394 | | | 0.314 | | |
| | (0.422) | (0.316) | | | (0.555) | | |
| Observations | 2099 | 2099 | 1851 | 1851 | 662 | 652 | 7944 |
| F | 95.449 | 112.489 | 61.331 | 63.221 | 28.746 | 16.715 | 73.954 |
| Kleibergen-Paap | 91.652 | 87.146 | 28.875 | 27.962 | 17.463 | 10.480 | 31.202 |
| F-stat. for excl. instr. | [ 104.18, p=0.00; | [100.69,p=0.0; | [28.58 ,p=0.0; | [27.79 ,p=0.0; | [38.47 ,p=0.0; | [18.02 ,p=0.0; | [36.41, p=0.0; |
| | 72.66 , p=0.00] | 73.68 ,p=0.00] | 60.75 ,p=0.00] | 60.92 ,p=0.00] | 6.02 ,p=0.00] | 7.95 ,p=0.00] | 24.38, p=0.0] |
| sample period | 1993-2006 | 1993-2006 | 1993-2006 | 1993-2006 | 1993-2006 | 1993-2006 | 1993-2006 |
| | 01 and 02 excl. | 01 and 02 excl. | 01 and 02 excl. | 01 and 02 excl. | 01 and 02 excl. | 01 and 02 excl. | 01 and 02 excl. |
| sector aggregation | HS6 | HS6 | HS6 | HS6 | HS4 | HS4 | HS4 |
| sector FE | No | No | Yes, HS4 | Yes,HS4 | No | Yes, HS2 | Yes,HS2 |
| Estimator | FD-2SLS | FD-2SLS | FD-2SLS | FD-2SLS | FD-2SLS | FD-2SLS | FE model-2SLS |

Standard errors in parentheses.They are clustered by sectors and adjusted for the heteroskedasticity
* p<0.1, ** p<0.05, *** p<0.01

A concern for my instrument strategy is that in some sectors, the import demand shocks may be correlated. If this correlation exists, using the average number of varieties imported in developed economies as an instrument may be problematic. These shocks should upward bias the results if the demand shocks were positive and vice versa if they were negative.
However, this concern is not very problematic. Indeed, first, I exclude the United States and Canada from the sample used to build the instrument. Second, as shown above, the group of developed countries selected to build the instruments includes non OECD countries which have evolved differently than the United States. Notwithstanding, to be sure to alleviate any contemporaneous correlation between demand shocks across countries, I use as an instrument the lag of the number of varieties imported by those countries. Therefore, it should alleviate any potential correlation between the demand shocks across developed economies. The only issue with this strategy is the limit imposed by the availability of data. Indeed, the imports are reported under the Harmonized System since 1989 and not all the developed economies reported the value of their imports from this year. Therefore, I use the fourth lag (the longest lag I can take) to build the instrument for the change in the extensive margin in column (1) and (2) in table 11. In columns (3) and (4), I show the results when I build another instrument with 2 lags starting from 1993 and keep the same sample of countries used to build the instrument

used above and cover the period from 1995 to 2006 in order to see whether the changes in the results with the instrument including a lag is driven either by the change in the sample or by the potential correlation between demand shocks across economies. The instruments pass the test. The results are similar than the one derived in table 10.

Finally, I follow Autor, Dorn and Hanson(2012) and I regress (19) after having dropped sectors in which correlated demand shocks may be likely. I dropped the sectors of steel, flat glass, and cement industries (those sectors may have been subject to a positive demand shock due to the housing boom) in columns (5) and(6), apparel, footwear, and textiles in columns (7) and (8).[19] Columns (9) and (10) show the results when all of those sectors are dropped out. The instruments passed the rule-of-thumb test.

All these estimations speak in favor of a causal effect of the changes in the number of varieties on the evolution of the U.S. pattern of consumption. Having established the robustness of the baseline specification, I conclude that the extensive margin has a strong positive effect on the expenditure share. Therefore, consumers seem to draw away resources from sectors subject to a small variety expansion towards sectors expanding in number of products.

The identification strategy enables to disentangle the price effect from the product diversity effect and reveals the stronger impact of the latter relative to the price effect and the income effect (captured through a fixed effect). It also provides the opportunity to assess how technological shocks in a particular economy affect its trading partners. Hsieh and Ossa (2011) analyze how the productivity's growth of a country affects individual regions as well as worldwide real income through international trade. They highlight a new channel through which international trade transmits those productivity shocks across countries; the home market effect and they find that China's productivity growth rises American welfare by only 0.4 percent. The derivation of their results requires a constant pattern of consumption. My identification strategy does not aim to capture the overall effect of a third country's productivity growth on the total welfare of one of its trading partners but highlights the impact of a productivity growth in a third country on the composition of the consumption on one of its trading partners. Therefore, the identification strategy identifies an additional effect of the productivity shocks' transmission across countries and it also questions the validity of the constant pattern of consumption's assumption.

[19]Computers are also considered as problematic. However, those sectors were already dropped because they have been modified by structural changes by the Harmonized System Comittee.

Table 11: Identification: Alleviate potential demand shocks across sectors

| | Using lag to instrument the EM | | | | Dropping potentially correlated sectors | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| | Δ ln(x) | Δln(q) | Δln(x) | Δln(q) | Δln(x) | Δln(q) | $\Delta ln(x)$ | Δ ln(q) | Δ ln(x) | Δ ln(q) |
| Δ ln(n) | 4.164*** | 4.321*** | 2.804*** | 2.825*** | 3.580*** | 3.477*** | 3.501*** | 3.380*** | 3.495*** | 3.372*** |
| | (0.950) | (0.965) | (0.639) | (0.735) | (0.523) | (0.562) | (0.536) | (0.582) | (0.535) | (0.581) |
| Δ ln(p) | 0.003 | -0.184 | -0.215** | -0.326*** | 0.012 | -0.189 | 0.018 | -0.186 | 0.018 | -0.186 |
| | (0.113) | (0.120) | (0.088) | (0.109) | (0.108) | (0.117) | (0.109) | (0.117) | (0.108) | (0.117) |
| $\ln(x_{1993})$ | -0.021 | | -0.081** | | -0.052 | | -0.058 | | -0.059 | |
| | (0.057) | | (0.038) | | (0.038) | | (0.039) | | (0.039) | |
| $\ln(q_{1993})$ | | -0.053 | | -0.126*** | | -0.090*** | | -0.100*** | | -0.101*** |
| | | (0.047) | | (0.032) | | (0.033) | | (0.035) | | (0.035) |
| Observations | 1735 | 1735 | 1735 | 1735 | 1849 | 1849 | 1782 | 1782 | 1780 | 1780 |
| F | 34.047 | 39.534 | 37.833 | 33.296 | 61.419 | 63.308 | 61.570 | 64.947 | 61.671 | 65.049 |
| Kleibergen-Paap | 10.101 | 9.905 | 15.115 | 14.316 | 28.879 | 27.990 | 26.942 | 25.917 | 26.944 | 25.942 |
| F-stat. for excl. instr. | [11.10, p=0.00; 59.12 , p=0.00] | [10.77, p=0.00; 59.28 , p=0.00] | [15.29, p=0.00; 40.84, p=0.00] | [14.63, p=0.00; 41.33, p=0.00] | [28.58, p=0.00; 60.76, p=0.00] | [27.81, p=0.00; 60.92, p=0.00] | [27.09, p=0.00; 59.52, p=0.00] | [26.24, p=0.00; 59.60, p=0.00] | [27.09, p=0.00; 59.52, p=0.00] | [26.26, p=0.00; 59.61, p=0.00] |
| sample period | 1993-2006 | 1993-2006 | 1995-2006 | 1995-2006 | 1993-2006 | 1993-2006 | 1993-2006 | 1993-2006 | 1993-2006 | 1993-2006 |
| | 01 and 02 excl. | 01 and 02 excl. | 01 and 02 excl. | 01 and 02 excl. | 01 and 02 excl. | 01 and 02 excl. | 01 and 02 excl. | 01 and 02 excl. | 01 and 02 excl. | 01 and 02 excl. |
| sector aggregation | HS6 | HS6 | HS6 | HS6 | HS6 | HS6 | HS6 | HS6 | HS6 | HS6 |
| HS4 FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| | 4 lags | 4 lags | 2 lags | 2 lags | - | - | - | - | - | - |
| | New | New | Identical | Identical | - | - | - | - | - | - |
| sector dropped | - | - | - | - | steel, flat, glasses | steel, flat, glasses | apparel, footwear | apparel, footwear | All | All |
| | - | - | - | - | and cement indust. | and cement indust. | and textile | and textiles | All | All |

Standard errors in parentheses.They are clustered by sector and adjusted for heteroskedasticity.

* p<0.1, ** p<0.05, *** p<0.01

# 5 Conclusion

Consumer spending has been a key element of the U.S. economic growth. Its composition has changed over time but little analysis has been done to study its causes in presence of differentiated products. This analysis aims to study whether the changes in the product diversity has driven the evolution in the U.S. expenditure share relative to the price effect and the income effect. The expansion of variety provides the opportunity to the consumers to better match their taste to their consumption. Therefore, if the variety expansion is asymmetric across sectors, one can expect the consumers to allocate more resources in sectors subject to relatively larger variety growth. In this study, I analyze whether the growth of international trade amplifies this channel. The development of emerging countries and incremental trade liberalization are examples of factors that increase the number of varieties available in a particular economy in an asymmetric way. Such phenomenon enables to assess whether trade pattern and structure of consumption are interlinked and how the gains from trade are allocated in the economy. Using public database and after capturing the potential identification issues, I show the significant and positive effect of the changes in the extensive margin on the changes in the expenditure share. This empirical analysis also shows that the product diversity effect is stronger than the price effect. A variety expansion by 10 % increases the expenditure share of this sector by 10 % while a decrease in the price by 10 % will only increases the expenditure share by 2.4 % controling for invariant sectorial shocks, initial size of the market as well as potential income effect. This study also assesses the dynamic of the changes in the representative consumer's behavior and find that the adaptation of the consumer is instaneous to the varieties expansion. Moreover,the identification strategy enables to highlight two additional relevant results. First, it shows that goods sectors exposed to trade has been subjected to negative demand shocks in the 90's. This reallocation of expenditure from goods sectors towards services sectors seems to have been driven by the fast growth in price in the medical care sector. According to me, such observation deserves a finer analysis given the increase in the median age consumer and the debate on the social security that drive the current headlines. Second, the identification strategy enables to identify a new channel towards which a productive shock in a country can affect its trading partners; the changes in the number of varieties traded between the countries.

## References


Antras, P., D. Chor, T. Fally, and R. Hillberry (2012, May). Measuring the upstreamness of production and trade flows. *American Economic Review 102*(3), 412–16.

Arkolakis, C., S. Demidova, P. J. Klenow, and A. Rodriguez-Clare (2008, May). Endogenous variety and the gains from trade. *American Economic Review 98*(2), 444–50.

Autor, D. H., D. Dorn, and G. H. Hanson (2012, May). The china syndrome: Local labor market effects of import competition in the united states. NBER Working Papers 18054, National Bureau of Economic Research, Inc.

Bernard, A. B., J. B. Jensen, S. J. Redding, and P. K. Schott (2009, May). The margins of us trade. *American Economic Review 99*(2), 487–93.

Bils, M. and P. J. Klenow (2001, May). The acceleration of variety growth. *American Economic Review 91*(2), 274–280.

Blonigen, B. A. and A. Soderbery (2010, November). Measuring the benefits of foreign product variety with an accurate variety set. *Journal of International Economics 82*(2), 168–180.

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

Broda, C. and D. W. Weinstein (2004, May). Variety growth and world welfare. *American Economic Review 94*(2), 139–144.

Cameron, A. C. and P. K. Trivedi (1998). *Regression Analysis of Count Data*. Cambridge: Cambridge University Press.

Colantone, I. and R. Crino (2011, September). New imported inputs, new domestic products. Development Working Papers 312, Centro Studi Luca dÁgliano, University of Milano.

Debaere, P. and S. Mostashari (2010, July). Do tariffs matter for the extensive margin of international trade? an empirical analysis. *Journal of International Economics 81*(2), 163–169.

di Giovanni, J., A. A. Levchenko, and J. Zhang (2012). The global welfare impact of china: Trade integration and technological change. Technical report.

Diewert, W. E. (1976, May). Exact and superlative index numbers. *Journal of Econometrics 4*(2), 115–145.

Dixit, A. K. and J. E. Stiglitz (1977, June). Monopolistic competition and optimum product diversity. *American Economic Review 67*(3), 297–308.

Falkinger, J. and J. Zweimuller (1996). The cross-country engel curve for product diversification. *Structural Change and Economic Dynamics 7*, 79–97.

Feenstra, R. C. (1994, March). New product varieties and the measurement of international prices. *American Economic Review 84*(1), 157–77.

Feenstra, R. C., J. Romalis, and P. K. Schott (2002, December). U.s. imports, exports, and tariff data, 1989-2001. NBER Working Papers 9387, National Bureau of Economic Research, Inc.

Feenstra, R. C. and D. E. Weinstein (2010, February). Globalization, markups, and the u.s. price level. NBER Working Papers 15749, National Bureau of Economic Research, Inc.

Foellmi, R., C. Hepenstrick, and J. Zweimuller (2010, August). Non-homothetic preferences, parallel imports and the extensive margin of international trade. CEPR Discussion Papers 7939, C.E.P.R. Discussion Papers.

Foellmi, R. and J. Zweimuller (2004, January). Inequality, market power, and product diversity. *Economics Letters 82*(1), 139–145.

Ghosh, M. and S. Rao (2010, May). Chinese accession to the wto: Economic implications for china, other asian and north american economies. *Journal of Policy Modeling 32*(3), 389–398.

Goldberg, P. K., A. K. Khandelwal, N. Pavcnik, and P. Topalova (2010, November). Imported intermediate inputs and domestic product growth: Evidence from india. *The Quarterly Journal of Economics 125*(4), 1727–1767.

Hausman, J., B. H. Hall, and Z. Griliches (1984, July). Econometric models for count data with an application to the patents-r&d relationship. *Econometrica 52*(4), 909–38.

Hsieh, C.-T. and P. J. Klenow (2009, November). Misallocation and manufacturing tfp in china and india. *The Quarterly Journal of Economics 124*(4), 1403–1448.

Hsieh, C.-T. and R. Ossa (2011, February). A global view of productivity growth in china. NBER Working Papers 16778, National Bureau of Economic Research, Inc.

Kehoe, T. J. and K. J. Ruhl (2009). How important is the new goods margin in international trade? Technical report.

Khandelwal, A. (2010, October). The long and short (of) quality ladders. *Review of Economic Studies 77*(4), 1450–1476.

Liu, X. (2009, 08). Gatt/wto promotes trade strongly: Sample selection and model specification. *Review of International Economics 17*(3), 428–446.

Lorenzen, G. (1989). Log-ratios and the logarithmic mean. *Statistical Papers 30*, 61–75.

McCully, C. (2011). Trends in consumer spending and personal saving, 19592009. Technical report, Bureau of Economic Analysis, U.S. Department of Commerce.

Murata, Y. (2009, 07). On the number and the composition of varieties. *Economic Journal 119*(539), 1065–1087.

Sato, K. (1976, May). The ideal log-change index number. *The Review of Economics and Statistics 58*(2), 223–28.

Trefler, D. (2004, September). The long and short of the canada-u. s. free trade agreement. *American Economic Review 94*(4), 870–895.

UNDP (1998). Human development report 1998. Technical report, United Nations.

Vartia, O. (1976). Ideal log-change index numbers. *Scandinavian Journal of Statistics 3*, 121–126.

Wooldridge, J. M. (1999, May). Distribution-free estimation of some nonlinear panel data models. *Journal of Econometrics 90*(1), 77–97.

Yu, D. (2008, February). The harmonized system - amendments and their impact on wto members schedules. Staff Working Paper ERSD-2008-02.

# A Evolution of the U.S. pattern of consumption

Table 12: Evolution of PCE from 1988 to 2007

| | Average annual growth (percent) (percent) | | | Share of current-dollar PCE (percent) | | |
|---|---|---|---|---|---|---|
| Category | Quantity | Price | Current dollars | 1988 | 2007 | change in share |
| Personal consumption expenditures (PCE) | 2.83 | 2.57 | 5.40 | 100 | 100 | |
| Goods | 3.28 | 1.20 | 4.48 | 40 | 34 | -6 |
| Durable goods | 5.05 | -0.75 | 4.30 | 14 | 12 | -2 |
| Other nondurable goods | 3.32 | 2.49 | 5.81 | 7 | 8 | 1 |
| Services | 2.60 | 3.35 | 5.95 | 60 | 66 | 6 |
| Household consumption expenditures | 2.43 | 3.50 | 5.93 | 58 | 63 | 5 |
| Health care | 2.59 | 4.38 | 6.97 | 12 | 15 | 3 |
| Recreation services | 3.43 | 3.36 | 6.79 | 3 | 4 | 1 |

Source:NIPAs, Bureau of Economic Analysis, U.S. Department of Commerce.

# B Description of variables

## B.1 Number of varieties

I follow the Armington assumption when I carry on the analysis at HS 6-digit level. In other words, one variety is defined as a HS 6-digit product imported by a country. For instance, if wine is a product, French wine is a variety. At the HS 4-digit level, I define a new variety as a HS 6-digit partnername product whose was not recorded before. In other words, this product can come from a country that already imports other HS 6-digit product in a particular HS 4-digit sector. An exporter can then be counted more than once within a HS 4-digit sector. The formalization of this definition is the following:

$$n_{hs4,t} = \sum_{c \in C} \sum_{i \in I} product_{it} \tag{21}$$

## B.2 Unit value

The price is only defined for varieties that were continuously traded since it captures the intensive margin. In the text, the price is proxied by the unit value (computed from the cif trade value). At HS 6-digit level, it is computed either as the average unit value charged across countries or the median of this same value. At the HS 4-digit level, it is aggregated at the HS 4-digit level by taking the median across partnername for each HS 6-digit variety and the median over the HS 6-digit sectors for each HS 4-digit level sector. The unit value is defined according to two different ways for each of the trade value's definition. The first one consists to take the average unit value within HS 4-digit industry. The second one consists to take the average across HS 6-digit products for each country and then take the average unit value over the countries within each HS 4-digit sector. A country is then not counted twice and each average unit value incorporates the information of the intensive margin imported by this particular country. The formalization of this definition is the following:

$$p_{hs4,ct} = \frac{1}{I}\sum_{i \in I} p_{it} \quad (22)$$

$$p_{hs4} = \frac{1}{C}\sum_{c \in C} p_{hs4,ct} \quad (23)$$

I will take the first difference either at the HS 4-digit level or at the HS 6-digit level. The specification is always defined. Finally, I aggregate by taking a Fischer index.

# C Measure of final goods

After merging both measures of consumption goods (the "End Use Commodity Category" by th BEA and the measure of upstreamess by Antras et al. (2012), sthe following HS 6-digit sectors were dropped because unlikely to be consumes as final use.

Table 13: Final goods: products dropped and their definition

| HS2 code | Definition |
|---|---|
| 12 | oil seeds/misc. Grains/med. Plants/straw |
| 13 | lac, gums, resins |
| 23 | residues from food industries, animal feed |
| 38 | miscellaneous chemical products |
| 56 | wadding, felt, nonwovens, special yarns, twine, cordage, ropes, cables articles |
| 79 | zinc |
| 90 | optical, photographic, cinematographic, measuring, checking, precision, medical or surgical instruments, accessories |