Document de travail du LEM / Discussion paper LEM
2020-11 version révisée

Trade, access to varieties and patterns of consumption

Amandine Aubry

LEM UMR 9221 / amandine.aubry@univ-lille.fr

Trade, access to varieties and patterns of consumption*

Amandine Aubry[†]

April 22, 2021

Abstract

This paper provides empirical evidence that product diversity drove a part of the evolution of consumers' spending in the U.S. over the 1993-2018 period. The change in the set of varieties available led consumers to increase by 1.19% the share of their budget allocated to a sector subject to the average variety expansion over that period. I exploit the exogenous change in the range of products available due to the growth of international trade to identify the causal relation. Using this identification strategy, I show that through changes in product diversity, international trade has a sizable effect on the evolution of patterns of consumption in a country, especially relative to the price effect.

JEL Classification : F12 F14 F41 D12

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

*The author is grateful to Michel Beine, Pierre-Philippes Combes, Nicolas Debarsy, Frédéric Docquier, Etienne Farvaque, Robert Feenstra, Julien Martin, Gianluca Orefice, Dominique Peeters, Giovanni Peri, Luca Pieroni and John Romalis for many invaluable comments. I also thank the participants of the GATE seminar, the UPPA seminar, the LEDI seminar, IRES Macrolunch seminar, the ADRES conference, the GEP postgraduate conference in Nottingham, the European Trade Study Group in Leuven for their helpful suggestions.

[†]Université de Lille- CNRS, IESEG School of Management, UMR 9221 - LEM - Lille Économie Management

1 Introduction

Consumer spending has been a key element of the U.S. economic growth. Its composition has changed over time (McCully (2011)). Figure 1 shows, for instance, that spending increased in the pharmaceutical or information media and equipments sectors between 1993 and 2018 to the detriment of sectors such as the food or the household utensils sectors.[1] Despite these transformations in the pattern of consumption, little analysis has been done to study the change in its composition in the presence of differentiated products. This paper fills the gap by assessing whether those changes in consumer spending have been partly driven by asymmetric shocks in the set of products available to consumers. Over time, the number of products available to consumers has changed in sectors. In sectors that have been subject to larger product expansions, consumers can better match their expenditure to their preferences. They may then allocate a larger share of their budget to these sectors to the detriment of sectors with a decrease in the products range. To identify the causal effect of a change in product diversity on consumption patterns, this paper uses exogenous changes in the range of products available to consumers due to the growth of international trade.

The scarcity of disaggregated data as well as the prevalence of the assumption of a constant pattern of consumption in models that include differentiated products are likely to explain why this channel of reallocation of resources has never been quantified. New trade theory frameworks that model an economy with differentiated products assume that the changes in income or in the expenditure (due to a taste change) drive the variations in the number of products (Falkinger and Zweimuller (1996), Foellmi, Hepenstrick, and Zweimuller (2010) and Foellmi and Zweimuller (2004)). However, a few papers such as Bils and Klenow (2001) provide evidence of the reverse relation: growth in the number of products available could be a cause of the evolution of the pattern of consumption. This paper studies the second relationship. By investigating the impact of product diversity on the expenditure share, this study contributes to the recent literature that analyzes how product diversity affects consumers'environment and consequently their welfare. Handbury and Weinstein (2014), for instance, highlight the difference in product availability across 49 U.S. cities and quantify how the spatial variation in the product availability affects spatial price index measurement. Handbury (2019) quantifies the welfare impact of the differences in the product assortment offered across poor and wealthy neighborhoods.

Given the two potential explanations, such analysis requires a disentangling of the effects of supply-side shocks from the ones on the demand-side on the structure of private consumption expenditures. This studies uses the growth in international trade to circumvent the data limitations and to identify an exogenous source of variation in the number of products across sectors. Trade flows have been

[1] For further information on the changes in the consumption pattern, refer to UNDP (1998) for a worldwide analysis, Konya and Ohashi (2007) for a study on high-income OECD countries and to McCully (2011) for the U.S. case.

expanding for many decades, and micro-level data have became available for researchers, uncovering new stylized facts. Among those, researchers have observed that the increase in international trade has been accompanied by a rise in the number of imported products.[2] Table 1 shows that in 2018, on average, seven more imported varieties were available by HS 6-digit sector to American consumers than in 1993.[3] Ninety percent of the HS 6-digit sectors in the sample have been subject to an increase in the number of varieties available for consumers over the 1993-2018 period. A large empirical field initiated by Feenstra (1994) and Broda and Weinstein (2006) has subsequently emerged to analyze the impact of globalization on product diversity (see, e.g., Debaere and Mostashari (2010), Arkolakis, Demidova, Klenow, and Rodriguez-Clare (2008), Kehoe and Ruhl (2013), Feenstra and Weinstein (2017), and Blonigen and Soderbery (2010)). Arkolakis, Costinot, and Rodríguez-Clare (2012) summarize the state of the art that quantifies the gains from trade. These studies quantify the gains from trade without considering their allocation in the economy. Atkin, Faber, and Gonzalez-Navarro (2018) and Fajgelbaum and Khandelwal (2016) go further and analyze the reallocation of these gains across consumers within a country. However, the authors assume that the level of expenditure drives the optimum product diversity.

Moreover, a new strand of literature explains how the growth in international trade and in product diversity is, in part, explained by economic reforms and technological shocks. Autor, Dorn, and Hanson (2013) show, for instance, that the growth in low-income country exports during the 1990s and the beginning of the 2000s was driven largely by China's transition to a market-oriented economy, by the access of Chinese industries to long-banned foreign technologies (Hsieh and Klenow (2009)), and by the successive trade reforms. Moreover, Hsieh and Klenow (2009) show that other emerging countries such as India have faced internal productivity growth induced by a better reallocation of inputs across firms, stimulating trade. The growth of trade flows from these countries due to technological shocks has been in part transmitted through a rise in the number of varieties exported (Colantone and Crinò (2014), Goldberg, Khandelwal, Pavcnik, and Topalova (2010)). Murata (2009) analyzes theoretically how the composition and the degree of diversity are driven by technological feasibility. Hsieh and Ossa (2016) highlight the extensive margin as a new channel through which the productivity shocks can be transmitted.

Figure 2 shows the ranking of the 20 most important importers to the U.S. They are ordered by the number of varieties they export to the U.S. Emerging economies such as China or India have risen dramatically in the ranking. Those countries contributed heavily to the increase in available varieties for American consumers. Between 1993 and 2018, the number of varieties exported from China to the U.S.

[2] I follow the Armington assumption to define the concept of "new imported variety". A variety is defined as an HS 6-digit good imported from a particular country. For instance, dark chocolate is a good whereas Belgian dark chocolate is one variety. A new variety is a variety that does not previously have a positive record in that HS 6-digit category.The literature uses the word variety instead of product. The two words are used interchangeably hereafter.

[3] Note the standard deviation in Table 1 shows the varieties expansion is heterogeneous across sectors.

more than doubled, while the number of varieties exported from India more than tripled. Such growth has not been uniform across sectors. Countries export more varieties in sectors in which they have a comparative advantage. Therefore, the growth of trade from these countries may explain in part the supply-driven heterogeneous changes in the optimum product diversity available in each sector. I exploit this exogenous source of variety to identify how the expenditure shares respond to an alteration of the consumer's environment.

To guide the empirical analysis, I first present a theoretical framework built on Feenstra (1994) and Redding and Weinstein (2020). The theoretical framework exactly rationalizes observed micro-economic data on prices and expenditure shares while permitting macroeconomic comparisons over time. To identify the causal effect of a change in the set of varieties on the evolution in the expenditure share, I exploit technological shocks, trade reforms, and trade liberalization occurring in emerging economies. These shocks have expanded the set of varieties produced and exported to the destination countries, including the U.S. This identification strategy speaks in favor of a causal effect of the changes in the number of varieties imported on expenditure shares. Focusing on American trade data over the 1993-2018 period, I show the alteration of consumers' environment in terms of the product diversity affected the expenditure share allocated to a particular sector. In particular, between 1993 and 2018, the average variety expansion led to an increase in the expenditure share by 1.19% in HS 6-digit sector, controlling for the price effect and demand shifts over time. The share of the volume imported rises by 1.12%, which then confirms the changes in the expenditures are not driven either by a fall in price or by measurement errors. This positive relationship proves strongly robust across several alternative specifications.

The paper is organized as follows. A theoretical framework is presented in section 2. The data are described in section 3. Section 4 describes the empirical methodology as well as the identification strategy. Section 5 provides evidence of a positive correlation between changes in the extensive margin and changes in the pattern of consumption and establishes the causal relation. In section 6, the results are assessed through alternative estimations. Finally, section 7 concludes the paper and discusses the implications of the results.

2 Theoretical Framework

In this section, I derive the expression for the changes in the sectoral expenditure share in response to changes in the set of varieties available for consumers in the sector, controlling for price and demand effects. I follow the methodology set up by Redding and Weinstein (2020) and use a CES framework to derive the relation between the expenditure share and the appearance and disappearance of varieties. The CES functional form is the preferred approach to model product varieties in the literature (Redding and Weinstein (2020)). A recent strand in the literature in international trade assumes an Almost Ideal Demand System (AIDS defined by Deaton and Muellbauer (1980)) to study the impact of income changes on expenditure shares (Hummels and Lee (2018)) or on inequality through the expenditure channel (Fajgelbaum and Khandelwal (2016), Liu and Meissner (2017)). The AIDS expenditure function allows for flexibility both in cross-price elasticities between goods and non-homothetic preferences. Moreover, it generates a demand system in which unobserved demand characteristics are additively separable from the price and income effects. Despite these useful characteristics, Chaudhuri, Goldberg, and Jia (2006) show the implementation of this demand system is difficult in the presence of a varying number of products because it was developed with broad commodity categories in mind, which are consumed by all consumers every period. Moreover, assuming homothetic preferences on the period analyzed in this study seems reasonable. Indeed, the lower level part in Table 1 shows that during the period analyzed in the study, no reallocation of expenditure occured, on average, at the aggregate level (i.e., HS 2-digit level in the dataset).

In a CES framework, the price index defines a relationship between the change in the product diversity and the variation in the expenditure share for each aggregate CES good. Redding and Weinstein (2020) develop a price index for CES aggregate goods (Unified Price Index, UPI hereafter) that accounts for changes in the varieties available for consumers as well as changes in the preferences towards these varieties. By allowing the demand parameter for each variety to change over time, the UPI exactly rationalizes the observed data on prices and expenditure shares while permitting exact aggregation and welfare comparison over time. The UPI nests all major price indexes used in welfare or demand system analysis.

The components of an aggregate price index can be explained from a description of the utility function. In this particular analysis, I follow Broda and Weinstein (2006) by describing the preferences of a representative consumer by a nested CES utility function. In particular, the preferences of a representative consumer are described by a third-level utility function. The upper-level utility function is specified as the following:

$$U_t = \left((\eta_{Dt} D_t)^{(\kappa-1)/\kappa} + (\eta_{Mt} M_t)^{(\kappa-1)/\kappa} \right)^{\kappa/(\kappa-1)}, \qquad \kappa > 1 \tag{1}$$

where M_t is the composite imported good at time t which is defined below, D_t is the domestic good, and κ is the elasticity of substitution between domestic goods and imports. η_{Dt} and η_{Mt} are preference ("demand") parameters for either domestic or import goods. This functional form creates a separability between imported and domestic goods that allows the import price index to be distinguished from the domestic one. The import price index is the one of interest because the empirical exercise only focuses on imported goods. In this analysis, the relationship between the expenditure share and the number of varieties is inferred from the one between the reallocation of expenditure across imports and the growth of imported varieties due to the scarcity of data on the number of varieties available to consumers.

For the sake of clarity, hereafter, I only describe the demand function for an import good. The composite imported good is defined as

$$M_t = \left(\sum_{g \in G} (\theta_{gt} M_{gt})^{\frac{\gamma-1}{\gamma}} \right)^{\frac{\gamma}{\gamma-1}}, \qquad \gamma > 1 \tag{2}$$

where M_{gt} is the sub-utility derived from the consumption of the imported good g at time t, γ denotes the elasticity of substitution among imported goods, and G is the set of all imported goods. G is kept constant over time because the aim of the analysis is to assess how the allocation of income changes across sectors g over time. θ_{gt} is a preference (demand) parameter.

M_{gt} is defined by a nonsymmetric CES function represented by

$$M_{gt} = \left(\sum_{c \in C_{gt}} (\varphi_{ct} M_{gct})^{\frac{\sigma_g - 1}{\sigma_g}} \right)^{\frac{\sigma_g}{\sigma_g - 1}}, \qquad \sigma_g > 1 \forall g \in G \tag{3}$$

where σ_g is the elasticity of substitution between varieties within a particular sector g. As defined in the previous section, a variety c of imported good g is defined by a good imported by a country of origin (or an exporting country) of this good. C_{gt} is the set of all countries exporting goods in sector g at time t. φ_{ct} is a preference (demand) parameter.

The expenditure share for a particular variety c within sector g and a particular level of expenditure, E_{gt}, is defined as

$$s_{gct} = \frac{p_{gct} M_{gct}}{E_{gt}} = \frac{\left(\frac{p_{gct}}{\varphi_{ct}} \right)^{1-\sigma_g}}{\sum_{i \in C_{gt}} \left(\frac{p_{git}}{\varphi_{git}} \right)^{1-\sigma_g}}, \tag{4}$$

where p_{gct} is the price of a variety c at time t.

For a given expenditure on imported category E_t, the expenditure share in sector g is defined as

$$s_{gt} = \frac{P_{gt} M_{gt}}{E_t} = \frac{\left(\frac{P_{gt}}{\theta_{gt}} \right)^{1-\gamma}}{\sum_{i \in G} \left(\frac{P_{it}}{\theta_{it}} \right)^{1-\gamma}}, \tag{5}$$

where P_{gt} is an exact price index for good g at time t, defined as

$$P_{gt} = \left[\sum_{i \in C_{gt}} \left(\frac{p_{git}}{\varphi_{git}} \right)^{1-\sigma_g} \right]^{\frac{1}{1-\sigma_g}}. \tag{6}$$

The change in the expenditure share is defined by taking the ratio of eq. (5) between two periods, $t-1$ and t:

$$\frac{s_{gt}}{s_{gt-1}} = \left(\frac{\theta_{gt}}{\theta_{gt-1}} \right)^{\gamma-1} \left(\phi^g_{t-1,t}\right)^{1-\gamma} \left(\phi_{t-1,t}\right)^{\gamma-1}, \tag{7}$$

where $\phi^g_{t-1,t}$ is the change in the cost of living from t-1 to t in a particular sector g and is defined as

$$\phi^g_{t-1,t} = \left[\frac{P_{gt}}{P_{gt-1}} \right] = \left[\frac{\sum_{c \in C_{gt}} \left(\frac{p_{gct}}{\varphi_{ct}} \right)^{1-\sigma_g}}{\sum_{c \in C_{gt-1}} \left(\frac{p_{gct-1}}{\varphi_{ct-1}} \right)^{1-\sigma_g}} \right]^{\frac{1}{1-\sigma_g}}, \tag{8}$$

and $\phi_{t-1,t}$, the change in the cost of living from t-1 to t, is defined as

$$\phi_{t-1,t} = \left[\frac{\sum_{i \in G} \left(\frac{P_{it}}{\theta_{it}} \right)^{1-\gamma}}{\sum_{i \in G} \left(\frac{P_{it-1}}{\theta_{it-1}} \right)^{1-\gamma}} \right]^{\frac{1}{1-\gamma}}. \tag{9}$$

The change in the set of varieties available in a particular sector g between $t-1$ and t is included in $\phi^g_{t-1,t}$. Indeed, the set of varieties in the numerator is not the same as that in the denominator. Feenstra (1994) shows that the assumptions of the CES utility function enable the "variety" effect to be disentangled from the price effect in the price index, $\phi^g_{t-1,t}$. To do so, Feenstra (1994) allows for at least one price to approach infinity for some periods, whereas the demand for this variety is positive (which can be interpreted as the unavailability of those varieties in that particular period). Therefore, the set of varieties includes varieties that do not necessarily exist in both period $t-1$ and t. Feenstra (1994) shows that even under those assumptions, the price index gets a reasonable limiting value.

The price effect captures the changes in the intensive margin (i.e., changes that occur for varieties available in both time periods). To disentangle the "variety effect" from the price effect, I first compute the expenditure on all common varieties by summing eq.(4) over the set of common varieties consumed in both period t and t-1, C (defined as $C_{t-1} \cap C_t$). I then write the expenditure on all common varieties as a share of total expenditure in periods t and t-1:

$$\lambda_{gt} = \frac{\sum_{c \in C} \left(\frac{p_{gct}}{\varphi_{ct}} \right)^{1-\sigma_g}}{\sum_{c \in C_{gt}} \left(\frac{p_{gct}}{\varphi_{ct}} \right)^{1-\sigma_g}} \quad \text{, and} \quad \lambda_{gt-1} = \frac{\sum_{c \in C} \left(\frac{p_{gct-1}}{\varphi_{ct-1}} \right)^{1-\sigma_g}}{\sum_{c \in C_{gt-1}} \left(\frac{p_{gct-1}}{\varphi_{ct-1}} \right)^{1-\sigma_g}}. \tag{10}$$

λ_{gt} and λ_{gt-1} are defined as long as some overlap occurs in the varieties available between the two periods ($C \neq \varnothing$). I also assume $\varphi_{ct} = 0$ for a good c before it enters and after it exits the market, as in Redding and Weinstein (2020).

To see how to decompose the price index ϕ^g_{it} into the "variety effect" and the price effect, I multiply the numerator and the denominator within the bracket defined in (8) by $\sum_{c\in C}(\frac{p_{gct}}{\varphi_{ct}})^{1-\sigma_g}$. I rearrange the terms in the bracket and rewrite $\phi^g_{t-1,t}$ as the following:

$$\phi^g_{t-1,t} = \left[\frac{1}{\lambda_{gt}} \frac{\sum_{c\in C}\left(\frac{p_{gct}}{\varphi_{ct}}\right)^{1-\sigma_g}}{\sum_{c\in C_{gt-1}}\left(\frac{p_{gct-1}}{\varphi_{ct-1}}\right)^{1-\sigma_g}}\right]^{\frac{1}{1-\sigma_g}}. \tag{11}$$

I then multiply the numerator and the denominator within the bracket defined in (11) by $\sum_{c\in C}\left(\frac{p_{gct-1}}{\varphi_{ct-1}}\right)^{1-\sigma_g}$. I rearrange the terms in the bracket and rewrite $\phi^g_{t-1,t}$ as the following:

$$\phi^g_{t-1,t} = \left[\frac{\lambda_{gt-1}}{\lambda_{gt}} \frac{\sum_{c\in C}\left(\frac{p_{gct}}{\varphi_{ct}}\right)^{1-\sigma_g}}{\sum_{c\in C}\left(\frac{p_{gct-1}}{\varphi_{ct-1}}\right)^{1-\sigma_g}}\right]^{\frac{1}{1-\sigma_g}} = \left(\frac{\lambda_{gt}}{\lambda_{gt-1}}\right)^{\frac{1}{\sigma_g-1}} \frac{P^*_{gt}}{P^*_{gt-1}}, \tag{12}$$

where P^*_{gt} and P^*_{gt-1} are the exact price indices for the common set of varieties at time t and t-1.

The term $\left(\frac{\lambda_{gt}}{\lambda_{gt-1}}\right)^{\frac{1}{\sigma_g-1}}$ captures the role of new and disappearing varieties. λ_{gt} can be interpreted as the period t expenditure on varieties in the common set, C, relative to the period t total expenditure. Therefore, $1-\lambda_{gt}$ (when $\lambda_{gt} \leq 1$) can be interpreted as the period t expenditure on new varieties relative to the period t total expenditure. The higher the expenditure share of new varieties in period t, the lower λ_{gt} is and the smaller the exact price index is. More varieties being available lowers the cost of reaching a certain level of satisfaction, and the magnitude of the decrease in the cost depends on the demand for these new varieties relatives to the ones already consumed and how similar the new varieties are to the ones already consumed.

I follow Redding and Weinstein (2020) to define $\frac{P^*_{gt}}{P^*_{gt-1}}$ in expression (12). I use the demand system in (4) to define the share of a particular good c consumed in both periods in expenditure on the set of common varieties, C:

$$s^*_{gct} = \frac{\left(\frac{p_{gct}}{\varphi_{ct}}\right)^{1-\sigma_g}}{\sum_{i\in C}\left(\frac{p_{git}}{\varphi_{git}}\right)^{1-\sigma_g}}, \quad c \in C. \tag{13}$$

I rearrange the terms in eq. (13) to define the exact price index for the common set of varieties:

$$\left(P^*_{gt}\right)^{1-\sigma_g} = \sum_{i\in C}\left(\frac{p_{git}}{\varphi_{git}}\right)^{1-\sigma_g} = \frac{1}{s^*_{gct}}\left(\frac{p_{gct}}{\varphi_{gct}}\right)^{1-\sigma_g}. \tag{14}$$

I then take the logarithm of both sides of eq. (14), the difference over time and the sum across all $c \in C$, and divide both sides by the number of common varieties. I can then write the logarithm change in the common-varieties price index as

$$ln\left(\frac{P^*_{gt}}{P^*_{gt-1}}\right) = \frac{1}{\sigma_g - 1} ln\left(\frac{\tilde{S}^*_{gt}}{\tilde{S}^*_{gt-1}}\right) + ln\left(\frac{\tilde{P}^*_{gt}}{\tilde{P}^*_{gt-1}}\right) - ln\left(\frac{\tilde{\varphi}^*_{gt}}{\tilde{\varphi}^*_{gt-1}}\right), \tag{15}$$

where a tilde over a variable denotes a geometric average. As in Redding and Weinstein (2020) , I assume a constant aggregate utility function, which requires that the following condition holds:

$$ln\left(\frac{\tilde{\varphi}^*_{gt}}{\tilde{\varphi}^*_{gt-1}}\right) = 0. \tag{16}$$

In other words, I allow time-varying demand shocks for each variety as long as the geometric mean of the consumer taste parameters is constant across common varieties. I use this assumption and substitute (15) into the overall CES price index, $\phi^g_{t-1,t}$, in eq. (12) to derive the "unified price index" defined by Redding and Weinstein (2020):

$$(\phi^g_{t-1,t})^{UPI} = \underbrace{\left(\frac{\lambda_{gt}}{\lambda_{gt-1}}\right)^{\frac{1}{\sigma_g-1}}}_{\text{variety adjustment}} \underbrace{\left(\frac{\tilde{P}^*_{gt}}{\tilde{P}^*_{gt-1}} \quad \left(\frac{\tilde{S}^*_{gt}}{\tilde{S}^*_{gt-1}}\right)^{\frac{1}{\sigma_g-1}}\right)}_{\text{Common goods price index}}. \tag{17}$$

The price index, $(\phi^g_{t-1,t})^{UPI}$, expresses the change in the cost of living as a function of a first term that captures the changes in the set of varieties available. This term captures how variety turnover, changes in the number of varieties, consumed and or the appearance or disappearance of products affect the cost of reaching a certain level of satisfaction. Holding other things constant, an expansion of the number of varieties available to consumers in a particular sector improves the consumers' ability to match their expenditure to their preferences, and then, consequently, decreases their cost of reaching a certain level of satisfaction. It then decreases the price index $(\phi^g_{t-1,t})^{UPI}$ by an amount that depends on the elasticity of substitution between varieties available in sector g. The second term of $(\phi^g_{t-1,t})^{UPI}$ measures how changes in prices, demand shifts, and product substitution for the set of varieties available in both periods affect the cost of reaching a certain level of satisfaction. This second term comprises two ratios. The first one, $\frac{\tilde{P}^*_{gt}}{\tilde{P}^*_{gt-1}}$, is the change in the geometric average of prices of varieties. The second term is the main innovation of the UPI's price index. It captures heterogeneity in expenditures across varieties consumed in both periods. It aims to capture whether the match between the expenditures and the demand changes increases. Demand shifts that raise the dispersion in expenditure lower the price index. This decrease in the price index reflects the greater satisfaction the consumers derive from the positive taste shifts for varieties that constitute big shares of expenditures.

I substitute eq. (17) into eq. (7) and log-linearize the changes in import demand in a particular sector g such that

$$\begin{aligned} ln\left(\frac{s_{gt}}{s_{gt-1}}\right) = & (\gamma-1) \quad ln\left(\frac{\theta_{gt}}{\theta_{gt-1}}\right)+\frac{1-\gamma}{\sigma_g-1} \quad ln\left(\frac{\lambda_{gt}}{\lambda_{gt-1}}\right)+(1-\gamma) \quad ln\left(\frac{\tilde{P}^*_{gt}}{\tilde{P}^*_{gt-1}}\right)+\frac{1-\gamma}{\sigma_g-1} \quad ln\left(\frac{\tilde{S}^*_{gt}}{\tilde{S}^*_{gt-1}}\right) \\ & +(\gamma-1) \quad ln\left(\phi_{t-1,t}\right), \end{aligned} \tag{18}$$

where the exact aggregate price index, $\phi_{t-1,t}$, can be derived in the same way I derive $\phi^g_{t-1,t}$. However, as explained above, the set of sectors, G , is fixed over time. Therefore, the variety correction term at that level of aggregation is equal to 1. $\phi_{t-1,t}$ is defined as the following:

$$\phi_{t-1,t} = \left(\frac{\tilde{P}^*_t}{\tilde{P}^*_{t-1}}\left(\frac{\tilde{S}^*_t}{\tilde{S}^*_{t-1}}\right)^{\frac{1}{\sigma_g-1}}\right). \tag{19}$$

Eq. (18) describes how expenditure shares change between periods t and $t-1$ in a particular sector g in response to a variation in the varieties available in sector g, in prices, and in demand shifts for the set of varieties available in both time periods. A positive correlation is expected between a change in the number of varieties and the evolution of the expenditure shares given that the presence of more varieties in a sector decreases the cost to reach a certain level of satisfaction for the consumer.[4] Consequently, one can then expect the representative consumer to allocate more resources in sectors with more varieties for a particular price and a particular demand-shift for the varieties available in both periods.

3 Data sources and variables definitions

Data on consumers spending and on the range of products available for consumption are scarce. To circumvent their unavailability, I focus the empirical analysis on trade data to analyze the relationship between expenditure shares and the set of varieties available. Trade data are available at a disaggregated level and can be compared across countries. I use these characteristics to identify exogenous sources of the expansion of varieties.

3.1 Trade data

The data on trade come from the UN Comtrade Statistics Database, from the United States International Trade Commission (USITC hereafter), and from Schott, Fuest, and O'Rourke (2008). These databases

[4]In the case of a variety expansion, $\frac{\lambda_{gt}}{\lambda_{gt-1}}$ is expected to be smaller than 1. The logarithm of a number smaller than 1 is negative. This negative value is then multiplied by $\frac{1-\gamma}{\sigma_g-1}$, which is also negative given that $\gamma > 1$.

report the trade value as well as the quantity imported in the Harmonized System (HS hereafter) classification. They are reported at the 6-digit level in the UN Comtrade Statistics database and at the 10-digit level in the USITC and in Schott et al. (2008)'s databases. The HS scheme is an internationally standardized system that theoretically covers all commodities in international trade since 1988. Whereas 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 these errors, I organize the sample 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 2002, *structural changes* were implemented. Those changes preclude a comparison of commodities over time because 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 the level of aggregation of the analysis, the HS 6-digit level.[5] I also disregard the imports from countries that have been divided, reunified after 1991, or reclassified in the database. Indeed, a new variety is defined as an HS 6-digit good imported from a country that did not previously export that good to the U.S. Therefore, any redefinition of a country would affect the computation of the set of new varieties. For this reason, the sample starts in 1993. This focus prevents the dissolution of the Soviet Union from affecting the results, because this historical event has artificially led to an increase in the number of countries. The sample ends in 2018, the last year for which the data are available.

The final database includes 1, 761 HS 6-digit sectors for which data were reported in 1993 and 2018 and 204 economic entities over the 1993-2018 period. Table 1 reports descriptive statistics on the sample and the variables used for the analysis that are described below.

3.2 Unit value

The datasets I use collect expenditure and quantity purchased, not prices. Unit values (computed from the Cost Insurance Freight (cif) trade value divided by quantity) provide the price measures for each HS 6-digit sector. To alleviate the measurement errors in the unit value, I computed it for each HS 6-digit sector using different methods. I follow the theoretical definition of the price index described in eq. (15) by taking the ratio of the geometric means of the unit values in each HS 6-digit sector. I computed the aggregate unit value (trade value/ quantity) at the HS 6-digit sector. I computed the weighted mean to capture the heterogeneity of spending allocated to each HS 6-digit good. I also computed the median unit value as in Colson-Sihra, Sousa, and Mayer (2020) or the unweighted mean unit value for each HS 6-digit sector. I follow the theoretical framework and used the ratio of the geometric mean of unit values

[5] Those modifications consist of merging, splitting categories, or both (called complex changes).

in the main analysis. The ratio of the geometric mean and the unweighted means are the price's measures that provide the most stable results in terms of coefficients across alternative samples and in terms of the significance of the price coefficients. In Table 7, I show the estimation of eq. (18) using alternative proxies of price and show the main conclusions are not affected by the use of different definitions for the price.

The first five rows of Table 1 show the mean unit values according to the several definitions described above. The other rows in part 1 of the table report the mean of the unit value, defined as the unweighted mean across varieties for each HS 6-digit sector by type of variety. The first column reports the average unit value in 1993 by set of goods (only consumed in 1993, consumed in both periods, or the average for all varieties imported in 1993). The varieties consumed in both periods are, on average, cheaper than products that were consumed in 1993 but were no longer consumed in 2018. They are also cheaper than the new varieties available in 2018. On average, a new variety costs US$293 in an HS 6-digit sector in 2018, whereas a variety already consumed in this sector in 1993 only costs US$143 in 2018.

3.3 Variety adjustment and demand heterogeneity terms

I use the CES demand system described in section 2 to express $\frac{\lambda_{gt}}{\lambda_{gt-1}}$ and $\frac{\tilde{S}^*_{gt}}{\tilde{S}^*_{gt-1}}$. To compute these ratios, I use the generic expression of the expenditures. The variety-adjustment term captures the sum of the expenditures allocated to the common varieties at current prices on the sum of the current expenditures on all varieties available at time t and t-1:

$$\lambda_{gt} = \frac{\sum_{c \in C} \text{trade value at time t}}{\sum_{c \in C_{gt}} \text{trade value at time t}} \quad \text{, and} \quad \lambda_{gt-1} = \frac{\sum_{c \in C} \text{trade value at time t-1}}{\sum_{c \in C_{gt-1}} \text{trade value at time t-1}}. \tag{20}$$

The heterogeneity term is computed by taking the geometric mean of the expenditures allocated to the set of common varieties at time t-1 and t, weighted by the number of common varieties :

$$\frac{\tilde{S}^*_{gt}}{\tilde{S}^*_{gt-1}} = ln \left(\frac{\left(\prod_{c \in C} \text{trade value at time t}\right)^{\frac{1}{C}}}{\left(\prod_{c \in C} \text{trade value at time t-1}\right)^{\frac{1}{C}}} \right) \tag{21}$$

Part 2 and Part 3 of Table 1 describe the evolution of the imports and these two terms. The first row of the second part of Table 1 shows the current dollar's imports grew, on average, by a factor of 17 for a particular HS 6-digit sector. However, the disparity between sectors is quite large. The second row shows the growth was much less important for goods that were already imported in 1993. On average, the import of varieties that were already imported in 1993 grew by a factor of 10 in a particular HS 6-digit sector. The third row reports that overall imports grew 1.70% faster than for common goods to the two time periods. The fourth row reports the variety-adjustment ratio. On average, the variety-adjustment term was 0.95 by sector. Expenditures in new varieties expanded (i.e., the ratio <1) in 68.00% of the HS-6 digit sectors included in the sample. The expenditures in new varieties were important in sectors such as "Cotton or man-made textiles materials" (0.08) or "different colors filaments of nylon (0.20)". By contrast, the diversity of varieties shrank over the period in sectors such as "some Fertilizers" (25.00) or "silk "(109). Part 3 of Table 1 reports some observations about the demand-heterogeneity term. On average, the expenditures across varieties consumed in both periods became slightly less dispersed in a particular HS 6-digit sector. However, the geometric mean of the expenditures decreased for some sectors such as some "Fabrics" (0.27), "Boilers" (0.27), "some Fertilizers" (0.33), or "Plasters" (0.34). This decrease over time means the consumption of varieties already largely consumed in 1993 increased over time. On the other hand, the geometric mean largely increased for other factors such as "some Synthetic" (4.70), "some meat preparations" (3.42) or "some blankets" (2.80). These examples show the disparity observed in the demand-heterogeneity term.

Table 1 shows that, on average, over the 1993-2018 period, sectors faced an increase in the expenditures in new varieties, leading to a decrease in the cost of reaching a certain level of satisfaction. However, the average hides disparities across HS 6-digit sectors.

4 Empirical Analysis

I aim to estimate eq. (18) to analyze how the expenditure shares respond to a change in the set of varieties available for consumers. The empirical work focuses on U.S. import trade observations, due to the absence of a dataset including the total number of varieties available to consumers. Before describing the equation estimated in section 4.2, I analyze over which timeline the econometric analysis is relevant in section 4.1. The last subsection describes the identification strategy.

4.1 Long differences vs. annual changes

I follow Bernard, Jensen, Redding, and Schott (2009) methodology to decompose the change in aggregate U.S. trade between period t-1 and t. Let $\triangle x_t$ denotes the change in total U.S. imports between t-1 and t. $\triangle x_t$ can be decomposed into the increase due to the entry of new varieties, the decrease due to the exit of existing importers, and the change due to increases or decreases in trade for the continuing firms:

$$\triangle x_t = \underbrace{\sum_{c \in N} x_{ct} - \sum_{c \in E} x_{ct-1}}_{\text{extensive margins}} + \underbrace{\sum_{c \in G} \triangle x_{ct}}_{\text{intensive margin}} \quad , \tag{22}$$

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 2 shows the decomposition of total U.S. import variation into the contribution of the margins described above from 1993 to 2018. The first part of the table reports annual changes whereas, the last part reports long differences (i.e., from 1993 to 2006, from 2006 to 2018, and from 1993 to 2018).

Table 2 shows positive growth in the number of varieties over time except for the period corresponding to the American economic recession (2001-2002) and the great recession (2008-2009). The growth in 1995-1996 was driven by the imports of vehicle and aircraft accessories, mainly from European countries, Mexico, and Taiwan.[6] The growth in 2006-2007 was driven by Chinese imports in the toys accessories industries. Finally, the large variation in the extensive margin in 2011-2012 is explained mainly by two factors. First, American economic growth decelerated that year. Second, The U.S. increased its consumption of domestically produced shale gas, resulting in a large decrease in its oil imports (UNCTAD (2013)).

Like Bernard et al. (2009), I find the short-run changes in the U.S. imports are largely accounted for by the intensive margin, whereas the long-run decomposition highlights large growth in the extensive margin (90.00% between 1993 and 2018). The results of the import lead me to focus the analysis on long differences instead of annual changes.

4.2 Baseline specification

The following equation is estimated for between 1993 and 2018:

$$ln\left(\frac{s_{g2018}}{s_{g1993}}\right) = \beta_0 + \beta_1 ln\left(\frac{\lambda_{g2018}}{\lambda_{g1993}}\right) + \beta_2 ln\left(\frac{\tilde{P}^*_{gt}}{\tilde{P}^*_{gt-1}}\right) + \beta_3 ln\left(\frac{\tilde{S}^*_{g2018}}{\tilde{S}^*_{g1993}}\right) + \epsilon_{gt}, \tag{23}$$

where g represents the sector, and s_{gt} measures the share of imports of product g in year t. The constant, β_0, captures the logarithm of the aggregate price index, $ln(\phi_{1993,2018})$, because it affects every expenditure share in an identical way. $\frac{\lambda_{g2018}}{\lambda_{g1993}}$ is the variety-adjustment term. Therefore, $\beta_1 = \frac{1-\gamma}{\sigma_g - 1}$ is the coefficient of interest. It captures the elasticity of the expenditure share with respect to the extensive margin. β_1 is expected to be negative. Indeed, a change in the number of varieties affects the expenditure share of sector g through the price index. An increase in the number of varieties available for consumers decreases the price index. An increase in the number of varieties is then captured by a ratio smaller than 1 leading to a negative value once I take the logarithm. Therefore, this decline multiplied by β_1

[6] Such large growth may have been driven by the European demand for aircrafts. Indeed, neither the MFN nor the Mexican tariffs in those sectors have fallen by a large amount. However, 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.

increases the expenditure share. $\tilde{P}^*_{gt}$ represents the geometric average of the unit values of the import of sector g at time t. $\frac{\tilde{S}^*_{g2018}}{\tilde{S}^*_{g1993}}$ captures the average demand shift between 1993 and 2018 for the set of common varieties. An increase in the demand dispersion decreases the demand-heterogeneity term and consequently the price index. Therefore, a decrease in this term is expected to increase the expenditure share in a particular sector g. The first difference allows me to eliminate any time-invariant preferences shocks but not the time-varying ones. $\epsilon_{gt} = (\gamma - 1) ln \left(\frac{\theta_{g2018}}{\theta_{g1993}}\right)$ captures time-varying demand shocks.

4.3 Identification strategy

This section describes the identification strategy used to treat the potential endogeneity issues that may impede the econometric analysis of eq. (23). The empirical setting described in the previous section may, first, suffer from reverse causality. Eq. (23) aims at studying whether the variety growth could be the cause of the evolution in the pattern of consumption as suggested by Bils and Klenow (2001). However, the traditional literature in new trade theory assumes the reverse relationship, namely, that changes in the expenditure drive the variations in the number of varieties. Second, the estimates β_1 and β_2 from eq. (23) are likely to be biased by unobserved time-variant demand shocks that are captured by ϵ_{gt} and that are also correlated with the changes in the number of varieties and the price changes. These endogeneity issues are controlled by instrumental variables. At least two instruments are required to identify variations in the changes in prices and in the changes in the number of varieties that are exogenous to U.S. demand shocks.

The first instrument used is the changes in variety-specific unit transportation cost for the U.S. as in Khandelwal (2010).[7] The instrument varies across countries, industries, and years, which makes it possible to use with long-difference. I compute the aggregate unit transportation costs following the different definitions used for the price's definition and described in section 3.2 at the HS 6-digit sector.

The second instrument is defined following the strategy in Autor et al. (2013) and Colantone and Crinò (2014).[8] I compute the ratio, $ln \left(\frac{\lambda_{g2018}}{\lambda_{g1993}}\right)$ based on the imports from low- and middle-income countries in a particular sector g for high-income economies.[9] I then take the average ratio across these economies. Formally, I construct the following variable:

$$\overline{ln \left(\frac{\lambda_{g2018}}{\lambda_{g1993}}\right)}^{inst} = \frac{1}{H} \sum_{h \in H} ln \left(\frac{\lambda^{\text{emerging}}_{g2018}}{\lambda^{\text{emerging}}_{g1993}}\right),$$

[7]The data are sourced from Feenstra, Romalis, and Schott (2002) and Schott et al. (2008).

[8]Auer and Fischer (2010) use a similar instrumentation strategy to identify comparative advantage induced supply shocks in low-wage countries.

[9]These countries are Australia, Denmark , Greece, Iceland, Ireland, Japan, Korea, the Netherlands, Portugal,Spain, Sweden, Switzerland, and the U.S. Different groups of countries are taken based on the analysis to guarantee the potential endogenous variables are well-instrumented. The group of countries is always described for each analysis.

where H is the group of high-income countries and $\overline{ln\left(\frac{\lambda_{g2018}}{\lambda_{g1993}}\right)^{inst}}$ is the average variety-adjustment term across countries $\in H$ in sector g computed based only on the imports from low and middle-income countries.[10] The underlying assumption behind this strategy is that the growth in the number of varieties exported by middle-income countries captures, in part, the productivity gains and institutional reforms occurring in these countries. Autor et al. (2013) show that the growth in middle-income country exports during the 1990s and the beginning of the 2000s is largely driven by China's transition to a market-oriented economy and by the access of Chinese industries to long-banned foreign technologies (Hsieh and Klenow (2009)). Successive trade reforms, which have increased linkages with other countries and have decreased trade costs, have also increased exports. Hsieh and Klenow (2009) show other emerging countries such as India have faced internal productivity growth induced by a better reallocation of inputs across firms, stimulating trade. The growth of these trade flows in low- and middle-income countries not only reflects the growth in the intensive margin but also the growth in the extensive margin as highlighted in Figure 2.[11]

The change in the extensive margin from these countries must capture a part of these economic and institutional reforms in these countries, but it can also be driven by a change in the U.S. demand for the products exported by these countries. To exclude this potential correlation, I use the global nature of the growth of exports from low- and middle-income countries (e.g., as di Giovanni, Levchenko, and Zhang (2014) underline with the Chinese exports), and consider the evolution of the average extensive margin imported in 13 high-income countries. The *average* of the variety-adjustment term across these countries is taken to alleviate the potential correlation of demand shocks between those countries.

A potential correlation could have come from a conversion of a part of the U.S. growth in the 1990s into an increase in the U.S. imports from all countries including high-income economies. Then, those economies may have allocated the incremental income induced by these additional exports towards the U.S. into the purchase of new varieties. In that case, the variety-adjustment term of these high-income economies should have been correlated with the American income increase (and potentially any change in expenditure shares if the preferences are non-homothetic). Table 3 shows an absence of correlation between the export growth in the U.S. over the 1993-2018s period (which proxies for U.S. economic growth) and the average variety-adjustment term for the high-income countries selected to build the instrument. An issue could have been the correlation between the U.S. economic growth and the expansion of exports in emerging economies. Indeed, the economic growth in the U.S. could have boosted the exports to China and been the source of a technological shock in emerging economies

[10]The selection of middle-income countries is based on the income classification of countries by the World Bank in 2018. I include countries classified as low-income, lower-middle-, and upper-middle-income countries.

[11]Those results are in line with the observations made by Feenstra (1994), Broda and Weinstein (2006), and Debaere and Mostashari (2010).

such as China (Goldberg et al. (2010), Colantone and Crinò (2014)). For this particular case, we know the U.S. is not an important exporter to China. Almost 60.00% of Chinese imports comes from other Asian countries (Ghosh and Rao (2010)). Additional testing has been done (see section 6) to guarantee unobserved demand shocks in the U.S. are not correlated with the instruments of the changes in the number of varieties in a particular sector.

5 Results

Tables 5 and Table 6 show the results for the estimation of eq. (23). The empirical setting enables the "variety effect" to be disentangled from the price and demand-heterogeneity effects on the expenditure share in a particular HS 6-digit sector. The errors are clustered at the sector level to account for serial correlation across sectors and to adjust for potential heteroscedasticity.

5.1 Main results

The first two columns in Table 5 report OLS estimators whereas the last two columns describe the results of two-stage least squares (2SLS) estimations of eq. (23). Table 4 reports the first-stage results for the the two potential endogenous variables. At the bottom of each table, statistical tests on the quality of the instrumentation strategy are reported. The Kleibergen-Paap Wald F statistic reveals the high predictive power of the instruments. It is above the critical value compiled by Stock and Yogo (2005), rejecting the assumption of weak instruments. An endogeneity test is also run and reported in order to control for the exogeneity of the instruments. The null hypothesis that the specified endogenous regressors can actually be treated as exogenous is not rejected.[12] Columns (1) and (3) report the results for the analysis of the impact of a change in the number of varieties available to consumers and the price on the expenditure share in a particular HS 6-digit sector assuming a constant demand parameter over time. Columns (2) and (4) show the results once I estimate the relationship assuming demand parameters may shift over time by including the demand-heterogeneity term.

The first two columns show a positive correlation between the change in the number of varieties and the change in the expenditure share. On average, a growth in the number of varieties leading to a 10.00% decrease in the price index in a particular HS-6 digit sector is associated with a growth in the expenditure share by 7.65%. The prices' coefficients are negative and significant as expected. The demand-heterogeneity term is also associated with a positive increase in the expenditure share. Over time, the sectors, in which preferences shifted towards largely consumed varieties in 1993, faced a larger

[12] Under conditional homoskedasticity, this endogeneity test statistic is numerically equal to a Hausman test statistic (Hayashi (2000)).

decrease in the cost of reaching a certain level of satisfaction. These results indicate a positive correlation between the change in the extensive margin and the expenditure share.

The results reported in columns (3) and (4) in Table 5 confirm the positive and significant impact of product diversity on the expenditure share. On average, the availability of new varieties to consumers between 1993 and 2018 led to an increase in the expenditure share by 1.19% (exp(-0.97 $\overline{ln\left(\frac{\lambda_{g2018}}{\lambda_{g1993}}\right)}$) in a particular HS 6-digit sector. Moreover, the impact of the extensive margin is larger than that of the price (and then the intensive-margin effect). However, the results found for the extensive margin are unexpected. Indeed, the coefficient is higher than the one derived with an OLS estimation. Therefore, the demand shock downward biased the effect. In other words, the sectors analyzed were affected by a negative demand sectoral taste shock that 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 in the import sector. Such a method limits the ability to observe 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. [13] Another explanation could be a reallocation of the expenditure from manufacturing sectors to services/non-tradable sectors. Between 1993 and 2018, the expenditure share in the service sector grew by 6 percentage points to the detriment of manufacturing sectors. However, this evolution was driven by the growth of the price while the manufacturing sector has been subject to a growth of consumption in terms of quantity. Therefore, the sectors considered for this analysis may have been subject to negative demand shocks and reallocation of their expenditure towards services such as medical care (i.e., hospital and nursing home services, outpatient services, and paramedical services).[14] Such an explanation would explain why the value of the estimates by OLS of the extensive margin are lower than the one derived by 2SLS.

The 2SLS estimation is necessary, and when the analysis is purged from demand variations, the conclusions still hold. The results reported in columns (3) and (4) in Table 5 suggest a causal effect of the changes in the number of varieties on the evolution of the U.S. pattern of consumption. Consumers draw resources away from sectors subject to a small variety expansion towards sectors expanding in the varieties available for consumers.

5.2 Analysis by type of goods

The results described above apply to all the tangible goods imported by the U.S. Assessing the relationship according to the type of goods is interesting. Two different types of good classifications are considered: the final use of the product and whether the product can be described as a differentiated

[13] This potential shock is capture by η_{Mt} in the theoretical framework.

[14] Those services have not become tradable over time, in contrast to other services. Therefore, the increase in consumption for domestic products is not driven by an increase in either tariff or non-tariff trade barriers.

product.

The Bureau of Economic Analysis (BEA) classifies goods into six principal "end-use" categories (foods, intermediate, capital (except automotive), automotive, consumer, and other goods). In the benchmark sample, out of the 1, 761 HS 6-digit sectors, 1, 251 are classified as "intermediate goods", 327 as "consumer goods", 142 as "food goods", 38 as "capital (except automotive), and 3 as "automotive". The first two columns in Table 6 report the results of the 2SLS estimation of eq. (23) on subsamples including either intermediate goods (in column (1)) or consumer goods (in column (2)). The coefficient of the variety-adjustment term for consumer goods is larger than that for intermediate goods. A decrease by 1.00% in the price index induced by an expansion of varieties for consumers goods increases, on average, the expenditure share in the sector by 1.21%. The demand-heterogeneity term is also larger for consumer goods than for intermediate goods, meaning that expenditures have increased for varieties that were already favored in 1993 for consumers goods.

The Rauch classification categorizes goods based on their degree of differentiation (Rauch (1999)). According to this classification, homogeneous commodities are divided into those whose reference prices are quoted on organized exchanges (essentially primary commodities) and those whose reference prices are quoted only in trade publications. A third category of products includes differentiated goods, that is those goods for which the information is not centralized into either prices or trade publications. The goods are classified based on the 2007 revision of the SITC coding system maintained by the United Nations. Based on the Rauch classification, HS 6-digit sectors could be classified: 704 are classified as "differentiated products", 670 as "reference priced goods" and 189 as "traded on organized exchanges". The last two columns in Table 6 describe the 2SLS estimations on a subsample including products that are referenced priced (in column (3)) and on a subsample including differentiated products (in column (4)). The results show that changing the product's environment of consumers has a stronger effect on the expenditure share for differentiated products than for homogeneous products.

5.3 A non-linear relationship between product diversity and expenditure share

Does too much diversity kill the "diversity gain" for the consumer? A large expansion of varieties could overwhelm the consumers, making the decision to buy more difficult.[15] To assess whether the effect of the variety expansion on the expenditure depends on the expansion size, I regress the expenditure share on bins representing quartiles of the distribution of the change in the number of varieties using an OLS

[15]In psychology, the paradox of choice analyzes the extent to which an increase in consumer choice can raise shoppers' anxiety. A meta-analysis has been conducted on the topic and highlights the moderating variables that can lead to such a situation (Chernev, Böckenholt, and Goodman (2015)).

estimator. The following equation is estimated :

$$ln(\frac{s_{g2018}}{s_{g1993}}) = \omega_0 + \sum_{i=[1,3]} \omega_i \left(ln\left(\frac{\lambda_{g2018}}{\lambda_{g1993}}\right) * bins \right) + \sum_{j=[4,6]} \omega_j bin_i + \omega_7 ln\left(\frac{p_{g2018}}{p_{g1993}}\right) + \omega_8 ln\left(\frac{\tilde{S}^*_{g2018}}{\tilde{S}^*_{g1993}}\right) + \zeta_{gt}. \tag{24}$$

The results are displayed in Figure 3. Figure 3 shows evidence that the correlation between variety expansion and the expenditure share depends on the size of the expansion. Until the third quartile of the distribution of the expansion in the number of varieties, the correlation between the decrease in the price index induced by an expansion of variety and the expenditure share increases. From the third quartile, the correlation is lower. This first evidence should be further investigated.

6 Sensitivity Analysis

In this section, different sensitivity tests assess the robustness of the main analysis.

6.1 Price analysis

As described in section 3.2, the price is proxied by the unit value. It may be contaminated by measurement errors on expenditure. Therefore, different methods have been used to compute the unit value. Columns (1), (3), (5), (7), and (9) in Table 7 report the results of the OLS estimations of eq. (23) using alternative price measures. Columns (2), (4), (6), (8), and (10) show the results of 2SLS estimations. Columns (1) and (2) report the benchmark results when the ratio of the geometric mean of the units values for the varieties consumed in both periods is taken. Columns (3) and (4) describe the results when the unweighted average unit value across varieties within each HS 6-digit sector is used. Columns (5) and (6) report the results when the average of the unit value across varieties weighted by the spending allocated to each variety is taken. Columns (7) and (8) report the results when the median unit value across varieties is taken to measure the price. Finally, columns (9) and (10) describe the results when the unit value is directly computed at the HS 6-digit level (i.e., HS 6-digit expenditure/ HS 6-digit quantity). Different definitions of transportation costs are used as an instrument. In columns (2) and (4), the unweighted average of the unit transportation cost is used. In columns (6) and (10), I use the weighted average unit transportation cost across varieties, whereas in column (8), the median unit transportation cost is used.

The main conclusions concerning the variety-adjustment term still hold regardless of the definition used to measure the price. However, the price's coefficient is not longer significant when the weighted mean, the median, and the HS 6-digit unit values are used. The non-significance of the coefficient is of a particular concern. Based on the demand theory, the price is expected to be correlated with the expenditure. To assess whether the non-significance of the estimates for the price's variables could result from measurement errors, I estimate eq. (23) using the quantity instead of the trade value. The quantity is not affected by potential measurement errors that may alter the relationship between the change in price and the one in the expenditure. To make products as comparable as possible, I run the estimation on a sample that only includes commodities expressed in kilos. The sample includes 1, 151 HS 6-digit sectors. Table 8 reports the results of the analysis. All price measures are significant. Except for the HS 6-digit unit value, they are all negative as expected. These results indicate the three price measures — the weighted average, the median value, and the HS 6-digit unit value — may be noisy proxies for the price in this analysis.

6.2 Potential correlation in demand shocks across countries

A concern about the instrumentation strategy could be that import demand shocks in some sectors may be correlated across high-income economies. If this correlation exists, using the average variety-adjustment term across these countries as an instrument may be problematic. I follow Autor et al. (2013) and regress eq. (23) on subsamples that exclude sectors in which demand shocks are likely to be correlated. The results are reported in Table 9. Column (1) reports results when eq. (23) is estimated on a subsample that excludes the sectors of steel, glass, and cement industries (which may have been subject to a positive demand shock due to the housing boom). Column (2) describes the results when the sectors of apparel, footwear, and textiles are excluded.[16] Column (3) reports the results when all of those sectors are dropped. The null hypothesis that the specified endogenous regressors can actually be treated as exogenous is not rejected. The effect of a change in the varieties available for consumers has a lower effect once I drop the clothing industry. However, dropping these sectors that may have been subject to unobserved demand shocks over that period does not affect the main result that the extensive margin has a strong positive effect on the expenditure share.

7 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 the 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 potential demand shifts. The expansion of variety provides the opportunity to better match their taste to their consumption. Therefore, if the variety

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

expansion is asymmetric across sectors, one can expect consumers to allocate more resources in sectors subject to relatively larger variety growth. In this study, I exploit the exogenous change in the range of products available to consumers due to the growth of exports in low- and middle-income countries in order to identify a causal relationship. 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. Using a 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 the product diversity effect is stronger than the price effect. Between 1993 and 2018, the increase in the availability of varieties decreased the cost of reaching a certain level of satisfaction for the consumer by 5.00%. This "gain from diversity" has led consumers to increase their spending in a sector subject to a variety expansion by 1.19% holding other things constant. The empirical analysis also highlights that over the same period and among products that were already consumed in 1993, consumers reallocated their spending towards goods they already favored in 1993 to the detriment of less appreciated goods. Some evidence also shows this average increase in the expenditure shares induced by an expansion of varieties hides disparities. The variety effect is stronger for consumer goods and differentiated products, as expected from the theoretical framework. Moreover, the average effect could hides some non-linearity. A "too large" expansion of varieties in a sector could overwhelm the consumer with choices, discouraging her from consuming in this sector. This potential effect is interesting and should be further investigated. The identification strategy enables me to highlight two additional relevant results. First, it shows that goods sectors exposed to trade were subjected to negative demand shocks in the 1990s and early 2000s. This reallocation of expenditure from goods sectors towards services sectors seems to have been driven by the fast growth in prices in the medical care sector. Second, the identification strategy points to a new channel through which economic reforms and technological shocks in a country can affect its trading partners. Finally, I note this analysis infers the reallocation of expenditure across sectors using import data. An analysis covering all final products would be preferable. However, the concordance between the classification of production data and the one of imports does not exist yet for the U.S. A recent work by Feenstra, Luck, Obstfeld, and Russ (2018) matched these two datasets for 191 goods, which is not enough for the present analysis. Working on the concordance of classifications between import products and the production of goods remains a fruitful avenue for future research.

References

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

Arkolakis C, Costinot A, Rodríguez-Clare A (2012) New trade models, same old gains? American Economic Review 102(1):94–130

Atkin D, Faber B, Gonzalez-Navarro M (2018) Retail globalization and household welfare: Evidence from mexico. Journal of Political Economy 126(1):1–73

Auer R, Fischer AM (2010) The effect of low-wage import competition on u.s. inflationary pressure. Journal of Monetary Economics 57(4):491–503

Autor DH, Dorn D, Hanson GH (2013) The china syndrome: Local labor market effects of import competition in the united states. American Economic Review 103(6):2121–68

Bernard AB, Jensen JB, Redding SJ, Schott PK (2009) The margins of us trade. American Economic Review 99(2):487–93

Bils M, Klenow PJ (2001) The acceleration of variety growth. American Economic Review 91(2):274–280

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

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

Chaudhuri S, Goldberg PK, Jia P (2006) Estimating the effects of global patent protection in pharmaceuticals: A case study of quinolones in india. American Economic Review 96(5):1477–1514

Chernev A, Böckenholt U, Goodman J (2015) Choice overload: A conceptual review and meta-analysis. Journal of Consumer Psychology 25(2):333–358

Colantone I, Crinò R (2014) New imported inputs, new domestic products. Journal of International Economics 92(1):147 – 165

Colson-Sihra E, Sousa JD, Mayer T (2020) Market integration and culturalhomogenization in consumption patterns, mimeo

Deaton A, Muellbauer J (1980) An almost ideal demand system. The American Economic Review 70(3):312–326

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

Fajgelbaum PD, Khandelwal AK (2016) Measuring the Unequal Gains from Trade *. The Quarterly Journal of Economics 131(3):1113–1180

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

Feenstra RC (1994) New product varieties and the measurement of international prices. American Economic Review 84(1):157–77

Feenstra RC, Weinstein DE (2017) Globalization, markups, and us welfare. Journal of Political Economy 125(4):1040–1074

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

Feenstra RC, Luck P, Obstfeld M, Russ KN (2018) In search of the armington elasticity. The Review of Economics and Statistics 100(1):135–150

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

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

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

di Giovanni J, Levchenko AA, Zhang J (2014) The global welfare impact of china: Trade integration and technological change. American Economic Journal: Macroeconomics 6(3):153–83

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

Handbury J (2019) Are poor cities cheap for everyone?non-homotheticity and the cost of living across u.s. cities, mimeo

Handbury J, Weinstein DE (2014) Goods Prices and Availability in Cities. The Review of Economic Studies 82(1):258–296

Hayashi F (2000) Econometrics, Princeton University Press, New York

Hsieh CT, Klenow PJ (2009) Misallocation and manufacturing tfp in china and india. The Quarterly Journal of Economics 124(4):1403–1448

Hsieh CT, Ossa R (2016) A global view of productivity growth in china. Journal of International Economics 102:209 – 224

Hummels D, Lee KY (2018) The income elasticity of import demand: Micro evidence and an application. Journal of International Economics 113:20 – 34

Kehoe TJ, Ruhl KJ (2013) How important is the new goods margin in international trade? Journal of Political Economy 121(2):358–392

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

Konya I, Ohashi H (2007) International consumption patterns among high-income countries: Evidence from the oecd data. Review of International Economics 15(4):744 – 757

Liu D, Meissner CM (2017) Geography, income, and trade in the 21st century. Working Paper 24121, National Bureau of Economic Research

McCully C (2011) Trends in consumer spending and personal saving, 1959–2009. Tech. rep., Bureau of Economic Analysis, U.S. Department of Commerce

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

Rauch JE (1999) Networks versus markets in international trade. Journal of International Economics 48(1):7–35

Redding SJ, Weinstein DE (2020) Measuring Aggregate Price Indices with Taste Shocks: Theory and Evidence for CES Preferences. The Quarterly Journal of Economics 135(1):503–560

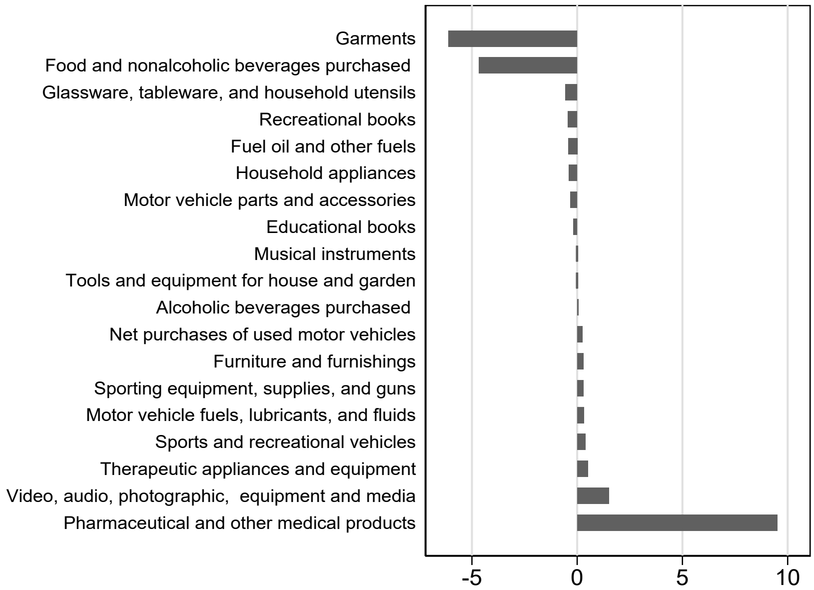
Schott PK, Fuest C, O'Rourke K (2008) The relative sophistication of chinese exports [with discussion]. Economic Policy 23(53):5–49

Stock J, Yogo M (2005) Testing for Weak Instruments in Linear IV Regression, Cambridge University Press, New York, pp 80–108

UNCTAD (2013) Trade and development report, 2013. Report

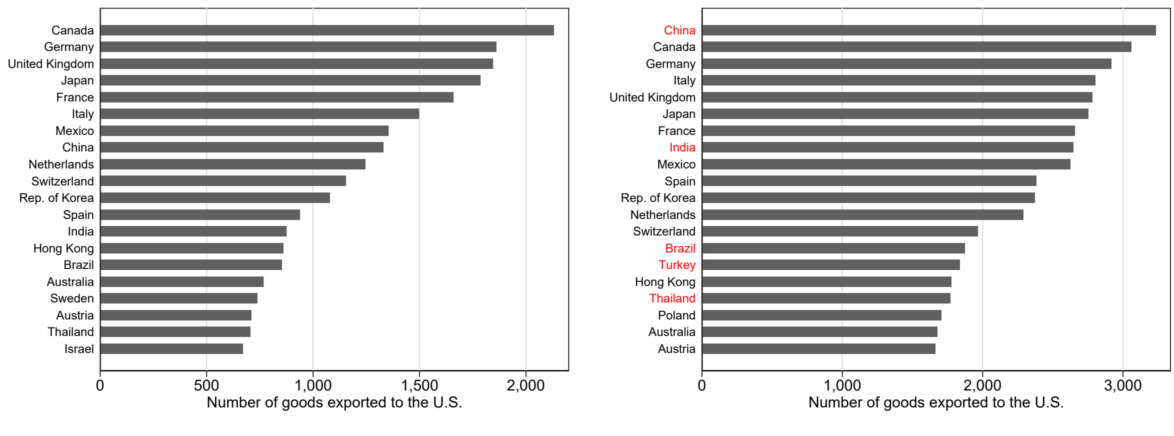
UNDP (1998) Human development report 1998. Tech. rep., United Nations

Figure 1: Changes in the pattern of expenditure in the consumption of goods in the U.S. between 1993 and 2018



Source: NIPAs, Bureau of Economic Analysis, U.S. Department of Commerce. The changes are in percentage points. The sectors are ranked from the largest decrease to the largest increase.

Figure 2: Ranking of the top 20 importers in terms of number of goods imported by the U.S. in 1993 and 2018



Emerging countries that have improved their position in the ranking between 1993 and 2018 are highlighted in red in Figure (b).

Figure 3: Non-linearity in the correlation between the variety expansion and the expenditure share

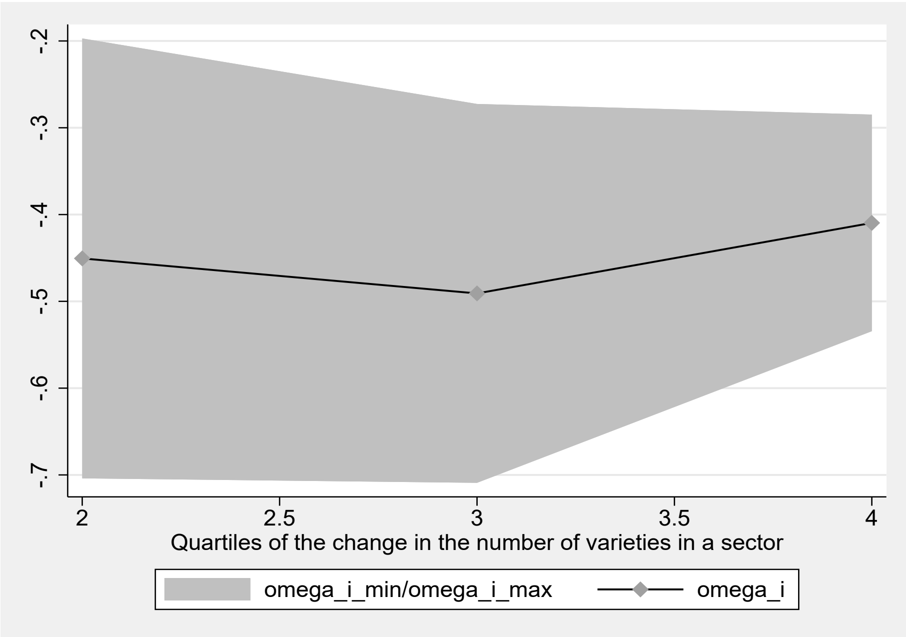


Table 1: Unit values, variety change, and demand-heterogeneity terms

HS 6-digit sectors				
Part 1 : Unit values				
	Mean	Std Dev.	Min	Max
$ln\left(\frac{\tilde{P}^*_{gt}}{\tilde{P}^*_{gt-1}}\right)$	0.48	0.71	-5.00	7.36
	Mean in 1993	Mean in 2018		
$\overline{price}$	US$ 205	US$148		
$price_w$	US$ 228	US$ 101		
$price_m$	US$ 55	US$ 100		
$price$	US$ 36	US$ 65		
Varieties existing only in 1993	US$ 222			
Varieties existing only in 2018		US$ 293		
Varieties consumed in both periods	US$ 205	US$ 148		
All varieties	US$ 213	US$ 220		
Part 2 : Variety changes by sector				
	Mean	Std Dev.	Min	Max
$\frac{\text{Total imports in 2018}}{\text{Total imports in 1993}}$	17.28	203.83	0.00	8028.24
$\frac{\text{Imports of the set of common goods in 2018}}{\text{Imports of the set of common goods in 1993}}$	10.06	62.23	0.00	1987.20
row 1/ row 2	1.70	-	-	4.04
Change in the number of varieties by HS 4-digit sector between 1993 and 2018	7.31	9.62	-29.00	59.00
$\frac{\lambda_{gt}}{\lambda_{gt-1}}$	0.95	0.22	0.08	4.74
Part 3 : Demand heterogeneity term by sector				
	Mean	Std Dev.	Min	Max
$\frac{\tilde{S}^*_{gt}}{\tilde{S}^*_{gt-1}}$	1.09	0.35	0.16	4.70
HS 2-digit sectors				
Percentage points difference in the expenditure share	0.0	0.02	-0.03	0.13

Source: UN Comtrade Statistics Database.

Table 2: Decomposition of the change of U.S. imports into the extensive and intensive margins over time

	93-94	**94-95**	**95-96**	**96-97**	**97-98**	**98-99**	**99-00**	**00-01**	**01-02**	**02-03**	**03-04**	**04-05**	**05-06**	**06-07**
Extensive Margin														
New varieties (%)	21.22	32.54	84.74	11.04	14.42	14.35	13.67	4.44	365.41	4.99	2.96	2.54	3.14	213.35
Exit (%)	-8.99	-6.80	-11.45	-6.48	-7.11	-13.75	-5.36	-5.06	395.36	-3.79	-1.48	-1.78	-2.02	-176.45
Net Entry (p.p.)	12.23	25.70	72.29	4.56	7.31	-0.60	8.31	-0.62	-29.95	1.20	1.48	0.76	1.12	36.89
Intensive margin	87.76	74.30	27.70	95.43	92.68	99.34	91.69	99.38	129.96	98.80	98.52	99.24	98.88	63.12
Import growth rate (%)	15.50	18.49	15.90	10.99	7.81	4.91	12.86	-6.16	1.89	8.24	16.69	13.69	11.02	5.31
	07-08	**08-09**	**09-10**	**10-11**	**11-12**	**12-13**	**13-14**	**14-15**	**15-16**	**16-17**	**17-18**	**93-06**	**06-18**	**93-18**
Extensive Margin														
New varieties (%)	3.32	0.66	1.42	3.55	112.05	90.50	7.25	5.46	7.14	40.11	3.07	92.77	67.30	92.41
Exit (%)	-3.37	-1.15	-1.22	-1.42	-114.13	-69.46	-5.48	-4.50	-6.30	-26.74	-2.11	-2.84	-38.62	-2.32
Net Entry (p.p.)	-0.05	-0.49	0.19	2.12	-2.08	21	1.77	0.96	0.84	13.37	0.96	89.93	28.68	90.10
Intensive margin	100.05	-99.51	99.80	97.87	102.08	-121	98.23	-100.98	-100.84	86.63	99.04	10.07	71.31	9.90
Import growth rate (%)	7.50	-25.57	22.56	15.48	3.04	-0.31	3.39	-4.13	-2.7	7.06	8.50	1100	36.96	1644

Source: UN Comtrade Statistics Database.

Table 3: Correlations between developed economies

	(1) Δ ln (U.S. exports)
$ln\left(\frac{\lambda_{g2018}}{\lambda_{g1993}}\right)^{inst}$	-0.022
	(0.039)
Constant	0.666***
	(0.038)
Observations	1735
R^2	0.0004

Robust standard errors, clustered at the HS 6-digit sector, are in in parentheses.
* p<0.1, ** p<0.05, *** p<0.01

Table 4: First-stage results

	(1)	(2)	(3)	(4)
	$\ln\left(\frac{\lambda_{g1993}}{\lambda_{g2006}}\right)$	$\ln\left(\frac{\tilde{P}^*_{gt}}{\tilde{P}^*_{gt-1}}\right)$	$\ln\left(\frac{\lambda_{g1993}}{\lambda_{g2006}}\right)$	$\ln\left(\frac{\tilde{P}^*_{gt}}{\tilde{P}^*_{gt-1}}\right)$
$ln\left(\frac{\lambda_{g2018}}{\lambda_{g1993}}\right)^{inst}$	0.220***	-0.015	0.215***	-0.017
	(0.026)	(0.021)	(0.026)	(0.021)
$\Delta\ \ln(\overline{\text{unit cost}})$	0.003	-0.074***	0.003	-0.074***
	(0.010)	(0.011)	(0.010)	(0.011)
$\ln\left(\frac{\tilde{S}^*_{g2018}}{\tilde{S}^*_{g1993}}\right)$			-0.139***	-0.054***
			(0.022)	(0.018)
Constant	-0.062***	0.486***	-0.090***	0.475***
	(0.022)	(0.019)	(0.024)	(0.020)
Observations	1761	1761	1761	1761
R^2	0.102	0.049	0.130	0.055

Robust standard errors, clustered at the HS 6-digit sector, are in parentheses.
* p<0.1, ** p<0.05, *** p<0.01
The sample of high-income economies used to build the instrument includes Australia, Denmark, Ireland, Iceland, Japan, Korea, Portugal, Sweden, Spain, Switzerland, and the U.S. The instrument is computed by only taking imports from low-and middle-income countries.The price index is measured by the ratio of the geometric means in 1993 and 2018 for the varieties consumed in both periods. The transportation cost for a particular HS 6-digit sector is computed by taking the average unit transportation cost across varieties.

Table 5: Main results

	Baseline specification		2SLS	
	(1)	(2)	(3)	(4)
	Δ ln (share)	Δ ln (share)	Δ ln (share)	Δ ln (share)
$\ln\left(\frac{\lambda_{g2018}}{\lambda_{g1993}}\right)$	-0.631***	-0.765***	-0.872***	-0.973***
	(0.055)	(0.049)	(0.148)	(0.144)
$\ln\left(\frac{\tilde{P}^*_{gt}}{\tilde{P}^*_{gt-1}}\right)$	-0.344***	-0.418***	-0.598**	-0.571**
	(0.069)	(0.062)	(0.245)	(0.227)
$\ln\left(\frac{\tilde{S}^*_{g2018}}{\tilde{S}^*_{g1993}}\right)$		-0.583***		-0.622***
		(0.035)		(0.046)
Constant	-0.064***	-0.160***	0.016	-0.129
	(0.047)	(0.042)	(0.113)	(0.101)
Observations	1761	1761	1761	1761
R^2	0.128	0.263	0.101	0.248
Kleibergen-Paap			22.53	22.45
Endogeneity test			0.2120	0.3316
Sample period	1993-2018	1993-2018	1993-2018	1993-2018
Estimator	FD-OLS	FD-OLS	FD-2SLS	FD-2SLS

Robust standard errors, clustered at the HS 6-digit sector, are in parentheses.
* p<0.1, ** p<0.05, *** p<0.01
The Stock-Yogo weak ID test critical value is 7.03 for 10% maximal IV size.
The sample of high-income economies used to build the instrument includes Australia, Denmark, Ireland, Iceland, Japan, Korea, Portugal, Sweden, Spain, Switzerland, and the U.S. The instrument is computed by only taking imports from low- and middle-income countries. The price index is measured by computing the ratio of the geometric means in 1993 and 2018 for the varieties consumed in both periods. The transportation cost for a particular HS 6-digit sector is computed by taking the average unit transportation cost across varieties.

Table 6: Analysis by type of goods

	BEA classification		Rauch classification	
	(1)	(2)	(3)	(4)
	Δ ln (share)	Δ ln (share)	Δ ln (share)	Δ ln (share)
$\ln\left(\frac{\lambda_{g2018}}{\lambda_{g1993}}\right)$	-0.680***	-1.212***	-0.619***	-1.308***
	(0.161)	(0.437)	(0.190)	(0.224)
$\ln\left(\frac{\tilde{S}^*_{g2018}}{\tilde{S}^*_{g1993}}\right)$	-0.461***	-0.830***	-0.439***	-0.655***
	(0.053)	(0.074)	(0.057)	(0.094)
Δ ln(*price*)	-0.252***	-0.450***	-0.258**	
	(0.069)	(0.150)	(0.105)	
Δ ln($\overline{price}$)				-0.604
				(0.501)
Constant	-0.120	0.560	-0.010	-0.551***
	(0.084)	(0.490)	(0.104)	(0.110)
Observations	1102	327	603	680
R^2	0.180	0.425	0.146	0.253
Kleibergen-Paap	26.36	10.50	13.16	9.13
Endogeneity test	0.1050	0.6360	0.2487	0.8510
Sample period	1993-2018	1993-2018	1993-2018	1993-2018
Estimator	FD-2SLS	FD-2SLS	FD-2SLS	FD-2SLS
types of product	Intermediate	Consumer goods	Reference priced goods	Differentiated goods

Robust standard errors, clustered at the HS 6-digit sector, are in parentheses.
* p<0.1, ** p<0.05, *** p<0.01
The Stock-Yogo weak ID test critical value is 7.03 for 10% maximal IV size.
price is computed by dividing the trade value by the quantity for each HS 6-digit sector. $\overline{price}$ is computed by taking the unweighted average unit value across varieties within each HS 6-digit sector. The sample of high-income economies used to build the instrument for intermediate goods includes Denmark, Greece, Ireland, Iceland, Japan, the Netherlands, Switzerland, and the U.S. I can compute the instrumental variable for 1, 102 sectors out of the 1, 251 intermediate sectors. The sample of high-income economies used to build the instrument for the estimations reported in the last three columns includes Australia, Ireland, Iceland, Portugal, and the U.S. The instrument is computed by only taking imports from low- and middle-income countries. I can compute the instrumental variable for 603 referenced goods out of the 670 in the sample and for 680 differentiated sectors out of the 704 included in the sample.The transportation cost for a particular HS 6-digit sector is computed by taking the median unit transportation cost across varieties for the estimations reported in columns (1), (2), and (4). The weighted average unit transportation cost is used in column (3).

Table 7: Price analysis: Alternative proxies of price

	(1) Δ ln (share) value	(2) Δ ln (share) value	(3) Δ ln (share) value	(4) Δ ln (share) value	(5) Δ ln (share) value	(6) Δ ln (share) value	(7) Δ ln (share) value	(8) Δ ln (share) value	(9) Δ ln (share) value	(10) Δ ln (share) value
$\ln\left(\frac{\lambda_{g2018}}{\lambda_{g1993}}\right)$	-0.765***	-0.973***	-0.757***	-0.930***	-0.761***	-0.877***	-0.733***	-0.917***	-0.705***	-0.811***
	(0.049)	(0.144)	(0.049)	(0.141)	(0.047)	(0.162)	(0.048)	(0.163)	(0.049)	(0.168)
$\ln\left(\frac{\tilde{S}^*_{g2018}}{\tilde{S}^*_{g1993}}\right)$	-0.583***	-0.622***	-0.605***	-0.642***	-0.503***	-0.591***	-0.545***	-0.574***	-0.576***	-0.601***
	(0.035)	(0.046)	(0.036)	(0.051)	(0.036)	(0.065)	(0.036)	(0.046)	(0.036)	(0.045)
$\ln\left(\frac{\tilde{P}^*_{g2018}}{\tilde{P}^*_{g1993}}\right)$	-0.418***	-0.571**								
	(0.062)	(0.227)								
$\Delta\ \ln(\overline{price})$			-0.265***	-0.333**						
			(0.031)	(0.137)						
$\Delta\ \ln(price_w)$					-0.280***	0.059				
					(0.050)	(0.294)				
$\Delta\ \ln(price_m)$							-0.094	-0.084		
							(0.060)	(0.240)		
$\Delta ln\,(price)$									-0.137***	-0.188
									(0.019)	(0.155)
Constant	-0.165***	-0.134	-0.259***	-0.267***	-0.249***	-0.408***	-0.319***	-0.360***	-0.169***	-0.123
	(0.042)	(0.101)	(0.035)	(0.059)	(0.038)	(0.108)	(0.039)	(0.085)	(0.042)	(0.224)
Observations	1761	1761	1761	1761	1761	1761	1761	1761	1761	1761
R^2	0.263	0.248	0.268	0.258	0.251	0.208	0.229	0.220	0.251	0.244
Kleibergen-Paap		22.45		38.43		14.11		20.55		9.37
Endogeneity test		0.3316		0.4438		0.1405		0.3879		0.5995
sample period	1993-2018	1993-2018	1993-2018	1993-2018	1993-2018	1993-2018	1993-2018	1993-2018	1993-2018	1993-2018
Estimator	FD-OLS	FD-2SLS	FD-OLS	FD-2SLS	FD-OLS	FD-2SLS	FD-OLS	FD-2SLS	FD-OLS	FD-2SLS

Robust standard errors, clustered at the HS 6-digit sector, are in parentheses.
* p<0.1, ** p<0.05, *** p<0.01
The Stock-Yogo weak ID test critical value is 7.03 for 10% maximal IV size.
$ln\left(\frac{\tilde{P}^*_{g2018}}{\tilde{P}^*_{g1993}}\right)$ is computed by taking the ratio of the geometric means in 1993 and 2018 for the varieties consumed in both periods. $\overline{price}$ is computed by taking the unweighted average unit value across varieties within each HS 6-digit sector. $price_w$ is computed by taking the average unit value across varieties within each HS 6-digit sector, weighted by the expenditure allocated to each variety. $price_m$ is computed by taking the median unit value across varieties within each HS 6-digit sector. $price$ is computed by dividing the trade value by the quantity for each HS 6-digit sector. The transport unit cost, one of the two instruments used to address the potential endogeneity issue, is computed in the same way as the specific unit value, except for the estimation in columns (2) and (10). In column (2), the unweighted mean of the transportation cost is taken. In column (10), the instrument is the average unit transportation cost across varieties weighted by the expenditure allocated to each variety in each HS 6-digit sector. The sample of high-income economies used to build the instrument includes Australia, Denmark, Ireland, Iceland, Japan, Korea, Portugal, Sweden, Spain, Switzerland, and the U.S. The instrument is computed by only taking imports from low- and middle-income countries.

Table 8: Price analysis: The quantity is used instead of the trade value to compute the share allocated to an HS 6-digit sector

	(1) Δ ln (share) quantity	(2) Δ ln (share) quantity	(3) Δ ln (share) quantity	(4) Δ ln (share) quantity	(5) Δ ln (share) quantity	(6) Δ ln (share) quantity	(7) Δ ln (share) quantity	(8) Δ ln (share) quantity	(9) Δ ln (share) quantity	(10) Δ ln (share) quantity
$ln\left(\frac{\lambda_{g2018}}{\lambda_{g1993}}\right)$	-0.776***	-0.827***	-0.741***	-0.703***	-0.801***	-0.889***	-0.772***	-0.940***	-0.714***	-0.818***
	(0.086)	(0.215)	(0.086)	(0.215)	(0.082)	(0.208)	(0.084)	(0.221)	(0.085)	(0.225)
$ln\left(\frac{\tilde{S}^*_{g2018}}{\tilde{S}^*_{g1993}}\right)$	-0.592***	-0.642***	-0.574***	-0.636***	-0.415***	-0.409***	-0.443***	- 0.456***	-0.486***	-0.501***
	(0.055)	(0.077)	(0.058)	(0.083)	(0.052)	(0.065)	(0.052)	(0.060)	(0.055)	(0.066)
$ln\left(\frac{\tilde{P}^*_{g2018}}{\tilde{P}^*_{g1993}}\right)$	-0.735***	-1.065***								
	(0.130)	(0.268)								
Δ ln($\overline{price}$)			-0.289***	-0.528***						
			(0.053)	(0.152)						
Δ ln($price_w$)					-0.631***	-0.770***				
					(0.099)	(0.277)				
Δ ln($price_m$)							-0.829***	-1.036***		
							(0.116)	(0.3 12)		
Δ ln($price$)									0.134***	0.301***
									(0.044)	(0.095)
Constant	0.406***	0.554***	0.157***	0.240***	0.302***	0.342***	0.440***	0.508***	-0.078	-0.276**
	(0.076)	(0.125)	(0.045)	(0.067)	(0.058)	(0.112)	(0.071)	(0.137)	(0.069)	(0.137)
Observations	1151	1151	1151	1151	1151	1151	1151	1151	1151	1151
R^2	0.277	0.254	0.219	0.179	0.293	0.286	0.319	0.304	0.178	0.153
Kleibergen-Paap		22.79		29.33		16.36		14.17		23.74
Endogeneity test		0.3108		0.1419		0.8116		0.6975		0.1336
sample period	1993-2018	1993-2018	1993-2018	1993-2018	1993-2018	1993-2018	1993-2018	1993-2018		
Estimator	FD-OLS	FD-2SLS	FD-OLS	FD-2SLS	FD-OLS	FD-2SLS	FD-OLS	FD-2SLS	FD-OLS	FD-2SLS

Robust standard errors, clustered at the HS 6-digit sector, are in parentheses.

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

The Stock-Yogo weak ID test critical value is 7.03 for 10% maximal IV size.

$ln\left(\frac{\tilde{P}^*_{g2018}}{\tilde{P}^*_{g1993}}\right)$ is computed by taking the ratio of the geometric means in 1993 and in 2018 for the varieties consumed in both periods. $\overline{price}$ is computed by taking the unweighted average unit value across varieties within each HS 6-digit sector. $price_w$ is computed by taking the average unit value across varieties within each HS 6-digit sector, weighted by the expenditure allocated to each variety. $price_m$ is computed by taking the median unit value across varieties within each HS 6-digit sector. $price$ is computed by dividing the trade value by the quantity for each HS 6-digit sector. The transport unit cost, one of the two instruments used to address the potential endogeneity issue, is computed in the same way as the specific unit value, except for the estimations in columns (2) and (10). In column (2), the unweighted mean of the transportation cost is taken. In column (10), the instrument is the average unit transport cost across varieties weighted by the expenditure allocated to each variety in each HS 6-digit sector. The sample of high-income economies used to build the instrument includes Australia, Denmark, Ireland, Iceland, Japan, Korea, Portugal, Sweden, Spain, Switzerland, and the U.S. The instrument is computed by only taking imports from low- and middle-income countries.

Table 9: Alleviate potential demand shocks across sectors

	(1) Δ ln (share)	(2) Δ ln (share)	(3) Δ ln (share)
$ln\left(\frac{\lambda_{g2018}}{\lambda_{g1993}}\right)$	-1.029***	-0.757***	-0.869***
	(0.157)	(0.174)	(0.187)
$ln\left(\frac{\tilde{P}^*_{g2018}}{\tilde{P}^*_{g1993}}\right)$	-0.415	-0.657***	-0.763***
	(0.279)	(0.199)	(0.213)
$ln\left(\frac{\tilde{S}^*_{g2018}}{\tilde{S}^*_{g1993}}\right)$	-0.654***	-0.499***	-0.575***
	(0.052)	(0.064)	(0.071)
Constant	-0.270**	0.318***	0.339***
	(0.115)	(0.093)	(0.096)
Observations	1567	1142	947
R^2	0.257	0.175	0.153
Kleibergen-Paap	15.80	27.80	22.58
Endogeneity test	0.1680	0.2100	0.1545
Sample period	1993-2018	1993-2018	1993-2018
Estimator	FD-2SLS	FD-2SLS	FD-2SLS
Sectors dropped	Steel, glass and cement indus.	Apparel, footwear and textiles	All

Robust standard errors, clustered at the HS 4-digit sector, are in parentheses.
* p<0.1, ** p<0.05, *** p<0.01
The Stock-Yogo weak ID test critical value is 7.03 for 10% maximal IV size.
The sample of high-income economies used to build the instrument includes Australia, Denmark, Ireland, Iceland, Japan, Korea, Portugal, Sweden, Spain, Switzerland, and the U.S. The instrument is computed by only taking imports from low- and middle-income countries. The transportation cost for a particular HS 6-digit sector is computed by taking the average unit cost from each exporter in column (1) and by taking the ratio of the geometric means of the unit transportation costs in columns (2) and (3).

A Analysis at the HS 4-digit level

In this section, the analysis is done at a higher level of aggregation. A good is defined as an HS 4-digit sector, and a variety is defined as an HS 6-digit good imported from a particular country within a particular HS 4-digit sector. A new product is a product that does not previously have a positive record in an HS 4-digit category. Using this definition, the variety-adjustment and the demand-heterogeneity terms are computed following the methodology described in section 3.3. On average, the variety-adjustment term is 0.97, which is similar to the statistics found at the HS 6-digit sector. On average, the demand has not shifted over time for a particular HS 4-digit sector. Interestingly, the ratio is much more centered toward 1 than at the HS 6-digit level. The standard deviation of the ratio is equal to 0.04 against 0.35 at the HS 6-digit level. The ratios range between 0.85 and 1.17, whereas it ranges between 0.16 and 4.70 at the HS 6-digit level. Within an HS 4-digit sector, some HS 6-digit products can have different units.To minimize measurement errors linked to the quantity units, I first convert units when possible (e.g. tons are converted to kilos). I then focus on HS 4-digit sectors such that the HS 6-digit products within them report quantity in the same unit of measure.

The price is also proxied by the unit value at the HS 4-digit level. As for the HS 6-digit sector, I computed and assessed the unit value for each HS 4-digit sector using the different methods described in section 3.2. In the analysis described in Table 10, I use the weighted average unit value because it provides the most stable results and is well instrumented across several estimations. The final database includes 454 HS 4-digit sectors and 196 economic entities over the 1993-2018 period.

I follow the same methodology and instrumentation strategy as the ones described in section 4. The first two columns in Table 10 report estimations of eq. (23) at the HS 4-digit level over the 1993-2018 period. Column (1) reports the results of an OLS estimation, whereas column (2) shows the coefficients computed from a 2SLS estimation. I find similar results to the ones at the HS 6-digit level. I find a positive and significant impact of the product diversity on the expenditure share, and a negative demand shock seems to have biased the correlation downward given the larger coefficient computed by 2SLS and reported in column (2). Contrary to the results found at the HS 6-digit level, the price's coefficients are not significant.[17] This result can be explained by the fact that the unit value is computed from trade values and by the level of aggregation. At the HS 4-digit level, the unit value is quite noisy. The demand-heterogeneity term is not significant neither. On average, over the 1993-2018 period for the 454 sectors included in the sample, small changes were observed in the preferences towards varieties that were already consumed in 1993 as described in the preceding section. Different explanations can

[17] Analyses using different price's measures have been done. The results do not differ from the ones reported in Table 10. Overall, the coefficients are usually not significant, in particular, when a 2SLS estimation is run.

clarify this absence of a demand effect. At the HS 4-digit level of aggregation, the average demand effect is small. Another potential explanation is that over 25 years, the average demand change may not affect the expenditure share. Columns (3) and (4) in Table 10 corroborate the second explanation more than the first one. An OLS and a 2SLS are run over the 1993-2006 period. Columns (3) and (4) show that over this period, the demand-heterogeneity term had a positive and significant effect on the expenditure share. The long-term effect of the demand term at that level of aggregation should be further investigated. Overall, the analysis shows that at both levels of aggregation, the product diversity leads to a reallocation of expenditure across sectors.

Table 10: 2SLS analysis at the HS 4-digit level

	(1)	(2)	(3)	(4)
	Δ ln (share)	Δ ln (share)	Δ ln (share)	Δ ln (share)
$\ln\left(\frac{\lambda_{gt}}{\lambda_{g1993}}\right)$	-0.862***	-1.524***	0.957***	1.072***
	(0.106)	(0.373)	(0.090)	(0.208)
Δ $\ln(price_w)$	-0.039	-0.057	0.021	0.059
	(0.033)	(0.090)	(0.023)	(0.056)
$\ln\left(\frac{\tilde{S}^*_{g2018}}{\tilde{S}^*_{g1993}}\right)$	0.040	0.061	0.638***	0.652***
	(0.038)	(0.050)	(0.071)	(0.070)
Constant	-0.716***	-0.924***	-0.660***	-0.765***
	(0.237)	(0.322)	(0.066)	(0.133)
Observations	454	454	454	454
R^2	0.241	0.100	0.455	0.449
Kleibergen-Paap		11.33		10.67
Endogeneity test		0.1035		0.6231
Sample period	1993-2018	1993-2018	1993-2006	1993-2006
Estimator	FD-2SLS	FD-2SLS	FD-2SLS	

Robust standard errors, clustered at the HS 4-digit sector, are in parentheses.
* p<0.1, ** p<0.05, *** p<0.01
The Stock-Yogo weak ID test critical value is 7.03 for 10% maximal IV size.
In column (2), the high-income economies used to compute the instrument are Australia, Ireland, Iceland, Portugal, Spain, Sweden, Switzerland, United Kingdom, and the U.S. The second instrument is computed by taking the average unit transportation cost weighted by the trade value allocated to each variety. The high-income economies considered in column (4) are Australia, Austria, Iceland, Ireland, New Zealand, and the U.S. The second instrument is computed by taking the median unit transportation cost across varieties within each HS 4-digit sector.