Aggregation and the Estimation of Quality Change: Application to the U.S. Import Prices*

Marco Errico† Danial Lashkari‡

September 24, 2024

Please click here for the most recent draft.

Abstract

Lack of detailed data on product characteristics and quality change poses a challenge for the accurate aggregation of changes in the real volumes of consumption, production, and trade flows. To tackle this problem, we propose a method that allows us to identify demand and infer unobserved quality change using data only on prices and market shares, without the need for external cost shock instruments or strong assumptions on the covariance between supply and demand shocks. We also characterize the contribution of changes in quality, price, and variety entry/exit to the aggregate price index for general invertible demand systems, generalizing the standard results derived in the CES case. We apply our strategy to compute the US import price index based on the Kimball demand, allowing for heterogeneity in substitutability across products. We find that quality change on average lowers the inflation in import prices by around 0.7% annually. To further validate our approach, we show that it estimates price elasticities and quality changes similar to those found by the standard mixed logit (BLP) demand in data on the US auto market, without relying on the information on product characteristics and price instruments.

JEL code: F1, F14, E31

*We thank Paul Grieco, Charles Murry, and Ali Yurukoglu for sharing their data and for insightful conversations, and Mary Amiti, Amit Khandelwal, Pete Klenow, Ariel Pakes, Peter Schott, Anson Soderbery, Jonathan Vogel, and Ali Yurukoglu for valuable comments and feedback. Alberto Cappello and Jiyuan Zhong provided excellent research assistance. Danial Lashkari completed part of this work as a visiting fellow at the Stanford Institute for Policy Research (SIEPR) and the Princeton University International Economics Section.

†Bank of Italy. Email: marco.errico@bc.edu.

‡FRB New York. Email: danial.lashkari@ny.frb.org. Website: www.daniallashkari.com.

1 Introduction

Aggregate price indices allow us to translate changes in the volumes of production, consumption, imports, or exports into welfare-relevant measures of real economic activity. The data used to construct these measures often record volumes and unit values at the level of industry or product classifications but fail to account for changes within these units, such as entry and exit of goods or shifts in their characteristics. While statistical agencies attempt to adjust for such within-unit *quality* changes, much of this variation remains unmeasured.[1] Consider, for example, the task of measuring the prices of U.S. imports, where it has proven hard to collect systematic data on product characteristics. The official BLS series for import prices suggests a nearly five-fold increase for computer and peripheral equipment goods relative to US producer prices in this sector between 1989 and 2018.[2] Over this period, US imports in this sector rapidly rose relative to domestic production, implying potential mismeasurement in import quality.

To address this problem, academic work has developed alternative methods to adjust for the entry and exit of imported varieties (Feenstra, 1994; Broda and Weinstein, 2006) and quality changes inferred from variations in residual demand (Khandelwal, 2010; Hallak and Schott, 2011). While widely adopted in many research applications, these approaches still suffer from two limitations. First, they heavily rely on the assumption of Constant Elasticity of Substitution (CES) demand, which rules out heterogeneity in patterns of cross-product substitutability. Moreover, these methods require identifying substitution elasticities without access to plausibly exogenous variations in prices, often assuming that supply and demand shocks are uncorrelated—a risky assumption if quality and production costs are in fact intertwined.

In this paper, we develop and implement novel strategies to construct aggregate price indices that account for quality and variety change without restrictive assumptions on the demand system or on the covariance structure between supply and demand shocks. Our first contribution is to generalize the recent unified CES price index of Redding and Weinstein (2020a) and the Feenstra (1994) variety correction beyond the CES model, showing how the heterogeneity in cross-product elasticities of substitution affects the contributions of quality and variety to changes in the aggregate price indices. Given the data limitations often faced, we further specialize our results to a family of income-invariant (homoth-

[1] This problem is sometimes referred to as the *quality change bias* in the measures of inflation in the cost-of-living (Boskin et al., 1998; Gordon and Griliches, 1997). For an overview of techniques used by statistical agencies to adjust for product quality, e.g., see Triplett (2004).

[2] See Figure D.5 in Appendix D.3.2. This observation was made Robert J. Gordon in a presentation at the conference in honor of Zvi Grilliches at the College de France in May, 2024.

etic) demand that allow some elasticity variation while reducing the data requirements for estimation (Matsuyama, 2022).[3] These preferences include specifications like Kimball (1995a), CRESH (Hanoch, 1971), and HSA (Matsuyama and Ushchev, 2017), which have been gaining traction in recent empirical work (e.g., Edmond et al., 2023; Berlingieri et al., 2022; Grossman et al., 2023).

Our second contribution is to propose a method for estimating demand systems using panel data on prices and market shares, without requiring detailed product characteristics or external cost instruments. The idea behind our approach is to apply the dynamic panel (DP) methods to the joint evolution of product-level prices and quality (demand) shocks. More specifically, we assume that future shocks to the quality of each product, conditional on current quality, are uncorrelated with current product prices.[4] We show that this assumption is satisfied in most standard models that rule out dynamic pricing (e.g., when prices are flexible and demand does not directly depend on past prices). In such settings, firms choose current prices to maximize current-period profits irrespective of future demand shocks. Accordingly, we can derive moment conditions that identify flexible demand systems in the presence of correlated supply and demand shocks. The only additional requirement is that product prices exhibit strong autocorrelation over time due to persistent cost shocks.

We use our strategy for measuring the price index of imports in the US. To account for the substitutability between domestic and imported varieties of products, we build a dataset disaggregating total consumption to those sourced from the US and each of its trading partners at the level of 156 (5-digit) NAICS industry codes. We assume a nested demand structure in which consumers evaluate the varieties of products sourced by different origin countries, including the US, using a CES or Kimball aggregator, where the latter allows heterogeneity in cross-product substitutability. We take unit values of imports as our measure of price in the customs records and the US producer price index (PPI) for each industry as the price of the corresponding domestic variety. We express the quality of the varieties supplied by all other origins relative to the US, assuming that the US PPI accurately accounts for quality improvements.[5]

[3]For another generalization of the Feenstra variety correction to alternative family of demand systems, and its application to the cereal market in the US, see Foley (2021).

[4]This strategy has been combined with complementary instrumental variables in estimating rich demand systems in several IO applications (e.g., Grennan, 2013; Lee, 2013; Sweeting, 2013), and in estimating firm-level production functions (Caliendo et al., 2020). We note that our assumptions about the dynamics of demand shocks are also in line with Redding and Weinstein (2020a), who find a strong persistence in demand shocks in the Nielsen barcode data.

[5]The BLS performs multiple quality-adjustment strategies for PPIs that are not used for import prices, e.g., hedonic regressions on product characteristics (see U.S. Bureau of Labor Statistics, 2024, chapter 14).

Using the resulting data (1989-2016), we estimate CES and Kimball demand systems in each industry. We find large magnitudes in the CES estimates based on our approach compared to those found using the conventional methods that rule out demand and supply shocks (Feenstra, 1994; Broda and Weinstein, 2006; Soderbery, 2015a). Thus, the conventional estimates likely suffer from an endogeneity bias due to a positive correlation between quality shocks and prices. Next, focusing on the case of Kimball demand (featuring heterogeneity in substitution elasticities), we find that quality improvements in the imported products lower the US import price index by over 20% (0.7% annually). Using CES demand suggests a lower contribution of around 16% (0.5% annually).[6]

We further validate our method using rich data from the automobile market (1980-2018) that includes detailed information on product characteristics which provide proxies for product quality. We first verify our identification assumption by showing that, controlling for current product characteristics, future characteristics are not correlated with current prices. Furthermore, for both CES and Kimball demands, we show that our identification strategy leads to similar estimates to those found using a standard cost shock instrument based on the real exchange rate (RER) variations between the US and each model's country of assembly.

Using the auto data, we further compare our estimates for CES and Kimball demand with the mixed logit demand featuring heterogeneous elasticities (Berry, 1994; Berry et al., 1995). The estimates based on the Kimball demand system are closely aligned with those of the mixed logit demand, while those of CES feature a downward *heterogeneity bias* in the magnitude of the demand elasticity.[7] Compared to the price indices constructed for the US auto industry based on rich demand systems such as mixed logit and mixed CES, the Kimball price index appears to provide a better approximation compared to the CES price index. We lastly examine our inferred measures of quality and show that they are correlated with characteristics valued by consumers.

Prior Work The role of product quality for the patterns of international trade and specialization, at the aggregate and at the firm level, has been the subject of a vast body of theoretical and empirical work (e.g., Linder, 1961, Flam and Helpman, 1987; Hummels and Skiba, 2004; Hallak, 2006; Verhoogen, 2008; Fajgelbaum et al., 2011; Baldwin

[6]Relying on the standard identification approach ruling out correlated supply and demand shocks, Berlingieri et al. (2018) also find that quality change accounts for the bulk of the gains from openness accruing from the trade agreements signed by the EU. Using scanner-level data, Redding and Weinstein (2020a) show that the quality bias is sizable relative to the variety bias.

[7]Appendix E.3 shows, both theoretically and empirically, that this bias emerges in the presence of a correlation between the magnitude of the elasticity and the price volatility across products.

and Harrigan, 2011; Kugler and Verhoogen, 2012; Manova and Zhang, 2012; Martin and Mejean, 2014; Dingel, 2017; Eaton and Fieler, 2022). Early empirical work on the importance of quality proxied product quality with unit values (e.g., Schott, 2004; Hummels and Klenow, 2005).[8] As already mentioned, we follow the approach pioneered by Khandelwal (2010) and Hallak and Schott (2011) in inferring quality through variations in demand conditional on price. Our emphasis on heterogeneity in the patterns of cross-product substitutability is in line with recent work that distinguishes such variations as indicative of horizontal rather than vertical (quality) differentiation (Di Comite et al., 2014).[9]

Our paper is closely related to Feenstra and Romalis (2014) who study quality variations in trade flows across different countries. Unlike ours, their approach requires strict parametric restrictions on the relationship between quality and income elasticity, on the production cost of quality, and on the distribution of product quality in order to infer quality measures. Our paper is also closely related to the recent paper by Redding and Weinstein (2024), who decompose the different margins of change in US imports, using a detailed nested CES structure that additionally accounts for firm heterogeneity. Relative to these studies, our contribution is to offer a novel identification strategy that only requires assumptions on the dynamics of demand shocks and, crucially, generalizes beyond CES demand to allow for heterogeneous elasticities.[10]

Our paper also contributes to the recent work on the importance of accounting for demand and taste shocks in cost-of-living indices (e.g., Gábor-Tóth and Vermeulen, 2018; Ueda et al., 2019; Redding and Weinstein, 2020a; Baqaee and Burstein, 2022).[11] In particular, using US retail scanner data where quality is arguably constant at the barcode-level, Redding and Weinstein (2020a) derive a formula for the price index under CES demand that accounts for additional variations in demand due to taste shocks. Our estimation strategy allows us to apply their approach to settings in which changes in demand partially reflect changes in product quality. We also show that the CES assumption may overstate the contribution of taste shocks to the indices of cost-of-living.

[8] Some studies use proxies for quality available for specific sets of products (e.g., wine as in Crozet et al., 2012) or indirect proxies such as the ISO 9000 management scores (e.g., Verhoogen, 2008).

[9] We emphasize that our import price index does not provide the full consumption-side welfare effects of rising imports, since the gains due to imports may partly be compensated by a substitution away from domestic consumption (see, e.g., Feenstra and Weinstein, 2017; Hsieh et al., 2020).

[10] In a recent study, Head and Mayer (2021) study counterfactual trade policy exercises in a models with CES and with mixed logit demand in the context of the original automobile market dataset of Berry et al. (1995). While they find similar results, they emphasize the importance of incorporating heterogeneity in pass-throughs through oligopolistic competition under the CES model.

[11] In addition to changes in taste, the dependence of demand on income (nonhomotheticity) also matters for the measurement of consumption gains. Here, we abstract from this consideration by focusing on homothetic demand. Jaravel and Lashkari (2021) provide a method for tackling this problem based on cross-sectional consumption data.

Finally, a growing body of work in trade and macro goes beyond the standard CES assumption and allows for variations in price elasticities through specifications such as Kimball and HSA demand to study variable markups and pass-through (e.g., Amiti et al., 2019, Baqaee and Farhi, 2020, Wang and Werning, 2020, Matsuyama and Ushchev, 2022).[12] Typically, prior work has inferred the parameters of these demand systems through calibration, by matching specific moments of interest in the data. To our knowledge, our paper is the first to identify the parameters of such demand systems using data on observed prices and market shares.[13]

Outline The paper is organized as follow. Section 2 presents the homothetic demand systems we consider, our approach to their identification, and our theoretical results on the change in their aggregate price index. Section 4 presents the results of our estimation approach in the benchmark setting of the US automobile market. Section 3 reports our empirical results from the trade data and quantifies the gains from quality. We conclude in Section 5.

2 Theory

2.1 Environment

We consider a setting in which we observe data on prices and expenditure shares (or quantities) in a set I of different products (goods or services) consumed in the aggregate, over a number of discrete time points $t \in \{0, \cdots, T-1\}$. Let $(\boldsymbol{s}_t)_{t=0}^{T-1}$ denote the sequence of expenditure shares, where $\boldsymbol{s} \equiv (s_i)_{i \in I}$ stands for the vector of expenditure shares chosen by the consumer(s). Similarly, let $(\boldsymbol{p}_t)_{t=0}^{T-1}$ denote the sequence of prices, where $\boldsymbol{p} \equiv (p_i)_{i \in I}$ stands for the vector of prices faced by the consumer(s) in the set of available products. We begin our analysis in Section 2.2, assuming that the set of availabe products remains constant throughout, but generalize our setting in Section 2.4.1 to allow for product entry and exit.

[12]For instance, allowing for variable markups, Feenstra and Weinstein (2017) and Edmond et al. (2015), among others, show that pro-competitive effects of trade liberalization are quantitatively relevant in the US and Taiwan, respectively. Since we use aggregate trade data, we cannot directly speak to this margin. However, when we apply our method at the firm-level, we can provide measures of markups based on our estimated price elasticities. In our application to the US auto market, we show that our estimated markups are in line with those found by Grieco et al. (2021) using BLP demand.

[13]For an alternative approach to the estimation of HSA demand, see Kasahara and Sugita (2021).

Assumptions and Definitions We assume a representative household choosing the quantities to consume at each point in time, given the observed price p_{it} for each product $i \in I$, as well as a vector of characteristics $\boldsymbol{x}_{it}$ that impacts their demand but remains unobserved in our data. We further assume that household preferences are characterized by a utility function $u = U(\boldsymbol{q}^{\varphi})$ where the vector $\boldsymbol{q}^{\varphi} \equiv (e^{\varphi_i} q_i)_{i \in I}$ denotes quality-adjusted quantities. The variable $\varphi_i \equiv \widetilde{\varphi}(\boldsymbol{x}_i)$ for each product i characterizes the quality of the product, defined as a function of product characteristics $\boldsymbol{x}_i$, and accounts for how quality shifts household demand.[14] Correspondingly, we define the vector of quality-adjusted prices $\boldsymbol{p}^{\varphi} \equiv (e^{-\varphi_i} p_i)_{i \in I}$ across products. Importantly, between each consecutive time period $t-1$ and t, we assume that there exists a set O_t of products in which the average product quality remains the same. Accordingly, let the vector $\boldsymbol{\varpi}_t \equiv (\varpi_{it})_{i \in I}$ with weights defined as $\varpi_{it} \equiv \frac{1}{|O_t|} \mathbb{I}\{i \in O_t\}$ characterize this set, such that $\sum_{i \in I} \varpi_{it} (\varphi_{it} - \varphi_{it-1}) = 0$.[15]

Based on the above assumptions, we can characterize household preferences using a Hicksian demand system $\boldsymbol{q}^{\varphi} = \widetilde{\boldsymbol{q}}(\boldsymbol{p}^{\varphi}; u)$ as a function of the vector of quality-adjusted prices and total expenditure y. Correspondingly, we define the vector of expenditure shares $\boldsymbol{s} = \widetilde{\boldsymbol{s}}(\boldsymbol{p}^{\varphi}; u)$ such that $s_{it} \equiv p_i q_i / \sum_{i'} p_{i'} q_{i'}$. Define the collection (Allen-Uzawa) cross-product elasticities of substitution between two distinct products i and j as

$$\sigma_{ij} \equiv \frac{1}{s_j} \frac{\partial \log \widetilde{q}_i}{\partial \log p_j}, \qquad i \neq j. \tag{1}$$

Since elasticities are symmetric ($\sigma_{ij} = \sigma_{ji}$), we can accordingly define a symmetric matrix $\boldsymbol{\Sigma} \equiv [\sigma_{ij}]$ with the additional definition

$$\sigma_{ii} \equiv \frac{\sum_{j \neq i} s_j \sigma_{ij}}{1 - s_i}. \tag{2}$$

As we will see, the matrix of cross-product elasticities $\boldsymbol{\Sigma}$ plays a key role in our approximations of the price index throughout the paper. In the special case of the CES demand system, the matrix of cross-product elasticities simplifies to a constant matrix $\boldsymbol{\Sigma} \equiv \sigma \, \mathbf{1}\mathbf{1}'$, where $\mathbf{1}$ denotes a unit vector.

[14] In the empirical settings considered here, each product encompasses a collection of similar goods and services, such as a product classification code in trade statistics, a commodity classification in macro data, or an automobile model with several potential trims (varieties). To the extent consumers value variety, changes in quality may, in part, result from variations in the set of finer varieties included within each product. We assume that this information is included in the vector of characteristics $\boldsymbol{x}_{it}$. We further differentiate our setting from those that consider products at the barcode level in retail scanner data, where product characteristics remain constant over time. In such cases, the demand shifter φ_{it} captures demand shocks driven by changes in product appeal (consumer taste) (e.g., Redding and Weinstein, 2020a).

[15] Note that we can easily generalize this assumption by considering the case in which the average quality change between the two periods over this set, $\sum_{i \in I} \varpi_{it} (\varphi_{it} - \varphi_{it-1}) = \Delta \overline{\varphi}_{ot}$ is some known value $\Delta \overline{\varphi}_{ot}$.

2.2 Price Index with Changing Quality under Income Invariance

We begin by characterizing the change in the price index between any two consecutive periods under the assumption of income invariance (homotheticity) in demand. In Section 2.2.1 below, we present our results for general income-independent preferences. As will be shown, these results require full knowledge of the matrix of cross-product elasticities of substitution, which may not be feasible in many macroeconomic settings. Therefore, in Section 2.2.2, we specialize our results to a family of income-invariant preferences that imposes simplifying constraints on the structure of this matrix, thereby facilitating their estimation using datasets commonly available in macroeconomic settings.

2.2.1 Characterziation for General Income-Invariant Preferences

Assuming income-invariant preferences, we can write the demand system $\widetilde{\boldsymbol{q}}\left(\boldsymbol{p}^{\varphi};u\right) = \underline{\widetilde{\boldsymbol{q}}}\left(\boldsymbol{p}^{\varphi}\right) Q\left(u\right)$, as the product of unit demand function $\underline{\widetilde{\boldsymbol{q}}}\left(\boldsymbol{p}^{\varphi}\right)$ and a monotonic function $Q(u)$ of utility (canonical cardinalization), and the expenditure share function $\widetilde{\boldsymbol{s}}\left(\boldsymbol{p}^{\varphi};u\right) = \widetilde{\boldsymbol{s}}\left(\boldsymbol{p}^{\varphi}\right)$, as a function independent of utility. We can also write the corresponding expenditure function as the product of a price index (unit expenditure) function $P\left(\boldsymbol{p}^{\varphi}\right)$ and the canonical cardinalization $Q(u)$ (Diewert, 1993). Moroever, in this context, the own-price elasticity is given by $\frac{\partial \log \widetilde{q}_i}{\partial \log p_i} = -\sigma_{ii}(1-s_i)$, and the matrix of cross-product elasticities $\boldsymbol{\Sigma}$ provides a complete local characterization of the demand system. This allows us to construct an approximation of the change in the price index between each two consecutive periods $t-1$ and t using this matrix and the changes in the observed expenditure shares across products, as presented in the following proposition.

Proposition 1. *(Approximate Price Index for Income-Invariant Preferences) Assume that the demand system is homothetic, satisfies the connected substitute property of Berry et al. (2013), the corresponding price index is continuously differentiable in prices, and all products remain available between periods periods $t-1$ and t. Then, the change in the log price index between the two periods can be approximated as*

$$\Delta \log P_t = \sum_i \varpi_{it} \Delta \log p_{it} + \sum_{i,j} \varpi_{it} \overline{\overline{\frac{1}{\sigma_{ii,t}-1}\,\Psi^{-1}_{ij,t}}}\,\Delta \log s_{jt} + O\left(\delta^3\right), \tag{3}$$

where $\Delta \log \boldsymbol{z}_t \equiv \left(\log z_{it} - \log z_{it-1}\right)$ and $\overline{\overline{\boldsymbol{z}_t}} \equiv \frac{1}{2}\left(\boldsymbol{z}_{t-1} + \boldsymbol{z}_t\right)$ denote the log difference and the mean of variable $\boldsymbol{z}_t$ between two consecutive periods, where $\boldsymbol{\varpi}_t$ is the vector of the weights corresponding to the base set O_t between these two periods, where we have defined the matrix $\boldsymbol{\Psi}_t$

through

$$\boldsymbol{\Psi}_t \equiv \boldsymbol{I} - \left(\boldsymbol{\Sigma}_t - \boldsymbol{\sigma}_t \boldsymbol{1}'\right) \, diag\left(\boldsymbol{s}_t\right) \left(\boldsymbol{\Sigma}_t^d - \boldsymbol{I}\right)^{-1}, \tag{4}$$

with $\boldsymbol{\sigma}_t$ and $\boldsymbol{\Sigma}_t^d$ denoting the vector of diagonal elements of $\boldsymbol{\Sigma}$ from Equation (2) and the corresponding diagonal matrix, respectively, and where $\boldsymbol{1}$ and $\boldsymbol{I}$ are the unit vector and matrix, respectively. The approximation error in Equation (3) is $\delta \equiv \max\left\{\max_{i\in O_t}\left\{|\Delta \log p_{it}|\right\}, \max_{i\in I}\left\{|\Delta \log s_{it}|\right\}\right\}$.

Proof. As the complete proof presented in Appendix B.1 (on page A18) shows, at any point $\tau \in (t, t-1)$ along smooth paths of quality adjusted prices $\boldsymbol{p}_\tau^\varphi$ between the two consecutive periods $t-1$ and t, the matrix $\boldsymbol{\Psi}_\tau$ satisfies $d\log \boldsymbol{s}_\tau = -\boldsymbol{\Psi}_\tau \left(\boldsymbol{\Sigma}_\tau^d - \boldsymbol{I}\right) \left(d \log \boldsymbol{p}_\tau - \Delta \boldsymbol{\varphi}_\tau\right)$. From Berry et al. (2013), we know that if the demand system satisfies the connected substitute property, the demand system - and thus also the matrix $\boldsymbol{\Psi}_\tau$ - is invertible. Applying the condition implied by the change in the quality of products in the base set, $\boldsymbol{\varpi}_t' \, d\boldsymbol{\varphi}_\tau = 0$, it follows that

$$d\log P_\tau = \sum_i \varpi_{it} d\log p_{i\tau} + \sum_{i,j} \varpi_{it} \frac{1}{\sigma_{ii,\tau} - 1} \Psi_{ij,\tau}^{-1} \, d\log s_{i\tau}. \tag{5}$$

The desired result follows as a second-order approximation of Equation (5). □

Let us consider CES preferences, where all cross-product elasticities of substitution are constant, $\sigma_{ij,t} \equiv \sigma$, and thus the matrix of cross-product elasticities is $\boldsymbol{\Sigma}_t \equiv \sigma \boldsymbol{1}\boldsymbol{1}'$. Substituting this in Equation (4), we obtain $\boldsymbol{\sigma}_t \equiv \sigma \boldsymbol{1}$, $\boldsymbol{\Sigma}_t^d = \sigma \boldsymbol{I}$, and $\boldsymbol{\Psi}_t = (\sigma - 1)\, \boldsymbol{I}$. In this case, the expression in Equation (5) simplifies to

$$d\log P_\tau = \sum_i \varpi_{it} d\log p_{i\tau} + \frac{1}{\sigma - 1} \sum_i \varpi_{it} \, d\log s_{i\tau},$$

which we can integrate to find an exact expression for the chage in the price index between times $t-1$ and t:

$$\Delta \log P_t = \sum_i \varpi_{it} \Delta \log p_{it} + \frac{1}{\sigma - 1} \sum_i \varpi_{it} \, \Delta \log s_{it}. \tag{6}$$

If we consider the base set to be the entire set of products $O_t \equiv I$ with constant weights $\varpi_{ij,t} \equiv \frac{1}{|I|}$, the above expression represents the logarithm of the CES unified price index (CUPI) defined by Redding and Weinstein (2020a) for the case with no product entry/exit. The first term represents the logarithm of the Jevons index, while the second term represents the logarithm of the geometric mean of the change in log expenditure shares. In Section 2.4.1 below, we will further generalize this result to account for the contributions of entry and exit.

Equation (3) highlights a key deviation from the CUPI when preferences feature heterogeneous cross-product elasticities of substitution. As the second term on the right hand side of Equation (3) shows, we must to adjust the weights in the weighted geometric mean of changes in the expenditure shares within the common set. This adjustment requires a weight proportional to $\frac{1}{2}\sum_i \varpi_{it}(\frac{1}{\sigma_{ii,t-1}-1}\Psi^{-1}_{ij,t-1} + \frac{1}{\sigma_{ii,t}-1}\Psi^{-1}_{ij,t})$ (instead of $\frac{1}{\sigma-1}\varpi_{jt}$) for each product $j \in I$.

The change in the price index can be decomposed into the contributions from changes in price and quality. From Shephard's lemma, along the smooth paths between $t-1$ and t, we have $d\log P_\tau = \sum_i s_{i\tau} d\log p_{i\tau} - \sum_i s_{i\tau} d\varphi_{i\tau}$. We can approximate the integral of the first term as the contribution of price changes $\Delta \log \Pi_t \equiv \sum_i \overline{\overline{s_{it}}}\,\Delta \log p_{it}$, leading to the following approximation for the contribution of quality:

$$\Delta \log \Phi_t \equiv \sum_i \left(\overline{\overline{s_{it}}} - \varpi_{it}\right)\,\Delta \log p_{it} - \sum_{ij} \varpi_{it} \overline{\overline{\frac{1}{\sigma_{ii,t}-1}\Psi^{-1}_{ij,t}}}\,\Delta \log s_{jt}. \tag{7}$$

The proof of the proposition further shows that we can approximate the change in the quality of product i between periods $t-1$ and t as

$$\Delta\varphi_{it} = \Delta \log p_{it} - \sum_i \varpi_{it}\,\Delta \log p_{it} + \sum_j \left(\overline{\overline{\frac{1}{\sigma_{ii,t}-1}\Psi^{-1}_{ij,t}}} - \sum_{i'} \varpi_{i't} \overline{\overline{\frac{1}{\sigma_{i'i',t}-1}\Psi^{-1}_{i'j,t}}} \right) \Delta \log s_{jt} + O\left(\delta^3\right). \tag{8}$$

To apply the above results to general income invariant demand systems, we need to specify the $\frac{1}{2}|I| \times (|I|-1)$-dimensional matrix of cross-product substitution elasticities $\mathbf{\Sigma}_t$ at each point in time. In most empirical applications involving trade or macro data, data limitations prevents us from estimating demand systems with sufficient richness to fully characterize these substitution patterns.[16] In Section 2.2.2 below, we specialize these results to a family of income invariant preferences that imposes specific restrictions on the matrix of cross-product substitution elasticities, facilitating estimation using commonly available trade or macro data. Nevertheless, these restrictions are much weaker than

[16]In our validation in Section 4, we consider information on a vector of product characteristics $\boldsymbol{x}_i$ for each product i, we can express rich patterns of cross-product elasticities of substitution in the space of product characteristics, whose dimensionality grows with the dimensionality of the characteristics space. For instance, the mixed logit demand system (McFadden, 1974; Berry, 1994) relies on product characteristics to define the expenditure-share functions as $\widetilde{s}_i(\boldsymbol{p}) \equiv \int \frac{\exp(-\alpha p_i + \boldsymbol{\beta}'\boldsymbol{x}_i)}{\sum_{i' \in I}\exp\left(-\alpha p_{i'} + \boldsymbol{\beta}'\boldsymbol{x}_{i'}\right)} dF(\alpha, \boldsymbol{\beta})$. Note that the mixed logit preferences are nonhomothetic.

those imposed by the conventional CES demand system, which assumes constant and identical patterns of cross-product substitution elasticities.

2.2.2 Preference Specifications

In this section, we introduce a family of income-invariant (homothetic) demand systems that generalizes the CES demand model in various dimensions and introduces some degree of heterogeneity in the matrix of cross-product elasticities of substitution.

Definition 1 (*Homothetic with Aggregator (HA) Demand*)**.** An income-invariant (homothetic) demand system belongs to this family if the expenditure share function can be written as

$$\widetilde{s}_i\left(\boldsymbol{p}^{\varphi}\right) \equiv \frac{\check{p}_i\, d_i\left(\check{p}_i\right)}{\sum_{i'} \check{p}_{i'}\, d_{i'}\left(\check{p}_{i'}\right)}, \tag{9}$$

where we have defined the normalized quality-adjusted price $\check{p}_i \equiv p_i^{\varphi}/\widetilde{h}\left(\boldsymbol{p}^{\varphi}\right)$ for a (homogenous of first degree) aggregator function $\widetilde{h}\left(\cdot\right)$, and where we have defined a collection of $|I|$ single-argument, positive-valued, and monotonic functions $d_i\left(p\right)$ decreasing over some interval $p \in \left(0, \underline{p}_i\right)$, and which satisfy $\lim_{p\to\underline{p}_i} d_i\left(p\right) = 0$ and $d_i\left(p\right) = 0$ for $p \geq \underline{p}_i$ where $\underline{p}_i \in \mathbb{R}_+ \cup \{\infty\}$. Two specific subfamilies of HA demand are as follows.

1. *Homothetic Implicit Additive (HIA).* This system is characterized by an aggregator function $h = \widetilde{h}\left(\boldsymbol{p}^{\varphi}\right)$, which is implicitly defined by the value of h that satisfies one of the two following conditions, depending on the type of HIA demand

$$1 = \begin{cases} \sum_{i\in I} \int_0^{d_i\left(p_i^{\varphi}/h\right)} d_i^{-1}\left(v\right) dv, & \text{directly additive (HDIA),} \\ \sum_{i\in I} \int_0^{p_i^{\varphi}/h} d_i\left(v\right) dv, & \text{indirectly additive (HIIA),} \end{cases} \tag{10}$$

   where each condition corresponds to one of the two types of HIA demand: directly or indirectly additive.

2. *Homothetic with a Single Aggregator (HSA).* This system is characterized by an aggregator function $h = \widetilde{h}\left(\boldsymbol{p}^{\varphi}\right)$, which is implicitly defined by the value of h that satisfies $1 = \sum_{i\in I} \frac{p_i^{\varphi}}{h} d_i\left(\frac{p_i^{\varphi}}{h}\right)$.

Definition 1 for the HA demand nests many well-known income-invariant demand systems commonly used in the literature, but does not ensure that they are rationalized by an underlying utility function. The rationalizability is ensured by the restrictions imposed by the choices of the HIA and HSA aggregator functions $h\left(\cdot\right)$, as shown by Matsuyama and Ushchev (2017) (see also Matsuyama, 2022). For instance, both the Kimball (e.g., Kimball, 1995b; Klenow and Willis, 2006) and the CRESH demand systems belong to the HDIA family. The HA family of demand systems departs from the CES model in two significant ways: 1) it introduces heterogeneity in the matrix of cross-product substitution elasticities, and 2) it goes beyond the highly restrictive independence of irrelevant alternatives (IIA) property of CES.

Assuming an HA demand system substantially simplifies the structure of the matrix of cross-product elasticities of substitution, as shown in the following lemma.

Lemma 1. *The matrix of cross-product elasticities of substitution for the HA demand systems introduced in Definition 1 satisfies*

$$\sigma_{ij} = \begin{cases} \frac{\varepsilon_i \varepsilon_j}{\bar{\varepsilon}}, & HDIA, \\ \varepsilon_i + \varepsilon_j - \bar{\varepsilon}, & HIIA, \qquad i \neq j, \\ 1 + \frac{(\varepsilon_i - 1)(\varepsilon_j - 1)}{\bar{\varepsilon} - 1}, & HSA, \end{cases} \tag{11}$$

where we have defined the elasticity of the corresponding demand function for product i as

$$\varepsilon_i \equiv \widetilde{e}_i \left(\frac{\boldsymbol{p}^{\varphi}}{\widetilde{h}\left(\boldsymbol{p}^{\varphi}\right)} \right) \equiv - \frac{\check{p}_i \, d_i' \left(\check{p}_i\right)}{d_i \left(\check{p}_i\right)} \Bigg|_{\check{p}_i = p_i^{\varphi} / \widetilde{h}(\boldsymbol{p}^{\varphi})}, \tag{12}$$

and where the expenditure-share-weighted mean of these elasticities are defined as $\bar{\varepsilon} \equiv \sum_i s_i \varepsilon_i$.

Proof. See Appendix B.1 on page A19. □

Lemma 1 shows that the $\frac{1}{2} |I| \times (|I| - 1)$ matrix of cross-product substitution elasticities for the HA demand systems only depends on $|I|$ variables, characterized by the vector of elasticities $\boldsymbol{\varepsilon} \equiv (\varepsilon_i)_{i \in I}$. Intuitively, these elasticities capture the degree of substitutability of each product with all other products, ruling out the possibility that particular pairs of products are more substitutable with one another relative to other pairs.[17] Neverthelss, as previously mentioned, the HA demand already allows for much richer patterns of heterogeneity in the matrix of cross-product substitution elasticities compared to the CES benchmark.

Lemma A.4 in Appendix A.5 shows that we can analytically invert the matrix $\boldsymbol{\Psi}_t$ for the family of HA preferences. This allows us to specialize the general expression for the approximate price index for income-independent preferences, as stated in Proposition 1, to this case. Proposition 2 below presents this result.

[17]In particular, for any combinations of four products $i \neq i' \neq j \neq j'$, we have the following constraints on the matrix of cross-product substitution elasticities: $\sigma_{ij} / \sigma_{ij'} = \sigma_{i'j} / \sigma_{i'j'}$ for HDIA, $\sigma_{ij} - \sigma_{ij'} = \sigma_{i'j} - \sigma_{i'j'}$ for HIIA, and $(\sigma_{ij} - 1)/(\sigma_{ij'} - 1) = (\sigma_{i'j} - 1)/(\sigma_{i'j'} - 1)$ for HSA.

Proposition 2. *(Approximate Price Index for H(S/I)A Preferences) For the three families of HA demand systems introduced in Definition 1, the change in the price index can be approximated as follows:*

$$\Delta \log P_t = \sum_i \varpi_{ti} \Delta \log p_{it} + \sum_i \varpi_{it} \overline{\overline{\tfrac{1}{\varepsilon_{it}-1}}} \Delta \log s_{it} - \sum_i \overline{\overline{s_{it} \bar{\iota}_t^o \left(\tfrac{1}{\varepsilon_{it}-1} - \overline{\left(\tfrac{1}{\varepsilon_{it}-1} \right)} \right)}} \Delta \log s_{it} + O\left(\delta^3\right), \tag{13}$$

where we have defined $\overline{\frac{1}{\varepsilon_{it}-1}} \equiv \sum_i s_{it} \frac{1}{\varepsilon_{it}-1}$, $\bar{\iota}_t^o \equiv \sum_i \varpi_{it}\, \iota_{it}$ *, and* $\bar{\iota}_{t-1}^o \equiv \sum_i \varpi_{it}\, \iota_{it-1}$*, with* $\iota_{it} \equiv \frac{1+1/(\varepsilon_{it}-1)}{\overline{1+1/(\varepsilon_{it}-1)}}$ *for HDIA,* $\iota_{it} \equiv \frac{1/(\varepsilon_{it}-1)}{\overline{1/(\varepsilon_{it}-1)}}$ *for HIIA, and* $\iota_{it} \equiv 1$ *for HSA, and where* δ *is defined as in Proposition 1.*

Proof. The result is a special case of Proposition A.4 in Appendix A.1.3 in the case of $I_t^* \equiv I$. □

Compared to the CUPI, the approximate price index for HA preferences shows several adjustments. Firstly, in the geometric mean (across the base set) of the changes in log expenditure shares, each product is weighted by its average love-of-variety index, $\overline{\overline{1/\left(\varepsilon_{it}-1\right)}}$, between the two periods. Moreover, the third term on the right hand side of Equation (13) accounts for the fact that a shift of expenditure shares within the common set of products toward those with higher love-of-variety indices $1/\left(\varepsilon_{it}-1\right)$ results in a lower price index for the consumer.

2.3 Dynamic Panel Estimation of Income-Invariant Demand

In this section, we present an approach for identifying a parametrized homothetic demand system, where data on expenditure shares and prices are assumed to follow Equation (9). Let us define the normalized quality-adjusted relative price of product i at time t as

$$\check{p}_{it} \equiv \frac{e^{-\varphi_{it}} p_{it}}{h_t}, \qquad h_t \equiv \widetilde{h}\left(\boldsymbol{p}_t^{\varphi}; \varsigma\right). \tag{14}$$

Note that the space of the quality-adjusted relative price vectors $\check{\boldsymbol{p}}_t$ at time t constitutes a $(|I|-1)$-dimensional manifold in $\mathbb{R}^{|I|}$, since all such vectors satisfy $\widetilde{h}\left(\check{\boldsymbol{p}}\right) = 1$. We now assume that the demand system satisfies the connected substitutes property as defined by Berry et al. (2013), making it a bijection from the space of quality-adjusted relative prices to the space of consumption expenditure shares. As a result, there exists an inverse demand function $\boldsymbol{\pi}\left(\cdot;\varsigma\right)$ such that $\check{p}_{it} = \pi_i\left(\boldsymbol{s}_t;\varsigma\right)$.

Using Equation (14) and the normalization of quality in the base product set, we can then write quality shocks as a function of observed expenditure shares and prices according to[18]

$$\varphi_{it} = \log \widehat{p}_{it} - \log \widehat{\pi}_i \left(\boldsymbol{s}_t; \varsigma\right), \qquad i \in V_t. \tag{15}$$

where we define the notation $\log \widehat{v}_{it} \equiv \log v_{it} - \sum_i \varpi_{it} \log \nu_{it}$ as the difference between the logarithm of variable v_{it} and its unweighted mean within the set of base products O_t.

Equation (15) presents a parametrized demand function that can be estimated in the data. Needless to say, the key challenge for the identification of this demand system is the potential correlation between the demand shock, log price, and expenditure shares. We now turn to our approach for tackling this problem.

2.3.1 Identification Assumptions

We begin by imposing the following restrictions on the stochastic dynamics of the quality shocks.

Assumption 1 (Dynamics of Demand Shocks)**.** *The following Markov process governs the dynamics of quality (demand) shocks φ_{it} for product i at time t:*

$$\varphi_{it} = g_i \left(\varphi_{it-1}; \boldsymbol{\varrho}\right) + u_{it}, \tag{16}$$

where u_{it} is a zero-mean i.i.d innovation to the demand shock, and $\boldsymbol{\varrho}$ is a vector of parameters characterizing the persistence of the demand shock process.[19]

Equation (16) implies that, despite potential persistence in the process of quality shocks, these shocks cannot be entirely predicted based on past realizations due to the arrival of innovations in each period. In our baseline model, we assume that the demand shock process is a stationary AR(1) process with a product-specific mean:[20]

$$g_i \left(\varphi_{it-1}; \boldsymbol{\varrho}\right) \equiv \rho \varphi_{it-1} + (1-\rho)\, \phi_i, \tag{17}$$

[18] By definition, we have $\log p_{it} - \log h_t - \varphi_{it} = \log \pi_i \left(\boldsymbol{s}_t; \varsigma\right)$. Using the condition $\sum_i \varpi_{it} \varphi_{it} = 0$ then leads to Equation (15).

[19] Note that we can generalize this condition to higher order Markov dynamics, for instance, assuming $\varphi_{it} = g_i \left(\varphi_{it-1}, \varphi_{it-2}, \cdots; \boldsymbol{\varrho}\right) + u_{it}$, where the contemporaneous demand shock further depends on its higher-order lags.

[20] This model can also account for a process with stationary growth, e.g., a model with $g_i \left(\varphi_{it-1}\right) \equiv \varphi_{it-1} + \gamma_i$, such that $\gamma_i \equiv \lim_{\rho \to 1} (1-\rho)\, \phi_i$.

where $\boldsymbol{\varrho} \equiv (\rho, \phi)$ is the vector of the parameters of the Markov process, and ϕ_i constitutes the expected long-run mean quality of product i.

Next, we introduce our main identification assumption, which excludes the dependence of past decisions by firms and consumers on the current innovation to the demand shock.

Assumption 2 (Identification Assumptions)**.** *Demand shock innovations are zero mean, conditional on lagged log prices (and potentially the latter's powers):*

$$\mathbb{E}\left[u_{it} \middle| \left(\log p_{it-1}\right)^m\right] = 0, \qquad 1 \leq m \leq D, \tag{18}$$

where $D \geq 1$ denotes the dimensionality of the parameters characterizing consumer demand. Moreover, we assume that the log price process has a nonzero autocorrelation, i.e., $\mathbb{E}\left[\log p_{it-1} \log p_{it}\right] \neq 0$.

In combination with Equations (15) and (16), we can use Equation (18) to derive a set of orthogonality conditions that allow us to estimate the vectors of parameters $\boldsymbol{\varsigma}$ and $\boldsymbol{\varrho}$. This leads to the following moment conditions

$$\mathbb{E}\left[\left(\log \widehat{p}_{i,t} - \log \widehat{\pi}_i\left(\boldsymbol{s}_t; \boldsymbol{\varsigma}\right) - g_i\left(\log \widehat{p}_{i,t-1} - \log \widehat{\pi}_i\left(\boldsymbol{s}_{t-1}; \boldsymbol{\varsigma}\right); \boldsymbol{\varrho}\right)\right) \times z_{it-1}\right] = 0, \tag{19}$$

where z_{it} is an instrument that is orthogonal to the value of the quality innovation u_{it} for product i at time t, as defined by the expression within the main parentheses. The instruments z_{it} include lagged values of different powers of log prices $\left(\log p_{it-1}\right)^m$ for $m \leq D$, a combination of lagged value of the quality shock φ_{it-1} (and potentially its powers) given by Equation (15), and product dummies, depending on the structure of the process $g_i\left(\cdot; \boldsymbol{\varrho}\right)$. For instance, in the AR(1) process considered in Equation (17), we use the lagged quality shocks φ_{it-1} and product dummies to identify ρ and ϕ_i's. The assumption of nonzero autocorrelation ensures that the lagged values of log prices provide meaningful instruments for the corresponding contemporaneous values of the same variables.

Example: CES Demand As an example, let us consider the case of CES demand where, as already mentioned, we have $\widetilde{s}_i\left(\check{\boldsymbol{p}}; \boldsymbol{\varsigma}\right) \equiv \check{p}_i^{1-\sigma}$, $\widetilde{h}\left(\boldsymbol{p}; \boldsymbol{\varsigma}\right) \equiv \sum_i p_i^{1-\sigma}$, and where $\boldsymbol{\varsigma} \equiv (\sigma)$. In this case, the inverse demand function can be analytically written as $\pi_i\left(\boldsymbol{s}; \sigma\right) \equiv s_i^{1/(1-\sigma)}$. According to Equation (15), we can write the quality shock as $\varphi_{it} = \log \widehat{p}_{it} + \frac{1}{\sigma-1} \log \widehat{s}_{it}$. Since a single parameter σ fully characterizes demand, we only need to use the case of

$D = 1$ in Equation (18), and thus use the orthogonality conditions $\mathbb{E}\left[u_{it} | \log p_{it-1}\right] = 0$, $\mathbb{E}\left[u_{it} | \varphi_{it-1}\right] = 0$, $\mathbb{E}\left[u_{it} | \varphi_{it-1}\right] = 0$, and $\mathbb{E}\left[u_{it}\right] = 0$ for each product i and each time t.

If we further consider the AR(1) assumption in Equation (17), we can leverage the log-linearity of the model to express the moment conditions in first-differences as

$$\mathbb{E}\left[\left(\Delta \log \widehat{p}_{it} + \tfrac{1}{\sigma-1} \log \widehat{s}_{it} - \rho\left(\Delta \log \widehat{p}_{it-1} + \tfrac{1}{\sigma-1} \log \widehat{s}_{it-1}\right)\right) \times z_{it}\right] = 0, \tag{20}$$

where $\Delta \log v_{it} \equiv \log v_{it} - \log v_{it-1}$ for any variable v_{it}, and the instruments z_{it} include *double* lagged log prices $\log p_{i,t-2}$ (case of $D = 1$ in Equation 18) and demand shocks $\varphi_{i,t-2}$. In this case, the two instruments allow us two identify the two parameters, demand elasticity parameter σ and the demand shock persistence ρ, and we do not need to estimate the long-run mean of product-level demand shocks ϕ in Equation (17).

2.3.2 Discussion

The Logic of Identification To gain more intuition about the assumption in Equation (18), we present an explicit model of firm price setting that satisfies this assumption.Consider the standard environment in which firms set prices flexibly to maximize contemporaneous profits. In this scenario, the price at a given point in time should depend solely on the current variables, and should not depend on the firm's information or forecasts regarding future product demand and quality. More specifically, let q_{it} denote the quantity of product i purchased by consumers. This scenario leads to the following process for the evolution of log prices:

$$\log p_{it} = \log mc_i\left(q_{it}, \varphi_{it}, w_{it}\right) + \log \mu_i\left(\boldsymbol{p}_t, \boldsymbol{s}_t, \boldsymbol{\varphi}_t\right) + v_{it}, \tag{21}$$

where $mc_i(\cdot, \cdot, \cdot)$ is the marginal cost function, which may depend on quantity q_{it}, quality φ_{it}, and exogenous cost shifters w_{it}; $\mu_i(\cdot, \cdot, \cdot)$ is the markup function, which may depend on the vector of current prices $\boldsymbol{p}_t$, market shares $\boldsymbol{s}_t$, and demand shocks $\boldsymbol{\varphi}_t$ of all products in the market; and v_{it} is the residual *i.i.d.* error, which accounts for measurement or firm pricing error. The price setting Equation (21) satisfies Equation (18) even if the firm knows its future demand shock innovation.[21]

More generally, we may consider a model of dynamic price setting where the log price

[21] Note that under the assumption of flexible pricing, our identification assumption is weaker compared to the typical assumptions in the application of the dynamic panel methods to production function estimation (see Ackerberg, 2016). In particular, we do not require the assumption that the innovation u_{it} does not belong to the information set of the firm at time $t-1$. With flexible pricing, even if the firm knows its future demand shock, it does not have an incentive to reflect that in its current pricing decision.

additionally depends on the expected value of future cost and demand shocks, as well as those of the competitors, conditional on the information set $\mathcal{I}_{it}$ of the firm at that moment in time. In this case, it is sufficient to assume that the firm does not have knowledge of the future demand shock innovation $u_{it} \notin \mathcal{I}_{it}$ to satisfy the assumption in Equation (18). Regardless of the underlying price-setting model, the orthogonality assumption allows us to rule out a *direct* functional dependence of the price p_{it} on the future demand shocks φ_{it+1}. Thus, all systematic correlations between log price and the future demand shocks φ_{it+1} are driven by the persistence of the demand shock process φ_{it}.

Deviating from the settings discussed above, our identification assumption may be violated in several different ways. First, as mentioned above, when the prices are sticky and firms know their future innovations to their demand shocks. Second, when prices or market shares are measured with autocrrelated error. Third, when demand shock for each product dynamically varies as a function of past consumption of that product. Fourth, when quality evolves endogenously as the result of investments in R&D. In the last case, our identification could be violated if contemporaneous costs to production simultaneously affect innovation costs.

Comparison with Alternative Approaches to Identification Despite potential limitations to our identification assumption, we believe our approach offers an improvement to the state-of-the-art demand estimation approaches in trade and macroeconomics. In these settings, we typically only have access to information on prices and quantities. This prevents us from applying the standard identification approaches that use exogenous cost shifters w_{it}, which affect prices through marginal cost as in Equation (21), as instruments to estimate Equation (15). Our identification assumption allows us to use the lagged values of log price as an instrument for current log price, after controlling for the expectation of the demand shock conditional on lagged prices.[22]

In the absence of external cost shock instruments, the conventional approach to estimating demand is that of Feenstra (1994) (based on the earlier work by Leamer, 1981). This approach rules out correlations between demand shocks φ_{it} and any shocks to prices that are not driven by quantity changes. In particular, this assumption is violated if there is any dependence of the marginal cost on quality in Equation (21), i.e., $\frac{\partial \log mc}{\partial \varphi} \neq 0$. Intuitively, improvements in quality are often associated with more costly inputs, suggesting that this assumption is likely violated in practice. Section A.6 in Appendix B provides a

[22] It is important to note that most cost shock instruments used in practice impact the price or costs of specific inputs. If firms adjust their input usage in response to these shocks, such substitution may additionally affect product quality, thereby violating the exogeneity of some cost shock instruments.

detailed discussion of how our assumptions on the dynamics of demand shocks allow us to estimate demand without relying on Feenstra (1994)'s identification assumption.

2.4 Extensions

2.4.1 Product Entry and Exit

In this section, we generalize the results in Section 2.2 to account for the scenario where the set $I_t \subset I$ of products *available* at time t varies over time. We define the set of continuing products between two consecutive periods as $I_t^* \equiv I_t \cap I_{t-1}$. We now let $\boldsymbol{p}_t$ and $\boldsymbol{s}_t$ denote the vectors of prices and expenditure shares over the extended set of all products I, letting prices and expenditure shares outside the set of available products to infinity and zero, respectively. Proposition 1 below generalizes Proposition 1 to accomodate product entry and exit.

Proposition 3. *(Approximate Price Index with Product Entry and Exit) Assume that the demand system is income-invariant, satisfies the connected substitute property of Berry et al. (2013), the corresponding expenditure function is twice continuously differentiable in prices, that Ψ_{ij}^{-1}/s_j remains everywhere bounded for all $i \neq j$, and that all products within the base set continue from period $t-1$ to period t, that is $O_t \subset I_t^*$. Then, for any income-invarient demand system, the change in the log price index between periods $t-1$ and t can be approximated as*

$$\Delta \log P_t = \sum_i \omega_{it} \Delta \log p_{it} + \sum_{i,j \in I_t^*} \omega_{it} \overline{\overline{\tfrac{1}{\sigma_{ii,t}-1} \Psi_{ij,t}^{-1}}} \Delta \log s_{jt}^* + \left(\sum_{i,j \in I_t^*} \omega_{it} \overline{\overline{\tfrac{1}{\sigma_{ii,t}-1} \Psi_{ij,t}^{-1}}} \right) \Delta \log \Lambda_t^*$$
$$+ \sum_{i \in I_t^*} \omega_{it} \left(\sum_{j \in I_t \setminus V_t^*} \tfrac{1}{\sigma_{ii,t}-1} \Psi_{ij,t}^{-1} - \sum_{j \in I_{t-1} \setminus V_t^*} \tfrac{1}{\sigma_{ii,t-1}-1} \Psi_{ij,t-1}^{-1} \right) + O\left(\delta^3\right), \quad (22)$$

where $\Lambda_t^* \equiv \sum_{i \in I_t^*} s_{it}$ is the expenditure share of the continuing products, $\boldsymbol{s}_t^* \equiv \boldsymbol{s}_t / \Lambda_t^*$ is the vector of expenditure shares within the continuing set, and the error is givven by $\delta \equiv \max \left\{ \max_{i \in O_t} \{ |\Delta \log p_{it}| \}, \max_{i \in I_t^*} \{ |\Delta \log s_{it}^*| \}, |\Delta \log \Lambda_t^*|, \max_{i \notin V_t^*} \left\{ |\Delta s_{it}|^{\frac{2}{3}} \right\} \right\}$.

Proof. See Appendix B.1 on page A21. □

Appendix A.1.2 also shows that the above result is exact in the CES case and leads to

$$\Delta \log P_t = \sum_i \omega_{it} \Delta \log p_{it} + \frac{1}{\sigma - 1} \sum_i \omega_{it} \Delta \log s_{it}^* + \frac{1}{\sigma - 1} \Delta \log \Lambda_t^*. \quad (23)$$

The above expression is the logarithm of the CES unified price index (CUPI) as defined by Redding and Weinstein (2020a), assuming that the set of base products corresponds to the set of continuing set, $O \equiv I_t^*$ with constant weights $\varpi_{i,t} \equiv \frac{1}{|I_t^*|}$. The first two terms remain the same as before, but the last term now accounts for the Feenstra (1994) CES correction for the contributions of product entry and exit.

For general income-invariant preferences, Equation (22) shows two additional deviations from the CUPI compared to that in Equation (3). First, as shown in the third term, we replace the CES love-of-variety index $\frac{1}{\sigma-1}$ with the mean love-of-variety index $\frac{1}{2}\sum_{ij} \varpi_{it} \left(\frac{1}{\sigma_{ij,t-1}-1}\Psi_{ij,t-1}^{-1} + \frac{1}{\sigma_{ij,t}-1}\Psi_{ij,t}^{-1}\right)$ between the two periods in accounting for the contribution of the change in the share of the continuing set Λ_t^*. Second, as indicated by the last term on the right-hand side, we additionally have to account for the potential mismatch between consumer's love-of-variety for exiting versus entering products. Specifically, the price index rises if the latter exceeds the former.

Appendix A.1 offers a step-by-step construction of the proof of the above proposition. In addition, Proposition A.1 in Appendix A.1.3 generalizes the decomposition of the change in price index provided in Equation (7) to the case involving product entry and exit, characterizing the contributions of changes in unit prices, quality, and product entry and exit. Furthermore, we offer an approximation of the change in the quality of continuing products, generalizing Equation (8). Finally, Proposition A.4 in the same appendix extends Proposition 2 to the case involving product entry and exit by considering the family of HA preferences in Definition 1.

2.4.2 Income Dependence (Nonhomotheticity)

In this section, we further generalize Proposition 3 to construction an expression for a *Universal Price Index (UPI)* that approximates changes in the cost of living for general preferences that may feature income dependence (nonhomotheticity). To achieve this, we extend the characterization of the demand system by defining the vector of income elasticities $\boldsymbol{\eta}_t \equiv (\eta_{it})$ where $\eta_{it} \equiv \partial \log \widetilde{q}_i^{uc}(\boldsymbol{p}_t; y_t) / d \log y_t$ and $\widetilde{q}_i^{uc}(\boldsymbol{p}; y)$ stands for the uncompensated (Marshallian) demand corresponding to the underlying preferences.

Proposition 4. *(Universal Price Index) Assume that the demand system satisfies the connected substitute property of Berry et al. (2013), the corresponding expenditure function is twice continuously differentiable in prices, that Ψ_{ij}^{-1}/s_j remains everywhere bounded for all $i \neq j$, and that all products within the base set continue from period $t-1$ to period t, that is $O_t \subset I_t^*$. Then, the quality-adjusted Divisia price index, the integral of the (expenditure-weighted) mean change in* *log quality-adjusted prices between the two consecutive periods,*[23] *can be approximated as*

$$\log P_t \equiv \int_{t-1}^{t} s_{i\tau} d\log p_{i\tau}^{\varphi} = \frac{\Delta \log P_t^{II} - \kappa_t \Delta \log y_t}{1-\kappa_t} + O\left(\delta^3\right), \tag{24}$$

where $\kappa_t \equiv \overline{\overline{\sum_{i,j} \varpi_{it} \frac{1}{\sigma_{ij,t}-1} \Psi_{ij,t}^{-1} \left(\eta_{jt}-1\right)}}$ *and* $\Delta \log P_t^{II}$ *is given by the expression for the income-invariant demand in Equation (22).*

Proof. See Appendix B.1 on page A22. □

Recent work has shown that in the presence of income-dependence, we cannot chain the conventional indices of the change in the cost of living, such as those constructed above, due to variations in the benchmark utility u_t^* over time (e.g., Baqaee and Burstein, 2022; Jaravel and Lashkari, 2024). In Appendix A.2 below, we show that our results easily generalize to the approach proposed by Jaravel and Lashkari (2024), which adjusts for the contributions of income dependence to the aggregation of prices.[24]

2.4.3 Category-Specific Price Indices

Let us now consider an arbitrary partitioning of the space of products into disjoint categories I^k, where each product in I belongs to some category $i \in I^k$ for $k \in \mathcal{K}$. For instance, we may consider two disjoint sets I^M and I^D, representing the sets of imported and domestic products, respectively. Alternatively, we may consider the disjoint sets of different sectors or industries K in the economy. The only restriction we impose on these categories is that at least some products from each category should be available to and selected by the consumer at all times, that is, $I^k \bigcap I_t \neq \emptyset$ for all k and all $0 \le t \le T-1$.

Assuming again income invariance, the price index between the two consecutive periods can be approximated as a Törqvist index of the changes in category-specific price indices as the following

$$\Delta \log P_t = \sum_{k \in K} \overline{\overline{s_t^k}} \Delta \log P_t^k + O\left(\delta^3\right), \tag{25}$$

where the category-specific price index is defined as $\Delta \log P_t^k \equiv \sum_{i \in I^k} \int_{t-1}^{t} s_{i\tau}/s_\tau \left(d\log p_{i\tau} - d\varphi_{it}\right)$

[23] As shown, among others, by Jaravel and Lashkari (2024), there exists a level of utility u_t^* between u_{t-1} and u_t such that the change in the cost of living for that level of utility corresponds to the integral of the Divisia index, that is, $P_t \equiv E\left(u_t^*; \boldsymbol{p}_t^{\varphi}\right) / E\left(u_t^*; \boldsymbol{p}_{t-1}^{\varphi}\right) + O\left(\delta^3\right)$.

[24] In practice, their approach requires access to the cross-sectional data on the composition of consumption across households, which is unfortunately not available in the case of our leading application to the measurement of the import price index presented in Section 3 below.

and $\delta \equiv \max_k \left\{\Delta \log P_t^k\right\}$ (see Appendix A.3.1). Crucially, the decomposition in Equation (25) does *not* assume that the the preferences are separable into category-specific nests. The decomposition holds under any arbitrary patterns of substitutability that products across different groups may have. Therefore, without loss of generality, we can, for instance, decompose the changes in the consumer price index into a domestic and an import price index, or into price indices corresponding to different sectors/industries, assuming we have estimates of quality change at the product level.

In the absence of observations on quality change, if we want to use the results of Section 2.2 to approximate the category-specific price indices using only observations of expenditure shares and unit prices, we need to know the entire matrix of elasticities of substitution across products in all product categories. To further simplify the structure of the matrix of cross-product elasticities of substitution, we can make the additional assumption that preferences are separable across different categories. Under this assumption, Proposition A.2 in Appendix A.3 shows that we can recover an approximation of the change in cateogry-specific price index $\Delta \log P_t^k$ by applying Proposition 3 within each category. Thus, in this case, we only need to know the within-category matrix of cross-product elasticities of substitution within each category to approximate the change in the aggregate price index.

2.4.4 Production-Based Price Aggregation

So far, we have approached the aggregation of prices from a consumption-based perspective. Alternatively, we can approach the aggregation of prices from a production-based approach, where we are interested in aggregating prices in order to construct production-based measures of real aggregate income (e.g., Diewert and Morrison, 1986). Since in Section 3 below we apply our approach to the construction of the import price indices, which covers a wide range of intermediate products, it is desirable to understand the extent to which our approach remains valid under this alternative approach. Appendix A.4 first revisits the production-based approach to the construction of the import price index and shows its role in the measurement of real GDP and real consumption growth. Moreover, it characterizes the assumptions needed under this approach for our income-independent (homothetic) specification of import demand to remain valid.

3 Application: the Price Index of US Imports

We now turn to evaluating the impact of the changes in the size, content, and composition of US imports for the welfare of consumers in the United States from 1989 to 2018, as captured by the price index of US import. First, we briefly outline a model of consumer demand for imports and define the corresponding price index, building on the results of Section 2.2. We then present the results of estimating the US import demand with the DP approach and discuss the resulting measures of the change in the price index of US import.

3.1 Aggregation and The Import Price Index

We assume an income-independent (homothetic) aggregator across all products, as defined in Section 2.1. Moreover, we consider a partitioning of these products across a number of separable categories of products following the discussions in Section 2.4.3, where each category corresponds to a given industry (see Appendix A.3 for more details). Within each industry, we identify varieties with the country of origin (Armington assumption), whether sourced domestically (from the US) or from various trade partners exporting their products to the US within that industry.

Within each industry, we consider a parameterized aggregator across these different varieties that belongs to the HDIA family, as defined by Equations (9) and (10). Specifically, we consider Kimball aggregators Q_t^k in each industry k, which aggregate the vector $\boldsymbol{q}_t^{\varphi,k}$ of quality-adjusted quantities for product varieties within that industry (Kimball, 1995a). This aggregator is implicitly defined according to the constraint

$$\sum_{i\in I^k} K\left(\frac{q_{it}^{\varphi}}{Q_t^k};\, \varsigma_k\right) = K\left(1;\, \varsigma_k\right), \tag{26}$$

where $K(\cdot;\varsigma_k)$ is the Kimball function, parameterized by a vector ς_k, which satisfies $K\left(\check{q}\right) \equiv \int_0^{\check{q}} d^{-1}\left(v;\varsigma\right) dv$ for the corresponding demand function $d\left(\cdot;\varsigma\right) \equiv d_i\left(\cdot\right)$ used in Equations Equations (9) and (10). We consider a number of different parameterizations of the Kimball function, characterized by the *Kimball elasticity* functions:

$$\widetilde{e}\left(\check{p};\varsigma\right) \equiv -\left.\frac{K'\left(\check{q};\varsigma\right)}{\check{q}\,K''\left(\check{q};\varsigma\right)}\right|_{\check{q}=d(\check{p})}, \tag{27}$$

where we have used the definition of the demand elasticity function $\widetilde{e}\left(\cdot\right)$ in Equation (12), along with the demand relations $K'(\check{q};\varsigma) = d^{-1}(\check{q};\varsigma)$ and $\check{q} = d\left(\check{p};\varsigma\right)$. Given our assumptions on the Kimball function $K(\cdot;\varsigma)$, the elasticity function $\widetilde{e}(\check{p})$ is positive-valued for all $\check{p} < \underline{\check{p}}$.[25] Appendix A.5.3 provides the different families considered, which include CES (as a benchmark with an isoelastic Kimball function) and three alternative families: Klenow and Willis (2006), Finite-Infinite Limit FIL), and Finite-Finite Limit (FFL). These vary depending on whether the elasticity function remains finite as relative quality-adjusted prices approach zero and/or infinity. Throughout, our baseline specification will be the FFL case, in which the Kimball elasticity converges to two finite constants in the limits as the quality-adjusted quantity approaches zero or inifinity.

To compute the aggregate import price index from Equation (25), we need to compute the change $\Delta \log P_t^k$ in the logarithm of the price index for each industry k, relying on the results of Section 2.2. As discussed below, we first estimate the Kimball demand system separately for each industry across all varieties, including the domestic ones, using the technique presented in Section 2.3. We then rely on the results of Section 2.4.3, which imply that the price index for each industry can be approximated to the second-order by the convex combination of the price indices of two product categories: domestic and imported, following $\Delta \log P_t^k \approx \overline{\overline{s_t^{k,D}}} \Delta \log P_t^{k,D} + \overline{\overline{s_t^{k,M}}} \Delta \log P_t^{k,M}$. As we will see, we can directly observe the price index of domestic goods $\Delta \log P_t^{k,D}$ in the data, allowing us to find the industry-specific import price index as

$$\Delta \log P_t^{k,M} \approx \frac{1}{\overline{\overline{s_t^{k,M}}}} \left(\Delta \log P_t^k - \overline{\overline{s_t^{k,D}}} \Delta \log P_t^{k,D} \right). \tag{28}$$

We then aggregate the industry-specific import price indices using Tornqvist weights.

3.2 Data and Estimation

Customs records provide information about the composition of US imports at a fairly detailed (10-digit) level of disaggregation. Unfortunately, information on domestic consumption is not available at the same level of detail. To solve this problem, following Broda and Weinstein (2006), most of the prior work assumes that domestic and imported components of consumption are separable and focus on characterizing the demand solely among the imported goods. However, in so far as we are interested in accounting for patterns substitutability to infer quaility change, we may not wish to abstract away from the potentially direct substitutability between domestic and imported varieties within each

[25]We may consider additional constraints that imply this function is also nonincreasing and is smaller than unity, implying price elasticities of demand that exceed unity and are nondecreasing in quantity (satisfying Marshall's Second Law of Demand).

product/industry. To account for such substitutability patterns, we use the available information on the composition of the US domestic consumption to construct a dataset that contains information on both domestic and imported sales in the US at a more aggregate level.

We measure domestic sales as the gap between domestic production and net exports. We first measure domestic production and prices using data from the NBER-CES Manufacturing Industry Database and the BEA's "Gross Output by Industry" tables, covering the period from 1989 to 2018. The NBER-CES Manufacturing Industry Database provides annual data on shipments (vship, total value of shipments) and prices (piship, sectoral price indices) for U.S. manufacturing sectors, classified at the 6-digit level of the North American Industrial Classification System (NAICS) (Becker et al., 2021). We aggregate this data to the 5-digit NAICS level using a standard Tornqvist formula to align it with the level of aggregation used in trade data. For non-manufacturing sectors such as farming and mining, we use gross output and corresponding price indices from the BEA.[26]

We then integrate the domestic production data with U.S. Census Bureau import and export data from 1989 to 2018 to calculate the domestic sales. The U.S. Census Bureau provides detailed information on trade flows by HS-country-year. A key challenge in combining the production and trade data is ensuring a time-consistent mapping between industry and product codes. To address this, we use the latest version of the Pierce and Schott (2012) algorithm to create time-consistent 10-digit HS codes. Following Amiti and Heise (2021), we map these HS codes to the 5-digit NAICS 2012 classification.

Using this concordance, we calculate the domestic sales of U.S. products for each industry by subtracting exports from total shipments. In cases where this results in negative values, we assume that all domestic consumption is imported and assign a value of zero to domestic sales of U.S. products. To measure domestic absorption, we subtract exports and add imports to total shipments.[27]

The final dataset covers 155 time-consistent industries across manufacturing, farming, and mining sectors from 1989 to 2018. We identify product varieties by country of origin within each 5-digit NAICS sector. This allows us to construct data on the set of imported varieties for each sector using U.S. Census Bureau import data. A detailed description of the construction of the variables used in this section, along with robustness

[26]Data on farming and mining sectors from BEA are reported at the "Summary" level, which is more aggregate relative the 5-digit NAICS level.

[27]We adjust our measure of domestic sales of U.S. firms for the presence of re-exports in order not to underestimate it. We measure U.S. exports using Domestic Exports as defined by the U.S. Census Bureau (Total Export minus Foreign Exports). Domestic absorption is not affect by this adjustment because re-exports also enters in imports, thus, not affecting the net exports. Independently of how it is constructed, domestic absorption is always positive, as expected.

Table 1: CES and Kimball Elasticities

	Kimball - DP	CES - DP	CES - FBW
Mean	17.0	6.31	3.40
	(0.120)	(0.009)	(0.005)
Median	4.70	4.27	2.58
	(0.023)	(0.072)	(0.057)
Weighted Mean	3.11	5.99	4.62
	(0.005)	(0.009)	(0.008)
5th percentile	1.82	1.72	1.70
25th percentile	3.07	2.93	2.11
75th percentile	9.16	7.44	3.64
95th percentile	48.5	18.3	6.54

Note: The table reports the mean, median, the expenditure-weighted average, and the 5th, 25th, 75th and 95th percentiles of the distribution of the elasticity of the demand function for both the Kimball and CES specifications. For the Kimball specification, we can compute the elasticity for each variety at each moment in time while, in the CES case, each variety-time pair is associated with the corresponding sectoral CES elasticity. For the CES case, we report the DP and the BW estimates. Standard errors are bootstrapped.

tests comparing our data to the more aggregate industry definitions from the BEA annual Input-Output tables, is provided in Appendix D.1.

We first estimate the CES elasticity of substitution across product varieties in each industry. We use our Dynamic Panel (DP) approach using the moment condition in Equation (20), and compare our estimates against those found using the conventional Feenstra (1994) and Broda and Weinstein (2006) estimator (henceforth FBW). We next apply the DP approach to the Finite-Finite Limit (FFL) specification of the Kimball preferences. We use the moment conditions in Equations (19) and (20) with lagged log prices and quantities, their quadratic and third power as instruments. For expressing the quality of the imported varieties and the computing the price index, we use the U.S. variety as basesline product.[28]

3.3 Estimates of the Elasticitiy of Substitution

Table 1 compares the elasticity of the HA demand function ε_i defined in Equation (12), as estimated by the two models (Kimball vs. CES) and the two different strategies (DP

[28] For the purpose of estimation, in case FBW fails to converge, we use any continuously imported variety over the period from 1989 to 2018 within each industry as the normalizing product. In practice, this restricts the possibility to the major advanced economies and few other exporters. In case DP fails to converge or the estimated values were not feasible, we use the moment conditions in Equations (19) and (20) with lagged log prices and quantities and their quadratic power as alternative set of instruments. When a different variety is used as normalizing product, quality is still expressed relative to the U.S. variety. In the few cases the domestic sales of U.S. products become negative and the U.S. variety disappears, we proceed as follow: we first use Canada as baseline product; we then estimate a trend in the U.S. quality; we then normalize quality using the U.S. as baseline product and project its quality for those years in which it does not appear.

vs. FBW) across different industries.[29] First, note that comparing the magnitudes across different identification methods for the CES case (in which $\varepsilon_i \equiv \sigma$ is the constant elasticity parameter), we find that the elasticities estimated using DP are larger compared to those obtained using the FBW method across all the moments considered. As discussed in Section 2.3.2, the FBW method assumes uncorrelated demand and supply shocks–an assumption that is likely to be violated when marginal cost depends on quality. The resulting positive correlation between demand and supply shocks should lead to a downward bias in the elasticities estimated by the conventional method, consistent with the results in Table 1.[30]

We now turn our attention to the estimated elasticities for the Kimball model and compare them to the corresponding CES estimates.[31] Table 1 compares different moments of the distribution of elasticities across varieties between Kimball and CES estimates.[32] We find that the estimates under the Kimball demand system are larger in terms of the mean, median, and all other moments of the distribution. This result suggests that ignoring the heterogeneity in elasticities across varieties leads to a bias in the estimated demand elasticity at the variety level. Figure 1 orders all industries from left to right based on the share-weighted mean elasticity under Kimball, reporting the estimated lower and upper limits of the Kimball specification, the expenditure share weighted Kimball elasticitiy, and the estimated CES elasticity for each industry. The solid red line shows a strong positive correlation between the expenditure-share weighted mean Kimball elasticity and the corresponding CES elasticity. However, the estimated lower and upper limits of the Finite-Finite specification indicate extensive heterogeneity in price elasticities across varieties within each sector, suggesting that the CES assumption may be a poor approximation of the elasticity for many individual varieties.[33]

[29]Table D.6 in Appendix D.4 reports the same moments for the own-price elasticities $\partial \log \widetilde{q}_i / \partial \log p_i$, and shows that the same conclusions qualitatively hold.

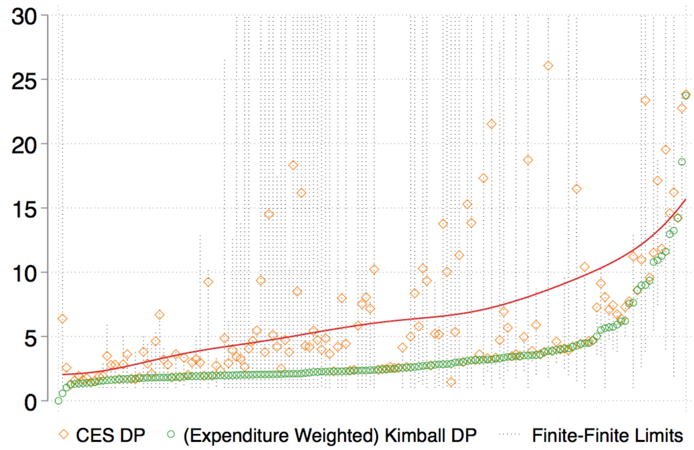
[30]Appendix D.2 provide a more extensive discussion of the differences between DP and previous identification methods (FBW and LIML) in the CES case. We use the data from Broda and Weinstein (2006) from 1989 to 2006 and show that also the DP method produces lower elasticities as we aggregate products into broader categories. Moreover, we confirm the existence of a downward bias in the elasticities estimated by the two conventional methods. Leveraging the Rauch (1999) classification, we also find that the bias is stronger for more differentiated products since quality should be more relevant for this type of products compared to more homogenous ones.

[31]Table D.5 in Appendix D.4 reports summary statistics of the distribution of the estimated Finite-Finite Kimball parameters.

[32]Recall that for the Kimball demand, we can compute the elasticity ε_{it} for each variety at each moment in time while in the CES case we only compute a common value across time and varieties within each industry. The moments for CES are computed assuming that each variety-time pair within the same sector has the same elasticity parameter ($\varepsilon_{it} \equiv \sigma$).

[33]Figure D.9 in Appendix D.4 illustrates the extent of the heterogeneity in elasticities for the motor vehicle parts manufacturing industry (NAICS number 33639). The figure reports the entire set of Kimball

Figure 1: Comparison with CES Elasticities



Note: In the figure we rank each 5-digit NAICS industry by the expenditure-share weighted mean Kimball elasticity. For each sector, it display the estimated lower and upper limits of the Finite-Finite Kimball specification (dotted line), the expenditure-share weighted mean Kimball price elasticity (green circles) and the corresponding CES estimate (orange diamonds). The upper limits are truncated at 30. The solid red line shows a fitted curve through the CES estimates.

In Appendix , we show that the heterogeneity bias in elasticity estimation arises depending on the covariance between the elasticity parameters and the of the cost shifter. Specifically, we find that the CES elasticity is below the Kimball mean of the elasticities if there is a positive covariance between the own-price elasticities of demand and the volatility of price changes. Figure E.1 confirms that products with higher own-price elasticities of demand are those that have more volatile prices, rationalizing the differences in estimated elasticity we get between CES and Kimball.

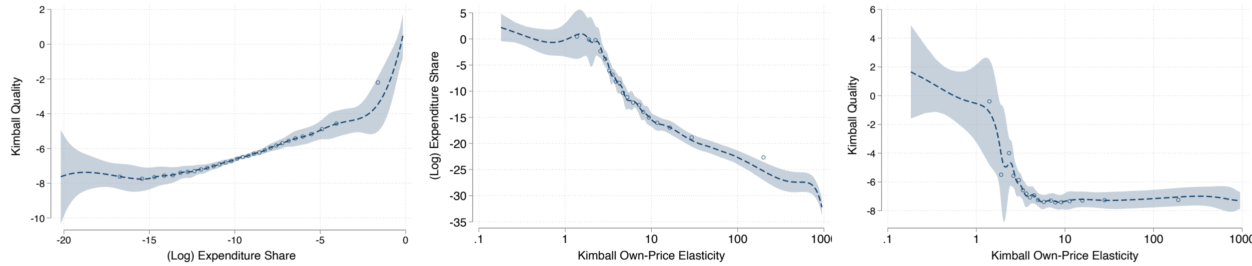
In Appendix E.3, we show that the heterogeneity bias in elasticity estimation depends on the covariance between the elasticity parameters and the cost shifter. Specifically, we show that the CES elasticity is lower than the average elasticity in a VES model when there is a positive covariance between the own-price elasticities of demand and the volatility of price changes. Figure E.1 confirms that products with higher Kimball own-price elasticities of demand tend to have more volatile prices, explaining the differences in the estimated elasticity in Table 1.

Lastly, we compute the own-price elasticities of demand and study the relationship between inferred quality, own-price elasticity, and expenditure share. We infer the measure of product quality for the Kimball case by inverting the Kimball demand using Equation (15).[34] As expected, Figure 2 shows that within each industry, varieties with higher

elasticities, their expenditure-share weighted and unweighted means, and the CES estimate. Even if the average Kimball elasticity is close to the CES estimate, the Kimball elasticities range from 1.5 to 4.5 and decrease with market share.

[34] See the discussion in Appendix B.2.2 for more details on inverting the Kimball demand.

Figure 2: Kimball Own-Price Elasticities and Implied Quality



Note: The left panel plots the binscattered relationship between (log) expenditure share of each variety-time observation and the inferred quality. The panel in the center plots the binscattered relationship between the Kimball own-price elasticity and the (log) expenditure share. The right panel directly plots the relationship between the inferred product quality and the Kimball own-price elasticity. In each panel, we use industry-time fixed effects and cluster the standard errors at the industry level.

inferred quality have higher expenditure shares and lower price elasticities.[35]

3.4 The Evolution of the US Import Price Index

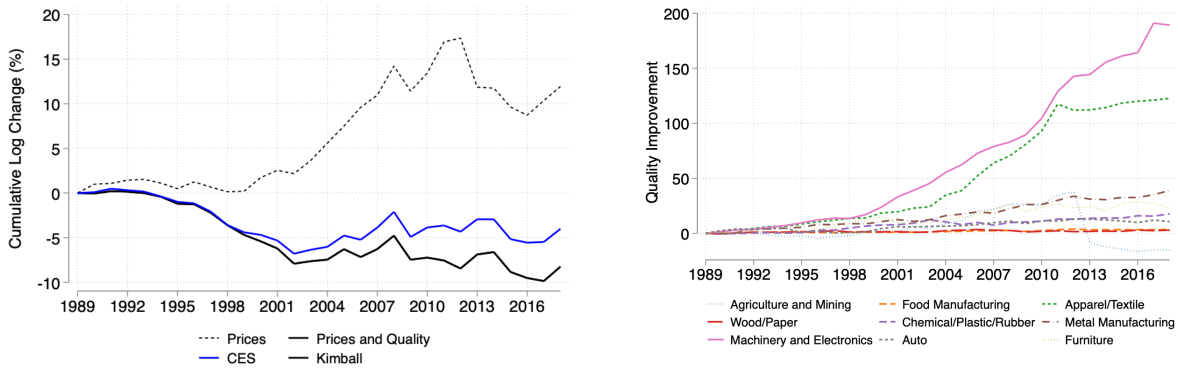
The left panel of Figure 3 and 2 report the cumulative and annual changes in the aggregate price of US imports from Equation (25), where the changes in the industry-level Kimball price indices are approximated using the expression in Proposition 2. Leveraging the result in Proposition 2, we quantify the contribution of quality changes to the dynamics of the aggregate price index of US imports. Improved product quality constitutes a major source of consumption gains from openness for the US, as quality improvements substantially reduce the increase in import prices. The import price index declined by more than 8% (approximately 0.3% annually) over the 1989-2018 period. In contrast, the index of unit prices of imported goods ($\Delta \log \Pi_t = \sum_i \overline{\overline{s_{it}}}\, \Delta \log p_{it}$) rose by 0.4% annually. For comparison, the official BEA Personal Consumption Expenditure (PCE) price index rose by over 1.9% annually during the same period.[36] However, the overall change in the price import price index is significantly lower when accounting for the improved quality of imported goods. Quality improvements reduced the aggregate import price index by 0.7% annually, suggesting that the official price index overestimate the degree of price inflation in imported goods.

Using CES preferences instead of Kimball reduces the consumption gains arising from the product quality channel by 20%, resulting in a substantial underestimation of the overall gains. The CES aggregate price index for imports shows a decline of around 4% (0.13% annually), 60% less than the Kimball case. The stark difference with respect to the

[35] Figure D.11 in Appendix D.4 shows that the same qualitative patterns hold when we control for variety fixed effects rather than industry-time fixed effects.

[36] Figure D.10 in Appendix D.4 shows that the year-to-year change in the price component of our aggregate import price index resembles the Import Price Index constructed by the BLS.

Figure 3: Dynamics of US Import Price Index



Note: The left panel plots the aggregate import price indices for both the CES and Kimball case and their decomposition into the price and quality components, according Proposition 2. The import price index is constructed using Equation (25). The solid lines represent the aggregate import price index including both the price and quality components. The dashed line represents the price component only. The black line refer to the Kimball model, while the blue line to the CES model estimated using DP. U.S. varieties are used to normalize the quality of imported goods. The right panel plots the quality improvement index at the sectoral level. The index is constructed using the inferred quality at the variety level aggregated using a tornqvist index.

Table 2: Change in the Import Price Index in the US, 1989–2018

	Total		PCE Index	Decomposition		
				Price	Quality	
	Kimball	CES			Kimball	CES
Cumulative Log Change (%)	-8.25	-4.01	57.8	11.9	-20.2	-15.9
Annual Change (%)	-0.28	-0.13	1.93	0.40	-0.67	-0.53

Note: The table reports the cumulative and average annual change in the aggregate import price indices constructed using Proposition 2 and Equation (25), and reported in Figure 3, and their decomposition. U.S. varieties are used to normalize the quality of imported goods. The third column reports the cumulative and annual change in the Personal Consumption Expenditure price index from the BEA.

Kimball aggregate price index arises from the different estimates of the role of quality upgrading. Whereas quality improvement reduces the Kimball aggregate import price by more than 20%, the corresponding contribution using CES is only 16%. This confirms the quantitative importance of departing from the constant elasticity assumption in the standard CES demand systems for evaluating the consumption gains from trade, both in terms of product quality and variety.

To better understand the drivers of the gap in the inferred price index under CES and Kimball, Equation (A.33) in Appendix A.1.3 provides a decomposition of this gap to a number of different components. Appendix D.3 uses this decomposition to show that the key reason for the underestimation of the price index under the CES specification is the heterogeneity in the matrix of cross-product elasticities of substitution, which is absent in the CES model.

Given that our data aggregates imports from product to the industry level, the entry of new products at finer levels of disaggregation appear in our estimates as within-industry quality improvements. For this reason, we have thus far abstracted away from the contribution of variety-level entry and exit to import prices. Leveraging the result in Proposition 3, Table D.7 in Appendix D.3.4 shows that accounting for the variety channel at this level of disaggregation makes a negligible impact on the import price index. Our estimates suggest that over the entire period (1989-2018), the aggregate import price index increased by 0.3% under Kimball and 0.06% under CES, respectively, due to the exit of some industry-level varieties.

Decomposition across Sectors and Exporters The right panel of Figure 3 reports the estimated quality improvements in US imports across different sectors. Leveraging the inferred quality at the variety level, we construct sector-specific Tornqvist indices of the changes in import quality. Imports of machinery and electrical equipments exhibit the strongest quality improvement, with a cumulative increase of approximately 200%, followed by goods in the textile and apparel sectors.[37] The left panel of Figure D.5 in Appendix D.5 highlights the relative importance of each sector in the aggregate price index of US imports, showing that 60% of the total gains from quality improvements can be attributed to quality improvements of goods in the machinery and electrical equipment industry. **Add literature**

The left panel of Figure D.6 in Appendix D.4 decomposes quality improvements in US imports across different sources countries, showing that Chinese products represent the single largest source of these gains. Approximately 35% of the total cumulative gains from quality improvement can be attributed to quality improvements of Chinese varieties alone, while the contribution of the OECD countries and all the other exporters to the overall quality improvement is about 7% and 59%, respectively.[38,39] This result is in line with the prior work documenting that the expansion of Chinese exports is not limited to the low-skill labor intensive and low-quality goods (Hsieh and Ossa, 2016). The right

[37] The right panel of Figure D.5 in Appendix D.5 shows that, in the Computer and Peripheral Equipment sector (NAICS 3341), accounting for quality improvements in imported goods resultsin a decreasing ratio of the import price index to the producer price index over time. This contrasts with the trends shown in Figure XX in the Introduction.

[38] Notice that Figure D.6 does not imply that the quality of OECD countries has not increased over time. The small contribution of OECD varieties reflects the fact that the quality of OECD countries relative to the U.S. has not changed over the last three decades.

[39] Figure D.12 in Appendix D.4 shows the same decomposition for the CES case. Chinese varieties still represent the major source of quality improvements, accounting for 35% of the aggregate quality improvement. The contribution of OECD varieties is small (2%), while other exporters' varieties account for the approximately 63% of the aggregate quality improvement. Departing from the constant elasticity assumption is important not only in evaluating the aggregate role of quality for the gains from trade, but also in decomposing its sources.

panel of Figure D.6 further shows that the quality upgrading accelerates after China's accession to the WTO, consistent with recent evidence for the substantial effect of the China's entry into the WTO on US prices particularly through the extensive margin of new firms (Amiti et al., 2020). In addition to this extensive margin, Redding and Weinstein (2024) also find that the appeal of the products of Chinese firms for US consumers rose in the 1998-2011 period. [40]

4 Validation: US Auto Data

In this section, we validate the Dynamic Panel (DP) approach for demand estimation by applying it to detailed data on the US automobile market and comparing the resulting estimates with those found using benchmark methods of demand estimation including the random coefficient logit model (Berry, 1994; Berry et al., 1995).

4.1 Data

We use data on the US automobile market from 1980 to 2018. The Wards Automotive Yearbooks contain information on specifications, list prices and sales by model for all cars, light trucks, and vans sold in the US.[41] Vehicle characteristics include horsepower, miles-per-dollar, miles-per-gallon, weight, width, height, style (car, truck, SUV, van, sport), and producer. Additional information such as the producer's region, whether the model is an electric vehicle, a luxurious brand, or a new design (redesign), complement the data from the yearbooks.[42] We perform standard cleaning to the data following Grieco et al. (2021) and Berry et al. (1995).[43] In addition, for the estimation of the Kimball specification, we exclude models that have an average price higher than \$100k over the entire time period and drop observations with a change in market share above (below) the 99th (1th) percentile within each year.

[40] This result is also consistent with the evidence of the effects of trade liberalization on firm performance. Prior work has documented that a reduction in (input and output) tariffs spurs innovation, productivity and product quality (see Shu and Steinwender (2019) for a survey, and see, among others, Brandt et al., 2017; Fan et al., 2015; Hsieh and Ossa, 2016 for discussions of the specific Chinese case). Schott (2008) show that Chinese products undertook a rapid process of sophistication. See Appendix **??** for further discussion.

[41] The Wards Automotive Yearbooks contain information for all trims (variants) of each model. Following standard practice, we aggregate all information at the model level based on the median across trims (Berry et al., 1995; Grieco et al., 2021).

[42] Table E.1 in Appendix E.1 provides additional details and displays summary statistics for our sample.

[43] Following Grieco et al. (2021), we drop models with unit price higher than \$100k. As in Berry et al. (1995), we define the new variable "space" as the product between length and width and exclude observations with a value larger than 6. Similarly, we define the ratio of horsepower per 10lbs and exclude observations with a value larger than 3.

We follow Grieco et al. (2021) and Goldberg and Verboven (2001) in the construction of an exogenous instrument for prices based on exchange rates. We use the lagged bilateral real exchange rate between the US and the country of assembly of each model, henceforth RER.[44] RER constitutes an arguably exogenous shifter of production costs capturing, in part, local labor market conditions in the country of assembly. This is because exogenous changes in local wages are reflected on the local price level and, in turn, on the real exchange rate. In addition, exogenous movements in the nominal exchange rate between the US and the country of assembly represents another source of variation for the RER as firms can lower their prices when the local currency depreciates.

Testing the Identification Assumption Before applying our methodology for demand estimation, we rely on the availability of product characteristics to directly test our identification assumption (Assumption 2) that lagged log prices are uncorrelated with innovations to quality. Appendix E.2, we show that lagged log prices are uncorrelated with current product characteristics after controlling for lagged product characteristics. In addition, product characteristics exhibit strong autocorrelations, supporting our Markov process assumption for the dynamics of product-level quality.

4.2 Demand Specifications and Benchmark Empirical Models

Our goal is validate two distinct aspects of the approach we proposed in Section 2: the effectiveness of the DP approach as an identification strategy, and the ability of a homothetic with aggregator (HA) demand system, e.g., the Kimball demand system from the HDIA family, to provide a satisfactory account of heterogeneity in price elasticities. First, to study the identification aspects, we estimate a standard CES specification using the DP approach and compare it against the standard instrumental variable approach that uses cost shocks (RER). Second, we consider Kimball aggregators and compare them against the current workhorse demand model for differentiated products, i.e., the random coefficient logit model (Berry, 1994; Berry et al., 1995). In this exercise, we also compare the estimates of the Kimball specification using the two alternative identification strategies: the DP approach and the standard cost shock IV approach. Below, we discuss the details of these alternative benchmark models.

[44] The RER is constructed as the ratio of the expenditure price levels between the assembly country and the US. The expenditure price levels are available from the Penn World Tables. See Grieco et al. (2021) for additional details.

CES Demand To study the properties of the DP identification strategy, we consider the CES specification that leads to a simple log-linear relationship between market shares and prices to estimate the elasticity of substitution σ:

$$\log s_{it} = -(\sigma - 1) \log p_{it} + \beta \boldsymbol{x}_{it} + make_i + \delta_t + \epsilon_{it}, \tag{29}$$

where $make_i$ specifies the producer of product i. Here, $\boldsymbol{x}_{it}$ stands for the vector of product characteristics, including footprint, horsepower, miles-per-dollar, curbweight, years since redesign, luxury brand, vehicle type (sport, electric, truck, suv, van). As mentioned, we can address the endogeneity of prices using a proxy for the costs of production, the real exchange rate (RER) in the assembly country, as a price instrument and also controlling for product characteristics and time and producer fixed effects. To compare the two identification methods, we additionally estimate Equation (29) with the DP approach, using the moment conditions in first-differences as in Equation (20) and relying on double-lagged prices and market shares as instruments, together with time and producer fixed effects.

Kimball Demand Next, we compare an HA demand system in the form of a Kimball specification against the empirical discrete choice model of differentiated products presented in Berry (1994) and Berry et al. (1995) (henceforth Mixed Logit). We estimate the three parametric families of Kimball functions presented in Equations (A.36), (A.35) and (A.34), using the moment condition in Equation (19). We estimate the Kimball specification using both the DP identification strategy and the RER as a cost-shock instrument.[45] Importantly, we define the base set O_t in each year to the set of continuing models that are *not* redesigned in that year, normalizing quality with respect to the average in this set. For the DP approach, we use lagged prices and their powers as instruments, as well as time and producer fixed effects. For the standard IV approach, we use log(RER) and their powers as the instrument.

Mixed Logit The Mixed Logit demand assumes heterogeneous consumers, whereby each consumer chooses one product that maximizes their random utility. The utility u^n_{it} of consumer n with total expenditure y^n_t from a product i with the vector $\boldsymbol{x}_{it}$ of product characteristics is given by $u^n_{it} = \frac{y^n_t - p_{it}}{p_{ot}} + \boldsymbol{x}'_{it}\widetilde{\boldsymbol{\beta}}^n + \widetilde{\xi}_{it} + \frac{1}{\alpha^n}\epsilon^n_{it}$, with p_{ot} being the price of the consumption of non-auto (outside) goods. The consumer-specific coefficients α^n and $\widetilde{\boldsymbol{\beta}}^n$ on price and on characteristics, respectively, depend on demographics and allow for

[45] Our algorithm for inverting Kimball demand in Appendix B.2.2 uses a reference product in each period. We choose the Chevrolet Corvette as this reference product.

observed and unobserved heterogeneity in consumer taste. We estimate the Mixed Logit model including the same set of product characteristics as in the CES specification, using the RER as a cost-shock instrument, and using additional micro moments on choice based on demographics and car purchases and second choices to estimate the parametric distributions of taste parameters as in Grieco et al. (2021). Moreover, we also estimate the Mixed Logit model using the DP approach, relying on on double-lagged prices and market shares as instruments instead of the RER as a cost-shock instrument.[46]

4.3 Comparison between Estimates

In Table 3, we report the estimated elasticities found by the different approaches for the whole sample. The first three columns show the estimates under the CES specification using OLS estimation, using the RER variable as the cost shock instrument (IV henceforth), and using our DP approach. The first row shows the estimated CES parameter while the remaining four columns display different moments of the distribution of the estimated own-price elasticities. The latter are also reported under the two models with variable elasticities, the Mixed Logit and the Kimball specifications. In each case, the table also shows the estimates when using the RER as the cost shock instrument and when using our DP approach.

As expected, we find that the OLS estimate of the CES price elasticity displays a bias towards zero due to the positive correlation between demand and price shocks. This is despite the fact that our specification in Equation (29) includes product characteristics to control for quality. When we use the cost shock instrument, the magnitude of the estimated CES elasticity rises relative to its OLS counterpart (1.35 from 4.67). This result confirms the need for price instruments to correct for the endogeneity bias in this setting.

Importantly, applying the DP approach to the CES specification delivers a CES elasticity of substitution of 4.51, close to the estimated elasticity obtained with the cost shock instrument. This suggests that our DP approach provides a solution for the endogeneity problem without relying on additional costs shocks, and even without controlling for product characteristics.

How important is accounting for heterogeneity in price elasticities? Comparing the estimates under the CES and the Mixed Logit models, we find that ignoring the heterogeneity in price elasticities leads to a bias toward zero under the former. The median, the unweighted, and the weighted means of the estimated elasticities are larger under the

[46] Additional details on the Mixed Logit specification, its estimation, and the definition of its exact price index in Appendix E.4.

Table 3: Comparing Own-Price Elasticities

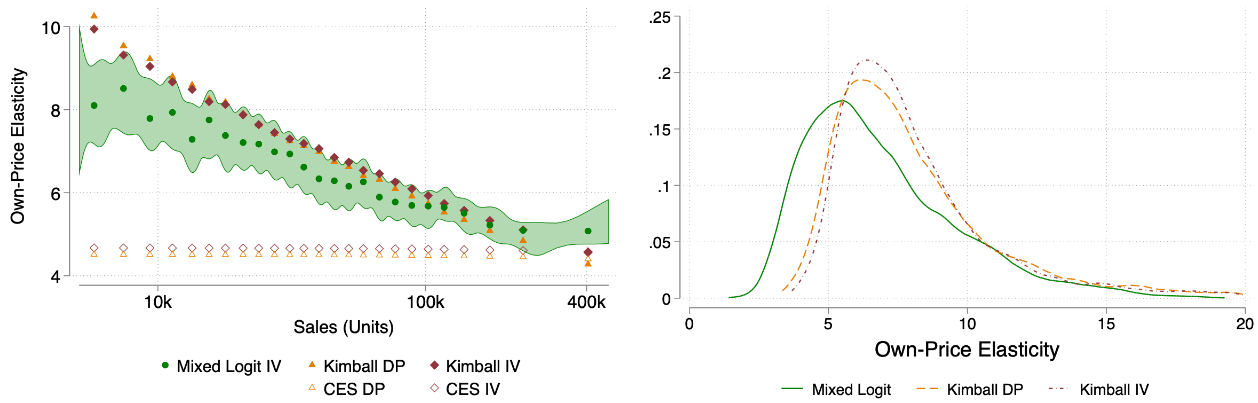
	CES			Mixed Logit		Kimball	
	OLS	IV	DP	IV	DP	IV	DP
σ	1.35	4.67	4.51				
	(0.25)	(1.47)	(0.13)				
Own-price Elasticity:							
Weighted Mean		4.62	4.46	6.15	6.97	5.60	5.79
		(0.00)	(0.00)	(0.02)	(0.03)	(0.01)	(0.01)
Mean		4.65	4.50	6.94	7.87	8.90	8.68
		(0.00)	(0.00)	(0.03)	(0.03)	(0.05)	(0.05)
Median		4.66	4.51	6.32	7.13	7.41	7.45
		(0.00)	(0.00)	(0.03)	(0.04)	(0.04)	(0.03)
IQR		0.02	0.02	3.57	4.01	3.54	3.16
		(0.00)	(0.00)	(0.05)	(0.05)	(0.07)	(0.06)

Note: The table reports the estimated own-price elasticities from the full sample. Each column corresponds to a different econometric model: CES OLS, CES IV, CES DP, Mixed Logit IV, Mixed Logit DP, Kimball IV, and Kimball DP. For the VES cases (Mixed Logit and Kimball) we report a set of moments from the distribution of the estimated own-price elasticities, while for the CES cases, we also report the estimated price coefficients. We report the mean and the median elasticity together with the expenditure weighted mean elasticity and the interquartile range. We consider the Finite-Finite case for the Kimball specification. For the Kimball specification, we compute the own-price elasticities using Equation (**??**) and (11). For the Mixed Logit specification, we report the demand elasticity defined as the percent change in sales for a one percent increase in price. We report bootstrapped standard errors for the set of moments from the distribution of the estimated own-price elasticities. For the CES price coefficeints, standard errors are clustered at product (model) level. All estimated quantities use the full sample.

Mixed Logit specification compared to the CES. Despite its simplicity, the Kimball specification also appears to allow for sufficient heterogeneity to circumvent this problem: all three moments of the distributions of the estimated own-price elasticities under Kimball are larger compared to the CES, and close to those under Mixed Logit, particularly the weighted mean and the median. Moreover, for both the Kimball and Mixed Logit specifications, we again find that the elasticites estimated using the cost shock instrument and using the DP approach are close, providing additional evidence of the validity of the DP approach.

In the Kimball case, the heterogeneity in elasticities is entirely due to the heterogeneity in market shares. In contrast, the heterogeneity in the elasticities estimated by the Mixed Logit may additionally stem from the heterogeneity in product characteristics as well. We next explore the relationship between the sales and the estimated own-price elasticities across products under the Mixed Logit and the Kimball models. The left panel of Figure 4 shows that this relationship is similar between the Mixed Logit specification and the Kimball specification, when estimated under both identification strategies (DP and IV). This result confirms that the Kimball specification can indeed account for the a similar relationship between sales and own-price elasticity as that uncovered by the Mixed Logit specification, and that the DP approach can identify this pattern without the use of any

Figure 4: Elasticity Heterogeneity in Kimball and Mixed Logit



Note: The left panel plots a binscatter representation of the relationship between expenditure share and the estimated own-price elasticitiess. Products with less than 5000 units sold in a year are excluded. We consider the set of elasticities estimated from: i) the Mixed Logit model (estimated with IV); ii) the Finite-Finite Kimball model using cost shocks (RER) as instruments (Kimball IV); iii) the Finite-Finite Kimball model using the DP approach (Kimball DP). We also report the CES price coefficient estimated using IV and DP. The shaded green area represents the confidence bands of the distribution of Mixed Logit elasticities. The right panel shows the distribution of own-price elasticities of all Kimball specifications (Finite-Finite, Finite-Infinite and Klenow-Willis) estimated using both the DP and IV instruments. The distribution of Mixed Logit elasticities is also reported. Values are truncated at 20.

additional information other than prices and market shares. Figure E.5 in Appendix E.7 shows that, as per the own-price elasticities, the matrix of cross-price elasticities is similar between the Mixed Logit specification and the Kimball specification.

The right panel of Figure 4 shows that the entire distribution of onw-price elasticities estimated by the Mixed Logit model is similar to those estimated under the different Kimball specifications and using the two different identification strategies. This result, in addition to the evidence on the similarity of the interquartile range values reported in Table 3, confirms that the heterogeneity in the price elasticities estimated under the Kimball specification bears a close resemblance to that under the Mixed Logit specification.[47]

Lastly, we show that the Kimball models resemble the Mixed Logit specifications because market shares carry a substantial amount of information on product characteristics. Figure E.4 in Appendix E.7 shows that there is a strong correlation between market shares and the first two principal components obtained from a singular value decomposition on the set of product characteristics. This explains how the Kimball models generates sufficient heterogeneity in elasticities to resemble a much richer specification as Mixed Logit by using information on market shares only.[48]

[47]See also Figure E.3 in Appendix E.7 for additional comparisons across Kimball specifications and identification strategies. We show that the distribution of elasticities, estimated using both the DP and the IV approaches, is robust to the choice of different families of the Kimball functions (Finite-Finite, Finite-Infinite and Infinite-Infinite).

[48]Table E.5 and Figure E.6 in Appendix E.7 show that the correlation between market shares and the principal components is stronger when we consider more homogeneous subsamples, such as products of

Inferred Quality and Product Characteristics Using detailed data on the US automobile market allows us to examine whether our approach retrieves meaningful measures of quality. Appendix E.6 examines this question by quantifying the correlation between our inferred measures of quality and the product characteristics valued by consumers available in our dataset, which is not feasible using standard customs data. In both CES and Kimball cases, we find find a strong correlation between the estimated values of quality φ_{it} and product characteristics such as horsepower, size, or whether a car is an SUV, truck, or a van. Importantly, however, the CES demand system attributes a much larger quality value to these characteristics compared to the Kimball demand. As we will see in Section 4.4 below, these differences translate into differences between the two models in the predicted change in the aggregate price index for the entire market.

Marginal Costs, Markups, and Quality If we assume that the market structure is characterized by monopolistic competition, the markup charged for each vehicle-year is given by $\mu_{it} = \frac{1}{\sigma_{it}-1}$, where σ_{it} is the estimated own-price elasticity for vehicle i at time t. Given this measure of markups, we infer the marginal cost of each vehicle to be $mc_{it} = \frac{p_{it}}{1+\mu_{it}}$. The right panel of Figure E.7 in Appendix E.7 shows that there is a strong positive relationship between a proxy of input cost, the weight of the vehicle multiplied by the price of steel, and our measure of inferred marginal cost, supporting the relevance of the latter. Figure E.7 shows that higher quality models have lower elasticities of demand and, thus, higher markups. Figure E.8 displays a positive relationship between inferred quality and the cost of production, in line with the findings of the prior literature on product quality (e.g., Verhoogen, 2008).[49]

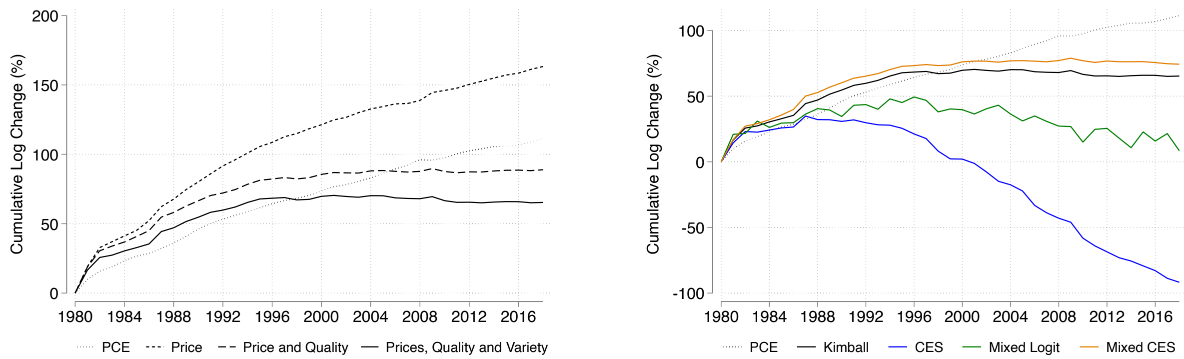
4.4 The Price Index of the US Auto Market

We construct the price index for the entire US auto market following Propositions 3 and A.1, quantifying the contribution of changes in unit price, quality, and the set of available models for consumers. Quality is normalized such that the average quality change in the set of continuing models that are not redesigned between each two consecutive years is zero (Grieco et al., 2021). As we explain in Appendix E.5, constructing a price index for the auto industry based on the Mixed Logit specification is quasi-linearity of this demand

the same style (i.e. car, SUV, truck, and van).

[49]Consistent with this evidence, Figure E.9 in Appendix E.7 shows that our measure of marginal costs is strongly correlated with the product characteristics consumers value (e.g. horsepower, footprint and miles-per-gallon). Moreover, these results are also consistent with Atkin et al. (2015), who show direct evidence for the relationship between markups and costs.

Figure 5: The Price Index for the US Auto Market



Note: The left panel plots the cumulative change in the price index for the auto market and its decomposition into the price, quality, and variety components for the Kimball specification. The solid line represents the price index including all three components. The dashed and dotted lines represent the price and quality components together and the price component only, respectively. We use the estimates from the Finite-Finite-Limit (FFL) Kimball specification, estimated using the DP approach. The dotted line represents the PCE index from BEA.
The right panel compares the cumulative change in the price index Kimball price index (black) to the CES (blue line), Mixed CES (orange lines), and Mixed Logit (green line) specifications. CES is also estimated using the DP approach. The price index for both the Kimball and the CES specifications is constructed using Proposition A.4. As reference, we report the price index for the Mixed CES case, constructed using the approximation in Proposition A.4, and the exact price index for the Mixed Logit specification (estimated using IV). In all cases, the measure of inferred quality is normalized such that the average change in quality of the set of continuing models that are not redesigned is zero.

system between automobile demand, which constitutes a very small share of the total consumer expenditure, and the rest of consumer expenditure. For this reason, Appendix E.5 sets up an alternative Mixed CES demand specification, which is equally as rich in terms of the patterns of cross-product substitution elasticities within the auto industry, but does not suffer from this limitation. We compare the dynamics of the price index under the Kimball and CES specifications to the Mixed Logit and Mixed CES price indices.

In the left panel of Figure 5 we plot the Kimball price index for the US auto market over the 1980-2018 period, highlighting the role of the price, quality, and variety channels. The aggregate price index increases by around 1.7% annually, relative to an annual increase in PCE of 2.9% over this period. Over the same period of time, unadjusted unit prices in the auto market increase by 4.30% annually. Quality improvement contributes substantially to attenuating the increase in the aggregate price index, accounting for an approximately 2% average annual decline. Figure 5 shows that the contribution of the availability of new models is marginal compared to the quality channel, accounting for a 0.6% annual drop in the aggregate price index.[50]

Table 4 also compares the price index for Kimball to the price index for the CES. The aggregate price index in the CES case decreases by 2.4% annually because the contribution of quality improvements is largely overestimated (6% in the CES case compared to

[50] Grieco et al. (2021) also attributes the bulk of the increase in consumer surplus in the auto industry to quality improvements, while a marginal role is played by the entry of new varieties.

Table 4: Price Index for the US Auto Market

	Total				Decomposition				
					Price	Quality		Variety	
	CES	Kimball	Mixed CES	Mixed Logit		Kimball	CES	Kimball	CES
Cumulative Log Change (%)	-91.76	65.31	74.35	8.44	163.26	-74.38	-226.36	-23.57	-28.66
Average Annual Change (%)	-2.41	1.72	1.96	0.22	4.30	-1.96	-5.96	-0.62	-0.75

Note: The Table reports the cumulative and the average annual change in the price indices for the auto market over the period 1980-2018 and its decomposition into the price, quality and variety channels. Quality is normalized such that the average change in quality of the set of continuing models that are not redesigned is zero. The price index is computed for both the Kimball and the CES specifications, estimated using the DP approach, constructed using Proposition A.4 and its decomposition. As reference, we report the price index for the Mixed CES case, constructed using the approximation in Proposition A.4, and the exact price index for the Mixed Logit specification (estimated using IV).

2% in the Kimball case). On the contrary, the Kimball aggregate price index is closer to the aggregate price indices for the variable elasticity models (1.7% in the Kimball case compared to 1.9% in the Mixed CES case and 0.2% in the Mixed Logit case), confirming that the distribution of elasticities is similar across these two specifications. We conclude that, despite its parsimony, the Kimball demand system captures the key sources of heterogeneity included a model with a rich pattern of cross-product elasticities of substitution such as Mixed CES or Mixed Logit.[51]

5 Conclusion

In this paper, we examined the role of quality improvements in shaping the variations in the price of US imports over the 1989-2018 period. We implemented a novel methodology to infer quality change as variations in residual demand in a flexible demand system using only data on prices and market shares. We presented approximate decompositions of the changes in the relative price of imports into the contribtuions of changes in prices, quality, and product variety. Moreover, we independently validated our approach in the context of the US auto market in which additional information on product chacteristics is available. Our baseline results suggest that, over the period from 1989 to 2018, quality improvements are responsible for an average 0.7 percentage points fall in import prices. By ignoring the heterogeneity in price elasticities, estimates based on CES demand provide a biased estimate of these gains. Applying our novel method to other specifications of demand, as well as to firm-level data to include pro-competitive effects and their interaction with quality, are promising venues for future research.

[51]We can use our estimation results to explore the evolution of markups and marginal cost in the US auto market. Figure E.10 in Appendix E.7 shows that marginal cost (markups) are increasing (decreasing) over the period 1980-2018, in line with previous work on this industry, Grieco et al. (2021).

References

Ackerberg, Daniel A, "Timing Assumptions and Efficiency: Empirical Evidence in a Production Function Context," 2016. 16

Amiti, Mary and Sebastian Heise, "US market concentration and import competition," 2021. 24, A43, A44

_, Mi Dai, Robert C Feenstra, and John Romalis, "How did China's WTO entry affect U.S. prices?," *Journal of International Economics*, 2020, *126* (71973013), 103339. 31

_, Oleg Itskhoki, and Jozef Konings, "International shocks, variable markups, and domestic prices," *The Review of Economic Studies*, 2019, *86* (6), 2356–2402. 6

Atkin, David, Azam Chaudhry, Shamyla Chaudry, Amit K. Khandelwal, and Eric Verhoogen, "Markup and Cost Dispersion across Firms: Direct Evidence from Producer Surveys in Pakistan," *American Economic Review: Papers & Proceedings*, 2015, *105* (5), 537–544. 37

Baldwin, Richard and James Harrigan, "Zeros, quality, and space: Trade theory and trade evidence," *American Economic Journal: Microeconomics*, 2011, *3* (May), 60–88. 4

Baqaee, David and Ariel Burstein, "Welfare and Output with Income Effects and Taste Shocks," *NBER Working Paper*, 2022, (1947611), No. 28754. 5, 20

Baqaee, David Rezza and Emmanuel Farhi, "Productivity and misallocation in general equilibrium," *The Quarterly Journal of Economics*, 2020, *135* (1), 105–163. 6

Becker, Randy A, Wayne B Gray, and Jordan Marvakov, "NBER-CES Manufacturing Industry Database (1958-2018, version 2021a)," 2021. 24, A44

Berlingieri, Giuseppe, Filippo Boeri, Danial Lashkari, and Jonathan Vogel, "Capital-skill complementarity in firms and in the aggregate economy," *Unpublished manuscript, Univ. Calif., Los Angeles*, 2022. 3

_, Holger Breinlich, and Swati Dhingra, "The impact of trade agreements on consumer welfareâEvidence from the EU common external trade policy," *Journal of the European Economic Association*, 2018, *16* (6), 1881–1928. 4

Berry, Steven, James Levinsohn, and Ariel Pakes, "Automobile prices in market equilibrium," *Econometrica: Journal of the Econometric Society*, 1995, pp. 841–890. 4, 5, 31, 32, 33

Berry, Steven T, "Estimating Discrete-Choice Models of Product Differentiation Authors," *The RAND Journal of Economics*, 1994, *25* (2), 242–262. 4, 10, 31, 32, 33

__, Amit Gandhi, and Philip A Haile, "Connected Substitutes and Invertibility of Demand," *Econometrica*, 2013, *81* (5), 2087–2111. 8, 9, 13, 18, 19, A3, A9, A24, A26

Boskin, Michael J., Ellen R. Dulberger, Robert J. Gordon, Zvi Griliches, and Dale W. Jorgenson, "Consumer Prices, the Consumer Price Index, and the Cost of Living," *Journal of Economic Perspectives*, 1998, *12* (1), 3–26. 2

Brandt, Loren, Johannes Van Biesebroeck, Luhang Wang, and Yifan Zhang, "WTO accession and performance of Chinese manufacturing firms," *American Economic Review*, 2017, *107* (9), 2784–2820. 31

Broda, Christian and David E Weinstein, "Globalization and the Gains from Variety," *The Quarterly journal of economics*, 2006, *121* (2), 541–585. 2, 4, 23, 25, 26, A1, A46, A47, A48, A49

Caliendo, Lorenzo, Giordano Mion, Luca David Opromolla, and Esteban Rossi-Hansberg, "Productivity and organization in portuguese firms," *Journal of Political Economy*, 2020, *128* (11), 4211–4257. 3

Comite, Francesco Di, Jacques-François Thisse, and Hylke Vandenbussche, "Verti-zontal differentiation in export markets," *Journal of International Economics*, 2014, *93* (1), 50–66. 5

Crozet, Matthieu, Keith Head, and Thierry Mayer, "Quality sorting and trade: Firm-level evidence for French wine," *The Review of Economic Studies*, 2012, *79* (2), 609–644. 5

Diewert, W Erwin and Catherine J Morrison, "Adjusting output and productivity indexes for changes in the terms of trade," *The Economic Journal*, 1986, *96* (383), 659–679. 21, A10

Diewert, Walter E, "The economic theory of index numbers: a survey," in Walter E Diewert and Alice O Nakamura, eds., *Essays in Index Number Theory*, Emerald Group Publishing Ltd, 1993, chapter 7, pp. 163–208. 8

Dingel, Jonathan I, "The determinants of quality specialization," *The Review of Economic Studies*, 2017, *84* (4), 1551–1582. 5

Dridi, Jemma and Kimberly Zieschang, "Export and import price indices," *IMF Staff Papers*, 2004, *51* (1), 157–194. A10

Eaton, Jonathan and Ana C Fieler, "The Margins of Trade," 2022. 5

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu, "Competition, markups, and the gains from international trade," *American Economic Review*, 2015, *105* (10), 3183–3221. 6

_ , _ , and _ , "How costly are markups?," *Journal of Political Economy*, 2023, *131* (7), 1619–1675. 3

Fajgelbaum, Pablo, Gene M Grossman, and Elhanan Helpman, "Income distribution, product quality, and international trade," *Journal of political Economy*, 2011, *119* (4), 721–765. 4

Fan, Haichao, Yao Amber Li, and Stephen R Yeaple, "Trade liberalization, quality, and export prices," *Review of Economics and Statistics*, 2015, *97* (5), 1033–1051. 31

Feenstra, Robert C, "New product varieties and the measurement of international prices," *The American Economic Review*, 1994, pp. 157–177. 2, 3, 4, 17, 18, 19, 25, A0, A5, A16, A17, A18, A47

_ , *Product variety and the gains from international trade*, Cambridge, MA: MIT Press, 2010. A17

_ and David E Weinstein, "Globalization, markups, and US welfare," *Journal of Political Economy*, 2017, *125* (4), 1040–1074. 5, 6

_ and John Romalis, "International prices and endogenous quality," *The Quarterly Journal of Economics*, 2014, *129* (2), 477–527. 5

_ , _ , and Peter K Schott, "US imports, exports, and tariff data, 1989-2001," 2002. A46

Flam, Harry and Elhanan Helpman, "Vertical product differentiation and advertising," *American Economic Review*, 1987, *77* (5), 810–822. 4

Foley, Conor, "Flexible Entry / Exit Adjustment for Price Indices," 2021. 3

Gábor-Tóth, Eniko and Philip Vermeulen, "The Relative Importance of Taste Shocks and Price Movements in the Variation of Cost-of-Living: Evidence From Scanner Data," 2018. 5

Goldberg, Pinelopi Koujianou and Frank Verboven, "The evolution of price dispersion in the European car market," *The Review of Economic Studies*, 2001, *68* (4), 811–848. 32

Gordon, Robert J and Zvi Griliches, "Quality change and new products," *The American Economic Review*, 1997, *87* (2), 84–88. 2

Grennan, Matthew, "Price discrimination and bargaining: Empirical evidence from medical devices," *American Economic Review*, 2013, *103* (1), 145–177. 3

Grieco, Paul LE, Charles Murry, and Ali Yurukoglu, "The evolution of market power in the US auto industry," Technical Report, National Bureau of Economic Research 2021. 6, 31, 32, 34, 37, 38, 39, A66, A69

Grossman, Gene M, Elhanan Helpman, and Hugo Lhuillier, "Supply chain resilience: Should policy promote international diversification or reshoring?," *Journal of Political Economy*, 2023, *131* (12), 3462–3496. 3

Hallak, Juan Carlos, "Product quality and the direction of trade," *Journal of International Economics*, 2006, *68* (1), 238–265. 4

_ **and Peter K Schott**, "Estimating cross-country differences in product quality," *The Quarterly journal of economics*, 2011, *126* (1), 417–474. 2, 5

Hanoch, Giora, "CRESH Production Functions," *Econometrica*, 1971, *39*, 695–712. 3

Head, Keith and Thierry Mayer, "Poor Substitutes? Counterfactual Methods in IO and Trade Compared," 2021. 5

Hsieh, Chang-Tai and Ralph Ossa, "A global view of productivity growth in China," *Journal of international Economics*, 2016, *102*, 209–224. 30, 31

_ **, Nicholas Li, Ralph Ossa, and Mu-Jeung Yang**, "Accounting for the new gains from trade liberalization," *Journal of International Economics*, 2020, *127*, 103370. 5

Hummels, David and Alexandre Skiba, "Shipping the good apples out? An empirical confirmation of the Alchian-Allen conjecture," *Journal of Political Economy*, 2004, *112* (6), 1384–1402. 4

_ **and Peter J Klenow**, "The variety and quality of a nation's exports," *American economic review*, 2005, *95* (3), 704–723. 5

Jaravel, Xavier and Danial Lashkari, "Nonparametric Measurement of Long-Run Growth in Consumer Welfare," 2021. 5

— **and** — , "Measuring Growth in Consumer Welfare with Income-Dependent Preferences: Nonparametric Methods and Estimates for the United States," *The Quarterly Journal of Economics*, 2024, *139* (1), 477–532. 20, A6, A7, A8

Kasahara, Hiroyuki and Yoichi Sugita, "Nonparametric Identification of Production Function, Total Factor Productivity, and Markup from Revenue Data," 2021. 6

Kehoe, Timothy J and Kim J Ruhl, "Are shocks to the terms of trade shocks to productivity?," *Review of Economic Dynamics*, 2008, *11* (4), 804–819. A11

Khandelwal, Amit, "The long and short (of) quality ladders," *The Review of Economic Studies*, 2010, *77* (4), 1450–1476. 2, 5

Kimball, Miles S, "The quantitative analytics of the basic neomonetarist model," 1995. 3, 22

— , "The quantitative analytics of the basic neomonetarist model," 1995. 12

Klenow, Peter J and Jonathan L Willis, "Real Rigidities and Nominal Price Changes," 2006. 12, 23, A16

Kohli, Ulrich, "Real GDP, real domestic income, and terms-of-trade changes," *Journal of International Economics*, 2004, *62* (1), 83–106. A10, A11

Kugler, Maurice and Eric Verhoogen, "Prices, plant size, and product quality," *Review of Economic Studies*, 2012, *79* (1), 307–339. 5

Leamer, Edward E, "Is it a Demand Curve, Or is It A Supply Curve? Partial Identification through Inequality Constraints," *Review of Economics and Statistics*, 1981, *63* (3), 319–327. 17, A16, A17, A18

Lee, Robin S, "Vertical integration and exclusivity in platform and two-sided markets," *American Economic Review*, 2013, *103* (7), 2960–3000. 3

Linder, Staffan B, *An Essay on Trade and Transformation*, Uppsala: Almqvist & Wiksells, 1961. 4

Manova, Kalina and Zhiwei Zhang, "Export prices across firms and destinations," *Quarterly Journal of Economics*, 2012, *127* (1), 379–436. 5

Martin, Julien and Isabelle Mejean, "Low-wage country competition and the quality content of high-wage country exports," *Journal of International Economics*, 2014, *93* (1), 140–152. 5

Matsuyama, Kiminori, "Non-CES aggregators: a guided tour," 2022. 3, 11

_ **and Philip Ushchev**, "Beyond CES: Three Alternative Classes of Flexible Homothetic Demand Systems," 2017. 3, 11

_ **and** _ , "Selection and Sorting of Heterogeneous Firms through Competitive Pressures," 2022. 6

McFadden, Daniel, "Conditional logit analysis of qualitative choice behavior," in P Zarembka, ed., *Frontiers in Econometrics*, Academic Press, 1974, pp. 105—-142. 10

Oulton, Nicholas, "The effect of changes in the terms of trade on GDP and welfare: A Divisia approach to the System of National Accounts," *The Manchester School*, 2023. A11

Pierce, Justin R and Peter K Schott, "A concordance between ten-digit US Harmonized System Codes and SIC/NAICS product classes and industries," *Journal of Economic and Social Measurement*, 2012, *37* (1-2), 61–96. 24, A43, A44

Rauch, James E, "Networks versus markets in international trade," *Journal of international Economics*, 1999, *48* (1), 7–35. 26, A47, A48, A49

Redding, Stephen J and David E Weinstein, "Measuring aggregate price indices with taste shocks: Theory and evidence for ces preferences," *The Quarterly Journal of Economics*, 2020, *135* (1), 503–560. 2, 3, 4, 5, 7, 9, 19

Redding, Stephen J. and David E. Weinstein, "Measuring Aggregate Price Indices with Taste Shocks: Theory and Evidence for CES Preferences," *Quarterly Journal of Economics*, 2020, *135* (1), 503–560. A4

Redding, Stephen J and David E Weinstein, "Accounting for trade patterns," *Journal of International Economics*, 2024, *150*, 103910. 5, 31

Reinsdorf, Marshall B, "Terms of trade effects: Theory and measurement," *Review of Income and Wealth*, 2010, *56*, S177–S205. A11

Schott, Peter K, "Across-product versus within-product specialization in international trade," *The Quarterly Journal of Economics*, 2004, *119* (2), 647–678. 5

_ , "The relative sophistication of Chinese exports," *Economic policy*, 2008, *23* (53), 6–49. 31, A43

Shu, P and C Steinwender, "The impact of trade liberalization on firm productivity and innovation. Innovation Policy and the Economy, Vol. 19," 2019. 31

Soderbery, Anson, "Estimating import supply and demand elasticities: Analysis and implications," *Journal of International Economics*, 2015, *96* (1), 1–17. 4

__ , "Estimating import supply and demand elasticities: Analysis and implications," *Journal of International Economics*, 2015, *96* (1), 1–17. A17, A47, A48

Sweeting, Andrew, "Dynamic Product Positioning in Differentiated Product Markets: The Effect of Fees for Musical Performance Rights on the Commercial Radio Industry," *Econometrica*, 2013, *81* (5), 1763–1803. 3

Triplett, Jack, "Handbook on hedonic indexes and quality adjustments in price indexes: Special application to information technology products," 2004. 2

Ueda, Kozo, Kota Watanabe, and Tsutomu Watanabe, "Product Turnover and the Cost-of-Living Index: Quality versus Fashion Effects," *American Economic Journal: Macroeconomics*, 2019, *11* (2), 310–347. 5

U.S. Bureau of Labor Statistics, *Handbook of Methods: Producer Price Index* U.S. Department of Labor 2024. `https://www.bls.gov/ppi/methodology.htm`. 3

Verhoogen, Eric A, "Trade, quality upgrading, and wage inequality in the Mexican manufacturing sector," *The Quarterly Journal of Economics*, 2008, *123* (2), 489–530. 4, 5, 37

Wang, Olivier and Iván Werning, "Dynamic oligopoly and price stickiness," Technical Report, National Bureau of Economic Research 2020. 6

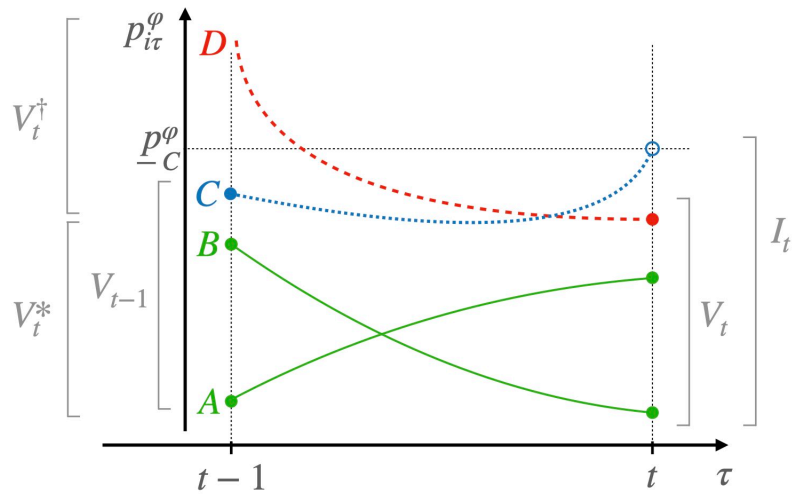
Appendix to "Aggregation and the Estimation of Quality Change: Application to the Case of the US Price Index"

Marco Errico, Bank of Italy
Danial Lashkari, Federal Reserve Bank of New York

September 2024

Contents

Figure A.1: Construction of the smooth paths of quality-adjusted prices



Note: Schematic depication of the paths of quality-adjusted prices connecting the observations between two time periods $t-1$ and t. Products A and B belong to the continuing set V_t^* and thus have nonzero consumption shares in both terminal periods. Product C exists in period t, in the sense that consumers cease to spend on it even though it remains available. In this case, the path converges to the quality-adjusted choke price $\underline{p}^{\varphi}_C$ at time t. Product D enters in period t, in the sense that it becomes available in this period. In this case, the path converges to infinity from below at time $t-1$.

A Additional Theoretical Results

A.1 Approximate Divisia Index with Income Dependence and Product Entry/Exit

A.1.1 Construction of the Smooth Paths of Quality-Adjusted Prices

Figure A.1 provides a schematic view of the construction of the smooth paths of prices and expenditure shares between periods $t-1$ and t. At time t, consumer(s) have nonzero expenditure on the *purchased set* $\overline{I}_t$ of products (so that $s_{it} = 0$, $i \notin \overline{I}_t$), which may potentially be different from the set of available products I_t. The constructed paths are the same as that in Section 2.2 for the *continuing set* $I_t^* \equiv \overline{I}_{t-1} \cap \overline{I}_t$ of products common between the purchased sets in the two periods: we simply consider any smooth interpolating paths of quality-adjusted prices $\boldsymbol{p}^{\varphi}_{t-1}$ to $\boldsymbol{p}^{\varphi}_t$ and total expenditure y_{t-1} to y_t that connect the two terminal periods. For the *entering/exiting set* $I_t^{\dagger} \equiv (I_{t-1} \cup I_t) \backslash I_t^*$ of products, we assume infinite prices if the product is unavailable, and a choke quality-adjusted price $\underline{p}^{\varphi}_i$ (above which the demand falls to zero) if the product is available but is not purchased. For instance, if a product i is unavailable in the end ($i \in I_{t-1}$ and $i \notin I_t$), its price satisfies $\lim_{\tau \downarrow t-1} p_{i\tau} = \infty$, and if it is unavailable in the beginning ($i \notin I_{t-1}$ and $i \in I_t$), it satisfies $\lim_{\tau \uparrow t} p_{i\tau} = \infty$. Note that the ultimate approximation is independent of the paths of prices considered.

Along the constructed paths, we can apply the definition of the demand system to define the corresponding paths of expenditure shares $s_{i\tau}$, cross-product substitution elasticities $\sigma_{ij,\tau}$, and the income elasticity $\eta_{i,\tau}$. We also define the total expenditure share of the continuing set as $\Lambda^*_\tau \equiv \sum_{i\in I^*_t} s_{i\tau}$, and the expenditure shares within the continuing set as $s^*_{i\tau} \equiv s_{i\tau}/\Lambda^*_\tau$ for $i \in I^*_t$. Similarly, we define the total expenditure share of the set of entering/exiting products as $\Lambda^\dagger_\tau \equiv \sum_{i\in I^\dagger_t} s_{i\tau} = 1 - \Lambda^*_\tau$, and the expenditure shares within this set $s^\dagger_{j\tau} \equiv s_{j\tau}/\Lambda^\dagger_\tau$ for $j \in I^\dagger_t$, denoting the distribution of expenditures within the set of entering/exiting products.

Lemma A.1. *Along the paths defined above, the vector of change in log expenditure shares and the vector of change in log quality-adjusted prices satisfy*

$$d\log s_{i\tau} = -\sum_{j\in I} \Xi_{ij,\tau} \left(d\log p_{j\tau} - d\varphi_{j\tau} - d\log P_\tau\right) + (\eta_{i\tau} - 1)\left(d\log y_\tau - d\log P_\tau\right), \quad \text{(A.1)}$$

where we have defined the instantaneous price index as

$$d\log P_\tau = \sum_i s_{i\tau}\left(d\log p_{i\tau} - d\varphi_{i\tau}\right), \quad \text{(A.2)}$$

and where the matrix mapping prices to expenditure shares satisfies $\boldsymbol{\Xi}_\tau \equiv \boldsymbol{\Psi}_\tau \left(\boldsymbol{\Sigma}^d_\tau - \boldsymbol{I}\right)$ with $\boldsymbol{\Psi}_\tau$ and $\boldsymbol{\Sigma}^d_\tau$ given as in Proposition 1.

Proof. See Appendix B.1 on page A23. □

A.1.2 Exact Characterization of the Price Index Change Along the Path

The following lemma characterizes the instantaneous change in the log price index anywhere along the path constructed in the previous section.

Lemma A.2. *Assume that the demand system is income invariant (homothetic) and satisfies the connected substitute property of Berry et al. (2013). Consider a path of quality-adjusted vector of prices $\boldsymbol{p}^\varphi_t$ and the corresponding expenditure shares constructed following Section A.1.1. Then, anywhere along such a path, the matrix $\boldsymbol{\Psi}_\tau \left(\boldsymbol{\Sigma}^d_\tau - \boldsymbol{I}\right)$ mapping relative prices to relative expenditure shares, with elements given by Equation (4), is invertible. Letting $\Psi^{-1}_{ij,\tau}$ denote the elements of the inverse matrix, we can alternatively write the change in the log price index*

$$d\log P_\tau = \sum_i \omega_{it} d\log p_{i\tau} + \sum_i \sum_{j\in I^*_t} \omega_{it} \tfrac{1}{\sigma_{ii,\tau}-1} \Psi^{-1}_{ij,\tau}\, d\log s^*_{j\tau}$$

$$+\left(\sum_i \sum_{j\in I_t^*} \tfrac{\omega_{it}}{\sigma_{ii,\tau}-1}\Psi_{ij,\tau}^{-1}\right) d\log\Lambda_\tau^* + \sum_i \sum_{j\in I_t\setminus I_t^*} \tfrac{\omega_{it}}{\sigma_{ii,\tau}-1}\Psi_{ij,\tau}^{-1}\, d\log s_{j\tau}, \tag{A.3}$$

where $(\omega_{it})_i$ are the weights corresponding to the base set O_t between periods $t-1$ and t. Moreover, the instantaneous change in the price index can be decomposed to the the contribution of unit prices and quality change in the set of continuing products, and product entry and exit $d\log X_\tau$ according to

$$d\log P_\tau = \sum_i s_{i\tau}^* d\log p_{i\tau} - \sum_i s_{i\tau}^* d\varphi_{i\tau} + d\log X_\tau, \tag{A.4}$$

where the latter two terms satisfy

$$\sum_i s_{i\tau}^* d\varphi_{i\tau} = \sum_i (s_{i\tau}^* - \omega_{it})\, d\log p_{i\tau} + \sum_{j\in I_t^*}\sum_i (s_{i\tau}^* - \omega_{it})\, \tfrac{1}{\sigma_{ii,\tau}-1}\Psi_{ij,\tau}^{-1}\left(d\log s_{j\tau}^* + d\log\Lambda_\tau^*\right)$$
$$+ \sum_{j\in I_t^\dagger}\sum_i (s_{i\tau}^* - \omega_{it})\, \tfrac{1}{\sigma_{ii,\tau}-1}\Psi_{ij,\tau}^{-1}\, d\log s_{j\tau}, \tag{A.5}$$

$$d\log X_\tau = \left(\sum_{i,j\in I_t^*} s_{i\tau}^* \tfrac{1}{\sigma_{ii,\tau}-1}\Psi_{ij,\tau}^{-1}\right) d\log\Lambda_\tau^* + \sum_{i,j\in I_t^*} s_{i\tau}^* \tfrac{1}{\sigma_{ii,\tau}-1}\Psi_{ij,\tau}^{-1} d\log s_{j\tau}^* + \sum_i \sum_{j\in I_t^\dagger} s_{i\tau}^* \tfrac{1}{\sigma_{ii,\tau}-1}\Psi_{ij,\tau}^{-1}\, d\log s_{j\tau}. \tag{A.6}$$

Proof. See Appendix B.1 on page A24. □

If we again consider CES preferences with elasticity parameter σ, and if we assume that $O_t \subset I_t^*$, the expression in Equation A.3 simplifies to

$$d\log P_\tau = \sum_i \omega_{it} d\log p_{i\tau} + \frac{1}{\sigma-1}\sum_i \omega_{it} d\log s_{i\tau}^* + \frac{1}{\sigma-1} d\log\Lambda_\tau^*.$$

In this case, we can integrate the expression between times $t-1$ and t to reach Equation (23).

Equation (A.4) expresses the inflation in the aggregate price index at any point along the path between periods $t-1$ and t as the sum of a number of different contributions. The first and the second terms are similar to those in Equations (**??**) and (7). The remaining terms account for the changes in the sets of products entering or exiting between the two periods. To unpack these three terms, let us first consider the special case of the CES demand system. In this case, all cross-product elasticities of substitution are identical and constant, $\eta_{ij,\tau} \equiv \sigma$ for all $i \neq j$ and τ. As a result, we find $\sigma_{i\tau} = \overline{\sigma}_\tau^* = \sigma$s and the covariance terms in Equation (A.6) vanish, leading to the standard result (e.g., Redding and Weinstein, 2020b):

$$d\log P_{\tau} = \sum_i s^*_{i\tau} d\log p_{i\tau} - \sum_i s^*_{i\tau} d\varphi_{i\tau} + \tfrac{1}{\sigma-1}\, d\log \Lambda^*_{\tau}. \tag{A.7}$$

The last term in Equation (A.7) is the Feenstra (1994) CES correction for the contributions of product entry and exit, given by the product of the CES love-of-variety parameter $\frac{1}{\sigma-1}$ and the change in the expenditure share of common set. If the share of continuing products in expenditure rises, it signifies that the consumer is switching their expenditure away from the entering/exiting products, implying a welfare loss captured by the rising price index.

More generally, Equations (A.4), (A.5), and (A.6) show how the presence of heterogeneity in cross-product substitution elasticities modifies the CES case in Equation (A.7). We will use this result in the derivation of Proposition A.1 below.

Finally, Lemma A.3 below generalizes this result to more general demand systems featuring potential income dependence (nonhomotheticity).

Lemma A.3. *Assume all the conditions in Lemma A.2 but relax the assumption of income invariant demand. Then, the quality-adjusted Divisia index, i.e., expenditure-share weighted mean of quality-adjusted prices $d\log P_\tau \equiv \sum_i s_{i\tau} d\log p^{\varphi}_{i\tau}$, satisfies*

$$d\log P_\tau \equiv \frac{d\log P^{II}_\tau - \left(\sum_{i,j} \frac{\omega_{it}}{\sigma_{ii,\tau}-1} \Psi^{-1}_{ij,\tau}\left(\eta_{j\tau}-1\right)\right) d\log y_\tau}{1-\sum_{i,j} \frac{\omega_{it}}{\sigma_{ii,\tau}-1} \Psi^{-1}_{ij,\tau}\left(\eta_{j\tau}-1\right)}, \tag{A.8}$$

where $d\log P^{II}_\tau$ is given by the expression for the income-invariant case provided in Equation (A.4).

Proof. See Appendix B.1 on page A26. □

A.1.3 Approximate Change in the Price Index and Quality

Proposition 3 provides an expression that approximately characterizes the change in the price index between each two consecutive periods in terms of the expenditure shares and prices.

Proposition A.1. *(Approximate Decomposition of the Price Index for Income-Invariant Preferences with Product Entry and Exit) Assume that the conditions in Proposition 3 hold and define δ accordingly. Up to the second order of approximation in δ, the change in the log price index between periods $t-1$ and t can be decomposed as $\Delta \log P_t = \sum_i \overline{\overline{s^*_{it}}} \Delta \log p_{it} - \Delta \log \Phi_t +$*

$\Delta \log X_t + O\left(\delta^3\right)$, *where the first term accounts for the contribution of unit prices in the set of continuing products, the second term* $\Delta \log \Phi_t$ *for quality change, and the last term* $\Delta \log X_t$ *for the contribution of product entry and exit. The latter two terms are given by*

$$\Delta \log \Phi_t = \sum_i \overline{\overline{s^*_{it}}} \Delta \log p_{it} - \sum_i \varpi_{it} \Delta \log p_{it} + \sum_{j \in I^*_t} \left(\overline{\overline{\sum_i s^*_{it} \Psi^{-1}_{ij,t}}} - \sum_i \varpi_{it} \overline{\overline{\Psi^{-1}_{ij,t}}} \right) \left(\Delta \log s^*_{jt} + \Delta \log \Lambda^*_t \right)$$
$$+ \sum_i \sum_{j \in I_t \setminus I^*_t} \left(s_{it} - \varpi_{it} \right) \Psi^{-1}_{ij,t} - \sum_i \sum_{j \in I_{t-1} \setminus I^*_t} \left(s_{it-1} - \varpi_{it} \right) \Psi^{-1}_{ij,t}, \tag{A.9}$$

$$\Delta \log X_t = \sum_{j \in I^*_t} \overline{\overline{\sum_i s^*_{it} \Psi^{-1}_{ij,t}}} \left(\Delta \log s^*_{jt} + \Delta \log \Lambda^*_t \right) + \sum_{j \in I_t \setminus I^*_t} s_{it} \Psi^{-1}_{ij,t} - \sum_{j \in I_{t-1} \setminus I^*_t} s_{it-1} \Psi^{-1}_{ij,t-1}. \tag{A.10}$$

Proof. See Appendix B.1 on page A26. □

The proof of the proposition shows that we can approximate the change in the quality of a continuting product $i \in V^*_t$ as follows

$$\Delta \varphi_{it} = \Delta \log p_{it} - \sum_i \varpi_{it} \Delta \log p_{it} + \sum_{j \in I^*_t} \left(\overline{\overline{\frac{1}{\sigma_{ii,t} - 1} \Psi^{-1}_{ij,t}}} - \sum_i \varpi_{it} \overline{\overline{\frac{1}{\sigma_{ii,t} - 1} \Psi^{-1}_{ij,t}}} \right) \left(\Delta \log s^*_{jt} + \Delta \log \Lambda^*_t \right)$$
$$+ \sum_{j \in I_t \setminus V^*_t} \left(\frac{1}{\sigma_{ii,t} - 1} \Psi^{-1}_{ij,t} - \sum_i \varpi_{it} \overline{\overline{\frac{1}{\sigma_{ii,t} - 1} \Psi^{-1}_{ij,t}}} \right) - \sum_{j \in I_{t-1} \setminus V^*_t} \left(\frac{1}{\sigma_{ii,t} - 1} \Psi^{-1}_{ij,t-1} - \sum_i \varpi_{it} \overline{\overline{\frac{1}{\sigma_{ii,t} - 1} \Psi^{-1}_{ij,t}}} \right) + O\left(\delta^3\right). \tag{A.11}$$

The contribution of unit price and quality change in the set of continuing products are approximately given by the Törqvist indices of price and quality, respectively. The residual term is given by Equation (A.10) and accounts for the contribution of product entry and exit.

A.2 From the Divisia to the Cost-of-Living Indices with Income Dependence

Our analysis so far has focused on recovering approximations of the Divisia index that account for the changes in quality and product entry and exit. As emphasized recently by Jaravel and Lashkari (2024), when the composition of consumption depends on income (income-dependence/nonhomotheticity), the Divisia index only provides a local approximation of the cost-of-living index and cannot be chained over time. They show how to apply corrections to the Divisia index to construct theoretically consistent cost-of-living indices that can be chained over time.

We can show that our results here easily generalize to their framework. To see this, let us follow their approach and define real consumption c and the mapping from real consumption to expenditure $\chi_t(\cdot)$ given the vector of quality-adjusted prices $\boldsymbol{p}_t^{\varphi}$ at time t and the choice of a base period b for measuring real consumption using the following two conditions

$$\chi_t^b(c) \equiv E\left(u; \boldsymbol{p}_t^{\varphi}\right), \tag{A.12}$$

$$c \equiv E\left(u; \boldsymbol{p}_b^{\varphi}\right), \tag{A.13}$$

expressed in terms of the expenditure function $E(u; \boldsymbol{p}^{\varphi})$. As such, we can define a corresponding price index as a function of real consumption

$$\widetilde{\mathcal{P}}_t^b(c) \equiv \frac{E\left(u; \boldsymbol{p}_t^{\varphi}\right)}{E\left(u; \boldsymbol{p}_b^{\varphi}\right)}, \tag{A.14}$$

where utility u again satisfies condition (A.13).

With the above definitions, we can generalize Proposition 1 of Jaravel and Lashkari (2024) to characterize the evolution of the mapping $\chi_\tau^b(\cdot)$ along a smooth path of quality-adjusted prices $\boldsymbol{p}_\tau^{\varphi}$ and total expenditures y_τ connecting the two periods $t-1$ and t as

$$\frac{\partial \log \chi_\tau^b(c)}{\partial \tau} = \log \mathcal{D}_\tau^{\varphi}\left(\chi_\tau^b(c)\right), \tag{A.15}$$

where we have defined the quality-adjusted Divida function $\mathcal{D}_\tau^{\varphi}(\cdot)$ of total expenditure y as

$$\log \mathcal{D}_\tau^{\varphi}(y) \equiv \sum_i \widetilde{s}_i^{uc}\left(y; \boldsymbol{p}_\tau^{\varphi}\right)\left(d\log p_{i\tau} - d\varphi_{it}\right), \tag{A.16}$$

where the uncompensated (marshallian) expenditure share function $\widetilde{s}_i\left(y; \boldsymbol{p}_\tau^{\varphi}\right)$ may depend on total expenditure y due to income-dependence in preferences. It then follows that the instantaneous change in the price index is given as a convex combination

$$\frac{\partial \log \widetilde{\mathcal{P}}_\tau^b(c)}{\partial \tau} = \left[1 - \left(\frac{\partial \log \chi_\tau^b(c_\tau)}{\partial \log c_\tau}\right)\right]^{-1} \frac{d\log y_\tau}{d\tau} + \left(\frac{\partial \log \chi_\tau^b(c_\tau)}{\partial \log c_\tau}\right) \log \mathcal{D}_\tau^{\varphi}(y),$$

of the growth in nominal expenditure $\frac{d\log y_\tau}{d\tau}$ and the quality-adjusted Divisa index $\log \mathcal{D}_\tau^{\varphi}(y)$. Our results in the paper focus on characterizing the latter, and are sufficient for recovering the change in the price index in the special case $\frac{\partial \log \chi_\tau^b(c_\tau)}{\partial \log c_\tau} \equiv 1$, which corresponds either to the case of income-dependent preferences, or to the knife-edge alternative in which the patterns of price inflation are orthogonal to the variations in income elasticity across products.

To empirically implement the strategy of Jaravel and Lashkari (2024), we need cross-sectional household-level data. In our main application of the import price index, unfortuantely we do not have access to such data and thus we have followed most of the prior work in choosing an income-invariant specification of the aggregate import demand.

A.3 Characterization of the Price Index Across Product Categories

A.3.1 General Results

Following the construction in Section 2.4.3, we can decompose the variations in the price index along the smooth paths of prices and qualities constructed in Section 2.1 between each two consecutive periods. It is easy to see that the instantaneous change in the price index can be decomposed into

$$d\log P_{\tau} = \sum_{k\in K} s_{\tau}^{k}\, d\log P_{\tau}^{k}, \qquad \tau\in(t-1,t)\,, \tag{A.17}$$

where $s_{\tau}^{k} \equiv \sum_{i\in I^{k}} s_{i\tau}$ is the total share of products in category k, and where we have defined the instantaneous change in the within-category price index as

$$d\log P_{\tau}^{k} = \sum_{i\in I^{k}} \frac{s_{i\tau}}{s_{\tau}^{k}}\left(d\log p_{i\tau} - d\varphi_{it}\right), \qquad \tau\in(t-1,t)\,. \tag{A.18}$$

Approximate integration of Equation (A.17) allows to write the change in the price index as in Equation (25).

A.3.2 Separable Demand Structure

Assume that the price index is separable across different product categories I^{k} defined in Section 2.4.3, in the sense that we can write the price index as

$$P\left(\boldsymbol{p}^{\varphi}\right) = \mathcal{P}\left(P^{1}\left(\boldsymbol{p}^{\varphi,1}\right),\cdots,P^{K}\left(\boldsymbol{p}^{\varphi,K}\right)\right), \tag{A.19}$$

for some 1st-degree homogeneous function $\mathcal{P}(\cdot)$ that constitutes the upper-level nest, and for a collection of K lower-level nests $P^{k}(\cdot)$, defined as a function of the vector of quality-adjusted prices for products in category k, that is, $\boldsymbol{p}^{\varphi,k} = \left(p_{i}^{\varphi}\right)_{i\in I^{k}}$. Applying Definition (1) to the upper-level price index function $\mathcal{P}(\cdot)$ and to the vector of price indices of categories $\boldsymbol{P} \equiv \left(P^{k}\right)$ where $P^{k} \equiv \mathcal{P}^{k}\left(\boldsymbol{p}^{k}\right)$, we can define the demand function for category k goods

$Q^k \equiv \partial \mathcal{P}/\partial P^k$, the corresponding expenditure-share function $\tilde{s}^k \equiv \partial \log \mathcal{P}/\partial \log P^k$, and cross-category elasticities of substitution $\sigma^C_{k\ell} \equiv (1/s^\ell)\partial \log Q^k/\partial \log P^\ell$. Similarly, applying the same definitions to the category-level price index function $P^k(\cdot)$, we define the within-category demand function for good i as $\tilde{q}^k_i \equiv \partial P^k/\partial p_i$, the corresponding expenditure-share function as $\tilde{s}^k_i \equiv \partial \log P^k/\partial \log p_i$, and within-category cross-category elasticities of substitution $\sigma^k_{ij} \equiv (1/\mathcal{S}^j)\partial \log Q^k_i/\partial \log p_j$. Accordingly, we define the cross-category and within-category/cross-product matrices of substitution elasticities $\mathbf{\Sigma}^N$ ($K \times K$-dimensional) and $\mathbf{\Sigma}^k$ ($|I^k| \times |I^k|$ dimensional), and the corresponding matrices $\mathbf{\Psi}^N$ and $\mathbf{\Psi}^k$, following Equation (4), where the expenditure share vectors are category-level shares and within-category-product-level shares, respectively. We similarly define the sets of available products I^k_t in period t, and the set of continuing products $I^{*,k}_t$ between periods $t-1$ and t. Finally, we assume that within each category, there exists a set O^k_t of products, whose quality on average remains constant from period $t-1$ to t. We let $\omega^k_t \equiv \frac{1}{|O^k_t|}\mathbb{I}\left\{i \in O^k_i\right\}$ denote the corresponding weights.

Using the above definitions, Proposition A.2 below shows that we only need within-category expenditure shares and cross-product elasticities of substitution to characterize the change in the category-level price indices.

Proposition A.2. *(Approximate Category-Specific Price Index for Separable Homothetic Preferences with Product Entry and Exit) Assume that the demand system is homothetic, satisfies the connected substitute property of Berry et al. (2013), the corresponding price index is continuously differentiable in prices and satisfies the separability condition in Equation (A.19), the elements of the inverse of the matrix $\mathbf{\Psi}^k$ defined as above are such that $\Psi^{k,-1}_{ij}/s_j$ remains everywhere bounded for all $i \neq j$, and that all products within the base set within category k continue from period $t-1$ to period t, that is $O^k_t \subset I^{k,*}_t$. Then, the change in the category-k log price index between periods $t-1$ and t can be approximated as*

$$\Delta \log P^k_t = \sum_i \omega^k_{it} \Delta \log p_{it} + \sum_i \omega^k_{it} \sum_{j \in I^{k,*}_t} \overline{\overline{\Psi^{k,-1}_{ij,t}}} \Delta \log s^*_{j\tau} + \sum_i \omega^k_{it} \sum_{j \in I^{k,*}_t} \overline{\overline{\Psi^{-1}_{ij,t}}} \Delta \log \Lambda^*_t$$
$$+ \left(\sum_i \omega^k_{it} \sum_{j \in I^k_t \setminus I^{k,*}_t} \Psi^{k,-1}_{ij,t} - \sum_i \omega^k_{it} \sum_{j \in I^k_{t-1} \setminus I^{k,*}_t} \Psi^{k,-1}_{ij,t-1} \right) + O\left(\left(\delta^k \right)^3 \right), \tag{A.20}$$

such that $\delta^k \equiv \max\left\{ \max_{i \in O^k_t} \{|\Delta \log p_{it}|\}, \max_{i \in V^*_t} \left\{ \left| \Delta \log s^{k,*}_{it} \right| \right\}, \left| \Delta \log \Lambda^{k,*}_t \right|, \max_{i \notin V^{k,*}_t} \left\{ \left| \Delta s^k_{it} \right|^{\frac{2}{3}} \right\} \right\}$.

Proof. See Appendix B.1 on page A27. □

A.4 Production-Based Price Import Index

In this section, we provide two sets of results. We first provide a number results on the roles of import (and export) price indices in measuring real GDP and welfare. Specifically, we show that a lower estimate for import prices (for instance due to quality adjustment) leads to lower estimates of real GDP growth while also raising our estimates of real consumption growth. Next, we use the same framework to offer a production-based derivation for the intermediate product components of the import price index.

A.4.1 GDP, Welfare, and Import-Export Price Indices

Following Diewert and Morrison (1986) (see also Kohli, 2004 and Dridi and Zieschang, 2004), we define the GDP function in terms of the vector of all quantity-adjusted prices $\boldsymbol{p}^{\varphi}$ and the vector of utilized factors $\boldsymbol{F}$ as

$$\mathcal{R}_t\left(\boldsymbol{p}^{\varphi};\boldsymbol{v}\right) \equiv \max_{\boldsymbol{q}^{\varphi}} \ \boldsymbol{p}^{\varphi,D}\cdot\boldsymbol{q}^{\varphi,D} + \boldsymbol{p}^{\varphi,X}\cdot\boldsymbol{q}^{\varphi,X} - \boldsymbol{p}^{\varphi,M}\cdot\boldsymbol{q}^{\varphi,M}, \qquad \text{such that} \quad \left(\boldsymbol{q}^{\varphi},\boldsymbol{F}\right)\in\Gamma_t, \tag{A.21}$$

where Γ_t indicates the technologically feasible set of quality-adjusted quantities and factor utilization, and where the set of all products $I = I^D \cup I^X \cup I^M$ is partitioned into three categories of domestically consumed final products I^D with the corresponding vectors of quality-adjusted quantities and prices $\boldsymbol{q}^{\varphi,D}$ and $\boldsymbol{p}^{\varphi,D}$, exported products I^X with the corresponding vectors $\boldsymbol{q}^{\varphi,X}$ and $\boldsymbol{p}^{\varphi,X}$, and imported intermediate products I^M with the corresponding vectors $\boldsymbol{q}^{\varphi,M}$ and $\boldsymbol{p}^{\varphi,M}$. Note that in this definition, the set of import products correspond to those imported products that are used as inputs in the production sector of the economy.

We again consider a setting in which we observe data on prices and quantities in each of the three sets, over a number of discrete time points $t \in \{0,\cdots,T-1\}$. In addition, we now also observe vectors of utilized factors $\boldsymbol{v}_t$, and assume that the set of feasible technologies Γ_t may also vary over time while remaining unobservable to us. Just like before, we also construct smooth paths along which the vector of quality-adjusted prices $\boldsymbol{p}^{\varphi}_{\tau}$, and now additionally the vectors of utilized factors $\boldsymbol{v}_{\tau}$ and the set of feasible technologies Γ_{τ}, smoothly evolve between each two time periods. Along such paths, we can write the growth in the nominal value of GDP as

$$
\begin{aligned}
d\log R_\tau = A_\tau\, d\tau + \sum_j s^j_{F,\tau}\, d\log F_{j\tau} \\
+ s^D_{G,\tau}\, d\log P^D_\tau + s^X_{G,\tau}\, d\log P^X_\tau - s^M_{G,\tau}\, d\log P^M_\tau,
\end{aligned}
\tag{A.22}
$$

where we have defined the effect of technological progress on GDP as $A_\tau \equiv \frac{\partial \log \mathcal{R}_\tau(\boldsymbol{p}^\varphi;\boldsymbol{F}_\tau)}{\partial \tau}$, the share of factor j in income as $S^j_{F,\tau} \equiv \frac{\partial \log \mathcal{R}_\tau(\boldsymbol{p}^\varphi;\boldsymbol{F}_\tau)}{\partial \log v_{j\tau}} = \frac{W_j F_{j\tau}}{R_\tau}$, the price index for the category k of products as

$$
d\log P^k_\tau \equiv \sum_{i\in I^k} \frac{p_{i\tau} q_{i\tau}}{\boldsymbol{p}^{\varphi,k}_\tau \cdot \boldsymbol{q}^{\varphi,k}_\tau}\, d\log p_{i\tau}, \qquad k \in \{D, X, M\}, \tag{A.23}
$$

and the share of the corresponding product category in GDP as $s^k_{G,\tau} \equiv \boldsymbol{p}^{\varphi,k}_\tau \cdot \boldsymbol{q}^{\varphi,k}_\tau / R_\tau$. We can define the real GDP growth as the component of the nominal GDP growth in Equation (A.22) that captures the contribution of changes in technology or factors used in production as[A1]

$$
d\log G_\tau \equiv d\log R_\tau - \underbrace{s^D_{G,\tau}\, d\log P^D_\tau + s^X_{G,\tau}\, d\log P^X_\tau - s^M_{G,\tau}\, d\log P^M_\tau}_{\text{GDP deflator}}. \tag{A.24}
$$

Accordingly, following the same arguments as before, we can approximate the GDP deflator according to

$$
\Delta \log P_{G,t} = \overline{\overline{s^D_{G,t}}}\, \Delta\log P^D_t + \overline{\overline{s^X_{G,t}}}\, \Delta\log P^X_t - \overline{\overline{s^M_{G,t}}}\, \Delta\log P^M_t + O\left(\delta^3\right), \tag{A.25}
$$

where we have $\delta \equiv \max_k \left\{\left|\Delta \log P^k_t\right|\right\}$, where $\Delta\log P^k_t$ is the change in the log price index corresponding to category k. Equation (A.25) shows that a rise in the import price index lowers the GDP deflator and thus, assuming a fixed growth in nominal GDP, leads to a rise in the GDP growth. We can use the same approach as that offered in the preceding sections to approximate the import price index in Equation (A.25).

As shown by Kohli (2004) (see also Kehoe and Ruhl, 2008, Oulton, 2023, and Reinsdorf, 2010), the notion of real GDP is not a satisfactory concept for the measurement of aggregate real income. In particular, it leads to counterintuitive results when evaluting

[A1] It is easy to see that the real GDP function defined in Equation (A.24) corresponds to the value of the GDP function for a fixed vector of prices $\boldsymbol{p}_b$, that is, $G_\tau \equiv \mathcal{R}_\tau\left(\boldsymbol{p}_b;\boldsymbol{F}_\tau\right)$, when the fixed vector of prices is evaluated (after taking the partial derivatives with respect to time) at the current value of the vector of prices as $\boldsymbol{p}_b = \boldsymbol{p}_\tau$. As such the growth in the GDP function only evaluates the contributions of changes in technology and factors of production in the value of production, keeping the prices as constant.

the effect of terms-of-trade shocks on real income. Since real GDP evaluates the new combination of imports and exports under a fixed set of prices, it can fall when the relative price of imports falls simply due the resulting rise on imports. This is in contrast to real consumption, which rises under this scenario since rising imports at lower prices bring a higher value to households. To capture this channel, let us again consider the income-independent (homothetic) preferences for households following Section 2.1, which implies the following cardinalized measure of welfare

$$U_t = \mathcal{V}\left(u_t\right) = \frac{R_t\left(\boldsymbol{p}_t^{\varphi}; \boldsymbol{F}_t\right)\left(1 + D_t\right)}{\mathcal{P}\left(\boldsymbol{p}_t^{\varphi}\right)}, \tag{A.26}$$

where D_t stands for the trade deficit (as a share of GDP), and where $\mathcal{P}\left(\boldsymbol{p}_t^{\varphi}\right)$ is the consumer price index. In Equation (A.26), we have slightly abused notation by using $\boldsymbol{p}_t^{\varphi}$ to indicate the entire vector of quality-adjusted products, including final and intermediate, domestic and imported products. The instantaneous change in the measure of welfare defined by Equation (A.26) is given by

$$\begin{aligned} d\log U_\tau = d\log G_\tau + \tfrac{D_\tau}{1+D_\tau} s_{G,\tau}^{D}\, d\log P_\tau^{D} + s_{G,\tau}^{X}\, d\log P_\tau^{X} \\ - \tfrac{1}{1+D_\tau} s_{G,\tau}^{F}\, d\log P_\tau^{F} - s_{G,\tau}^{M}\, d\log P_\tau^{M} + \tfrac{1}{1+D_\tau} dD_\tau, \end{aligned}$$

where the price indices P_τ^{j} for $j \in \{D, X, M\}$ are defined following the production-side definition in Equation (A.23), and where the change in the price index of imported final goods $d\log P_\tau^{F}$ can be defined using definition in Equation (A.18). Approximate integration of this result leads to the following second-order approximation

$$\begin{aligned} \Delta\log U_t = \Delta\log G_t + \overline{\overline{\left(\tfrac{D_t}{1+D_t} s_{G,t}^{D}\right)}}\, \Delta\log P_t^{D} + \overline{\overline{s_{G,t}^{X}}}\, \Delta\log P_t^{X} \\ - \overline{\overline{\left(\tfrac{1}{1+D_\tau} s_{G,\tau}^{F}\right)}}\, \Delta\log P_t^{F} - \overline{\overline{s_{G,t}^{M}}}\, \Delta\log P_t^{M} + \overline{\overline{\left(\tfrac{1}{1+D_t}\right)}} \Delta D_t + O\left(\delta^3\right), \end{aligned} \tag{A.27}$$

where $\delta \equiv \max\left\{\max_{k\in\{D,X,F,M\}}\left\{\left|\Delta\log P_t^{k}\right|\right\}, \Delta D_t\right\}$. Equation (A.27) shows that, keeping domestic technology and factor inputs as fixed, lowering the prices of final or intermediate imported goods leads to welfare gains. Moreover, in the special case of balanced trade $D \equiv 0$, Equation (A.27) simplifes to the standard result that decomposes changes of welfare into changes in domestic technology, factor inputs, and terms-of-trade shocks

$$\Delta\log U_t = \Delta\log G_t + \overline{\overline{s_{G,t}^{X}}}\, \Delta\log P_t^{X} - \overline{\overline{s_{G,t}^{Imp}}}\, \Delta\log P_t^{Imp} + O\left(\delta^3\right),$$

where $s_{G,t}^{Imp}$ is the total share of all imported goods in GDP, and where the price index of imports satisfies $\Delta \log P_t^{Imp} \equiv \overline{\overline{\left(s_{G,\tau}^F / s_{G,t}^{Imp}\right)}} \Delta \log P_t^F + \overline{\overline{\left(s_{G,t}^M / s_{G,t}^{Imp}\right)}} \Delta \log P_t^M$.

A.4.2 Intermediate Import Demand

Our construction for the import price indices following the setting in Sections 2.1 and 2.4.3 relies on a consumption-based approach. We can use the production-based approach used in this section to provide an alternative derivation for the components of the import price index that are used as intermediate inputs in the production sector of the economy. We can use the definition of the GDP function in Equation (A.21), along with the envelope theorym, to derive the demand for an imported good $i \in I^D$ as $q_{it} = \partial \mathcal{R}_t / \partial p_{it}$. However, unlike our construction in Section 2, this specification does not explicitly account for the potential substitutability between domestic and imported varieties of intermediate products. To derive a specification of import demand that allows us to explicitly account for the substitutability between domestic and imported goods, we need to specify input-output structure of the domestic production technology Γ_t in terms, not only of the factors used and products produced, but also of intermediate products used in the process. The following lemma sets up an example of the types of such input-output structures that are compatible with the homothetic demand systems considered in Sections 2.1 and 2.4.3.

Proposition A.3. *(Intermediate Import Price Index) Let $I^I \equiv I^H \cup I^M$denote the set of all intermediate products, including domestically produced set I^H and the imported intermediate inputs I^M. Assume a collection of price taking production units $n \in \mathcal{N}$ with production technologies characterized by*

$$y_n = A_t Z_n \left[\mathcal{F} \left(\boldsymbol{F} \right)^{1-\alpha} \mathcal{F}_n^I \left(\boldsymbol{q}_n^{\varphi,I} \right)^{\alpha} \right]^{\gamma}, \qquad 0 < \alpha < 1, \gamma < 1,$$

where $y_n = q_i^n$ is the output of unit n in a specific product $i \in I^D \cup I^X$ produced domestically by this unit, and where $\mathcal{F}(\cdot)$ and $\mathcal{F}_n^I(\cdot)$ are constant-returns-to-scale aggregators of factor and material inputs, respectively, with the latter potentially varying by unit. Under the above assumptions, the demand for intermediate goods I^I can be characterized by a homothetic demand system defined in Section 2.1, in the sense that the expenditure share $\mathcal{S}_i^I \left(\boldsymbol{p}^{\varphi,I} \right)$ of product i is only a function of quality-adjusted vector of product prices $\boldsymbol{q}_n^{\varphi,I}$ and is characterized by a scale-invariant demand system as that characterized in Section 2.2.

Proof. See Appendix B.1 on page A29. □

A.5 More Details on Homothetic with Aggregator Family

A.5.1 Local Demand Inversion

Another important feature of the H(S/I)A demand systems presented in Definition 1 is that we can analytically invert the matrix linking relative prices to relative expenditure shares. The following lemma characterizes this inverted matrix.

Lemma A.4. *The elements of the inverse of the matrix $\left[\Psi_{ij}\right]$ mapping relative prices to relative expenditure shares for the HA demand systems introduced in Definition 1 are given by*

$$\Xi_{ij}^{-1} = \mu_i \mathbb{I}_{ij} - \begin{cases} s_j \frac{1+\mu_i}{1+\overline{\mu}} \left(\mu_j - \overline{\mu}\right), & HDIA, \\ s_j \frac{\mu_i}{\overline{\mu}} \left(\mu_j - \overline{\mu}\right), & HIIA, \\ s_j \left(\mu_j - \overline{\mu}\right), & HSA, \end{cases} \tag{A.28}$$

where $\boldsymbol{\Xi} \equiv \boldsymbol{\Psi}\left(\boldsymbol{\Sigma}^d - \boldsymbol{I}\right)$, where $\mathbb{I}_{ij} = 1$ if $i = j$ and $\mathbb{I}_{ij} = 0$ otherwise, and where the love-of-variety index μ_i for each product is defined by $\mu_i \equiv \frac{1}{\varepsilon_i - 1}$ with ε_i given by Equation (12).

Proof. See Appendix B.1 on page A31. □

A.5.2 Approximate Price Index with Product Entry/Exit

Moreover, we use Lemma A.4 to specialize the general expression for the approximate price index for homothetic preferences in Proposition 3 to the family of HA preferences. Proposition A.4 below states this result.

Proposition A.4. *(Approximate Price Index for H(S/I)A Preferences) For the three families of HA demand systems introduced in Definition 1, assuming all products within the base set continue from period $t-1$ to period t, that is $O_t \subset V_t^*$, we have the following approximation for the change in the price index*

$$\begin{aligned} \Delta \log P_t = {} & \sum_i \varpi_{ti} \Delta \log p_{it} + \sum_i \varpi_{it} \overline{\overline{\frac{1}{\varepsilon_{it}-1}}} \Delta \log s_{it} - \sum_i \overline{\overline{\Lambda_t^* s_{it}^* \bar{\iota}_t^o \left(\frac{1}{\varepsilon_{it}-1} - \overline{\left(\frac{1}{\varepsilon_{it}-1}\right)}\right)}} \Delta \log s_{it}^* \\ & + \left(\sum_i \varpi_{it} \overline{\overline{\frac{1}{\varepsilon_{it}-1}}}\right) \Delta \log \Lambda_t^* - \Delta \left(\bar{\iota}_t^o \left(\overline{\frac{1}{\varepsilon_{it}-1}} - \overline{\frac{1}{\varepsilon_{it}-1}}^*_t\right)\right) + O\left(\delta^3\right), \end{aligned} \tag{A.29}$$

where we have defined $\overline{\frac{1}{\varepsilon_{it}-1}} \equiv \sum_i s_{it} \frac{1}{\varepsilon_{it}-1}$, $\overline{\frac{1}{\varepsilon_{it}-1}}^* \equiv \sum_i s_{it}^* \frac{1}{\varepsilon_{it}-1}$, $\bar{\iota}_t^o \equiv \sum_i \varpi_{it} \iota_{it}$, , and $\bar{\iota}_{t-1}^o \equiv \sum_i \varpi_{it} \iota_{it-1}$, with $\iota_{it} \equiv \frac{1+1/(\varepsilon_{it}-1)}{1+\overline{\overline{1/(\varepsilon_{it}-1)}}}$ for HDIA, $\iota_{it} \equiv \frac{1/(\varepsilon_{it}-1)}{\overline{\overline{1/(\varepsilon_{it}-1)}}}$ for HIIA, and $\iota_{it} \equiv 1$ for HSA, and *where δ is defined as in Proposition 3. The contribution of quality change $\Delta \log \Phi_t$ in the set of continuing products, and product entry and exit $\Delta \log X_t$ satisfy:*

$$\Delta \log \Phi_t = \sum_i \left(\overline{\overline{s_{it}^*}} - \omega_{it} \right) \Delta \log p_{it} + \sum_{j \in I_t^*} \left(\overline{\overline{s_{it}^*}} - \omega_{it} \right) \overline{\overline{\frac{1}{\varepsilon_{it}-1}}} \left(\Delta \log s_{jt}^* + \Delta \log \Lambda_t^* \right)$$
$$- \sum_i \overline{\overline{\Lambda_t^* s_{it}^* \left(\bar{\iota}_t^* - \bar{\iota}_t^o \right) \left(\frac{1}{\varepsilon_{it}-1} - \overline{\left(\frac{1}{\varepsilon_{it}-1} \right)} \right)}} \Delta \log s_{it}^* - \Delta \left(\left(\bar{\iota}_t^* - \bar{\iota}_t^o \right) \left(\overline{\frac{1}{\varepsilon_{it}-1}} - \overline{\frac{1}{\varepsilon_{it}-1}}_t^{\,*} \right) \right), \tag{A.30}$$

$$\Delta \log X_t = \sum_{j \in I_t^*} \overline{\overline{s_{it}^*}} \overline{\overline{\frac{1}{\varepsilon_{it}-1}}} \left(\Delta \log s_{jt}^* + \Delta \log \Lambda_t^* \right) - \sum_i \overline{\overline{\Lambda_t^* s_{it}^* \bar{\iota}_t^* \left(\frac{1}{\varepsilon_{it}-1} - \overline{\left(\frac{1}{\varepsilon_{it}-1} \right)} \right)}} \Delta \log s_{it}^* \tag{A.31}$$
$$- \Delta \left(\bar{\iota}_t^* \left(\overline{\frac{1}{\varepsilon_{it}-1}} - \overline{\frac{1}{\varepsilon_{it}-1}}_t^{\,*} \right) \right) - \Delta \left(\mathbb{E}_i^{s_t^*} \left[\iota_{it} \right] \left(\bar{\mu}_t - \bar{\mu}_t^* \right) \right), \tag{A.32}$$

where $\bar{\iota}_t^* \equiv \sum_i s_{it}^* \iota_{it}$.

Proof. See Appendix B.1 on page A35. □

Finally, we can provide a decomposition of the gap in the inferred price index under the HA family and under a CES specification. Assume that the true underlying preferences belong to a member of the HA family, in line with Proposition A.4. Assume now that we have used a misspecified CES model, estimated to have an elasticity of substitution $\hat{\sigma}$. The gap between the change in the true price index and that implied by the estimated CES model is given by

$$\Delta \log P_t - \Delta \log \widehat{P}_t^{CES} = \overbrace{\left(\sum_i \omega_{it} \overline{\overline{\mu_{it}}} - \frac{1}{\hat{\sigma}-1} \right) \left(\sum_i \omega_{it} \Delta \log s_{i\tau}^* + \Delta \log \Lambda_t^* \right)}^{\text{Gap in mean elasticities}}$$
$$\underbrace{+ Cov_i^{\varpi_t} \left(\overline{\overline{\mu_{it}}}, \Delta \log s_{i\tau}^* \right) - \sum_i \overline{\overline{\Lambda_t^* s_{it}^* \bar{\iota}_t^* \left(\frac{1}{\varepsilon_{it}-1} - \overline{\left(\frac{1}{\varepsilon_{it}-1} \right)} \right)}} \Delta \log s_{it}^* - \Delta \left(\bar{\iota}_t^* \left(\overline{\frac{1}{\varepsilon_{it}-1}} - \overline{\frac{1}{\varepsilon_{it}-1}}_t^{\,*} \right) \right)}_{\text{Heterogeneity in cross-product elasticities}}, \tag{A.33}$$

where $Cov_i^{\varpi_t}(\cdot, \cdot)$ denotes the covariance operator under the distribution ϖ_t. The term on the first line accounts for the contribution of the gap between the mean of the love of variety indices across base products and the one implied by the CES model, and the second line accounts for the contribution of heterogeneity in the matrix of cross-product elasticities of substitution, which is absent in the CES model.

A.5.3 Specifications for the Kimball Aggregator

We recover standard CES preferences by choosing Kimball function $K(\check{q};\varsigma) \equiv \check{q}^{1-1/\sigma}$ in Equation (**??**) with the corresponding choice of parameterization $\varsigma \equiv (\sigma)$. Below, we consider three additional parametric families of Kimball functions $\mathcal{K}(\cdot;\varsigma)$, each characterized by a corresponding family of elasticity functions $\widetilde{e}(\cdot;\varsigma)$, defined through the dual function $e(\check{q};\varsigma) \equiv 1/\widetilde{e}\left((K')^{-1}(\check{q})\right)$ as follows.

1. *Klenow and Willis (2006).* This case involves an elasticity function

$$e(\check{q};\varsigma) \equiv \frac{\check{q}^{\theta}}{\sigma}, \qquad \varsigma \equiv (\sigma, \theta), \tag{A.34}$$

   that goes from zero (corresponding to infinite price elasticity) to infinity as the normalized quantity goes from zero to infinity.

2. *Finite–Infinite Limits:* This case involves an elasticity function

$$e(\check{q};\varsigma) \equiv \frac{1}{\sigma + (\sigma_o - \sigma)\,\check{q}^{-\theta}}, \qquad \sigma < \sigma_o,\ \theta > 0,\ \varsigma \equiv (\sigma, \sigma_o, \theta), \tag{A.35}$$

   that goes from zero (corresponding to infinite price elasticity) to a finite value $1/\sigma$ as the normalized quantity goes from zero to infinity.

3. *Finite–Finite Limits:* This case involves an elasticity function

$$e(\check{q};\varsigma) \equiv \frac{1}{\sigma_o} + \left(\frac{1}{\sigma} - \frac{1}{\sigma_o}\right)\frac{\check{q}^{\theta}}{1+\check{q}^{\theta}}, \qquad \sigma < \sigma_o,\ \theta > 0,,\ \varsigma \equiv (\sigma, \sigma_o, \theta), \tag{A.36}$$

   that goes from a finite value $1/\sigma_o$ to another finite value $1/\sigma$ as the normalized quantity goes from zero to infinity.[A2]

Appendix B.2.1 below derives the Kimball functions $K(\cdot;\varsigma)$ corresponding to each of the three cases above.

A.6 Comparison of the Dynamic Panel Identification with Feenstra (1994)

In this section, we provide a brief comparison of the conceptual distinction between our approach and that of Feenstra (1994), which in turn builds on earlier insights of Leamer

[A2] In the first and the last cases, the marginal utility of consuming every product at a zero level of consumption ($\widetilde{q}_i = 0$) is infinity. Therefore, the demand takes a finite, nonzero value for every finite value of price. In contrast, in the second case, the marginal utility of consuming every product at a zero level of consumption ($\widetilde{q}_i = 0$) is finite. As a result, there is a finite choke price for any product, above which the consumption falls to zero.

(1981). For this purpose, let us consider a CES demand specification presented in Section 2.3.1, which leads to the following simple specification of demand

$$\Delta \log \widehat{q}_{it} = -\sigma \Delta \log \widehat{p}_{it} + \Delta \varphi_{it},$$

where we have defined log quantity and price relative to the base product $\widehat{q}_{it} \equiv q_{it}/q_{ot}$ and $\widehat{p}_{it} \equiv p_{it}/p_{ot}$ in a simple setting where the set of base products is a singleton $O \equiv \{o\}$, and where, as before, φ_{it} stands for the demand shock. The Leamer–Feenstra approach to identification begins with positing a supply relationship of the form

$$\Delta \log \widehat{p}_{it} = \zeta \log \Delta \widehat{q}_{it} + \Delta \xi_{it}, \tag{A.37}$$

where $\zeta > 0$ stands for the supply elasticity. The first key identification assumption is that the supply and demand shocks are uncorrelated $\mathbb{E}\left[\Delta \xi_{it} \Delta \varphi_{it}\right] = 0$. If we know the supply elasticity ζ, then this assumption leads to a synthetic instrument $z_{it}^{F-L}(\zeta) \equiv \Delta \log \widehat{p}_{it} - \zeta \log \Delta \widehat{q}_{it}$ that allows us to identify σ through the moment condition

$$\mathbb{E}\left[\left(\Delta \log \widehat{q}_{it} + \sigma \Delta \log \widehat{p}_{it}\right) \times z_{it}^{F-L}(\zeta)\right] = 0. \tag{A.38}$$

As shown in Feenstra (2010), the second key identification assumption is that there exists at least two products i and j for which the ratio of the variances of demand schock and supply shocks are not identical ($\mathbb{V}\left[\Delta \varphi_{it}\right]/\mathbb{V}\left[\Delta \xi_{it}\right] \neq \mathbb{V}\left[\Delta \varphi_{jt}\right]/\mathbb{V}\left[\Delta \xi_{jt}\right]$).[A3] We can think of the role of this additional *identification by heteroskedasticity* assumption as that of identifying the supply elasticity ζ, which would then enable condition (A.38) to identify the price elasticity of demand σ. In practice, the estimation strategy combines these identification assumptions to simultaneously estimate both ζ and σ.

Now, let us compare Equation (A.37) with our pricing Equation (21). Assuming small relative changes in all variables, we can write the change in log price in terms of the change in log quantity and other variables as:

$$\Delta \log p_{it} \approx \underbrace{\frac{\frac{\partial \log mc_{it}}{\partial \log q_{it}} + \frac{\partial \log \mu_{it}}{\partial \log q_{it}}}{1 - \frac{\partial \log \mu_{it}}{\partial \log p_{it}}}}_{\equiv \zeta_{it}} \Delta \log q_{it} + \underbrace{\frac{\frac{\partial \log mc_{it}}{\partial \varphi_{it}} + \frac{\partial \log \mu_{it}}{\partial \varphi_{it}}}{1 - \frac{\partial \log \mu_{it}}{\partial \log p_{it}}} \Delta \varphi_{it} + \frac{\frac{\partial \log mc_{it}}{\partial w_{it}}}{1 - \frac{\partial \log \mu_{it}}{\partial \log p_{it}}} \Delta w_{it} + \Delta v_{it}}_{\equiv \Delta \xi_{it}}.$$

We can make two observations. First, in general the supply elasticity may vary over time

[A3]See also Soderbery (2015b) for a detailed discussion of how this condition helps identify the elasticities using specific examples from trade data.

and across products. Second, and more importantly, there are two potential grounds for the violations of the Leamer–Feenstra identification assumption $\mathbb{E}\left[\Delta\xi_{it}\Delta\varphi_{it}\right]=0$. First, to the extent that marginal cost depends on quality, i.e., $\frac{\partial \log mc_{it}}{\partial \varphi_{it}} \neq 0$, there is a mechanical correlation between supply shocks $\Delta\xi_{it}$ and demand shocks $\Delta\varphi_{it}$. In addition, to the extent that shocks to production costs Δw_{it} leads to endogenous responses in product quality, we find another potential source of correlation between supply and demand shocks.

In contrast, our approach begins by assuming a simple dynamic process like that of Equation (17) on demand shocks. The same pricing Equation (21) now implies that $\mathbb{E}\left[\Delta u_{it} \log p_{it-2}\right]$, which leads to the following moment condition:

$$\mathbb{E}\left[\left(\Delta \log \widehat{q}_{it} + \sigma \Delta \log \widehat{p}_{it} - \rho \left(\Delta \log \widehat{q}_{it-1} + \sigma \Delta \log \widehat{p}_{it-1}\right)\right) \times \log p_{it-2}\right] = 0.$$

If we know ρ, the term $\rho\left(\Delta \log \widehat{q}_{it-1} + \sigma \Delta \log \widehat{p}_{it-1}\right)$ gives us a control function that accounts for the potential persistence between lagged price and current change in demand shocks, allowing us to identify the price elasticity σ. To recover the persistence parameter ρ, the same Equation (17) also implies that $\mathbb{E}\left[\Delta u_{it}\varphi_{it-2}\right]$ leading to another moment condition

$$\mathbb{E}\left[\left(\Delta \log \widehat{q}_{it} + \sigma \Delta \log \widehat{p}_{it} - \rho \left(\Delta \log \widehat{q}_{it-1} + \sigma \Delta \log \widehat{p}_{it-1}\right)\right) \times \varphi_{it-2}\right] = 0.$$

Just like the Leamer–Feenstra approach, we also combine the moment conditions in a GMM framework to jointly estimate both σ and ρ.

To summarize, our approach averts the need to make the counterfactual assumption that marginal costs do not depend on product quality by relying on the panel structure of the data and imposing restrictions on the dynamics of demand shocks.

B Proofs and Derivations

B.1 Proofs

Proof for Proposition 1. This result is a special case of ProÏposition 3 for the case in which the set of continuing products is the same as the entire set of products, $I_t^* \equiv I$. Note that in the proof of Proposition 3, included below, the condition in the statement of the proposition that Ψ_{ij}^{-1}/s_j remains everywhere bounded for all $i \neq j$ is only used in deriving the expression for the contribution of product entry and exit. As such, this condition is not required in the special case of Proposition 1 without product entry and exit. □

Proof for Lemma 1. First, note that for any income-invariant demand system we have

$$\frac{\partial \log \widetilde{s}_i}{\partial \log p_j} = \frac{\partial \log \left(\frac{p_i \widetilde{q}_i}{P\,Q(u)}\right)}{\partial \log p_j} = s_j \left(\sigma_{ij} - 1\right), \qquad i \neq j. \tag{B.1}$$

Next, we use the definition in Equation (9), letting $A = \sum_i \check{p}_i d_i \left(\check{p}_i\right)$ and $s_i = \check{p}_i d_i \left(\check{p}_i\right) / A$, we find

$$\begin{aligned} \frac{\partial \log \widetilde{s}_i}{\partial \log p_j} &= \frac{\partial \log \left[\check{p}_i d_i \left(\check{p}_i\right)\right]}{\partial \log p_j} - \frac{\partial \log A}{\partial \log p_j}, \\ &= \left(\varepsilon_i - 1\right) \frac{\partial \log \widetilde{h}\left(\boldsymbol{p}^{\varphi}\right)}{\partial \log p_j} - \frac{\partial \log A}{\partial \log p_j}, \qquad i \neq j, \end{aligned} \tag{B.2}$$

where we have used the fact that $\frac{\partial \log \check{p}_i}{\partial \log p_j} = -\frac{\partial \log \widetilde{h}(\boldsymbol{p}^{\varphi})}{\partial \log p_j}$ for $i \neq j$. Next, we compute the elasticity of the aggregator A with respect to price p_j :

$$\begin{aligned} \frac{\partial \log A}{\partial \log p_j} &= \frac{\partial \log \left[\sum_{i'} \check{p}_{i'} d_{i'} \left(\check{p}_{i'}\right)\right]}{\partial \log p_j}, \\ &= \frac{\sum_{i' \neq j} \check{p}_{i'} d_{i'} \left(\check{p}_{i'}\right) \frac{\partial \log\left[\check{p}_{i'} d_{i'}\left(\check{p}_{i'}\right)\right]}{\partial \log p_j} + \check{p}_j d_j \left(\check{p}_j\right) \frac{\partial \log\left[\check{p}_j d_j\left(\check{p}_j\right)\right]}{\partial \log p_j}}{\sum_{i'} \check{p}_{i'} d_{i'} \left(\check{p}_{i'}\right)}, \\ &= \frac{\sum_{i' \neq j} \check{p}_{i'} d_{i'} \left(\check{p}_{i'}\right) \left(\varepsilon_{i'} - 1\right) \frac{\partial \log \widetilde{h}}{\partial \log p_j} + \check{p}_j d_j \left(\check{p}_j\right) \left(\varepsilon_j - 1\right) \left(\frac{\partial \log \widetilde{h}(\boldsymbol{p}^{\varphi})}{\partial \log p_j} - 1\right)}{\sum_{i'} \check{p}_{i'} d_{i'} \left(\check{p}_{i'}\right)}, \\ &= \frac{\sum_{i'} \check{p}_{i'} d_{i'} \left(\check{p}_{i'}\right) \left(\varepsilon_{i'} - 1\right)}{\sum_{i'} \check{p}_{i'} d_{i'} \left(\check{p}_{i'}\right)} \frac{\partial \log \widetilde{h}}{\partial \log p_j} - \frac{\check{p}_j d_j \left(\check{p}_j\right) \left(\varepsilon_j - 1\right)}{\sum_{i'} \check{p}_{i'} d_{i'} \left(\check{p}_{i'}\right)}, \\ &= \left(\bar{\varepsilon} - 1\right) \frac{\partial \log \widetilde{h}}{\partial \log p_j} - s_j \left(\varepsilon_j - 1\right). \end{aligned}$$

Substituting this in Equation (B.2), we find

$$\frac{\partial \log \widetilde{s}_i}{\partial \log p_j} = s_j \left(\varepsilon_j - 1\right) + \left(\varepsilon_i - \bar{\varepsilon}\right) \frac{\partial \log \widetilde{h}}{\partial \log p_j}, \qquad i \neq j. \tag{B.3}$$

We now compute $\frac{\partial \log \widetilde{h}(\boldsymbol{p}^{\varphi})}{\partial \log p_j}$ for each of the three specifications in Definition 1. For HSA, from $A = 1$, we have

$$0 = \frac{\partial \log A}{\partial \log p_j} = \left(\bar{\varepsilon} - 1\right) \frac{\partial \log \widetilde{h}}{\partial \log p_j} - s_j \left(\varepsilon_j - 1\right),$$

leading to

$$\frac{\partial \log \widetilde{h}}{\partial \log p_j} = s_j \frac{\varepsilon_j - 1}{\bar{\varepsilon} - 1}.$$

Combining this result with Equations (B.1) and (B.3), we find

$$s_j \left(\varepsilon_j - 1 + (\varepsilon_i - \bar{\varepsilon}) \frac{\varepsilon_j - 1}{\bar{\varepsilon} - 1} \right) = s_j \left(\sigma_{ij} - 1 \right),$$

which leads to the desired result in Equation (11) for the case of HSA.

For the choice of HDIA demand, based on Equation (10), we can derive the price derivatives of the aggregator $\widetilde{h}$ as follows

$$\begin{aligned}
0 &= \frac{\partial \widetilde{h}}{\partial p_j} \sum_i \frac{\partial}{\partial h} \left[\int_0^{d_i\left(\frac{e^{-\varphi_i} p_i}{h}\right)} d_i^{-1}(v)\, dv \right] + \frac{\partial}{\partial p_j} \left[\int_0^{d_j\left(\frac{e^{-\varphi_j} p_j}{H}\right)} d_j^{-1}(v)\, dv \right], \\
&= \frac{\partial \widetilde{h}}{\partial p_j} \sum_i \frac{\partial d_i \left(\frac{e^{-\varphi_i} p_i}{h} \right)}{\partial h} \frac{e^{-\varphi_i} p_i}{h} + \frac{\partial d_j \left(\frac{e^{-\varphi_j} p_j}{h} \right)}{\partial p_j} \frac{e^{-\varphi_j} p_j}{h}, \\
&= -\frac{\partial \widetilde{h}}{\partial p_j} \sum_i \frac{e^{-\varphi_i} p_i}{h^2} d_i' \left(\frac{e^{-\varphi_i} p_i}{h} \right) \frac{e^{-\varphi_i} p_i}{h} + \frac{e^{-\varphi_j}}{h} d_j' \left(\frac{e^{-\varphi_j} p_j}{h} \right) \frac{e^{-\varphi_j} p_j}{h},
\end{aligned}$$

leading to the following result

$$\begin{aligned}
\frac{\partial \log \widetilde{h}}{\partial \log p_j} &= \frac{\frac{e^{-\varphi_j} p_j}{h} d_j \left(\frac{e^{-\varphi_j} p_j}{h} \right) \times \frac{\frac{e^{-\varphi_j}}{h} d_j' \left(\frac{e^{-\varphi_j} p_j}{h} \right)}{d_j \left(\frac{e^{-\varphi_j} p_j}{h} \right)}}{\sum_i \frac{e^{-\varphi_i} p_i}{h} d_i \left(\frac{e^{-\varphi_i} p_i}{h} \right) \times \frac{\frac{e^{-\varphi_i}}{h} d_i' \left(\frac{e^{-\varphi_i} p_i}{h} \right)}{d_j \left(\frac{e^{-\varphi_i} p_i}{h} \right)}}, \\
&= \frac{s_j \varepsilon_j}{\bar{\varepsilon}},
\end{aligned}$$

Combining this result with Equations (B.1) and (B.3), we find

$$s_j \left(\varepsilon_j - 1 + (\varepsilon_i - \bar{\varepsilon}) \frac{\varepsilon_j}{\bar{\varepsilon}} \right) = s_j \left(\sigma_{ij} - 1 \right),$$

which leads to the desired result in Equation (11) for the case of HDIA.

For the HIIA specification, based on Equation (10), we can derive the price derivatives of the aggregator $\widetilde{h}$ as follows

$$\begin{aligned}
0 &= \frac{\partial \widetilde{h}}{\partial p_j} \sum_i \frac{\partial}{\partial h} \left[\int_0^{\frac{e^{-\varphi_i} p_i}{h}} d_i\left(v\right) dv \right] + \frac{\partial}{\partial p_j} \left[\int_0^{\frac{e^{-\varphi_j} p_j}{H}} d_j\left(v\right) dv \right], \\
&= \frac{\partial \widetilde{h}}{\partial p_j} \sum_i \frac{\partial \left(\frac{e^{-\varphi_i} p_i}{h} \right)}{\partial h} d_i \left(\frac{e^{-\varphi_i} p_i}{h} \right) + \frac{\partial \left(\frac{e^{-\varphi_j} p_j}{h} \right)}{\partial p_j} d_j \left(\frac{e^{-\varphi_j} p_j}{h} \right), \\
&= -\frac{\partial \widetilde{h}}{\partial p_j} \sum_i \frac{e^{-\varphi_i} p_i}{h^2} d_i \left(\frac{e^{-\varphi_i} p_i}{h} \right) \frac{e^{-\varphi_i} p_i}{h} + \frac{e^{-\varphi_j}}{h} d_j \left(\frac{e^{-\varphi_j} p_j}{h} \right),
\end{aligned}$$

leading to the following result

$$\begin{aligned}
\frac{\partial \log \widetilde{h}}{\partial \log p_j} &= \frac{\frac{e^{-\varphi_j} p_j}{h} d_j \left(\frac{e^{-\varphi_j} p_j}{h} \right)}{\sum_i \frac{e^{-\varphi_i} p_i}{h} d_i \left(\frac{e^{-\varphi_i} p_i}{h} \right)}, \\
&= s_j,
\end{aligned}$$

Combining this result with Equations (B.1) and (B.3), we find

$$s_j \left(\varepsilon_j - 1 + \varepsilon_i - \bar{\varepsilon} \right) = s_j \left(\sigma_{ij} - 1 \right),$$

which leads to the desired result in Equation (11) for the case of HIIA. □

Proof for Proposition 3. First, we consider the paths of quality-adjusted prices along the two consecutive periods $t-1$ and t as constructed as in Appendix A.1.1 in the general case with possible product entry/exit. The instantaneous change in the price index $d \log P_\tau$ over the period $\tau \in (t, t-1)$ along such paths is characterized in Lemma A.2 in Appendix A.1.2. To derive the desired result in Proposition 3, we approximately integrate the instantaneous change in the price index $d \log P_\tau$ over the period $\tau \in (t, t-1)$ along the paths of quality-adjusted prices.

To apply the approximate integration, we use the following standard result on the error of the first and second order integration rules:

$$I \equiv \int_{v_{t-1}}^{v_t} f\left(v\right)\, dv = \sum_j f'\left(v_{t-1}\right) \left(v_t - v_{t-1}\right) + \frac{1}{2} f''\left(v_1^\dagger\right) \left(v_t - v_{t-1}\right)^2, \tag{B.4}$$

$$= \sum_j f'\left(v_t\right) \left(v_t - v_{t-1}\right) - \frac{1}{2} f''\left(v_2^\dagger\right) \left(v_t - v_{t-1}\right)^2, \tag{B.5}$$

$$= \sum_j \tfrac{1}{2} \left(f\left(v_{t-1}\right) + f\left(v_{t-1}\right)\right)\left(v_t - v_{t-1}\right) - \frac{1}{12} f'''\left(v_3^\dagger\right)\left(v_t - v_{t-1}\right)^3, \tag{B.6}$$

for some $v_1^\dagger, v_2^\dagger, v_3^\dagger \in [v_{t-1}, v_t]$. From this, it then follows that $I = f\left(v_{t-1}\right)\Delta v_t + O\left(\left|\Delta v_t\right|^2\right) = f\left(v_t\right)\Delta v_t + O\left(\left|\Delta v_t\right|^2\right) = \overline{\overline{f\left(v_t\right)}}\Delta v_t + O\left(\left|\Delta v_t\right|^3\right)$. Now, we integrate Equation (A.3) from Lemma A.2 in Appendix A.1.2 and apply Equation (B.6) to find

$$\begin{aligned}
\Delta \log P_t &= \int_{t-1}^{t} d\log P_\tau, \\
&= \sum_{i\in O_t} \omega_{it} \int_{t-1}^{t} d\log p_{i\tau} + \sum_{i\in O_t}\sum_{j\in I_t^*} \omega_{it} \int_{t-1}^{t} \tfrac{1}{\sigma_{ii,\tau}-1}\, \Psi_{ij,\tau}^{-1}\, d\log s_{j\tau}^* \\
&\quad + \sum_{i\in O_t}\sum_{j\in I_t^*} \omega_{it} \int_{t-1}^{t} \tfrac{1}{\sigma_{ii,\tau}-1}\, \Psi_{ij,\tau}^{-1}\, d\log \Lambda_\tau^* \\
&\quad + \sum_{i\in O_t}\sum_{j\in I_{t-1}\setminus I_t^*} \omega_{it} \int_{t-1}^{t} \tfrac{1}{\sigma_{ii,\tau}-1}\, \frac{\Psi_{ij,\tau}^{-1}}{s_{j\tau}}\, ds_{j\tau} \\
&\quad + \sum_{i\in O_t}\sum_{j\in I_t\setminus I_t^*} \omega_{it} \int_{t-1}^{t} \tfrac{1}{\sigma_{ii,\tau}-1}\, \frac{\Psi_{ij,\tau}^{-1}}{s_{j\tau}}\, ds_{j\tau}, \\
&\approx \sum_{i\in O_t} \omega_{it}\Delta\log p_{it} + \sum_{i\in O_t} \omega_{it} \sum_{j\in I_t^*} \tfrac{1}{2}\left(\tfrac{1}{\sigma_{ii,t-1}-1}\, \Psi_{ij,t-1}^{-1} + \tfrac{1}{\sigma_{ii,t}-1}\, \Psi_{ij,t}^{-1}\right)\Delta\log s_{j\tau}^* \\
&\quad + \sum_{i\in O_t} \omega_{it} \sum_{j\in I_t^*} \tfrac{1}{2}\left(\tfrac{1}{\sigma_{ii,t-1}-1}\, \Psi_{ij,t-1}^{-1} + \tfrac{1}{\sigma_{ii,t}-1}\, \Psi_{ij,t}^{-1}\right)\Delta\log \Lambda_t^* \\
&\quad + \sum_{i\in O_t}\sum_{j\in I_{t-1}\setminus I_t^*} \omega_{it} \tfrac{1}{\sigma_{ii,t-1}-1}\, \frac{\Psi_{ij,t-1}^{-1}}{s_{jt-1}}\left(-s_{jt-1}\right) + \sum_{i\in O_t}\sum_{j\in I_t\setminus I_t^*} \omega_{it} \tfrac{1}{\sigma_{ii,t}-1}\, \frac{\Psi_{ij,t}^{-1}}{s_{jt}} s_{jt} + O\left(\delta^3\right),
\end{aligned}$$

where we have used the approximation in Equation (B.6) in evaluating the first three integrals, and the first-order approximations in Equations (B.4) and (B.5), as well as the assumption that Ψ_{ij}^{-1}/s_j is always everywhere bounded for all i and j, in evaluating the fourth and fifth integrals. Note that we need this assumption only for the terms involving the set $I_t^\dagger$ of entering and exiting products and, as such, it is not needed for the special case of the proposition presented in Proposition 1. □

Proof for Proposition 4. The steps closely follow those used in the proof of Proposition 3 above, but instead rely on the approximate integration of the expression for the Divisia index given by Lemma A.3 in Equation (A.8). Multiplying both sides by the denominator of the right hand side, first we note that the integral of $d\log P_\tau^{II}$ is given as before. The only difference compared to the case of Proposition 3 is that we now have to two following additional terms

$$\int_{t-1}^{t} \sum_{i,j} \frac{\omega_{it}}{\sigma_{ii,\tau}-1} \Psi_{ij,\tau}^{-1} \left(\eta_{j\tau}-1\right) d\log P_{\tau} = \overline{\overline{\sum_{i,j} \frac{\omega_{it}}{\sigma_{ii,t}-1} \Psi_{ij,t}^{-1} \left(\eta_{jt}-1\right)}} \Delta \log P_{t},$$
$$\int_{t-1}^{t} \sum_{i,j} \frac{\omega_{it}}{\sigma_{ii,\tau}-1} \Psi_{ij,\tau}^{-1} \left(\eta_{j\tau}-1\right) d\log y_{\tau} = \overline{\overline{\sum_{i,j} \frac{\omega_{it}}{\sigma_{ii,t}-1} \Psi_{ij,t}^{-1} \left(\eta_{jt}-1\right)}} \Delta \log y_{t},$$

leading to the desired result. □

Proof for Proposition A.1. First, we compute the cross-product elasticities for the uncompensated (Marshallian) demand $\widetilde{q}_i^{uc}(\boldsymbol{p}, y)$ as follows. Letting $V(\boldsymbol{p}, y)$ denote the indirect utility function, we have

$$\begin{aligned}
\frac{\partial \log \widetilde{q}_i^{uc}(\boldsymbol{p}, y)}{\partial \log p_j} &= \frac{\partial \log \widetilde{q}_i(\boldsymbol{p}, V(\boldsymbol{p}, y))}{\partial \log p_j}, \\
&= \frac{\partial \log \widetilde{q}_i}{\partial \log p_j} + \frac{\partial \log \widetilde{q}_i}{\partial u} \frac{\partial V}{\partial \log p_j}, \\
&= s_j \sigma_{ij} + \frac{\partial \log \widetilde{q}_i}{\partial u} p_j \left(-\frac{\partial V}{\partial y}\right) q_j, \\
&= s_j \sigma_{ij} - \frac{p_j q_j}{y} \frac{\partial \log \widetilde{q}_i}{\partial \log y}, \\
&= s_j \left(\sigma_{ij} - \eta_i\right). \qquad \text{(B.7)}
\end{aligned}$$

where η_i denotes the income elasticity of demand. For the own-price elasticity, we have

$$\frac{\partial \log \widetilde{q}_i^{uc}}{\partial \log p_i} = \frac{\partial \log \widetilde{q}_i}{\partial \log p_i} + \frac{\partial \log \widetilde{q}_i}{\partial u} \frac{\partial V}{\partial \log p_i} = \left(s_i - 1\right) \sigma_{ii} - s_i \eta_i. \qquad \text{(B.8)}$$

Next, we can write the change in the expenditure of good i at any moment τ along the path as

$$\begin{aligned}
d\log s_{i\tau} = &\sum_{j \neq i} \frac{\partial \log \widetilde{q}_{j\tau}^{uc}}{\partial \log p_{j\tau}} \left(d\log p_{j\tau} - d\varphi_{j\tau}\right) + \left(1 + \frac{\partial \log \widetilde{q}_{i\tau}^{uc}}{\partial \log p_{i\tau}}\right) \left(d\log p_{i\tau} - d\varphi_{i\tau}\right) \\
&+ \left(\frac{\partial \log \widetilde{q}_{i\tau}^{uc}}{\partial \log y_{\tau}} - 1\right) d\log y_{\tau}, \\
= &\sum_{j \neq i} s_{j\tau} \left(\sigma_{ij,\tau} - \eta_{i\tau}\right) \left(d\log p_{j\tau} - d\varphi_{j\tau}\right) \\
&+ \left(1 + \left(s_{i\tau} - 1\right) \sigma_{ii,\tau} - s_{i\tau} \eta_{i\tau}\right) \left(d\log p_{i\tau} - d\varphi_{i\tau}\right) + \left(\eta_{i\tau} - 1\right) d\log y_{\tau},
\end{aligned}$$

$$
\begin{aligned}
&= -\left(\sigma_{ii,\tau}-1\right)\left(d\log p_{i\tau}-d\varphi_{i\tau}\right)+\sum_{j} s_{j\tau}\sigma_{ij,\tau}\left(d\log p_{j\tau}-d\varphi_{j\tau}\right)\\
&\quad -\eta_{i\tau}\sum_{j} s_{j\tau}\left(d\log p_{j\tau}-d\varphi_{j\tau}\right)+\left(\eta_{i\tau}-1\right)d\log y_{\tau},\\
&= -\left(\sigma_{ii,\tau}-1\right)\left(d\log p_{i\tau}-d\varphi_{i\tau}\right)+\sum_{j} s_{j\tau}\sigma_{ij,\tau}\left(d\log p_{j\tau}-d\varphi_{j\tau}\right)\\
&\quad -\eta_{i\tau}\,d\log P_{\tau}+\left(\eta_{i\tau}-1\right)d\log y_{\tau},\\
&= -\left(\sigma_{ii}-1\right)\left(d\log p_{i}-d\varphi_{i}-d\log P\right)+\sum_{j} s_{j\tau}\left(\sigma_{ij,\tau}-\sigma_{ii,\tau}\right)\left(d\log p_{j\tau}-d\varphi_{j\tau}-d\log P_{\tau}\right)\\
&\quad +\left(\eta_{i\tau}-1\right)\left(d\log y_{\tau}-d\log P_{\tau}\right).
\end{aligned}
$$

where in the second equality we have used Equations (B.7) and (B.8), and in the fourth and the last equalities we have used the definition in Equation (A.2). We can rewrite the above equation as in Equation (A.1) with the definition

$$
\begin{aligned}
\boldsymbol{\Xi} &\equiv \boldsymbol{\Sigma}_{\tau}^{d}-\boldsymbol{I}+\left(\boldsymbol{\Sigma}_{\tau}-\boldsymbol{\sigma}_{\tau}\mathbf{1}'\right)\operatorname{diag}\left(\boldsymbol{s}_{\tau}\right),\\
&= \left[\boldsymbol{I}+\left(\boldsymbol{\Sigma}_{\tau}-\boldsymbol{\sigma}_{\tau}\mathbf{1}'\right)\operatorname{diag}\left(\boldsymbol{s}_{\tau}\right)\left(\boldsymbol{\Sigma}_{\tau}^{d}-\boldsymbol{I}\right)^{-1}\right]\left(\boldsymbol{\Sigma}_{\tau}^{d}-\boldsymbol{I}\right).
\end{aligned}
$$

□

Proof for Lemma A.2. First, we construct the smooth paths of quality-adjusted prices and total expenditure between each two consecutive periods following the setup in Appendix A.1.1. Lemma A.1 characterizes the change log expenditure shares as a function of changes in quality-adjusted prices and total expenditure everywhere along this path. We can accordingly write the change in log expenditure shares from Equation (A.1) under the income invariant case ($\eta_{i\tau}\equiv 1$) as:

$$
d\log s_{i\tau} = -\sum_{j}\Psi_{ij,\tau}\left(\sigma_{jj,\tau}-1\right)\left(d\log p_{j\tau}-d\varphi_{j\tau}-d\log P_{\tau}\right).
$$

If the demand system satisfies the connected substitute property of Berry et al. (2013), we know that the demand system, and thus this relationship, is invertible and the inverse of the matrix $\boldsymbol{\Psi}_{\tau}\left(\boldsymbol{\Sigma}_{\tau}^{d}-\boldsymbol{I}\right)$ exists for all τ. Inverting the relationship in Equation (A.1) gives us

$$
d\log p_{i\tau}-d\varphi_{i\tau}-d\log P_{\tau} = -\sum_{j\in I}\frac{1}{\sigma_{ii,\tau}-1}\Psi^{-1}_{ij,\tau}d\log s_{j\tau}. \tag{B.9}
$$

First, averaging over the products the base set O_t using the distribution ϖ_{it}, we find

$$\begin{aligned}\sum_i \varpi_{it} d\log p_{i\tau} - \sum_i \varpi_{it} d\varphi_{i\tau} - d\log P_\tau &= -\sum_i \varpi_{it} \sum_{j\in I} \tfrac{1}{\sigma_{ii,\tau}-1} \Psi^{-1}_{ij,\tau} d\log s_{j\tau}, \\ &= -\sum_i \varpi_{it} \sum_{j\in I_t^*} \tfrac{1}{\sigma_{ii,\tau}-1} \Psi^{-1}_{ij,\tau} \left(d\log s^*_{j\tau} + d\log \Lambda^*_\tau\right) \\ &\quad -\sum_i \varpi_{it} \sum_{j\in I_t \setminus I_t^*} \tfrac{1}{\sigma_{ii,\tau}-1} \Psi^{-1}_{ij,\tau} d\log s_{j\tau},\end{aligned}$$

which leads to the desired result using the assumption $\sum_i \varpi_{it} d\varphi_{i\tau} = 0$. Note that we have used the fact that $s_{i\tau} = s^*_{i\tau}\Lambda^*_\tau$ for all $i \in I_t^*$.

In the special case in which $I_t^* \equiv I$, we have:

$$\sum_i \varpi_{it} \log p_{i\tau} - d\log P_\tau = -\sum_i \varpi_{it} \sum_{j\in I} \tfrac{1}{\sigma_{ii,\tau}-1} \Psi^{-1}_{ij,\tau} d\log s_{j\tau},$$

leading to the result in Proposition 1.

Next, using Equation (B.9), we can write

$$\begin{aligned} d\varphi_{i\tau} &= d\log p_{i\tau} - d\log P_\tau + \sum_{j\in \bar{I}_t} \tfrac{1}{\sigma_{ii,\tau}-1} \Psi^{-1}_{ij,\tau} d\log s_{j\tau}, \\ &= d\log p_{i\tau} - d\log P_\tau + \sum_{j\in I_t^*} \tfrac{1}{\sigma_{ii,\tau}-1} \Psi^{-1}_{ij,\tau} \left(d\log s^*_{j\tau} + d\log \Lambda^*_\tau\right) \\ &\quad + \sum_{j\in I_t^\dagger} \tfrac{1}{\sigma_{ii,\tau}-1} \Psi^{-1}_{ij,\tau}\, d\log s_{j\tau}, \qquad \text{(B.10)} \\ &= d\log p_{i\tau} - \sum_i \varpi_{it} d\log p_{i\tau} + \sum_{j\in I_t^*} \left(\tfrac{1}{\sigma_{ii,\tau}-1} \Psi^{-1}_{ij,\tau} - \sum_i \tfrac{\varpi_{it}}{\sigma_{ii,\tau}-1} \Psi^{-1}_{ij,\tau}\right) \left(d\log s^*_{j\tau} + d\log \Lambda^*_\tau\right) \\ &\quad + \sum_{j\in V_t^\dagger} \left(\tfrac{1}{\sigma_{ii,\tau}-1} \Psi^{-1}_{ij,\tau} - \sum_i \tfrac{\varpi_{it}}{\sigma_{ii,\tau}-1} \Psi^{-1}_{ij,\tau}\right) d\log s_{j\tau}, \qquad \text{(B.11)}\end{aligned}$$

where in the third equality, we have substituted for $d\log P_\tau$ from Equation (A.3). Equation (A.5) then follows from Equation (B.9). Substituting Equation (B.10) in Equation (A.4), we find

$$d\log X_\tau = \sum_{i,j\in I_t^*} \tfrac{s^*_{i\tau}}{\sigma_{ii,\tau}-1} \Psi^{-1}_{ij,\tau} \left(d\log s^*_{j\tau} + d\log \Lambda^*_\tau\right) + \sum_{i\in I_t^*} \tfrac{s^*_{i\tau}}{\sigma_{ii,\tau}-1} \sum_{j\in I_t^\dagger} \Psi^{-1}_{ij,\tau}\, d\log s_{j\tau}.$$

□

Proof for Lemma A.3. Once again, we construct the smooth paths of quality-adjusted prices and total expenditure between each two consecutive periods following the setup in Appendix A.1.1. Lemma A.1 characterizes the change log expenditure shares as a function of changes in quality-adjusted prices and total expenditure everywhere along this path. If the demand system satisfies the connected substitute property of Berry et al. (2013), the inverse of the matrix $\boldsymbol{\Psi}_\tau\left(\boldsymbol{\Sigma}^d_\tau - \boldsymbol{I}\right)$ exists for all τ, which allows us to write

$$d\log p_{i\tau} - d\varphi_{i\tau} - d\log P_\tau = -\sum_{j\in I^*_t\cup I^\dagger_t} \tfrac{1}{\sigma_{ii,\tau}-1}\Psi^{-1}_{ij,\tau}\left(d\log s_{j\tau} - (\eta_{i\tau}-1)\left(d\log y_\tau - d\log P_\tau\right)\right). \tag{B.12}$$

Averaging over the products the base set O_t using the distribution ϖ_{it} and using the assumption $\sum_i \varpi_{it} d\varphi_{i\tau} = 0$, we find

$$\begin{aligned}\sum_i \varpi_{it} d\log p_{i\tau} - d\log P_\tau = &-\sum_i \varpi_{it} \sum_{j\in I^*_t\cup I^\dagger_t} \tfrac{1}{\sigma_{ii,\tau}-1}\Psi^{-1}_{ij,\tau} d\log s_{j\tau}\\ &+\sum_i \varpi_{it}\sum_{j\in I^*_t} \tfrac{1}{\sigma_{ii,\tau}-1}\Psi^{-1}_{ij,\tau}(\eta_{i\tau}-1)\left(d\log y_\tau - d\log P_\tau\right),\end{aligned}$$

which leads to the desired result. □

Proof for Proposition A.1. We apply the second-order approximate integration of Equation (B.6) to the results of Lemma A.2 to find the contribution of prices

$$\sum_{i\in I^*_t}\int_{t-1}^{t} s^*_{i\tau} d\log p_{i\tau} = \sum_{i\in I^*_t} \overline{\overline{s^*_{it}}}\,\Delta\log p_{i\tau} + O\left(\delta^3\right).$$

wSimilarly, we integrate Equation (A.5) using the approximation in Equation (B.6) in evaluating the first three integrals, and the first-order approximations in Equations (B.4) and (B.5) for the last two

$$\begin{aligned}\int_{t-1}^{t}\sum_i s^*_{i\tau} d\varphi_{i\tau} &= \sum_{i\in I^*_t}\int_{t-1}^{t}\left(s^*_{i\tau} - \varpi_{it}\right) d\log p_{i\tau} + \sum_{j\in I^*_t}\int_{t-1}^{t}\left(s^*_{i\tau} - \varpi_{it}\right)\Psi^{-1}_{ij,\tau}\left(d\log s^*_{j\tau} + d\log\Lambda^*_\tau\right)\\ &\quad+\sum_{j\in V^\dagger_t}\int_{t-1}^{t}\left(s^*_{i\tau} - \varpi_{it}\right)\frac{\Psi^{-1}_{ij,\tau}}{s_{j\tau}} ds_{j\tau},\\ &= \sum_{i\in I^*_t}\left(\tfrac{1}{2}\left(s^*_{it-1} + s^*_{it}\right) - \varpi_{it}\right)\Delta\log p_{i\tau}\\ &\quad+\sum_{i\in I^*_t}\sum_{j\in I^*_t}\left(\tfrac{1}{2}\left(s^*_{it-1}\Psi^{-1}_{ij,t-1} + s^*_{it}\Psi^{-1}_{ij,t}\right) - \varpi_{it}\tfrac{1}{2}\left(\Psi^{-1}_{ij,t-1} + \Psi^{-1}_{ij,t}\right)\right)\left(\Delta\log s^*_{jt} + \Delta\log\Lambda^*_t\right)\end{aligned}$$

$$+\sum_{i\in I_t^*}\sum_{j\in I_t\setminus I_t^*}\left(s_{it}^*-\varpi_{it}\right)\frac{\Psi_{ij,t}^{-1}}{s_{jt}}\left(s_{jt}\right)+\sum_{i\in I_t^*}\sum_{j\in I_{t-1}\setminus I_t^*}\left(s_{it-1}^*-\varpi_{it}\right)\frac{\Psi_{ij,t-1}^{-1}}{s_{jt-1}}\left(-s_{jt-1}\right)+O\left(\delta^3\right)$$

which leads to Equation (A.9). Applying similar arguments, we can approximate the integral of Equation (A.6):

$$\begin{aligned}\int_{t-1}^{t} d\log X_\tau &= \sum_{j\in I_t^*}\int_{t-1}^{t} s_{i\tau}^*\Psi_{ij,\tau}^{-1}\left(d\log s_{j\tau}^*+d\log\Lambda_\tau^*\right)\\ &\quad+\sum_{j\in V_t^\dagger}\int_{t-1}^{t} s_{i\tau}^*\frac{\Psi_{ij,\tau}^{-1}}{s_{j\tau}}\,ds_{j\tau},\\ &=\sum_{j\in I_t^*}\tfrac{1}{2}\left(s_{it-1}^*\Psi_{ij,t-1}^{-1}+s_{it}^*\Psi_{ij,t}^{-1}\right)\left(\Delta\log s_{jt}^*+\Delta\log\Lambda_t^*\right)\\ &\quad+\sum_{j\in I_t\setminus I_t^*} s_{it}^*\frac{\Psi_{ij,t}^{-1}}{s_{jt}}\left(s_{jt}\right)+\sum_{j\in I_{t-1}\setminus I_t^*} s_{it-1}^*\frac{\Psi_{ij,t-1}^{-1}}{s_{jt-1}}\left(-s_{jt-1}\right)+O\left(\delta^3\right),\end{aligned}$$

leading to Equation (A.10).

Finally, we can also integrate Equation (B.11) to approximate the change in quality for a given product $i\in I_t^*$ between the two periods:

$$\begin{aligned}\Delta\varphi_{it} &= \int_{t-1}^{t} d\varphi_\tau,\\ &=\left(\int_{t-1}^{t} d\log p_{i\tau}-\sum_{i'\in I_t^*}\varpi_{i't}\int_{t-1}^{t} d\log p_{i'\tau}\right)\\ &\quad+\sum_{j\in I_t^*}\int_{t-1}^{t}\left(\Psi_{ij,\tau}^{-1}-\sum_{i'\in I_t^*}\varpi_{i't}\Psi_{i'j,\tau}^{-1}\right)\left(d\log s_{j\tau}^*+d\log\Lambda_\tau^*\right)\\ &\quad+\sum_{j\in I_t\setminus I_t^*}\int_{t-1}^{t}\left(\Psi_{ij,\tau}^{-1}-\sum_{i\in I_t^*}\varpi_{i't}\Psi_{i'j,\tau}^{-1}\right)\frac{ds_{j\tau}}{s_{j\tau}},\end{aligned}$$

leading to Equation (A.11). □

Proof for Proposition A.2. Based on the separability assumption in Equation (A.19), we can write the Hicksian demand for product i as

$$q_i=\frac{\partial\mathcal{P}}{\partial P^k}\frac{\partial P^k}{\partial p_j},$$

$$= \frac{q_i}{Q^k} Q^k,$$

where Q^k in the quantity aggregate of the demand for products in category k. Accordingly, we can write the cross-product elasticities of substitution between products i and j in terms of the cross-category elasticities of substitution and within-category elasticities of substitution as

$$\begin{aligned}
\sigma_{ij} &= \frac{1}{s_j} \frac{\partial \log q_i}{\partial \log p_j} = \frac{1}{s_j} \frac{\partial \log Q^k}{\partial \log P^\ell} \frac{\partial \log P^\ell}{\partial \log p_j}, \qquad i \in I^k, j \in I^\ell, k \neq \ell \\
&= \frac{1}{s_j} \frac{s_j}{s^\ell} \frac{\partial \log Q^k}{\partial \log P^\ell} = \sigma^C_{k\ell}, \qquad \text{(B.13)} \\
\sigma_{ij} &= \frac{1}{s_j} \frac{\partial \log q_i}{\partial \log p_j} = \frac{1}{s^k} \frac{1}{s_j / s^k} \frac{\partial \log \left(q_i / Q^k\right)}{\partial \log p_j} + \frac{1}{s_j} \frac{\partial \log P^k}{\partial \log p_j} \frac{\partial \log Q^k}{\partial \log P^k}, \qquad i, j \in I^k, \\
&= \frac{1}{s^k} \sigma^k_{ij} + \frac{1}{s_j} \frac{s_j}{s^k} \frac{\partial \log Q^k}{\partial \log P^k}, \\
&= \frac{1}{s^k} \left(\sigma^k_{ij} - \sigma^C_{kk} \left(1 - s^k\right) \right), \qquad \text{(B.14)}
\end{aligned}$$

where $\sigma^C_{k\ell}$ is the cross-category elasticity of substitution at the level of the upper nest and where σ^k_{ij} is defined as the within-category elasticity of substitution between products i and j as in Section A.3. Using the above results, we can also write the own price elasticity of product i aspects

$$\begin{aligned}
\frac{\partial \log \widetilde{q}_i}{\partial \log p_i} &= -\sum_{j \neq i} s_j \sigma_{ij} = -\sum_{j \in I^k, j \neq i} s_j \sigma_{ij} - \sum_{\ell \neq k} \sum_{j \in I^\ell} s_j \sigma_{ij}, \\
&= \frac{\sigma^C_{kk}}{s^k} \left(1 - s^k\right) \left(s^k - s_i\right) - \sum_{j \in I^k, j \neq i} \frac{s_j}{s^k} \sigma^k_{ij} - \sum_{\ell \neq k} s^\ell \sigma^C_{k\ell}, \\
&= \sigma^C_{kk} \left(1 - s^k\right) \left(1 - s^k_i\right) - \sigma^k_{ii} \left(1 - s^k_i\right) - \sigma^C_{kk} \left(1 - s^k\right), \\
&= -\sigma^C_{kk} \left(1 - s^k\right) s^k_i - \sigma^k_{ii} \left(1 - s^k_i\right). \qquad \text{(B.15)}
\end{aligned}$$

Using the above results, and assuming income invariance, we can rewrite the relationship between changes in the log expenditure share of product i and changes in log relative quality-adjusted prices of all other products as

$$d \log s_{i\tau} = \left(1 - \frac{\partial \log \widetilde{q}_i}{\partial \log p_i}\right) \left(d \log p_{i\tau} - d\varphi_{i\tau}\right) + \sum_{j \neq i \in I^k} s_{j\tau} \sigma_{ij,\tau} \left(d \log p_{j\tau} - d\varphi_{j\tau}\right)$$

$$
\begin{aligned}
&\quad + \sum_{\ell \neq k} \sum_{j \in I^{\ell}} s_{j\tau} \sigma_{ij,\tau} \left(d \log p_{j\tau} - d\varphi_{j\tau} \right) - d \log P_{\tau}, \\
&= \left[1 - \sigma^{C}_{kk,\tau} \left(1 - s^{k}_{\tau} \right) s^{k}_{i\tau} - \sigma^{k}_{ii,\tau} \left(1 - s^{k}_{i\tau} \right) \right] \left(d \log p_{i\tau} - d\varphi_{i\tau} \right) \\
&\quad + \sum_{j \neq i \in I^{k}} \frac{1}{s^{k}_{\tau}} \left(\sigma^{k}_{ij,\tau} - \sigma^{C}_{kk,\tau} \left(1 - s^{k}_{\tau} \right) \right) s_{j\tau} \left(d \log p_{j\tau} - d\varphi_{j\tau} \right) \\
&\quad + \sum_{\ell \neq k} s^{\ell}_{\tau} \sigma^{C}_{k\ell,\tau} \sum_{j \in I^{\ell}} s^{\ell}_{j\tau} \left(d \log p_{j\tau} - d\varphi_{j\tau} \right) - d \log P_{\tau}, \\
&= - \sum_{j \in I^{k}} \Psi^{k}_{ij,\tau} \left(d \log p_{j\tau} - d\varphi_{j\tau} \right) \\
&\quad - \sigma^{C}_{kk,\tau} \left(1 - s^{k}_{\tau} \right) \sum_{j \in I^{k}} s^{k}_{j\tau} \left(d \log p_{j\tau} - d\varphi_{j\tau} \right) \\
&\quad + \sum_{\ell \neq k} s^{\ell}_{\tau} \sigma^{C}_{k\ell,\tau} \, d \log P^{\ell}_{\tau} - d \log P_{\tau}, \\
&= - \sum_{j \in I^{k}} \Xi^{k}_{ij,\tau} \left(d \log p_{j\tau} - d\varphi_{j\tau} \right) - d \log P^{k}_{\tau} \\
&+ \left(1 - \frac{\partial \log Q^{k}_{\tau}}{\partial \log P_{\tau}} \right) d \log P^{k}_{\tau} + \sum_{\ell \neq k} s^{\ell}_{\tau} \sigma^{C}_{k\ell,\tau} \left(d \log P^{\ell}_{\tau} - d \log P_{\tau} \right), \\
&= \sum_{j \in I^{k}} \Psi^{k}_{ij,\tau} \left(d \log p_{j\tau} - d\varphi_{j\tau} - d \log P^{k}_{\tau} \right) - d \log s^{k}_{\tau},
\end{aligned}
$$

where in the second equality, we have substituted from Equations (B.13), (B.14), and (B.15) and in the third equality, we have substituted the expression for the within category $\Xi^{k}_{\tau} \equiv \boldsymbol{\Psi}^{k}_{\tau} \left(\boldsymbol{\Sigma}^{k,d}_{\tau} - \boldsymbol{I} \right)$, noting that $d \log P^{k}_{\tau} \equiv \sum_{j \in I^{k}} s^{k}_{j\tau} \left(d \log p_{j\tau} - d\varphi_{j\tau} \right)$. From the above result, it follows that

$$
d \log s^{k}_{i\tau} = \sum_{j \in I^{k}} \Xi^{k}_{ij,\tau} \left(d \log p_{j\tau} - d\varphi_{j\tau} - d \log P^{k}_{\tau} \right).
$$

Accordingly, we can apply the same arguments as in Proposition 3 to each product category to derive the corresponding approximate price index $\Delta \log P^{k}_{\tau}$ between each two consecutive periods. □

Proof for Proposition A.3. Let us first solve for the cost minimization problem of each unit of production

$$
\begin{aligned}
C_{nt} &= \min \sum_{j} W_{jt} F^{n}_{jt} + \sum_{i} p^{\varphi}_{nit} q^{\varphi}_{nit}, \qquad \text{s.t. } y_{nt} \geq A_t \, Z_n \left[\mathcal{F} \left(\boldsymbol{F}_{nt} \right)^{1-\alpha} \mathcal{F}^{I}_{n} \left(\boldsymbol{q}^{\varphi,I}_{nt} \right)^{\alpha} \right]^{\gamma}, \\
&= \min \sum_{j} W_{jt} F^{n}_{jt} + \sum_{i} p^{\varphi}_{nit} q^{\varphi}_{nit}, \;\; \text{s.t. } \; y_{nt} \geq A_t \, Z_n \left[\left(F^{F}_{nt} \right)^{1-\alpha} \left(F^{I}_{nt} \right)^{\alpha} \right]^{\gamma},
\end{aligned}
$$

and such that $F^F_{nt} \geq \mathcal{F}\left(\boldsymbol{F}_{nt}\right)$ and $F^I_{nt} \geq \mathcal{F}^I_n\left(\boldsymbol{q}^{\varphi,I}_{nt}\right)$. Since functions $\mathcal{F}$ and $\mathcal{F}^I_n$ are homogenous of degree 1, we can write their cost functions as $\sum_j W_{jt} F^n_{jt} = F^F_{nt}\,\mathcal{P}^F\left(\boldsymbol{W}_t\right)$ and $\sum_i p^{\varphi}_{njt} q^{\varphi}_{njt} = F^I_{nt}\,\mathcal{P}^I_n\left(\boldsymbol{p}^{\varphi,M}_t\right)$, thus simplifying the problem as

$$C_{nt} = F^F_{nt}\,\mathcal{P}^F\left(\boldsymbol{W}_t\right) + F^I_{nt}\,\mathcal{P}^I_n\left(\boldsymbol{p}^{\varphi,I}_t\right), \qquad \text{s.t. } y_{nt} \geq A_t\,Z_n\left[\left(F^F_{nt}\right)^{1-\alpha}\left(F^I_{nt}\right)^{\alpha}\right]^{\gamma}.$$

Taking the first-order conditions leads to

$$\begin{aligned}
\mathcal{P}^F\left(\boldsymbol{W}_t\right) &= \lambda_{nt}\,\gamma\left(1-\alpha\right)\,\frac{y_{nt}}{F^F_{nt}},\\
\mathcal{P}^I_n\left(\boldsymbol{p}^{\varphi,I}_t\right) &= \lambda_{nt}\,\gamma\alpha\,\frac{y_{nt}}{F^I_{nt}},
\end{aligned}$$

which gives us $C_{nt} = \lambda_{nt}\,\gamma\,y_{nt}$. Finally, we compute λ_{nt} by substituting the above expressions for F^F_{nt} and F^I_{nt} in the constraint

$$\begin{aligned}
y_{nt} &= A_t\,Z_n\left[\left(\frac{\lambda_{nt}\,\gamma\left(1-\alpha\right)y_{nt}}{\mathcal{P}^F\left(\boldsymbol{W}_t\right)}\right)^{1-\alpha}\left(\frac{\lambda_{nt}\,\gamma\alpha\,y_{nt}}{\mathcal{P}^I_n\left(\boldsymbol{p}^{\varphi,I}_t\right)}\right)^{\alpha}\right]^{\gamma},\\
&= \left(\gamma\lambda_{nt}y_{nt}\right)^{\gamma}\,A_t Z_n\left[\left(\frac{1-\alpha}{\mathcal{P}^F\left(\boldsymbol{W}_t\right)}\right)^{1-\alpha}\left(\frac{\alpha}{\mathcal{P}^I_n\left(\boldsymbol{p}^{\varphi,I}_t\right)}\right)^{\alpha}\right]^{\gamma},
\end{aligned}$$

leading to the following result

$$\lambda_{nt} y_{nt} = \frac{1}{\gamma}\left(\frac{\mathcal{P}^F\left(\boldsymbol{W}_t\right)}{1-\alpha}\right)^{1-\alpha}\left(\frac{\mathcal{P}^I_n\left(\boldsymbol{p}^{\varphi,I}_t\right)}{\alpha}\right)^{\alpha}\left(\frac{y_{nt}}{A_t Z_n}\right)^{\frac{1}{\gamma}}. \tag{B.16}$$

Substituting this in $C_{nt} = \lambda_{nt}\,\gamma\,y_{nt}$ gives us the following cost function

$$C_{nt} = \left(\frac{y_{nt}}{A_t Z_n}\right)^{\frac{1}{\gamma}}\left(\frac{\mathcal{P}^F\left(\boldsymbol{W}_t\right)}{1-\alpha}\right)^{1-\alpha}\left(\frac{\mathcal{P}^I_n\left(\boldsymbol{p}^{\varphi,I}_t\right)}{\alpha}\right)^{\alpha}, \tag{B.17}$$

for unit-level production technologies. □

Equalizing marginal costs λ_{nt} across firms from Equation (B.16) implies that the relative size of output across two units n and n' satisfies

$$\frac{y_{nt}}{y_{n't}} = \left(\frac{Z_n}{Z_{n'}}\right)^{\frac{1}{1-\gamma}} \left(\frac{\mathcal{P}^I_{n'}\left(\boldsymbol{p}^{\varphi,I}_t\right)}{\mathcal{P}^I_n\left(\boldsymbol{p}^{\varphi,I}_t\right)}\right)^{\alpha\frac{\gamma}{1-\gamma}}.$$

Substituting this expression in Equation (B.17) implies that the ratio of costs of two units n and n' are given by

$$\begin{aligned}\frac{C_{nt}}{C_{n't}} &= \left(\frac{Z_{n'}}{Z_n}\right)^{\frac{1}{\gamma}} \left(\frac{\mathcal{P}^I_n\left(\boldsymbol{p}^{\varphi,I}_t\right)}{\mathcal{P}^I_{n'}\left(\boldsymbol{p}^{\varphi,I}_t\right)}\right)^{\alpha} \times \left(\frac{Z_n}{Z_{n'}}\right)^{\frac{1}{\gamma(1-\gamma)}} \left(\frac{\mathcal{P}^I_{n'}\left(\boldsymbol{p}^{\varphi,I}_t\right)}{\mathcal{P}^I_n\left(\boldsymbol{p}^{\varphi,I}_t\right)}\right)^{\frac{\alpha}{1-\gamma}}, \\ &= \left(\frac{Z_n}{Z_{n'}}\right)^{\frac{1}{1-\gamma}} \left(\frac{\mathcal{P}^I_{n'}\left(\boldsymbol{p}^{\varphi,I}_t\right)}{\mathcal{P}^I_n\left(\boldsymbol{p}^{\varphi,I}_t\right)}\right)^{\frac{\alpha\gamma}{1-\gamma}}.\end{aligned}$$

Thus, we find that costs and output are propotional across firms, that is, we have $C_{nt} = B_t\left(\boldsymbol{W}_t\right)\widetilde{y}_{nt}\left(\boldsymbol{p}^{\varphi,I}_t\right)$. Moreover, letting the unit-level share of production i in expenditure on intermediate products be denoted $\mathcal{S}^I_{ni}\left(\boldsymbol{p}^{\varphi,I}_t\right) \equiv \frac{\partial \log \mathcal{P}^I_n\left(\boldsymbol{p}^{\varphi,I}_t\right)}{\partial \log p_{it}}$, Equation (B.17) implies the share of product i in the costs of unit n at time t is given by $\frac{p^{\varphi}_{it}q^{\varphi}_{it}}{C_{nit}} = \alpha\,\mathcal{S}^I_{ni}\left(\boldsymbol{p}^{\varphi,I}_t\right)$. This result, when combined with Equation (B.17) implies that each unit's share of aggregate costs $\Omega_{nt} \equiv \frac{C_{nt}}{\sum_{n'} C_{n't}}$ only varies as a function of the vector of prices of intermediate inputs $\boldsymbol{p}^{\varphi,I}_t$.

Together, these results imply that the share of intermediate product i in aggregate spending on intermediate products is given by

$$\mathcal{S}^I_i\left(\boldsymbol{p}^{\varphi,I}_t\right) = \sum_n \Omega_n\left(\boldsymbol{p}^{\varphi,I}_t\right)\mathcal{S}^I_{ni}\left(\boldsymbol{p}^{\varphi,I}_t\right),$$

only as a function of the vector of quality-adjusted products $\boldsymbol{p}^{\varphi,I}_t$. This leads to

$$\mathcal{Q}^I_i\left(\boldsymbol{p}^{\varphi,I}_t\right) = \alpha\frac{\sum_n C_{nt}}{p_{it}}\mathcal{S}^I_i\left(\boldsymbol{p}^{\varphi,I}_t\right) = \alpha B_t\left(\boldsymbol{W}_t\right)\frac{\sum_n \widetilde{y}\left(\boldsymbol{p}^{\varphi,I}_t\right)}{p_{it}}\mathcal{S}^I_i\left(\boldsymbol{p}^{\varphi,I}_t\right),$$

which is the product of a term that only depends on factor prices and a term that only depends on the prices of intermediate products $\boldsymbol{p}^{\varphi,I}_t$, thus leading to our desired result.

Proof for Lemma A.4. Let us start with HDIA and define, for now, $\sigma_{ii,\tau} \equiv \varepsilon_i^2/\overline{\varepsilon}$, so that we can rewrite Equation (A.1) as follows

$$d\log s_i = \left(1-\sum_{j\neq i} s_j\sigma_{ij}\right)\left(d\log p_j - d\varphi_j - d\log P\right) + \sum_j s_j\sigma_{ij}\left(d\log p_j - d\varphi_j - d\log P\right),$$

$$= \sum_j \left[\left(1-\sum_j s_j\sigma_{ij}\right)\mathbb{I}_{ij} + s_j\sigma_{ij}\right]\left(d\log p_j - d\varphi_j - d\log P\right), \tag{B.18}$$

where, as before, $\mathbb{I}_{ij}=1$ if $i=j$ and $\mathbb{I}_{ij}=0$, otherwise. Note that Equation (B.18) holds for any choice of σ_{ii}. In particular, the proof here simplifies if we choose $\sigma_{ii}=\varepsilon_i$. Using Equation (11) for HDIA, we now find

$$d\log s_i = (1-\varepsilon_i)\left(d\log p_{i\tau} - d\varphi_{i\tau} - d\log P_\tau\right) + \frac{\varepsilon_i}{\bar{\varepsilon}}\sum_j s_j\varepsilon_j\left(d\log p_{j\tau} - d\varphi_{j\tau} - d\log P_\tau\right),$$

where we have used the fact that $\sum_j s_{j\tau}\left(d\log p_{j\tau} - d\varphi_{j\tau} - d\log P_\tau\right)=0$. We can rewrite this as relationship as

$$d\log p_{j\tau} - d\varphi_{j\tau} - d\log P_\tau = \frac{1}{1-\varepsilon_i}\left(d\log s_i - \frac{\varepsilon_i}{\bar{\varepsilon}}d\log B\right), \tag{B.19}$$

where we have let $d\log B \equiv \sum_j s_j\varepsilon_j\left(d\log p_{j\tau} - d\varphi_{j\tau} - d\log P_\tau\right)$. Substituting Equation (B.19) in the definition of $d\log B$, we find

$$d\log B = -\sum_j s_j\frac{\varepsilon_j}{\varepsilon_j-1}\left(d\log s_j - \frac{\varepsilon_j}{\bar{\varepsilon}}d\log B\right),$$

$$= -\sum_j s_j\left(1+\frac{1}{\varepsilon_j-1}\right)d\log s_j + d\log B\sum_j s_j\left(1+\frac{1}{\varepsilon_j-1}\right)\frac{\varepsilon_j}{\bar{\varepsilon}},$$

$$= -\sum_j s_j\frac{1}{\varepsilon_j-1}d\log s_j + d\log B\left(1+\frac{1}{\bar{\varepsilon}}\sum_j s_j\left(1+\frac{1}{\varepsilon_j-1}\right)\right),$$

$$= -\sum_j s_j\mu_j d\log s_j + d\log B\left(1+\frac{1+\bar{\mu}}{\bar{\varepsilon}}\right),$$

$$= \frac{\bar{\varepsilon}}{1+\bar{\mu}}\sum_j s_j\mu_j d\log s_j.$$

Substituting the above expression for $d\log B$ in Equation (B.19) gives us

$$
\begin{aligned}
d\log p_{j\tau} - d\varphi_{j\tau} - d\log P_\tau &= -\frac{1}{\varepsilon_i - 1} d\log s_i + \frac{1}{\bar{\varepsilon}} \frac{\varepsilon_i}{\varepsilon_i - 1} d\log B, \\
&= -\frac{1}{\varepsilon_i - 1} d\log s_i + \frac{1}{\bar{\varepsilon}} \left(1 + \frac{1}{\varepsilon_i - 1}\right) d\log B, \\
&= -\mu_i d\log s_i + \frac{1+\mu_i}{1+\bar{\mu}} \sum_j s_j \mu_j d\log s_j,
\end{aligned}
$$

leading to the desired result, using again the fact that $\sum_j s_{j\tau}\left(d\log p_{j\tau} - d\varphi_{j\tau} - d\log P_\tau\right) = 0$.

Next we consider HIIA and re-define, this time, $\sigma_{ii} \equiv 2\varepsilon_i - \bar{\varepsilon}$, so that we can again write Equation (B.18). Using Equation 11 for HIIA, we now find

$$
\begin{aligned}
d\log s_i &= \sum_j \left[\left(1 - \sum_j s_{j\tau}\sigma_{ij,\tau}\right)\mathbb{I}_{ij} + s_{j\tau}\sigma_{ij,\tau}\right]\left(d\log p_{j\tau} - d\varphi_{j\tau} - d\log P_\tau\right), \\
&= (1-\varepsilon_i)\left(d\log p_{i\tau} - d\varphi_{i\tau} - d\log P_\tau\right) \\
&\quad + \sum_j s_j\left(\varepsilon_i + \varepsilon_j - \bar{\varepsilon}\right)\left(d\log p_{j\tau} - d\varphi_{j\tau} - d\log P_\tau\right), \\
&= (1-\varepsilon_i)\left(d\log p_{i\tau} - d\varphi_{i\tau} - d\log P_\tau\right) \\
&\quad + \sum_j s_j\varepsilon_j\left(d\log p_{j\tau} - d\varphi_{j\tau} - d\log P_\tau\right),
\end{aligned}
$$

where we have used the fact that $\sum_j s_{j\tau}\left(d\log p_{j\tau} - d\varphi_{j\tau} - d\log P_\tau\right) = 0$. We can rewrite this as relationship as

$$
d\log p_{j\tau} - d\varphi_{j\tau} - d\log P_\tau = \frac{1}{1-\varepsilon_i}\left(d\log s_i - d\log B\right), \tag{B.20}
$$

where we have let $d\log B \equiv \sum_j s_j\varepsilon_j\left(d\log p_{j\tau} - d\varphi_{j\tau} - d\log P_\tau\right)$. Substituting Equation (B.20) in the definition of $d\log B$, we find

$$
\begin{aligned}
d\log B &= -\sum_j s_j \frac{\varepsilon_j}{\varepsilon_j - 1}\left(d\log s_j - d\log B\right), \\
&= -\sum_j s_j\left(1 + \frac{1}{\varepsilon_j - 1}\right) d\log s_j + d\log B \sum_j s_j\left(1 + \frac{1}{\varepsilon_j - 1}\right),
\end{aligned}
$$

$$
\begin{aligned}
&= -\sum_j s_j \frac{1}{\varepsilon_j - 1} d\log s_j + d\log B \left(1 + \sum_j s_j \frac{1}{\varepsilon_j - 1}\right), \\
&= -\sum_j s_j \mu_j d\log s_j + d\log B \; (1 + \overline{\mu}), \\
&= \frac{1}{\overline{\mu}} \sum_j s_j \mu_j d\log s_j.
\end{aligned}
$$

Substituting the above expression for $d\log B$ in Equation (B.20) gives us

$$
\begin{aligned}
d\log p_{j\tau} - d\varphi_{j\tau} - d\log P_\tau &= -\frac{1}{\varepsilon_i - 1} d\log s_i + \frac{1}{\overline{\mu}} \frac{1}{\varepsilon_i - 1} \sum_j s_j \mu_j d\log s_j, \\
&= -\mu_i d\log s_i + \frac{\mu_i}{\overline{\mu}} \sum_j s_j \mu_j d\log s_j,
\end{aligned}
$$

leading to the desired result, using again the fact that $\sum_j s_{j\tau} \left(d\log p_{j\tau} - d\varphi_{j\tau} - d\log P_\tau\right) = 0$.

Finally, we consider HIIA and re-define, this time, $\sigma_{ii} = 1 + \frac{(\varepsilon_i - 1)^2}{\overline{\varepsilon} - 1}$, so that we can again write Equation (B.18). Using Equation 11 for HSA, we now find

$$
\begin{aligned}
d\log s_i &= \sum_j \left[\left(1 - \sum_j s_{j\tau}\sigma_{ij,\tau}\right) \mathbb{I}_{ij} + s_{j\tau}\sigma_{ij,\tau}\right] \left(d\log p_{j\tau} - d\varphi_{j\tau} - d\log P_\tau\right), \\
&= -\left(\varepsilon_i - 1\right) \left(d\log p_{i\tau} - d\varphi_{i\tau} - d\log P_\tau\right) \\
&\quad + \left(\varepsilon_i - 1\right) \sum_j s_{j\tau} \frac{\varepsilon_j - 1}{\overline{\varepsilon} - 1} \left(d\log p_{j\tau} - d\varphi_{j\tau} - d\log P_\tau\right), \\
&= \left(1 - \varepsilon_i\right) \Bigg\{ \left(d\log p_{i\tau} - d\varphi_{i\tau} - d\log P_\tau\right) \\
&\qquad - \tfrac{1}{\overline{\varepsilon} - 1} \sum_j s_{j\tau}\varepsilon_j \left(d\log p_{j\tau} - d\varphi_{j\tau} - d\log P_\tau\right) \Bigg\},
\end{aligned}
$$

where we have used the fact that $\sum_j s_{j\tau} \left(d\log p_{j\tau} - d\varphi_{j\tau} - d\log P_\tau\right) = 0$. We can rewrite this as relationship as

$$
d\log p_{j\tau} - d\varphi_{j\tau} - d\log P_\tau = -\frac{1}{\varepsilon_i - 1} d\log s_i + \frac{1}{\overline{\varepsilon} - 1} d\log B, \tag{B.21}
$$

where we have let $d\log B \equiv \sum_j s_j \varepsilon_j \left(d\log p_{j\tau} - d\varphi_{j\tau} - d\log P_\tau\right)$. Substituting Equation

(B.21) in the definition of $d \log B$, we find

$$\begin{aligned}
d \log B &= -\sum_j s_j \varepsilon_j \left(\frac{1}{\varepsilon_j - 1} d \log s_j - \frac{1}{\bar{\varepsilon} - 1} d \log B \right), \\
&= -\sum_j s_j \left(1 + \frac{1}{\varepsilon_j - 1} \right) d \log s_j + \frac{\bar{\varepsilon}}{\bar{\varepsilon} - 1} d \log B, \\
&= -\sum_j s_j \mu_j \, d \log s_j + \left(1 + \frac{1}{\bar{\varepsilon} - 1} \right) d \log B, \\
&= (\bar{\varepsilon} - 1) \sum_j s_j \mu_j d \log s_j.
\end{aligned}$$

Substituting the above expression for $d \log B$ in Equation (B.21) gives us

$$\begin{aligned}
d \log p_{j\tau} - d\varphi_{j\tau} - d \log P_\tau &= -\frac{1}{\varepsilon_i - 1} d \log s_i + \sum_j s_j \mu_j d \log s_j, \\
&= -\mu_i d \log s_i + \sum_j s_j \mu_j d \log s_j,
\end{aligned}$$

leading to the desired result, using again the fact that $\sum_j s_{j\tau} \left(d \log p_{j\tau} - d\varphi_{j\tau} - d \log P_\tau \right) = 0$. □

Proof for Proposition A.4. We let $\mu_i \equiv \frac{1}{\varepsilon_i - 1}$ and rely on Equation (A.28), but choose $\overline{\mu}^*$ instead of $\overline{\mu}$ as the constant to find:[A4]

$$\Xi^{-1}_{ij,\tau} = \mu_{it} \mathbb{I}_{ij} - \iota_{it} s_{jt} \left(\mu_{jt} - \overline{\mu}^*_t \right),$$

where $\iota_{it} \equiv \frac{1+\mu_{it}}{1+\overline{\mu}_t}$ for HDIA, $\iota_{it} \equiv \frac{\mu_{it}}{\overline{\mu}_t}$ for HIIA, and $\iota_{it} \equiv 1$ for HSA. We then evaluate each term in Equation (13) separately.

For the second term in Equation (13), we have

$$\begin{aligned}
\mathbb{E}^{\varpi_t}_i \left[\sum_i \sum_{j \in I^*_t} \omega_{it} \overline{\overline{\Xi^{-1}_{ij,t}}} \, \Delta \log s^*_{jt} \right] &= \sum_i \omega_{it} \overline{\overline{\mu_{it}}} \, \Delta \log s^*_{it} \\
&\quad - \sum_i \omega_{it} \sum_{j \in I^*_t} \frac{1}{2} \left(\iota_{it-1} \, s_{jt-1} \, \left(\mu_{jt-1} - \overline{\mu}^*_{t-1} \right) + \iota_{it} \, s_{jt} \, \left(\mu_{jt} - \overline{\mu}^*_t \right) \right) \Delta \log s^*_{jt}, \\
&= \sum_i \omega_{it} \overline{\overline{\mu_{it}}} \, \Delta \log s^*_{it}
\end{aligned}$$

[A4]The proof of Lemma A.4 shows that any constant instead of $\overline{\mu}$ also allows us to invert the relationship between changes in quality-adjusted relative log prices and changes in log expenditure shares.

$$
\begin{aligned}
&-\tfrac{1}{2}\left(\sum_i \varpi_{it}\iota_{it-1}\right)\Lambda^*_{t-1}\sum_{j\in I^*_t} s^*_{jt-1}\left(\mu_{jt-1}-\overline{\mu}^*_{t-1}\right)\Delta\log s^*_{jt}\\
&-\tfrac{1}{2}\left(\sum_i \varpi_{it}\iota_{it}\right)\Lambda^*_{t}\sum_{j\in I^*_t} s^*_{jt}\left(\mu_{jt}-\overline{\mu}^*_{t}\right)\Delta\log s^*_{jt}.
\end{aligned}
$$

For the third term, we find

$$
\begin{aligned}
\sum_i \varpi_{it}\sum_{j\in I^*_t}\overline{\overline{\left(\Xi^{-1}_{ij,t}\right)}} &= \sum_{i\in O_t}\varpi_{it}\overline{\overline{(\mu_{it})}}\\
&\quad -\sum_{i\in O_t}\varpi_{it}\sum_{j\in I^*_t}\frac{1}{2}\left(\iota_{it-1}\,s_{jt-1}\,\left(\mu_{jt-1}-\overline{\mu}^*_{t-1}\right)+\iota_{it}\,s_{jt}\,\left(\mu_{jt}-\overline{\mu}^*_{t}\right)\right),\\
&=\sum_i \varpi_{it}\overline{\overline{\mu_{it}}}.
\end{aligned}
$$

For the last term, given that $O_t\subset I^*_t$, we find

$$
\begin{aligned}
\sum_i \varpi_{it}\left(\sum_{j\in I_t\setminus I^*_t}\Xi^{-1}_{ij,t}-\sum_{j\in I_{t-1}\setminus I^*_t}\Xi^{-1}_{ij,t-1}\right) &= -\sum_i \varpi_{it}\sum_{j\in I_t\setminus I^*_t}\iota_{it}s_{jt}\left(\mu_{jt}-\overline{\mu}^*_t\right)\\
&\quad +\sum_i \varpi_{it}\sum_{j\in I_{t-1}\setminus I^*_t}\iota_{it-1}s_{jt-1}\left(\mu_{jt-1}-\overline{\mu}^*_{t-1}\right),\\
&= -\left(\sum_i \varpi_{it}\iota_{it}\right)\left(1-\Lambda^*_t\right)\left(\overline{\mu}^-_t-\overline{\mu}^*_t\right)\\
&\quad +\left(\sum_i \varpi_{it}\iota_{it-1}\right)\left(1-\Lambda^*_{t-1}\right)\left(\overline{\mu}^+_{t-1}-\overline{\mu}^*_{t-1}\right),
\end{aligned}
$$

where $\overline{\mu}^-_t$ and $\overline{\mu}^+_{t-1}$ satisfy

$$
\begin{aligned}
\overline{\mu}_t &= \Lambda^*_t\overline{\mu}^*_t+\left(1-\Lambda^*_t\right)\overline{\mu}^-_t,\\
\overline{\mu}_{t-1} &= \Lambda^*_{t-1}\overline{\mu}^*_{t-1}+\left(1-\Lambda^*_{t-1}\right)\overline{\mu}^+_t.
\end{aligned}
$$

The above relations allow us to rewrite the terms inside the parentheses as

$$
\begin{aligned}
\overline{\mu}^-_t-\overline{\mu}^*_t &= \frac{1}{1-\Lambda^*_t}\left(\overline{\mu}_t-\Lambda^*_t\overline{\mu}^*_t\right)-\overline{\mu}^*_t,\\
&=\frac{1}{1-\Lambda^*_t}\left(\overline{\mu}_t-\overline{\mu}^*_t\right),
\end{aligned}
\tag{B.22}
$$

$$\overline{\mu}_{t-1}^{+} - \overline{\mu}_{t-1}^{*} = \frac{1}{1-\Lambda_{t-1}^{*}}\left(\overline{\mu}_{t-1} - \Lambda_{t-1}^{*}\overline{\mu}_{t-1}^{*}\right) - \overline{\mu}_{t-1}^{*},$$
$$= \frac{1}{1-\Lambda_{t-1}^{*}}\left(\overline{\mu}_{t-1} - \overline{\mu}_{t-1}^{*}\right). \tag{B.23}$$

Together, the above relationships imply

$$\sum_i \omega_{it}\left(\sum_{j\in I_t\backslash I_t^*}\Psi_{ij,t}^{-1} - \sum_{j\in I_{t-1}\backslash I_t^*}\Psi_{ij,t-1}^{-1}\right) = -\left(\bar{\iota}_t^o\left(\overline{\mu}_t - \overline{\mu}_t^*\right) - \bar{\iota}_{t-1}^o\left(\overline{\mu}_{t-1} - \overline{\mu}_{t-1}^*\right)\right).$$

If we are willing to make assumptions about the asymptotic values of μ_{it}, we can also provide a slightly more accurate approximation. In particular, we have

$$\sum_{i\in O_t}\sum_{j\in I_t\backslash I_t^*}\omega_{it}\int_{t-1}^{t}\frac{\Psi_{ij,\tau}^{-1}}{s_{j\tau}}\,ds_{j\tau} = -\frac{1}{2}\sum_{i\in O_t}\omega_{it}\iota_{it}\sum_{j\in I_t\backslash I_t^*}\overline{\overline{\left(\mu_{jt}-\overline{\mu}_t^*\right)}}\,s_{jt},$$
$$\sum_{i\in O_t}\sum_{j\in I_{t-1}\backslash I_t^*}\omega_{it}\int_{t-1}^{t}\frac{\Psi_{ij,\tau}^{-1}}{s_{j\tau}}\,ds_{j\tau} = \frac{1}{2}\sum_{i\in O_t}\omega_{it}\iota_{it}\sum_{j\in I_{t-1}\backslash I_t^*}\overline{\overline{\left(\mu_{jt}-\overline{\mu}_t^*\right)}}\,s_{jt-1},$$

where we have defined $\mu_{jt} \equiv \lim_{\check{p}_{jt}\to\infty}\frac{1}{e(\check{p}_{jt})-1}$ for $j \in I_{t-1}\backslash I_t^*$ and $\mu_{jt-1} \equiv \lim_{\check{p}_{jt-1}\to\infty}\frac{1}{e(\check{p}_{jt-1})-1}$ for $j \in I_t\backslash I_t^*$. This leads to the following result:

$$\Delta\log P_t = \sum_i \omega_{it}\Delta\log p_{it} + \sum_i \omega_{it}\overline{\overline{\mu_{it}}}\left(\Delta\log s_{i\tau}^* + \Delta\log\Lambda_t^*\right) - \sum_i \overline{\overline{\Lambda_t^* s_{it}^*\bar{\iota}_t^o\left(\frac{1}{\varepsilon_{it}-1} - \overline{\left(\frac{1}{\varepsilon_{it}-1}\right)}\right)}}\Delta\log s_{it}^*$$
$$-\frac{1}{2}\bar{\iota}_t^o\left(\Lambda_t^\dagger\sum_j s_{jt}^\dagger\left(\overline{\overline{\mu_{jt}-\overline{\mu}_t^*}}\right) - \Lambda_{t-1}^\dagger\sum_j s_{jt-1}^\dagger\left(\overline{\overline{\mu_{jt}-\overline{\mu}_t^*}}\right)\right) + O\left(\delta^3\right), \tag{B.24}$$

for $\delta \equiv \max\left\{\max_{i\in O_t}\{|\Delta\log p_{it}|\},\max_{i\in I_t^*}\{|\Delta\log s_{it}^*|\},|\Delta\log\Lambda_t^*|,\max_{i\notin I_t^*}\{|\Delta s_{it}|\}\right\}$.

To derive Equation (A.32), we need to evalue the third term on the right hand side of Equation (A.32) as

$$\sum_{j\in I_t^*}\overline{\overline{\sum_i s_{it}^*\Xi_{ij,t}^{-1}}}\Delta\log s_{jt}^* = \sum_{j\in I_t^*}\left(\sum_{i\in I_t^*}\mathbb{I}_{ij}\tfrac{1}{2}\left(s_{it}^*\mu_{it} + s_{it-1}^*\mu_{it-1}\right)\right)\Delta\log s_{j\tau}^*$$
$$-\sum_{j\in I_t^*}\sum_{i\in I_t^*}\frac{1}{2}\left(\iota_{it-1}\,s_{jt-1}\,\left(\mu_{jt-1}-\overline{\mu}_{t-1}^*\right) + \iota_{it}\,s_{jt}\,\left(\mu_{jt}-\overline{\mu}_t^*\right)\right)\Delta\log s_{jt}^*,$$
$$= \sum_{i\in I_t^*}\overline{\overline{s_{it}^*\mu_{it}}}\,\Delta\log s_{i\tau}^*$$

$$-\tfrac{1}{2}\left(\sum_{i\in I_t^*} s_{it-1}^{*}\iota_{it-1}\right)\Lambda_{t-1}^{*}\sum_{j\in I_t^*} s_{jt-1}^{*}\left(\mu_{jt-1}-\overline{\mu}_{t-1}^{*}\right)\Delta\log s_{jt}^{*}$$
$$-\tfrac{1}{2}\left(\sum_{i\in I_t^*} s_{it}^{*}\iota_{it}\right)\Lambda_{t}^{*}\sum_{j\in I_t^*} s_{jt}^{*}\left(\mu_{jt}-\overline{\mu}_{t}^{*}\right)\Delta\log s_{jt}^{*},$$

where in the last equality, we have used the fact that $\overline{\overline{s_{it}^{*}\mu_{it}}} = \overline{\overline{s_{it}^{*}}}\cdot\overline{\overline{\mu_{it}}} + \tfrac{1}{4}\Delta s_{it}^{*}\cdot\Delta\mu_{it}$. Using a similar argument, we can also simplify the fourth term as

$$\begin{aligned}\sum_{j\in I_t^*}\overline{\overline{\sum_i s_{it}^{*}\Xi_{ij,t}^{-1}}} &= \sum_{j\in I_t^*}\left(\sum_{i\in I_t^*}\mathbb{I}_{ij}\tfrac{1}{2}\left(s_{it}^{*}\mu_{it}+s_{it-1}^{*}\mu_{it-1}\right)\right)\\ &\quad-\sum_{j\in I_t^*}\sum_{i\in I_t^*}\frac{1}{2}\left(\iota_{it-1}\,s_{jt-1}\,\left(\mu_{jt-1}-\overline{\mu}_{t-1}^{*}\right)+\iota_{it}\,s_{jt}\,\left(\mu_{jt}-\overline{\mu}_{t}^{*}\right)\right),\\ &=\sum_i\overline{\overline{s_{it}^{*}}}\,\overline{\overline{\mu_{it}}}.\end{aligned}$$

For the last term, we find

$$\begin{aligned}\sum_i s_{it}^{*}\sum_{j\in I_t\setminus I_t^*}\Psi_{ij,t}^{-1}-\sum_i s_{it-1}^{*}\sum_{j\in I_{t-1}\setminus I_t^*}\Psi_{ij,t-1}^{-1} &= -\sum_i s_{it}^{*}\sum_{j\in I_t\setminus I_t^*}\iota_{it}s_{jt}\left(\mu_{jt}-\overline{\mu}_{t}^{*}\right)\\ &\quad+\sum_i s_{it-1}^{*}\sum_{j\in I_{t-1}\setminus I_t^*}\iota_{it-1}s_{jt-1}\left(\mu_{jt-1}-\overline{\mu}_{t-1}^{*}\right),\\ &=-\overline{\iota}_t^{*}\,\left(1-\Lambda_t^{*}\right)\left(\overline{\mu}_t^{-}-\overline{\mu}_t^{*}\right)\\ &\quad+\overline{\iota}_{t-1}^{*}\,\left(1-\Lambda_{t-1}^{*}\right)\left(\overline{\mu}_{t-1}^{+}-\overline{\mu}_{t-1}^{*}\right),\\ &=-\overline{\iota}_t^{*}\left(\overline{\mu}_t-\overline{\mu}_t^{*}\right)+\overline{\iota}_{t-1}^{*}\left(\overline{\mu}_{t-1}-\overline{\mu}_{t-1}^{*}\right),\end{aligned}$$

where we have again used Equations (B.22) and (B.23).

The result for the case of Equation (A.30) also follows from the above result. □

B.2 Derivations for the Kimball Aggregators

B.2.1 Derivations for Kimball Specifications

Below, we derive the Kimball functions corresponding to each of the three cases discussed in Section 4.2. We have that $\tilde{e}(\check{p}) \equiv -d\log K'(\check{q})/d\log\check{q}\big|_{\check{q}=d(\check{p})}$. This allows us to integrate the function $\mathcal{E}(\cdot)$ twice to arrive at $K(\cdot)$.

Klenow-Willis In this case, we have:

$$\begin{aligned}\psi\left(\log \check{q}\right) \equiv \log K'\left(\check{q}\right) &= -\frac{1}{\sigma}\int_{-\infty}^{\log \check{q}} e^{\theta v} dv, \\ &= -\frac{1}{\sigma\theta}\check{q}^{\theta},\end{aligned}$$

for any constant ξ. Integrating this expression again, we find:

$$\begin{aligned}K\left(\check{q}\right) &= -\int_{\log \check{q}}^{\infty} e^{-v^{\theta}/\sigma\theta} dv, \\ &= \left(\sigma\theta\right)^{\frac{1}{\theta}}\frac{1}{\theta}\Gamma\left(\frac{1}{\theta}, \frac{1}{\sigma\theta}\check{q}^{\theta}\right),\end{aligned}$$

where $\Gamma\left(\cdot,\cdot\right)$ is the incomplete Gamma function.

Finite-Infinite Limits (FIL) In this case, we have:

$$\begin{aligned}\psi\left(\log \check{q}\right) \equiv \log K'\left(\check{q}\right) &= -\int_{-\infty}^{\log \check{q}} \frac{dv}{\sigma + \left(\sigma_o - \sigma\right)e^{-\theta v}}, \\ &= -\frac{1}{\sigma}\log \check{q} - \frac{1}{\sigma\theta}\log\left(\frac{\sigma}{\sigma_o - \sigma} + \check{q}^{-\theta}\right).\end{aligned}$$

Next, we integrate to find the expression for $K\left(\cdot\right)$:

$$\begin{aligned}K\left(\check{q}\right) &= \int_{0}^{\log \check{q}} \left(\frac{\sigma v^{\theta} + \sigma_o - \sigma}{\sigma_o - \sigma}\right)^{-\frac{1}{\sigma\theta}} dv, \\ &= \check{q}\cdot {}_2F_1\left(\frac{1}{\theta}, \frac{1}{\sigma\theta}; 1 + \frac{1}{\theta}; -\frac{\sigma}{\sigma_o - \sigma}\check{q}^{\theta}\right),\end{aligned}$$

where ${}_2F_1$ is the hypergeometric function. The functional form above implies the following expression for log demand:

$$\begin{aligned}d\left(\log \check{p}\right) &\equiv \psi^{-1}\left(\log \check{p}\right), \\ &= \frac{1}{\theta}\log\left[\frac{\sigma_o - \sigma}{\sigma}\left(e^{\theta\sigma\left(\xi - \log \check{p}\right)} - 1\right)\right].\end{aligned}$$

In this case, there exists a finite chocke price for any product, above which demand drops to zero.

Finite-Finite Limits (FFL) In this case, we have:

$$\psi\left(\log\check{q}\right) \equiv \log\mathcal{K}'\left(\check{q}\right) = -\int_{-\infty}^{\log\check{q}}\left[\frac{1}{\sigma_o}+\left(\frac{1}{\sigma}-\frac{1}{\sigma_o}\right)\frac{e^{\theta_o}e^{\theta v}}{1+e^{\theta_o}e^{\theta v}}\right]dv,$$
$$= -\frac{1}{\sigma_o}\log\check{q}-\left(\frac{1}{\sigma}-\frac{1}{\sigma_o}\right)\frac{1}{\theta}\log\left(1+e^{\theta_o}\check{q}^{\theta}\right).$$

Finally, we integrate to find the expression for $K\left(\cdot\right)$:

$$K\left(\check{q}\right) = \int_0^{\check{q}} v^{-\frac{1}{\sigma_o}}\left(1+e^{\theta_o}v^{\theta}\right)^{-\left(\frac{1}{\sigma}-\frac{1}{\sigma_o}\right)\frac{1}{\theta}}dv,$$
$$= \frac{\sigma_o}{\sigma_o-1}\check{q}^{1-\frac{1}{\sigma_o}}\cdot {}_2F_1\left(\left(1-\frac{1}{\sigma_o}\right)\frac{1}{\theta},\left(\frac{1}{\sigma}-\frac{1}{\sigma_o}\right)\frac{1}{\theta};1+\left(\frac{1}{\sigma_o}+1\right)\frac{1}{\theta};-e^{\theta_o}\check{q}^{\theta}\right),$$

where ${}_2F_1$ is the hypergeometric function.

B.2.2 Inverting Kimball Demand

We implement the demand inversion through the dual problem, meaning that we map the vector of observed expenditure shares $\boldsymbol{s}_t$ to a corresponding vector of normalized quantities $\check{\boldsymbol{q}}_t$. Formally, we solve for the function $d\left(\pi_i\left(\cdot;\varsigma\right);\varsigma\right)$ corresponding to the definition (26) and (27) with $K'(\check{q};\varsigma) = d^{-1}(\check{q};\varsigma)$ and $\check{q} = d\left(\check{p};\varsigma\right)$.

To invert the demand, for any collection of $\left(\boldsymbol{p}_t,\boldsymbol{s}_t\right)$ at time t, we need to solve for the vector $\left(\log\check{q}_{it}\right)_i$, such that:

$$\log s_{it} = \log\check{q}_{it}+\psi\left(\log\check{q}_{it}\right)-\log\left[\sum_{j\in V_t}\exp\left(\log\check{q}_{jt}+\psi\left(\log\check{q}_{jt}\right)\right)\right], \qquad \forall i\in V_t, \tag{B.25}$$

$$k\left(1\right) = \log\left[\sum_{i\in V_t}\exp\left(k\left(\log\check{q}_{it}\right)\right)\right], \tag{B.26}$$

where $k\left(\cdot\right)\equiv\log K\left(\exp\left(\cdot\right)\right)$ and $\psi\left(\cdot\right)\equiv\log K'\left(\exp\left(\cdot\right)\right)$. We can rewrite Equation (B.25) as (assuming $O\equiv\{o\}$):

$$\log\left(\frac{s_{it}}{s_{ot}}\right) = \log\left(\frac{\check{q}_{it}}{\check{q}_{ot}}\right)+\psi\left(\log\check{q}_{it}\right)-\psi\left(\log\check{q}_{ot}\right), \qquad \forall i\in V_t. \tag{B.27}$$

Using the identity

$$k'\left(\log\check{q}\right) = \exp\left(\log\check{q}+\psi\left(\log\check{q}\right)-k\left(\log\check{q}\right)\right),$$

we can substitute Equation (B.27) in Equation (B.26), we find:

$$
\begin{aligned}
k\left(1\right) &= \log\left[\sum_{i\in V_t}\exp\left(k\left(\log\check{q}_{it}\right)\right)\right], \\
&= \log\left[\sum_{i\in V_t}\exp\left(\log\check{q}_{it}+\psi\left(\log\check{q}_{it}\right)-k'\left(\log\check{q}_{it}\right)\right)\right], \\
&= \log\left[\sum_{i\in V_t}\exp\left(\log\check{q}_{ot}+\psi\left(\log\check{q}_{ot}\right)+\log\left(\frac{s_{it}}{s_{ot}}\right)-k'\left(\log\check{q}_{it}\right)\right)\right], \\
&= \log\check{q}_{ot}+\psi\left(\log\check{q}_{ot}\right)+\log\left[\sum_{i\in V_t}\frac{s_{it}}{s_{ot}}\exp\left(-k'\left(\log\check{q}_{it}\right)\right)\right], \\
&= k\left(\log\check{q}_{ot}\right)+\log\left[\sum_{i\in V_t}\frac{s_{it}}{s_{ot}}\exp\left(k'\left(\log q_{ot}\right)-k'\left(\log\check{q}_{it}\right)\right)\right]. \qquad \text{(B.28)}
\end{aligned}
$$

We use an iterative approach: starting with some initial guess for $\check{q}_{ot}$, we iterate between updating values of $\check{q}_{it}$ for $i \neq o$ from Equation (B.27) and updating the value of $\check{q}_{ot}$ from Equation (B.28).

C Simulation Exercise

In this appendix, we discuss a Monte Carlo simulation exercise to examine the behavior of our identification strategy for demand estimation.

C.1 Data Generating Process

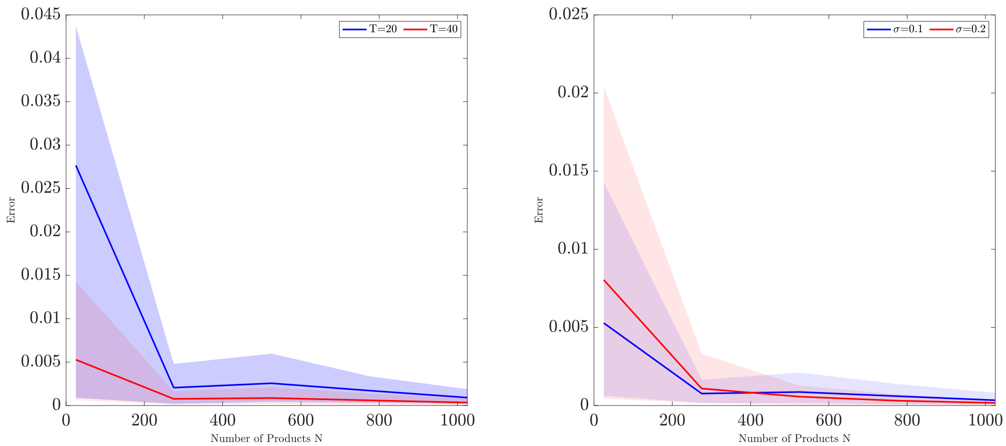
We use the following data generating process to create data on prices and expenditure shares $(\boldsymbol{p}_t, \boldsymbol{s}_t)_{t=0}^{T-1}$. First, we generate data on the vector of cost shifters $\boldsymbol{w}_t$, quality $\boldsymbol{\varphi}_t$, and prices $\boldsymbol{p}_t$ using the following dynamic AR(1) Markov model

$$
\begin{aligned}
\log w_{it} &= \rho_w \log w_{it} + (1-\rho_w)\log\overline{w}_i + \kappa_w v^{w}_{it}, \\
\varphi_{it} &= \rho\,\varphi_{it-1} + (1-\rho)\,\overline{\varphi}_i + \kappa_\varphi v^{\varphi}_{it}, \\
\log p_{it} &= \log w_{it} + \gamma_\varphi\,\varphi_{it} + \kappa_p v^{p}_{it},
\end{aligned}
$$

along with a similar process for the total expenditure

$$
\log y_t = \rho_y \log y_{t-1} + (1-\rho)\,\overline{y} + \kappa_y v^{y}_{it},
$$

Figure C.1: Montecarlo Simulation - Kimball Elasticities



Note:

where ρ_w, ρ, and ρ_y are the persistence paramters for the cost shifter, quality, and total expenditure, where $\log \overline{w}_i$, $\overline{\varphi}_i$, and $\log \overline{y}$ are the long-run values of the mean logarithm of the cost shifter, quality, and the logarithm of the total expenditure, and where v_{it}^{w}, v_{it}^{φ}, v_{it}^{p}, and v_{it}^{y} are normally distributed *i.i.d.* shocks to the logarithm of the cost shifter, quality, the logarithm of price, and the logarithm of total expenditure, with κ_w, κ_φ, κ_p, and κ_y denoting the corresponding standard deviations.

Note that the above process features a correlation between the quality (demand) shocks through $\mathbb{E}\left[\varphi_{it} \log p_{it}\right] = \gamma_\varphi \mathbb{E}\left[\varphi_{it}^2\right]$.

Given the vector of prices and qualities $(\boldsymbol{p}_t, \boldsymbol{\varphi}_t)$ at each point in time, we then compute the expenditure shares for each good $s_{it} = \widetilde{s}_i\left(\boldsymbol{p}^{\varphi}\right)$ following Equation (9) with the choice of Finite-Finite-Limit (FFL) Kimball aggregator defined in Equations (26), (27), and (A.36).

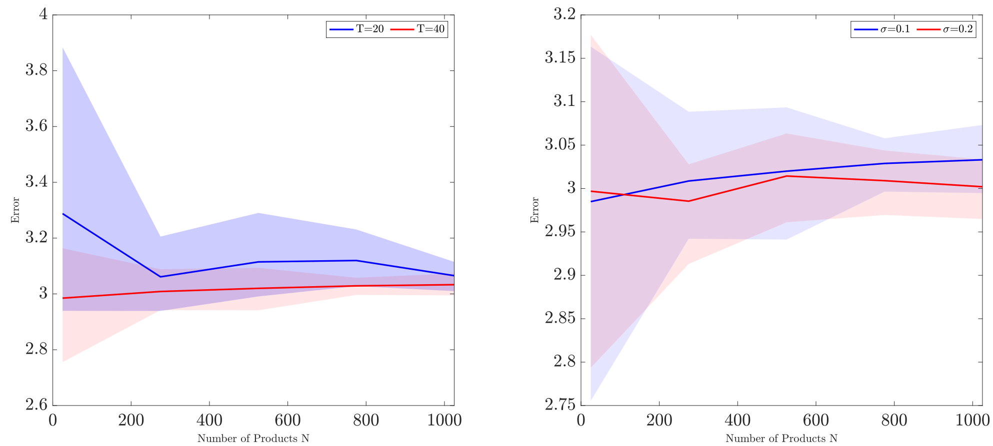
C.2 Estimation Results

D Details on the Application to the Price Index of US Imports

D.1 US Data Construction

We rely on several datasets from the Bureau of Economics Analysis (BEA), the U.S. Census Bureau, and the National Bureau of Economic Research (NBER). The first datasets are the HS-level U.S. import and export data from 1989 to 2018. We construct domestic absorp-

Figure C.2: Montecarlo Simulation - CES Elasticities



Note:

tion at the 5-digit NAICS level by matching the trade flow data to domestic production data from the NBER-CES Manufacturing Industry Database and Gross Output from the BEA. Our analysis covers 156 time-consistent industries for manufacturing, farming and mining sectors spanning from 1989 to 2018. We verify the robustness of our data construction procedure by comparing our data to the more aggregate industry definitions from the BEA annual Input-Output.

US Imports and Exports We use the U.S. import and export data at the level of 10-digit codes of the Harmonized System (henceforth HS10) over the period 1989-2018 from the U.S. Census Bureau and maintained by Peter K. Schott (Schott, 2008).[A5] The data provide information on HS-country-year import and export trade flows for the U.S. We use General Import and Total Export to measure the value of goods traded for imports and exports, respectively, and the primary quantity to measure the physical amount of goods traded. We use FOB import prices and FAS export prices to be consistent with the price definitions used by the U.S. Bureau of Labor Statistics (BLS) in the creation of the import and export price indices. We drop all import and export flows that report negative or missing value and quantity.

To map the trade data to the data on domestic sales, we follow Amiti and Heise (2021) and construct our concordance from time-consistent HS10 to time-consistent 5-digit NAICS classification codes (2012 vintage). For both imports and exports, we first use the most updated version of the Pierce and Schott (2012) algorithm to create time-

[A5]Data are available at https://faculty.som.yale.edu/peterschott/international-trade-data/.

consistent HS10 codes.[A6] This allows us to map and match obsolete and new HS10 codes over time and create combined HS10 codes. Next, we map the HS10 codes to time-consistent 5-digit NAICS using the concordances developed by Pierce and Schott (2012) and Amiti and Heise (2021).[A7] Since both concordance tables are updated until 2012, we manually adjust revisions over time in the HS10 and inconsistent mappings from HS10 to NAICS. There are about 100 problematic HS10 codes that map into different 5-digit NAICS codes. We manually check each case and make adjustment as follows: i) we assign each HS10 code to only one NAICS code after inspecting the code description, whenever possible; ii) in other cases, we avoid making many arbitrary assignments when a large number of HS10 codes map to the same set of NAICS code and decide to directly combine multiple NAICS codes; iii) in an handful of cases, we remove HS10 codes from a combined HS10 codes created using the Pierce and Schott (2012) algorithm and assign it to a more appropriate group after inspecting the code description. Following these steps, we obtain approximately 11,000 HS10 import and export codes mapped to one of 193 time-consistent 5-digit NAICS codes, which we use for our analysis.

Lastly, we perform additional cleaning on the import data before estimation. Specifically, we define within-industry variety labels as the country of origin-NAICS pairs and drop observations that exhibit change in unit value above the 99th percentile and below the 1st percentile within each NAICS. We also drop observations with reported quantity of one or below.

Domestic Production and Absorption We use the NBER-CES Manufacturing Industry Database and the BEA Gross Output by industry to measure domestic productcion and construct domestic absorption. The NBER-CES Manufacturing Industry Database dataset contains annual data on shipments and prices from the United States manufacturing sector for the period from 1958 to 2018 at the 6-digit 2012 NAICS level (Becker et al., 2021).[A8] We use the variable *vship* (total value of shipment), constructed using data from the Census Bureau's Annual Survey of Manufactures (ASM) and Census of Manufactures (CMF), to measure U.S. firms' total sales. The variable *piship*, constructed using detailed deflators from BEA and/or BLS, is used to measure domestic price changes at the industry level.[A9] Data are aggregated at the 5-digit NAICS level to match the level of aggregation of import

[A6]We use version 2019.07.12 of the algorithm. The concordance is available from Peter Schott's website (https://sompks4.github.io/sub_data.html).

[A7]The concordance by Amiti and Heise (2021) is available at https://www.sebastianheise.com.

[A8]The dataset is available at https://www.nber.org/research/data/nber-ces-manufacturing-industry-database.

[A9]Additional details on NBER-CES construction can be found at https://data.nber.org//nberces/nberces5818v1/nberces5818v1_technical_notes_Mar2021.pdf.

and export trade flows. We use a standard Tornqvist formula to aggregate 6-digit price indices into 5-digit ones, where shares are computed using total shipment values.

The NBER-CES database provides information only on manufacturing industries (NAICS 31-33). Beyond this sector, farming and mining industries also exhibit positive trade flows. Thus, we complement the information on the manufacturing sectors using BEA data for farming and mining industries. Specifically, we use the annual Gross Output measured by the BEA and the corresponding price index.[A10] BEA data are reported at the "Summary" level, forcing us to aggregate the 5-digit NAICS farming and mining industries into more aggregate industries.[A11]

We then compute the domestic sales of U.S. firms in each industry by subtracting exports from total shipments. We also compute domestic absorption as total shipments minus exports plus imports.[A12] As expected the latter is always positive, while the former exhibits a few negative values in selected industries that have experienced an increasingly negative balance of paymennt surplus such as Apparel, Iron and Steel, and Computer Manufacturing.

Aggregation and Comparison to Input-Output Data We show that the data we construct are highly correlated with official industry-level statistics from the BEA Input-Output tables. We use the annual Summary-level Use table at producer's prices from the BEA Input-Output accounts and aggregate our data on imports and exports, domestic production and absorption at the same level of aggregation.[A13] Figure D.1 shows that the data we construct are highly correlated with official national accounts from the BEA Input-Output tables. Specifically, the correlation between our measures of sectoral (Summary-level) domestic production, imports and exports with the BEA counterparts are 98%, 98%, and 93%, respectively.[A14] Discrepancies in domestic production measures

[A10] Data are available on the BEA website at https://apps.bea.gov/iTable/?reqid=147&step=2.

[A11] The 5-digit NAICS industries starting with 111 and 112 are aggregated into a unique Farming industry. Similarly, NAICS starting with 113, 114 and 115 are grouped into a unique Forestry and Fishing industry. Lastly, all mining activies other than oil and gas (NAICS starting 212) are aggregated into a unique industry.

[A12] We adjust our measure of domestic sales of U.S. firms for the presence of re-exports in order not to underestimate it. We measure U.S. exports using Domestic Exports as defined by the U.S. Census Bureau (Total Export minus Foreign Exports). Domestic absorption is not affect by this adjustment because re-exports also enters in imports, thus, not affecting the net exports.

[A13] BEA Input-Output accounts can be found at https://www.bea.gov/industry/input-output-accounts-data#supplemental-estimate-tables. Concordance between the BEA Summary-level and NAICS classifications is constructed by the BEA.

[A14] We also construct exports and imports from the U.S. Census data excluding re-exports. Excluding re-exports means dropping the item called "Foreign Export" and focusing only "Domestic Exports." On the imports side, we measure imports with "Imports for Consumption" only, rather than "General Imports." The correlation with BEA accounts does not change when we exclude re-exports. Figure D.1 also shows the strong correlation between industry-level re-exports data from the U.S. Census Bureau (i.e. Foreign Export) shrink if we further aggregate sectors that are close in their description, suggesting that the differences are due to the allocation of output and shipments in similar sectors.[A15] Similarly, discrepancies between our U.S. Census based measures of imports and exports and the BEA Input-Output accounts is most likely due to adjustments that the latter performs on the U.S. Census data. For instance, the BEA measures of imports tend to be larger than our U.S. Census based because the former are reported at foreign port value (i.e. FOB) while BEA uses domestic port value (approximately CIF) to evaluate imports. BEA exports are also larger than U.S. Census data because of statistical adjustments performed by the BEA, such as adding net export under merchanting, defined as the purchase of goods by a resident from a nonresident combined with the subsequent resale of the same goods to another nonresident without the goods being present in the U.S., and thus, not recorded by the U.S. Customs.[A16,A17]

D.2 Further Examination of CES Estimates and Comparison to Original Broda and Weinstein (2006)

Data and Estimation We use product-level data on US imports from 1989 to 2006 compiled originally by Feenstra et al. (2002) and used in Broda and Weinstein (2006). These data record US imports at the HS10 level, reporting also the corresponding SITC classification. We define a good to be an HS10 category and we follow the standard approach to identify varieties with the country of origin, e.g., an origin country-HS10 pair. A variety's unit value is defined as the sum of the value, total duties, and transportation costs divided by the import quantity. To minimize the effects of noise in the data, we trim the data as follows: we exclude all varieties that report a quantity of one unit or less than the 5th percentile within each HS10 product category; we remove varieties with an annual unit value increase that fall below the 5th percentile or above the 95th percentile within each HS10 product category.

We estimate the CES elasticity of substitution across product varieties at the HS10

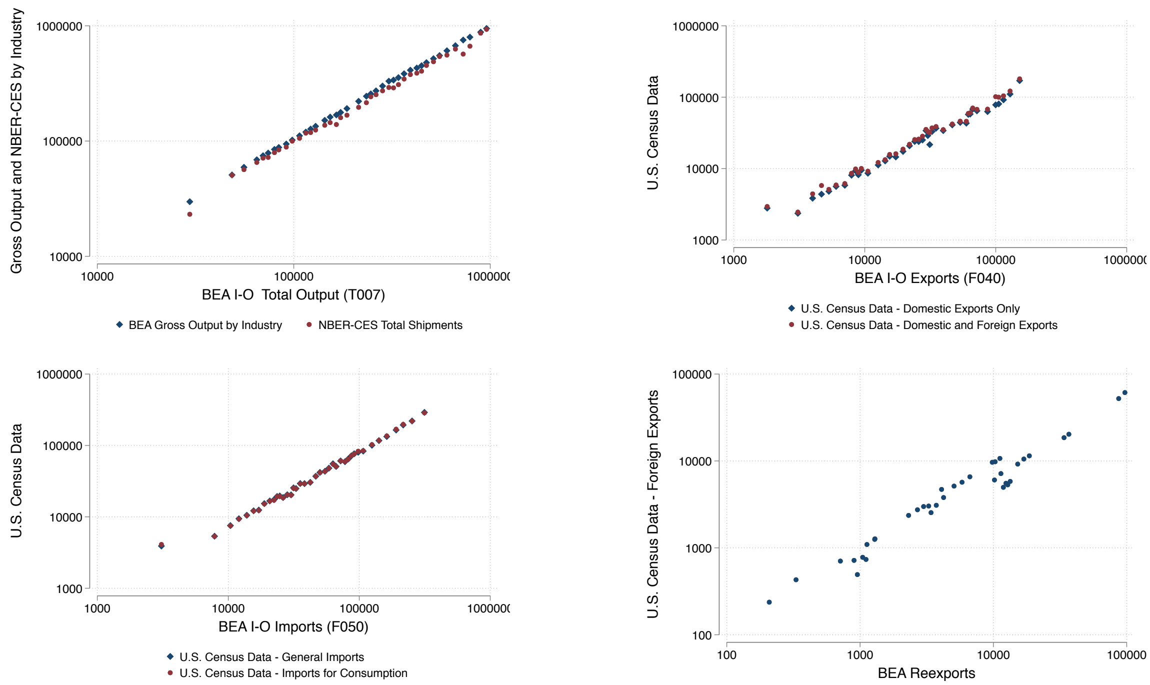
and BEA (about 98%). The BEA data on re-exports are only available from 2015.

[A15] This is the case in Apparel and Textile, Oil extraction and Refinery, Computers and Electric equipment, and Paper and Printing.

[A16] BEA collects information on merchanting trade in its Quarterly Survey of Transactions and provides statistics at aggregate level, not allowing us to gauge its relevance at the NAICS level.

[A17] The largest discrepancy concerns the exports of the Aircraft industry, which have been suppressed by the U.S. Census and aggregated in a manner that prevents disclosure of confidential information. More information on the adjustments can be found here: the explanatory notes to the FT900 release (https://www.census.gov/foreign-trade/Press-Release/current_press_release/ft900.pdf) and the BEA IEA and I-O manuals (https://www.bea.gov/international/concepts_methods.htm, https://www.bea.gov/resources/methodologies/concepts-methods-io-accounts).

Figure D.1: Comparison to BEA I-O Data



Note: The top right panel compares the NBER Total Shipments and the BEA Gross Output measures to the Total Output measure from the BEA I-O. Unit of observation is a sector-year, where sectors are defined at the I-O Summary-level. NBER Total Shipments are aggregated accordingly. The top left, bottom right, and bottom left panels compare Total Exports, Total Import and Reexports, respectively, from the U.S. Census Data to the corresponding measures from the BEA I-O. Census Data are aggregated to match the I-O Summary-level definitions. Both domestic and foreign exports, and general imports and imports for consumption are considered. BEA data on reexports are only available from 2015.

level, together with the 5, 4, and 3-digit SITC levels of aggregation (SITC5, SITC4 and SITC3, respectively). The SITC4 level allows us to map our data to the Rauch product classification (Rauch, 1999). We use our Dynamic Panel (DP) approach using the moment condition in Equation (20) with double lagged (log) prices and market shares as instruments. For the purpose of estimation, we use any continuously imported variety over the period from 1989 to 2006 within each product classification as the baseline product to infer quality in Equation (15). We compare our estimates against those found using the conventional Feenstra (1994) and Broda and Weinstein (2006) estimator (henceforth FBW), and as well as the more recent Limited Information Maximum Likelihood estimation approach (Soderbery, 2015b, henceforth LIML).

Price Elasticities Across Different Levels of Aggregation Table D.1 reports the mean and the median of the estimated elasticities using the DP approach for three different levels of product aggregation. As expected, we find lower elasticities when we aggregate products in broader categories. The average elasticity is 4.5 at the SITC3 level and it increases to 5.6 at the HS10 level. Even if the differences appear small, we can statistically

Table D.1: CES Elasticities based on the DP Approach at Different Levels of Aggregation

	HS10	SITC5	SITC3
Mean	5.65	5.09	4.49
(SE)	(0.09)	(0.21)	(0.40)
Median	3.37	3.13	2.87
(SE)	(0.05)	(0.10)	(0.23)
N	8508	1296	147
T-statistics		2.493	2.836
Pearson χ^2 p-value		0.043	0.025

Note: Mean and median of the elasticities of substitution estimated with the DP approach for the products defined at the HS10, SITC5 and SITC3 levels of aggregation. Only feasible estimates are reported. Values above 130 are censored. Standard errors for each statistics are bootstrapped. T-statistics refer to a t-test for differences in mean with respect to the HS10 level; p-values for Pearson difference in median tests with respect to the HS10 level.

reject the null hypothesis that the mean elasticities are the same across all level of aggregations. Note also that the median elasticities of substitution exhibit the same qualitative pattern, as their values increase from 2.9 to 3.4. The median estimates at more aggregate levels (three and five digit) statistically differ from the most disaggregated level.[A18]

Comparison to Broda and Weinstein (2006) and Soderbery, 2015b Table D.2 compares the price elasticities estimated by the different strategies across different product classifications. First, note that the magnitude of the estimated price elasticities falls as we estimate them across more aggregated varieties, as varieties become less substitutable at these more aggregated levels. Comparing the magnitudes across different methods, we find that the elasticities estimated using DP are larger compared to those obtained using the FBW or LIML methods, in both mean and median terms, at all levels of aggregation. For instance, at the three-digit level, the mean elasticity for DP is 4.5, 50% greater than the number for FBW and more than twice that for LIML. Similarly, the median elasticity for DP is 2.8, while the value is 2.3 and 1.2 for the conventional methods FBW and LIML, respectively. We can easily reject the hypothesis that the means and the medians are the same. Figure D.8 in Appendix D.4 shows the strong correlation among the estimates found by the three methods.

Price Elasticities Across Different Rauch (1999) Product Classes Intuitively, we expect the magnitude of the price elasticities to be higher among more homogenous goods compared to more differentiated ones, since these homogenous goods should be more substi-

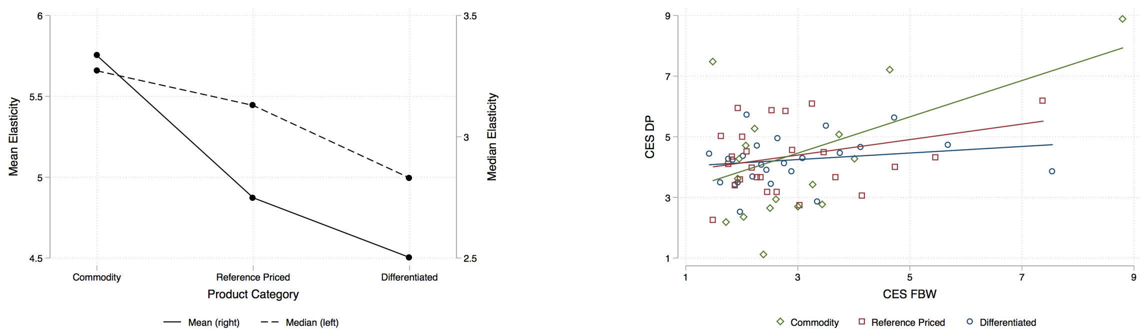
[A18]In contrast to the case of the mean estimates, we cannot statistically reject the hypothesis that the medians are the same at the SITC3 and SITC5 level.

Table D.2: Comparison between DP, FBW and LIML

	HS 10			SITC 5			SITC 3		
	DP	BW	LIML	DP	BW	LIML	DP	BW	LIML
Mean	5.70	4.64	4.50	5.09	3.44	3.21	4.49	2.97	1.70
(SE)	(0.15)	(0.09)	(0.11)	(0.23)	(0.13)	(0.15)	(0.45)	(0.39)	(0.11)
Median	3.35	2.74	2.10	3.08	2.43	1.65	2.79	2.29	1.23
(SE)	(0.05)	(0.02)	(0.02)	(0.10)	(0.04)	(0.04)	(0.25)	(0.08)	(0.03)
T-statistics		7.89	8.08		6.40	6.91		2.56	6.06
Pearson χ^2 p-value		0.00	0.00		0.00	0.00		0.03	0.00
N	7283	7283	7283	1140	1140	1140	127	127	127

Note: Mean and median of the elasticities of substitution estimated with the DP, FBW and LIML methods for the HS10, SITC5 and SITC3 levels of aggregation. Only feasible estimates for common products are reported. Values above 130 are censored. Standard errors for each statistics are bootstrapped. For each level of aggregation, T-statistics refer to a t-test for differences in mean with respect to DP; p-values for Pearson difference in median tests with respect to DP.

Figure D.2: DP Elasticities and Rauch Conservative Classification



Note: The left panel displays the mean and the median of the elasticities of substitution estimated with the DP approach for each category of the Rauch Conservative Classification at the SITC4 level of aggregation. The right panel shows the correlation between the DP and FBW estimates for each category of the Rauch Conservative Classification at the SITC4 level of aggregation.

tutable (Broda and Weinstein, 2006). We use the Rauch (1999) classification to distinguish products at the SITC4 level into three categories: commodities, referenced priced, and differentiated goods. Rauch (1999) provides two distinct classifications, "Liberal" and "Conservative", that only differ in a few products that can be classified in multiple ways. The left panel of Figure D.2 shows both the mean and the median elasticity for each Rauch Conservative category. Both these statistics are ranked in increasing order between commodities, referenced priced, and differentiated products, as expected. We can reject the hypothesis that the combined set of commodities and referenced priced goods have the same mean or median than differentiated products.[A19] Table D.3 reports the corresponding values and their standard errors for Figure D.2 and show that qualitative results holds also for the Liberal version of the classification.

In addition, again using the classification proposed by Rauch (1999), we can show

[A19]We statistically test the difference between differentiated products and the remaining categories pooled together. Differences are not statistically significant if the two categories are considered individually.

that the quality bias in the conventional estimates is stronger among more differentiated products. Intuitively, quality differentiation is less likely among homogeneous goods, suggesting that the DP estimates in this case should on average be closer to, and more correlated with, the conventional estimates. Consistently with this intuition, the right panel of Figure D.2 shows that the correlation between DP and FBW is stronger for commodities and the average difference between the two sets of estimates is smaller. As we consider less homogenous categories, referenced priced and differentiated products, the average quality bias increases while the correlation decreases.[A20] Figure D.3 shows that the qualitative pattern is robust to how products are grouped between homogenous and differentiated.

Table D.3: DP Estimates: Rauch Classifications

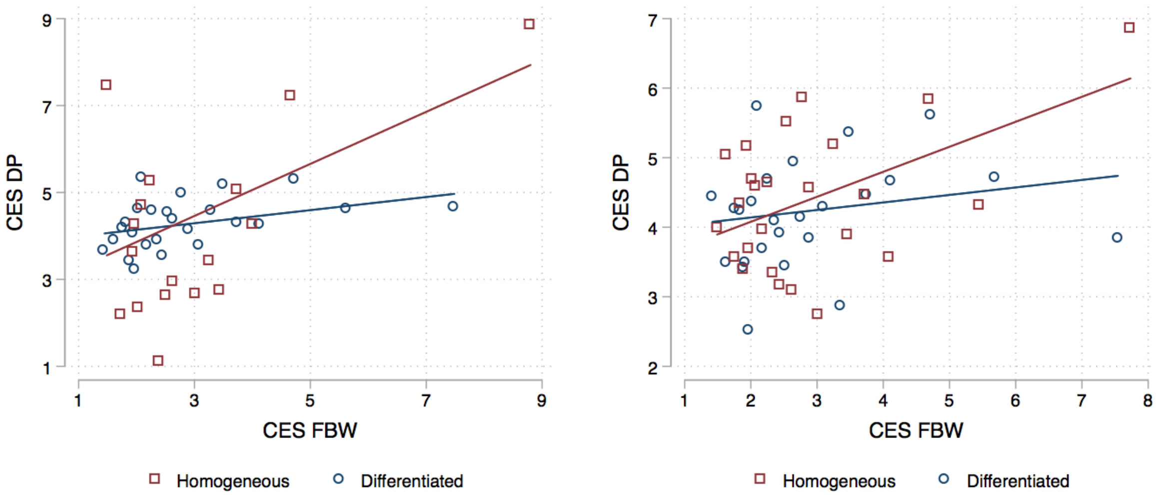
	Commodity	Reference Priced	Differentiated
Mean	5.75	4.87	4.50
(SE)	(0.86)	(0.42)	(0.25)
Median	3.27	3.13	2.83
(SE)	(0.69)	(0.18)	(0.18)
N	50	168	317

	Commodity	Reference Priced	Differentiated
Mean	5.28	4.77	4.58
(SE)	(0.63)	(0.42)	(0.27)
Median	3.24	3.10	2.82
(SE)	(0.37)	(0.18)	(0.21)
N	75	162	298

Note: For each category of the Rauch Classification (commodity, reference priced and differentiated), the tables report the mean and the median CES elasticity estimated using the DP approach at the SITC4 level. The left panel refers to the Conservative version of the classification (corresponding to Figure D.2 in the main text) while the right one to the Liberal version. It can be show that differences in mean and median are statistically significant at standard levels if the more homogeneous categories (commodities and reference priced) are pooled together and compared to differentiated products.

[A20]The average difference between group captures the average quality bias and is represented by the intercept of a linear regression (fitting line). The slope would capture instead the correlation across estimates.

Figure D.3: Correlation DP and FBW, Different Pooling of Rauch Categories



Note: The figure shows the correlation between the estimated elasticities using the DP and FBW methods at the SITC4 level using alternative breakdowns across products. Conservative Rauch classification is used. In the left panel, homogeneous products are defined as commodities only while, in the right panel, they include commodieties and reference priced goods.

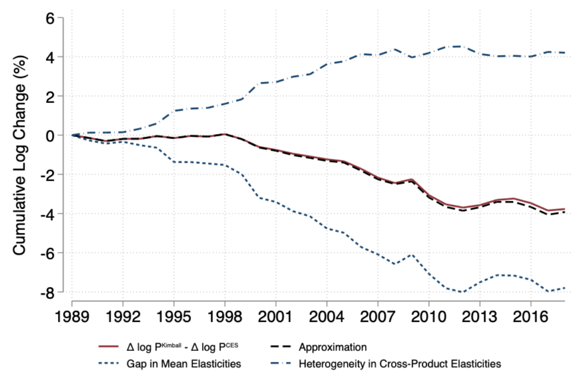
D.3 Further Results on Price and Quality Decomposition

D.3.1 Bias in Inferred Price Index: CES vs. Kimball

We can analyse the implications of the bias in elasticity estimation under the CES model for the construction of price indices. Equation (A.33) derived from Proposition A.4 in Appendix A.1.3 provides a decomposition of the gap between what Kimball and CES demand systems predict about the change in the aggregate price index. This gap is the sum of two terms: the gap between the mean of the love of variety indices across base products and the one implied by the CES model, and the contribution of heterogeneity in the matrix of cross-product elasticities of substitution, which is absent in the CES model.

Figure D.4 shows the cumulative gap between Kimaball and CES import price indices, and its decomposition into the two components based on Equation (A.33). The contribution of the second term, the heterogeneity in cross-product elasticity, is negative and explains more than 100% of the gap. The reason is the shift of expenditure shares within the common set of products away from those with higher love-of-variety indices. This negative covariance lowers the price index for the consumer, but cannot be captured in the CES model because cross-product elasticities are identical. The contribution of the first term, the gap in mean elasticity, is instead positive, partially offsetting the contribution of the heterogeneity in cross-product elasticity. Since the market share of the base

Figure D.4: Decomposition of the Gap in the Inferred Price Index - Kimball vs CES

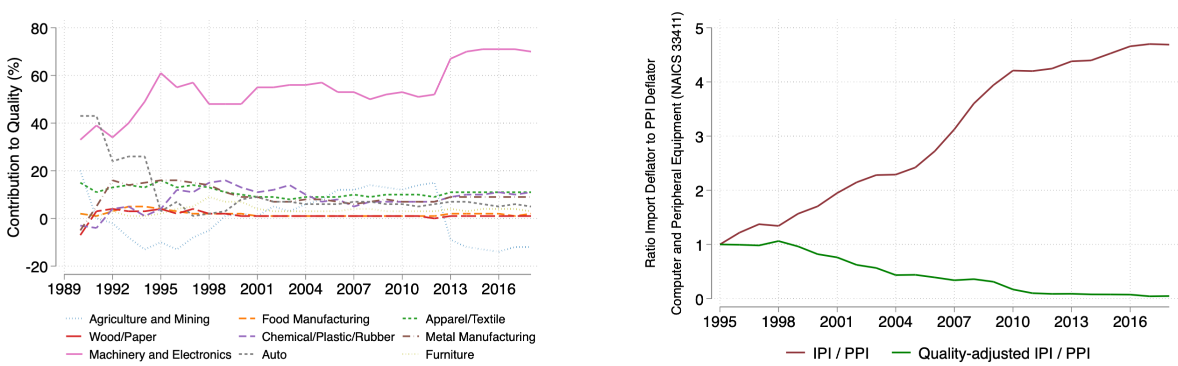


Note: The figure plots the decomposition of the gap between the inferred import price index under Kimball and that under CES. The solid red line represents the estimated difference between the import price index under Kimball and under CES. Import price indices are constructed according to Proposition A.4. The dashed black line represents the approximation of the gap according Equation (A.33). The approximation is the sum of two components: the gap in mean elasticity (dotted blue line), and the heterogeneity in cross-product elasticities (dashed blue line).

products, i.e. the U.S. variety, is falling over time, the key reason for the gap in mean elasticity being positive is simply that the estimated elasticity under Kimball (for the base products) is lower than the estimated CES elasticity. The dashed black line shows the sum of all the two terms in the approximation, is fairly close to the overall gap implied by the estimated Kimball and CES specifications (solid red line).

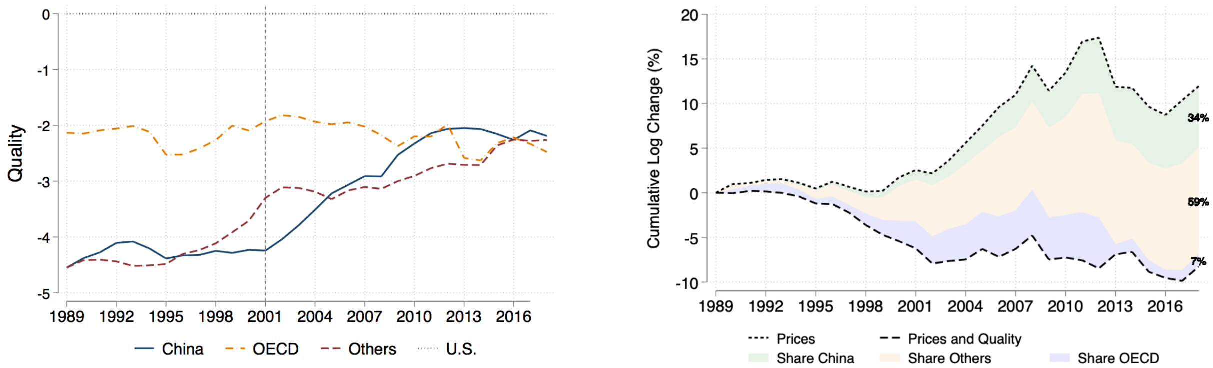
D.3.2 Quality Decomposition across Sectors

Figure D.5: Decomposition of Quality across Sectors



Note: The left panel plot the contribution of quality across industries in the Kimball model relative to the aggregate quality improvement. The quality contribution is computed using the inferred quality from the Kimball specification and constructing a Tornqvist index of quality changes at the sectoral and aggregate level. The right panel plot the ratio between the import price index and the producer price index for the Computer and Peripheral Equipment sector (NAICS 3341). The producer price index is from the BLS. The red line uses the official import price index from BLS, while the green line uses the import price index adjusted for the inferred quality from the Kimball specification.

Figure D.6: Decomposition of Quality across Countries



Note: The left panel decomposes the contribution of quality across countries in the Kimball model. The dashed line shows the price component of the aggregate import price index. The solid line shows the price component together with the quality component of the aggregate import price index. The quality contribution is computed using the inferred quality from the Kimball specification. Price indices and their decomposition are constructed from Proposition 2 and Equation (25). The difference between these two lines quantifies the role of quality changes and is decomposed into the role of Chinese varieties (green area), OECD varieties (purple area) and all other varieties pooled together (OECD area). The right panel shows the evolution of the (expenditure weighted) average quality of each (group of) exporter(s), China, OECD economies and rest of the world. Quality of imported varieties is expressed relative to the quality of the U.S. variety.

D.3.3 Quality Decomposition across Exporters

Figure D.6 shows the evolution of the expenditure-weighted quality for each (group of) exporter(s), China, OECD economies and all other countries. Recall that the quality of imported varieties is relative to the quality of the U.S. variety, which is used as base product. The (expenditure-weighted) average quality of Chinese varieties has increased since 1989 relative to the U.S. quality, which is normalized to zero over the entire time period. Notice also that the annual increase in quality is larger after China joined the WTO in 2001, suggesting that the trade liberalization shock boosted the sophistication process even more. This supports the extensive evidence that Chinese goods have undergone a sophistication process, catching up with more advanced economies and largely contributing to the aggregate quality improvement of U.S. imports.

Table D.4 shows that import quality has doubled over the time period from 1989 to 2018, increasing by almost 100%. This increase is exclusively driven by a rise in quality within each (group of) exporter(s) while compositional changes between exporters partially offset the within forces. This is consistent with the fact that Chinese products gained market share over the time period but still have lower quality compared to other exporters, even if they are catching up with the frontier.

D.3.4 Gains from Variety

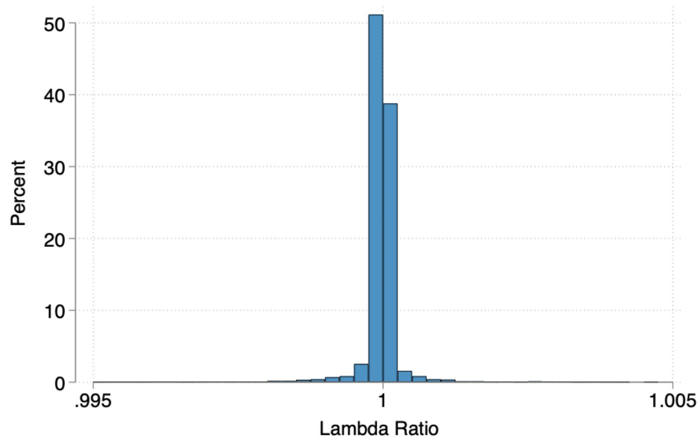
We can use Proposition A.4 to account for the contribution of new varieties to the changes in the aggregate import price index. At the level of disaggregation considered in our data (5-digit NAICS codes), we find that adjusting for this margin makes a negligible impact on the aggregate price indices. Improved product quality still constitutes the major driver of change in import prices. In fact, we find the overall impact of new varieties to be in the direction of raising prices. Over the period 1989-2018, the aggregate import price index increased by 0.3% under Kimball and 0.06% under CES, respectively, due to the exit of varieties. Figure D.7 shows that the distribution of Feenstra's λ ratios at the sector-year level is highly concentrated around the value of one. This is due to the aggregate nature of the dataset which defines sectors at the 5-digit NAICS level and varieties at level of the country of origin, limiting the scope of the variety channel.

Table D.4: Between and Within Decomposition

	$\Delta\varphi$	Δ within	Δ between
Full Sample	0.900	1.207	-0.306
Before 2001	0.477	0.657	-0.180
After 2001	0.423	0.549	-0.126

Note: The Table shows a decomposition of the growth in aggregate product quality between and within exporters. We consider China, OECD economies and all the other exporters pool together. For each exporter, we compute the aggregate product quality as the expenditure-weighted average across varieties.

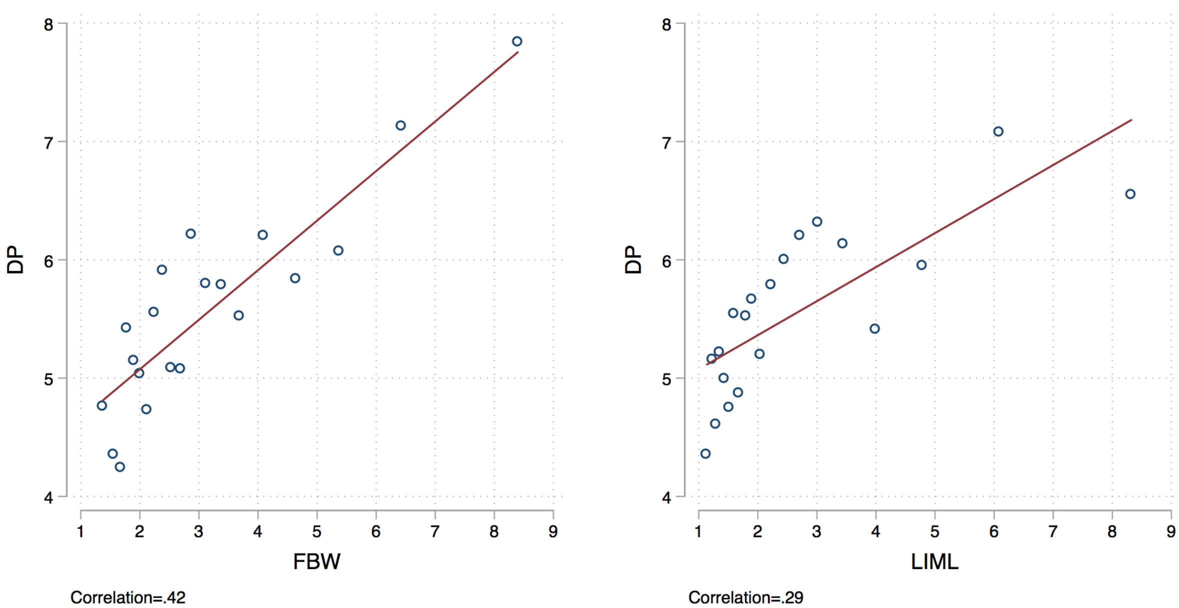
Figure D.7: Feenstra's Lambda Ratio



Note: The figure shows the histogram of the sector-year Feenstra's Λ ratios. The Λ ratio is computed as $\lambda_t = \frac{\Lambda^\star_t}{\Lambda^\star_{t-1}}$, where $\Lambda^\star$ is the total expenditure share of the continuing set.

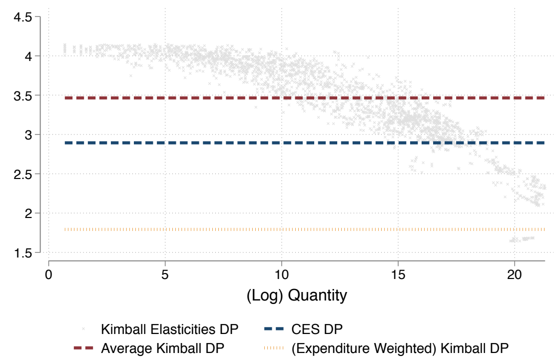
D.4 Additional Tables and Figures

Figure D.8: Correlation between DP and FBW or LIML Estimates, HS10 level



Note: The figure shows the binscatter plot of the relationship between the estimated elasticities using the DP approach and conventional methods like FBW (right panel) and LIML (left panel). The figures refers to the set of estimates at the HS10 level. Elasticities are censored at 10.

Figure D.9: CES - Kimball Elasticity



Note: The figure shows the entire set of Kimball demand elasticities as defined in Equation (12) of each variety-time pair as a function of the (log) quantity for the motor vehicle parts manufacturing industry, NAICS number 33639 (grey dots). The dot orange line represents the expenditure-weighted mean Kimball elasticity, while the dash red line represents the average Kimball elasticity. The dash blue line represents the CES estimated elasticity for the sector.

Table D.5: Kimball Parameters

	Lower bound σ	Upper bound σ	θ
Mean	2.36	245.7	6.95
	(0.23)	(33.1)	(3.08)
Median	1.15	7.34	0.33
	(0.16)	(1.38)	(0.040)
25th percentile	1.01	3.45	0.15
75th percentile	2.53	231.6	1.30

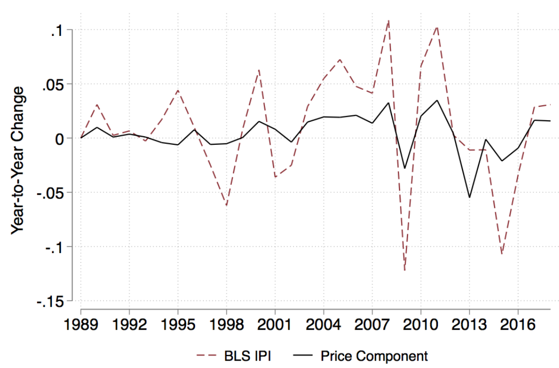
Note: The table displays the mean, the median, the 25th and the 75th percentile across all 5-digit NAICS industries, with the corresponding bootstrapped standard errors, of the estimated parameters of the Finite-Finite Kimball specification.

Table D.6: CES and Kimball Own-Price Elasticities

	Kimball - DP	CES - DP	CES - FBW
Mean	17.0	6.21	3.35
	(0.120)	(0.009)	(0.005)
Median	4.69	4.21	2.56
	(0.023)	(0.031)	(0.018)
Weighted Mean	1.26	1.96	1.57
	(0.005)	(0.005)	(0.005)
5th percentile	1.72	1.68	1.64
25th percentile	3.03	2.89	2.07
75th percentile	9.10	7.32	3.63
95th percentile	48.5	17.3	6.54

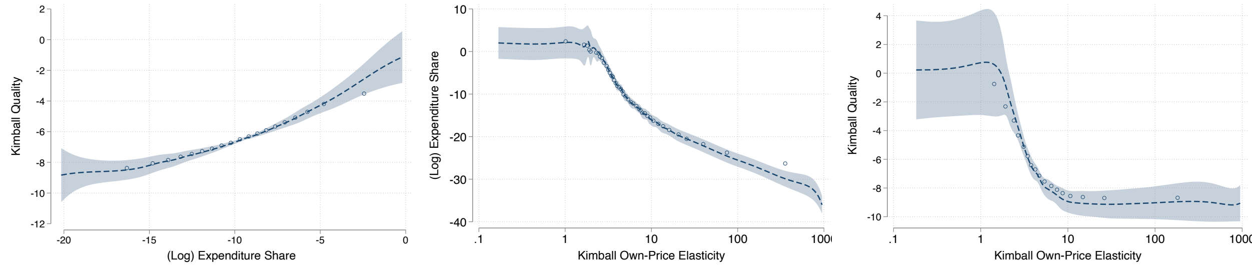
Note: The table reports the mean, median, the expenditure-weighted average, and the 5th, 25th, 75th and 95th percentiles of the distribution of own-price elasticities for both the Kimball and CES specifications. For the Kimball specification, we can compute the elasticity for each variety at each moment in time while, in the CES case, each variety-time pair is associated with the corresponding sectoral CES elasticity. For the CES case, we report the DP and the BW estimates. Standard errors are bootstrapped.

Figure D.10: Comparison with BLS Import Price Index



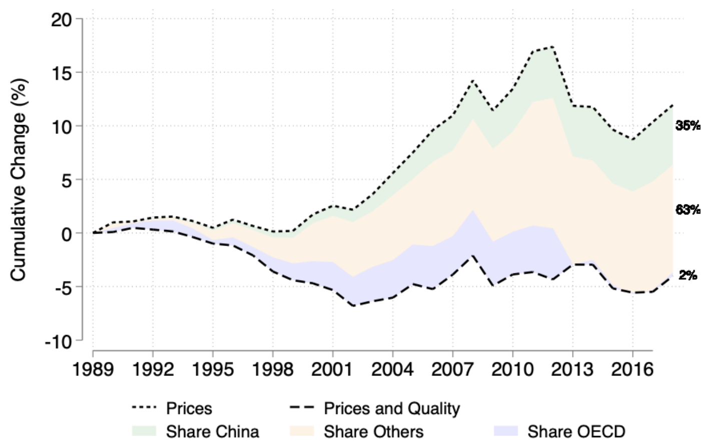
Note: The figure plots the year-to-year change in the BLS Import Price Index and a the price component of the aggregate import price constructed using Proposition A.4.

Figure D.11: Kimball Own-Price Elasticities and Implied Quality



Note: The left panel plots the binscattered relationship between (log) expenditure share of each variety-time observation and the inferred quality. The panel in the center plots the binscattered relationship between the Kimball own-price elasticity and the (log) expenditure share. The right panel directly plots the relationship between the inferred product quality and the Kimball own-price elasticity. In each panel, we use variety fixed effects and cluster the standard errors at the industry level.

Figure D.12: Price Index, Decomposition of Quality across Countries: CES case



Note: The dashed line figure shows the price component of the aggregate import price index. The solid line shows the price component and the quality component of the aggregate import price index. The quality contribution is computed using the inferred quality from the CES specification. The difference between these two lines quantifies the role of product quality change and is decomposed into the role of Chinese varieties (green area), OECD varieties (purple area) and all other varieties pooled together (orange area).

Table D.7: Change in the Import Price Index in the US, 1989–2018

	Total		PCE Index	Decomposition				
				Price	Quality		Variety	
	Kimball	CES			Kimball	CES	Kimball	CES
Cumulative Log Change (%)	-7.96	-3.96	57.8	11.9	-20.2	-16.0	0.30	0.061
Annual Change (%)	-0.27	-0.13	1.93	0.40	-0.67	-0.53	0.010	0.0020

Note: The table reports the cumulative and average annual change in the aggregate import price indices constructed using Proposition A.4 and Equation (25), and their decomposition. U.S. varieties are used to normalize the quality of imported goods. The third column reports the cumulative and annual change in the Personal Consumption Expenditure price index from the BEA.

E Details on the Validation using the US Auto Data

E.1 Data

Table E.1: Summary Statistics

	Mean	Std. Dev.	Min	Max
Sales	59106.19	86938.79	10.00	891482.00
Price	36.05	17.13	11.14	99.99
Miles/Gallon	20.94	6.58	10.00	50.00
Horsepower	192.18	83.88	44.00	645.00
Height	60.95	8.41	43.50	107.50
Footprint	13392.63	1968.92	6514.54	21821.86
Curbweight	3561.21	897.77	1113.00	8550.00
US Brand	0.44	0.50	0.00	1.00
Luxury	0.30	0.46	0.00	1.00
Electric	0.02	0.13	0.00	1.00
Sport	0.11	0.31	0.00	1.00
SUV	0.23	0.42	0.00	1.00
Truck	0.07	0.26	0.00	1.00
Van	0.07	0.25	0.00	1.00
N	9694			

Note: The table displays summary statistics of the main variables of our sample from the Wards Automotive Yearbooks. An observation is defined as a model-year pair. Prices are in thousands of 2015 US Dollars. Physical dimensions are in inches and curbweight is in pounds. The Electric dummy refers to EV (electric vehicles) and PHEV (plug-in hybrid electric vehicles). Observations with price higher than 100 thousands dollars are dropped.

E.2 Testing the Identification Assumption

We are able to test the identification assumption in Equation (18) leveraging the additional data on product characteristics available for the US auto market. The identification assumption relies on the orthogonality between demand shocks innovations, u_{it}, and lagged log prices and quantities. Under the assumption in Equation (16), the identification assumption between demand shocks innovations and lagged log prices can be rewritten as:

$$\mathbb{E}\left[\varphi_{it} | g_i\left(\varphi_{it-1}; \boldsymbol{\varrho}\right), \log p_{it-1}\right] \;=\; g_i\left(\varphi_{it-1}; \boldsymbol{\varrho}\right) + \alpha \, \log p_{it-1}.$$

where α is expected to be equal to zero when the orthogonality condition holds. Under the assumption that the demand shock process is a stationary AR(1) process, $g_i\left(\varphi_{it-1}; \boldsymbol{\varrho}\right) \equiv \rho\varphi_{it-1} + (1-\rho)\,\phi_i$ as in Equation (17), we use the set of characteristics available in our dataset as a proxy for φ_{it} and test whether the current value of product characteristics are correlated to lagged log prices after controlling for lagged characteristics. In other words, for each characteristic k, we estimate the following specification:

$$x_{kit} = \alpha \log p_{it-1} + \boldsymbol{\rho}_k' \boldsymbol{x}_{it-1} + \eta_t + \gamma_i + \epsilon_{it}, \tag{E.1}$$

where $\boldsymbol{x}_{it-1}$ is the entire set of lagged product characteristics. Table E.2 reports the set of coefficients estimated using Equation (E.1). No estimated $\widehat{\alpha}$ coefficients are statistically different from zero at standard level of significance, validating our identification assumption.[A21] Moreover, all product characteristics exhibit a strong degree of autocorrelation, supporting our choice for the process of demand shocks.[A22] We also standardize all variables and re-estimate Equation (E.1) in order to compare the coefficient of lagged price to the coefficients of lagged characteristics in terms of magnitude. Table E.3 shows that lagged product characteristics still exhibit strong and significant correlations, while lagged prices are not correlated to current product characteristics.

[A21] The only exception is Years since Design, which exhibits a statistically significant positive correlation with lagged price at the 5% significance level.

[A22] The only exception is Truck, which exhibits a weak autocorrelation.

Table E.2: Testing the Identification Assumption - Level Variables

	Horse Power	Footprint	Curbweight	Miles/Dollar	Years Since Design	Suv	Van	Truck
Lagged Price	0.043	-0.007	0.004	0.012	-0.043	0.028	-0.016	0.006
	(0.022)	(0.007)	(0.010)	(0.012)	(0.018)	(0.031)	(0.009)	(0.008)
Lagged Horse Power	0.560	0.015	0.009	-0.003	-0.010	-0.019	0.016	-0.009
	(0.023)	(0.006)	(0.007)	(0.011)	(0.011)	(0.015)	(0.012)	(0.008)
Lagged Footprint	0.087	0.547	0.144	-0.108	-0.010	-0.030	0.047	0.008
	(0.039)	(0.045)	(0.036)	(0.033)	(0.051)	(0.061)	(0.043)	(0.028)
Lagged Miles/Dollar	0.038	-0.005	-0.040	0.495	-0.047	0.014	0.005	0.024
	(0.029)	(0.010)	(0.013)	(0.026)	(0.015)	(0.022)	(0.009)	(0.018)
Lagged Curbweight	0.057	0.069	0.511	-0.128	-0.031	0.010	-0.013	0.046
	(0.031)	(0.020)	(0.035)	(0.026)	(0.022)	(0.033)	(0.018)	(0.047)
Lagged Years Since Design	-0.002	0.001	-0.001	-0.004	0.712	-0.008	-0.003	0.000
	(0.002)	(0.001)	(0.001)	(0.002)	(0.004)	(0.002)	(0.002)	(0.001)
Lagged Suv	-0.013	0.008	0.003	0.019	-0.036	0.257	0.030	-0.035
	(0.010)	(0.005)	(0.005)	(0.011)	(0.018)	(0.065)	(0.025)	(0.015)
Lagged Van	-0.020	0.004	0.000	0.014	-0.004	0.006	0.185	-0.060
	(0.018)	(0.003)	(0.004)	(0.009)	(0.012)	(0.034)	(0.075)	(0.034)
Lagged Truck	-0.001	0.016	0.020	0.006	0.019	-0.161	0.155	0.195
	(0.016)	(0.010)	(0.022)	(0.019)	(0.025)	(0.079)	(0.072)	(0.103)
N	6981	6981	6981	6981	6387	6981	6981	6981

Note: The table reports the coefficients estimated using Equation (E.1). Each column refers to a given product characteristics. We consider horsepower, footprint, miles-per-dollar, curbweight, years since design, truck, van and suv. All continuous variables are in logs. All regressions include model and year fixed effects. Standard errors are clustered at the producer level.

Table E.3: Testing the Identification Assumption - Zscored Variables

	Horse Power	Footprint	Curbweight	Miles/Dollar	Years Since Design	Suv	Van	Truck
Lagged Price	0.058	-0.026	0.010	0.020	-0.036	0.040	-0.038	0.013
	(0.030)	(0.029)	(0.023)	(0.019)	(0.015)	(0.045)	(0.022)	(0.020)
Lagged Horse Power	0.560	0.046	0.016	-0.004	-0.006	-0.020	0.028	-0.015
	(0.023)	(0.017)	(0.012)	(0.013)	(0.007)	(0.015)	(0.021)	(0.014)
Lagged Footprint	0.029	0.547	0.086	-0.042	-0.002	-0.011	0.028	0.005
	(0.013)	(0.045)	(0.021)	(0.013)	(0.010)	(0.022)	(0.025)	(0.016)
Lagged Miles/Dollar	0.033	-0.014	-0.061	0.495	-0.024	0.013	0.007	0.035
	(0.025)	(0.025)	(0.020)	(0.026)	(0.008)	(0.020)	(0.014)	(0.027)
Lagged Curbweight	0.032	0.116	0.511	-0.084	-0.011	0.006	-0.013	0.045
	(0.017)	(0.034)	(0.035)	(0.017)	(0.008)	(0.019)	(0.018)	(0.046)
Lagged Years Since Design	-0.003	0.006	-0.003	-0.007	0.712	-0.014	-0.008	0.000
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.006)	(0.004)
Lagged Suv	-0.012	0.021	0.006	0.021	-0.021	0.257	0.052	-0.057
	(0.009)	(0.013)	(0.008)	(0.013)	(0.010)	(0.065)	(0.043)	(0.024)
Lagged Van	-0.011	0.007	0.000	0.009	-0.001	0.004	0.185	-0.059
	(0.010)	(0.005)	(0.004)	(0.006)	(0.004)	(0.020)	(0.075)	(0.033)
Lagged Truck	-0.001	0.027	0.021	0.004	0.007	-0.098	0.160	0.195
	(0.009)	(0.017)	(0.023)	(0.013)	(0.009)	(0.048)	(0.074)	(0.103)
N	6981	6981	6981	6981	6387	6981	6981	6981

Note: The table reports the coefficients estimated using Equation (E.1). Each column refers to a given product characteristics. We consider horsepower, footprint, miles-per-dollar, curbweight, years since design, truck, van and suv. For each characteristic, Equation (E.1) is estimated using z-score variables, i.e. after standardizing all variables. All continuous variables are in logs. All regressions include model and year fixed effects. Standard errors are clustered at the producer level.

E.3 Heterogeneity Bias in Elasticity Estimation

E.3.1 A Theory of Heterogeneity Bias

We can build some intuition about the drivers of the heterogeneity bias in the estimation of demand elasticities in the following simplified setting. Consider a stylized demand model featuring heterogeneity in elasticities of substitution of the form

$$\Delta \log q_{it} = -\sigma_{it} \Delta \log p_{it} + \Delta\varphi_{it},$$

with an extral cost shifter w_{it} such that $\Delta \log p_{it} = \zeta w_{it} + v_{it}$ where $\mathbb{E}[w_{it}v_{it}] = 0$ and $\mathbb{E}[w_{it}\Delta\varphi_{it}] = 0$. Consider an estimator that aims to estimate the mean elasticity using a CES demand approximation of this model with the following estimator

$$\widehat{\sigma} \equiv -\frac{\mathbb{E}[w_{it} \Delta \log q_{it}]}{\mathbb{E}[w_{it} \Delta \log p_{it}]}.$$

If the underlying demand system is indeed CES, this estimator is unbiased. However, in the presence of heterogeneity, we have

$$\begin{aligned}
\widehat{\sigma} &= \frac{\mathbb{E}[\sigma_{it} w_{it} \Delta \log p_{it}]}{\mathbb{E}[w_{it} \Delta \log p_{it}]}, \\
&= \mathbb{E}[\sigma_{it}] - \frac{\mathbb{C}(\sigma_{it}, w_{it}\Delta \log p_{it})}{\mathbb{E}[w_{it} \Delta \log p_{it}]}, \\
&= \mathbb{E}[\sigma_{it}] - \frac{\mathbb{C}\left(\sigma_{it}, w_{it}^2\right)}{\mathbb{E}\left[w_{it}^2\right]} - \frac{\mathbb{C}(\sigma_{it}, w_{it}v_{it})}{\mathbb{E}\left[w_{it}^2\right]}
\end{aligned}$$

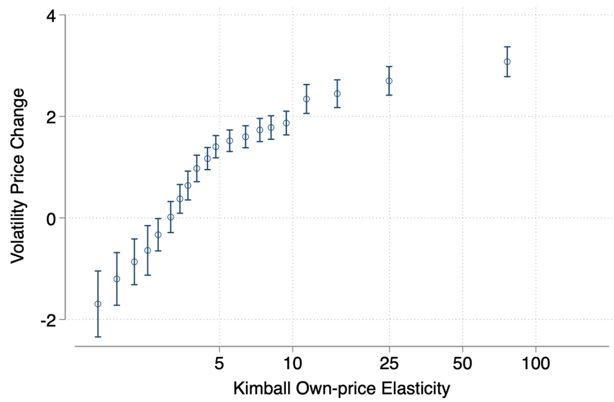
To simplify the setting, let us assume that $\mathbb{C}(\sigma_{it}, w_{it}v_{it}) = 0$ to find

$$\widehat{\sigma} = \mathbb{E}[\sigma_{it}] - \frac{\mathbb{C}\left(\sigma_{it}, w_{it}^2\right)}{\mathbb{E}\left[w_{it}^2\right]}.$$

We find that the estimated CES elasticity is below the population mean of the elasticities if there is a positive covariance between the elasticity parameters and the *magnitude* of the cost shifter. This would indeed be the case if products with higher own-price elasticities of demand are those that have more volatile prices.

We test whether the data we use in our setting are consistent with the negative heterogeneity bias we document. Figure E.1 below confirms the presence of a strong positive covariance between the estimated own-price elasticity in the Kimball model and the volatility of price changes, after absorbing for variety and year fixed effects (robust to

Figure E.1: Covariance Own-price Elasticity - Price Change Volatility



Note: The figure reports the binscatter relationship between the estimated own-price elasticities of the Finite-Finite Kimball (estimated using DP) and the volatility in price changes. We absorb for variety and year fixed effects. We plot the 95

alternative set of fixed effects, such as sector-year).

E.3.2 Evidence in the Data

E.4 Details on the Mixed Logit Specification

E.4.1 Mixed Logit Demand

The standard Mixed Logit model provides a usefull benchmark setting that allows for variable and heterogeneous cross-product substitution elasticities. We provide a definition of this model that is consistent with a corresponding market demand system rationalizable along the lines considered in Section 2. We consider a differentiated product demand in which consumers make a discrete choice among the car models available j and an outside "no-purchase" option. For each household n, we have that the utility from consuming product i is:

$$u_t^n = \max_j \left\{ \frac{y_t^n - p_{it}}{p_{ot}} + \left(\widetilde{\boldsymbol{\beta}}^n\right)' \boldsymbol{x}_{it} + \widetilde{\xi}_{it}^n + \frac{1}{\alpha^n}\, \epsilon_{it}^n \right\}, \tag{E.2}$$

where p_{ot} is the price of the consumption of non-auto (outside) goods, and $x_{ot} = 0$ is the characteristic of buying no automobile, $\frac{y_t^n - p_{it}}{p_{ot}}$ is the quantity of outside good purchased after selecting automobile model i,$\boldsymbol{x}_{it}$ is a vector of vehicle attributes, p_{it} is the price of automobile model i, $\widetilde{\xi}^n_{it}$ is an unobserved vehicle-specific term with $\xi^n_{ot} \equiv \gamma_t$ for the demand shifter for the outside good (no purchase), and ϵ^n_{it} is an idiosyncratic consumer-vehicle specific term distributed according to a type-II extreme value distribution. As is standard, we model consumer heterogeneity using parametric distributions allowing for unobservable heterogeneity. The corresponding uncompensated (Marshallian) demand is given by

$$\widetilde{q}_i^{uc}\left(\boldsymbol{p}_t; y_t\right) = \sum_n \omega^n \frac{\exp\left(-\alpha^n \frac{p_{it}}{p_{ot}} + \boldsymbol{x}'_{jt}\boldsymbol{\beta}^n + \xi^n_{it}\right)}{\exp\left(\gamma_t\right) + \sum_{i' \neq o} \exp\left(-\alpha^n \frac{p_{i't}}{p_{ot}} + \boldsymbol{x}'_{i't}\widetilde{\boldsymbol{\beta}}^n + \xi^n_{i't}\right)}, \qquad i \neq o, \tag{E.3}$$

where ω^n is the weight corresponding to household n, where we have defined $\boldsymbol{\beta}^n \equiv \widetilde{\boldsymbol{\beta}}^n/\alpha^n$ and $\xi^n_{it} \equiv \widetilde{\xi}^n_{it}/\alpha^n$.

E.4.2 Estimation

We closely follow Grieco et al. (2021) for the estimation of this demand system. In the first step, we leverage household demographics and second choice moments to estimate consumer heterogeneity and the mean valuation. In the second step, we use the estimated mean valuation to estimate the mean taste for product characteristics, employing an IV regression where prices are instrumented with the real exchange rate or using DP instruments. We differ from Grieco et al. (2021) in the treatment of the outside option. We proxy the price of the consumption of non-auto (outside) option, p_{ot}, as the Personal Consumption Expenditure (PCE) price index from the BEA, net of the price for the purchases on the auto market. In other words, we define the outside good as the PCE consumption basket, excluding the expenditure on the auto market. We then divide the prices of the car models, p_{it}, by the price of the consumption of non-auto (outside) option p_{ot} as in the expression in Equation (E.3).

E.4.3 Mixed Logit Price Index for the Auto Industry

We derive the analytical exact price index of the preferences in Equation (E.2). Let $E\left(u; \boldsymbol{p}_t, \boldsymbol{\xi}_t, \boldsymbol{x}_t\right)$ be the expenditure function for the demand system, which we can write as

$$E\left(u; \boldsymbol{p}_t, \boldsymbol{\xi}_t, \boldsymbol{x}_t\right) = p_{ot}\left[u - \left(\sum_n \frac{\omega^n}{\alpha^n}\right)\gamma_t - \sum_n \frac{\omega^n}{\alpha^n} \log\left(1 + \sum_{j \neq o} \exp\left(\delta_{jt} + \mu^n_{jt}\right)\right)\right], \tag{E.4}$$

where we have defined δ_{it} and μ_{it}^n as

$$\delta_{it} \equiv -\overline{\alpha} p_{it} + \overline{\boldsymbol{\beta}}' \boldsymbol{x}_{it} + \overline{\xi}_{it} - \gamma_t,$$
$$\mu_{it}^n \equiv -\left(\alpha^n - \overline{\alpha}\right) \frac{p_{it}}{p_{ot}} + \boldsymbol{x}_{jt}' \left(\boldsymbol{\beta}^n - \overline{\boldsymbol{\beta}}\right) + \xi_{it}^n - \overline{\xi}_{it}.$$

where $\overline{\alpha} \equiv \sum_n \omega^n \alpha^n$, $\overline{\boldsymbol{\beta}} \equiv \sum_n \omega^n \boldsymbol{\beta}^n$, and $\overline{\xi}_{it} \equiv \sum_n \omega^n \xi_{it}^n$ denote the household means, and γ_t denotes the quality of the outside good.

We construct a price index for the auto industry based on the Mixed Logit demand. Following the construction in Section 2.4.3, the set of automobiles constitute a subcategory of all products in the demand system that excludes the outside good. The Divisia index for the entire demand system between two consecutive periods can be written as follow

$$\log D_t \equiv \log \frac{E\left(u_{t-1}; \boldsymbol{p}_t, \boldsymbol{\xi}_t, \boldsymbol{x}_t\right)}{E\left(u_{t-1}; \boldsymbol{p}_{t-1}, \boldsymbol{\xi}_{t-1}, \boldsymbol{x}_{t-1}\right)}, \tag{E.5}$$

$$= \log\left(\frac{p_{o,t}}{p_{o,t-1}}\right) + \log\left(1 - \frac{\left(\sum_n \frac{\omega^n}{\alpha^n}\right)\left(\gamma_t - \gamma_{t-1}\right) + \sum_n \frac{\omega^n}{\alpha^n} \log\left(\frac{1+\sum_{j\neq o} \exp\left(\delta_{jt} + \mu_{jt}^n\right)}{1+\sum_{j\neq o} \exp\left(\delta_{jt-1} + \mu_{jt-1}^n\right)}\right)}{y_{t-1}/p_{o,t-1}}\right). \tag{E.6}$$

which is found using Equation (E.4) and noting that

$$u_{t-1} = \frac{y_{t-1}}{p_{o,t-1}} + \left(\sum_n \frac{\omega^n}{\alpha^n}\right)\gamma_{t-1} + \sum_n \frac{\omega^n}{\alpha^n} \log\left(1 + \sum_{j\neq o} \exp\left(\delta_{jt-1} + \mu_{jt-1}^n\right)\right).$$

Equation (E.6) is not directly comparable to the price indices constructed using Proposition A.4 for the Kimball and CES cases as it accounts for the dynamics of the price and quality of the outside good. We can define the Divisia index for the auto-industry as

$$\begin{aligned}
\log D_t^{auto} &\equiv \log \frac{E\left(u_{t-1}; \boldsymbol{p}_t^{o,+}, \boldsymbol{\xi}_t^{o,+}, \boldsymbol{x}_{t-1}\right)}{E\left(u_{t-1}; \boldsymbol{p}_{t-1}, \boldsymbol{\xi}_{t-1}, \boldsymbol{x}_{t-1}\right)}, \\
&= \log\left(\frac{p_{o,t}}{p_{o,t-1}}\right) \\
&\quad + \log\left(1 - \frac{\left(\sum_n \frac{w^n}{\alpha^n}\right)\left(\gamma_t - \gamma_{t-1}\right) + \sum_n \frac{w^n}{\alpha^n} \log\left(\frac{1+\sum_{j\neq o} \exp\left(\delta_{jt-1} + \mu_{jt-1}^n - \left(\gamma_t - \gamma_{t-1}\right) - \alpha^n p_{i,t-1}\left(\frac{1}{p_{o,t}} - \frac{1}{p_{o,t-1}}\right)\right)}{1+\sum_{j\neq o} \exp\left(\delta_{jt-1} + \mu_{jt-1}^n\right)}\right)}{y_{t-1}/p_{o,t-1}}\right),
\end{aligned}$$

where $\boldsymbol{p}_{t-1}^{o+} \equiv \left(p_{ot}, p_{1,t-1}, \cdots, p_{J,t-1}\right)$ and $\boldsymbol{\xi}_{t-1}^{o+} \equiv \left(\gamma_t, \xi_{1,t-1}, \cdots, \xi_{J,t-1}\right)$ are the vectors of prices and demand shifters that are only adjusted for the next-period outside good. We can now back out the contribution of the auto market to the aggregate price index net of the dynamics of the outside good:

$$\Delta \log P_t^{auto} = \frac{1}{1 - s_{o,t}} \left(\log D_t - \log D_t^{auto}\right), \tag{E.7}$$

where s_{ot} is the expenditure share of non-auto (outside) goods in the total consumption basket, which is measured by the per-capita personal consumption expenditure net of purchases in the auto market using data from the BEA.

E.5 Details on the Mixed CES Specification

We face several challenges in using the construction in Section E.4.3 to build a price index for the auto industry. The Mixed Logit demand is quasi-linear and specifies the demand for automobiles relative to an outside good. As Equation (E.7) shows, we need to determine the price and the share of consumption expenditure on this outside good to construct the desired price index. The simplest choice, the one we make here, is that the outside good is the remainder of household consumption, with a price p_{ot} corresponding to PCE net of automobile prices. However, in this setting, it is not clear that the patterns of substitutability implied by the Mixed Logit model between each product and the outside good capture that between new automobile purchases with alternative modes of transportation/commuting, or with the rest of household expenditure. More importantly, since the share of automobile expenditure in total household expenditure is fairly small, the price index in Equation (E.6) is dominated by the price of the oustide good and Equation (E.7) is highly sensitive to the fluctuations in the share of automobile purchases in total expenditure, as is evident from Figure 5.

To address these issues we additionally introduce an alternative Mixed CES demand system, which closely parallels the construction and the estimation of the Mixed Logit demand but allows us to assume separability between the demand between the auto industry and the remainder of the consumer expenditure.

E.5.1 Mixed CES Demand

We define the Mixed CES demand as an income-invariant system defined only over the bundle of new automobile purchases. We assume that all consumers dedicate the same level y^{auto} to their expenditure on the auto sector. Allowing for non-integer quantities, we assume that once a consumer chooses an automobile model, they spend the budget they have allocated to their automobile purchases to this model, buying y_t^{auto}/p_{it} units of that car. Accordingly, we assume that the utility of consumer n at time t from their automobile purchases is given by

$$u_t^n = \max_i \left\{ \log\left(\frac{y_t^{auto}}{p_{it}}\right) + \left(\widetilde{\boldsymbol{\beta}}^n\right)' \boldsymbol{x}_{it} + \widetilde{\xi}_{it}^n + \frac{1}{\alpha^n}\epsilon_{it}^n \right\},$$

where all the other variables are defined as in Equation (E.2). The probability of choosing item i by household n is then given by

$$\mathbb{P}_t^n(i) = \frac{\exp\left(-\alpha^n \log p_{it} + \boldsymbol{x}_{jt}'\boldsymbol{\beta}^n + \xi_{it}^n\right)}{\sum_{i'} \exp\left(-\alpha^n \log p_{i't} + \boldsymbol{x}_{i't}'\widetilde{\boldsymbol{\beta}}^n + \xi_{i't}^n\right)}, \tag{E.8}$$

while the quantity purchased across all households is given by $q_{it} = y_t^{auto}/p_{it} \cdot \sum_n \omega^n \mathbb{P}_t^n(i)$. The corresponding compensated (Hicksian) demand is given by the following income-invariant expression

$$\widetilde{s}_i(\boldsymbol{p}_t; u) = \sum_n \omega^n \frac{\exp\left(-\alpha^n \log p_{it} + \boldsymbol{x}_{jt}'\boldsymbol{\beta}^n + \xi_{it}^n\right)}{\sum_{i'} \exp\left(-\alpha^n \log p_{i't} + \boldsymbol{x}_{i't}'\widetilde{\boldsymbol{\beta}}^n + \xi_{i't}^n\right)},$$

where ω^n is the weight corresponding to household n, where we have defined $\boldsymbol{\beta}^n \equiv \widetilde{\boldsymbol{\beta}}^n/\alpha^n$ and $\xi_{it}^n \equiv \widetilde{\xi}_{it}^n/\alpha^n$. Note that the probability of choosing item i by household n is

E.5.2 Estimation

In terms of estimation, the key difference from the Mixed Logit model lies in the observation that, in this model, the probability of choosing item i at time t is the same as the expenditure share of the model, i.e.,$s_{it} = \sum_i \omega^n \mathbb{P}_t^n(i)$. Accordingly, we adapt the estimation strategy of Grieco et al. (2021), modifying the first step to recover mean valuations using products' expenditure shares rather than their sales shares. Since the estimation requires us to specify the expenditure share of the oustide option, we need to determine the expenditure share of the outside good, or equivalently, the price of the outside option, p_{ot}. We assume the price of the outside transportation option to be \$5,000, which allows us to generate a distribution of expenditure shares mirroring the distribution of sales shares.

Table E.4 presents the estimated elasticities from the Mixed CES model and compares them to those from alternative specifications. The first three columns report the estimated price coefficients under the CES specification: (i) using OLS estimation, (ii) using the RER variable as a cost shock instrument (henceforth IV), and (iii) applying our DP approach. The remaining columns show various moments of the distribution of the estimated own-price elasticities for the two models with variable elasticities—namely, the Mixed CES and Kimball specifications. Consistent with the findings from the Mixed Logit model, we observe that all three moments of the distribution of the estimated own-price elasticities in the Kimball model are similar to those in the Mixed CES model.

Table E.4: Comparing Own-price Elasticities

	CES			Mixed CES	Kimball	
	OLS	IV	DP	IV	IV	DP
σ	1.35	4.67	4.51			
	(0.25)	(1.47)	(0.13)			
Own-price Elasticity:						
Weighted Mean		4.62	4.46	8.09	5.60	5.79
		(0.00)	(0.00)	(0.02)	(0.01)	(0.01)
Mean		4.65	4.50	8.54	8.90	8.68
		(0.00)	(0.00)	(0.02)	(0.05)	(0.05)
Median		4.66	4.51	8.42	7.41	7.45
		(0.00)	(0.00)	(0.03)	(0.03)	(0.03)
IQR		0.02	0.02	2.60	3.54	3.16
		(0.00)	(0.00)	(0.03)	(0.07)	(0.06)

Note: The table reports the estimated own-price elasticities from the full sample. Each column corresponds to a different econometric model: CES OLS, CES IV, CES DP, Mixed CES IV, Kimball IV, and Kimball DP. For the VES cases (Mixed CES and Kimball) we report a set of moments from the distribution of the estimated own-price elasticities, while for the CES cases, we also report the estimated price coefficients. We report the mean and the median elasticity together with the expenditure weighted mean elasticity and the interquartile range. We consider the Finite-Finite case for the Kimball specification. For the Kimball and CES specification, we compute the own-price elasticities using Equation (**??**) and (11). We report bootstrapped standard errors for the set of moments from the distribution of the estimated own-price elasticities. For the CES price coefficeints, standard errors are clustered at product (model) level. All estimated quantities use the full sample.

E.5.3 Mixed CES Price Index for the Auto Industry

Given the separability between auto and non-auto purchases and the income-invariance properties of the Mixed CES demand system, it is straightforward to construct the price index. This price index is given by

$$\log P_t = \sum_n \frac{\omega^n}{\alpha^n} \log \left(\sum_i \exp \left(-\alpha^n \log p_{it} + \boldsymbol{x}'_{jt}\boldsymbol{\beta}^n + \xi^n_{it} \right) \right).$$

We can also use the approximation by computing the matrix $\mathbf{\Sigma}_t$, and applying the results in Proposition A.4 to construct the price index for the Mixed CES case.

E.6 Inferred Quality and Product Characteristics

In this section, we quantify the correlation between our inferred measures of quality and the product characteristics valued by consumers available in our dataset, which is not feasible using standard customs data. We again compare the results of our DP approach for the CES specification to alternative estimation strategies such as OLS and the standard IV approach using RER. We also explore the implications of accounting for heterogeneity in price elasticities for the inferred quality (compared to the standard CES case).

Using detailed data on the US automobile market allows us to examine whether our approach retrieves meaningful measures of quality. We examine this question by quantifying the correlation between our inferred measures of quality and the product characteristics valued by consumers available in our dataset, which is not feasible using standard customs data. We again compare the results of our DP approach for the CES specification to alternative estimation strategies such as OLS and the standard IV approach using RER. We also explore the implications of accounting for heterogeneity in price elasticities for the inferred quality (compared to the standard CES case).

In the CES case, the inferred quality of each product i at time t is computed according Equation (15) in which we use the elasticity estimated using the DP approach and reported in Table 3. Similarly, inverting the Kimball demand, we infer the measure of product quality for the Kimball case using Equation (15).[A23] We then study the correlation between the quality measure φ_{it} (inferred using either the CES or Kimball estimates) and a subset of product characteristics tightly linked to product quality in this specific market, e.g., horsepower, footprint, miles-per-dollar and style (i.e. truck, suv, and van):

$$\varphi_{it} = \beta \boldsymbol{x}_{it} + \eta_t + \gamma_i + \epsilon_{it}, \tag{E.9}$$

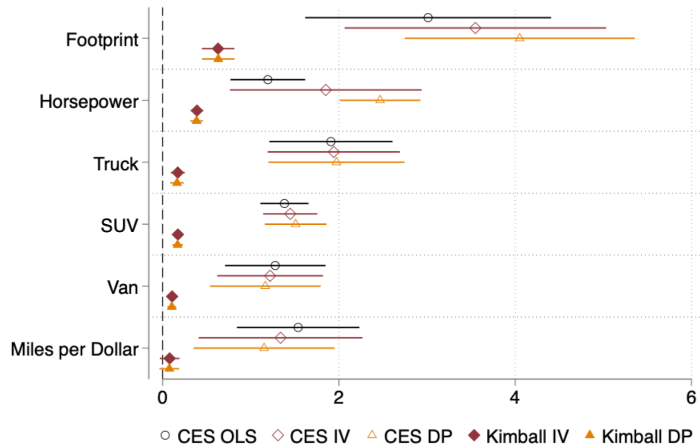
where $\boldsymbol{x}_{it}$ is the set of characteristics listed above. The correlation coefficients estimated from regression (E.9) are compared against the coefficients estimated from Equation (29) above.[A24]

Figure E.2 shows that the inferred quality estimated using DP and using the cost shock (RER) identification are related to product characteristics almost identically, in both the

[A23]See the discussion in Appendix B.2.2 for more details on inverting the Kimball demand.

[A24]In other words, we re-estimate Equation (29) above using the same set of product characteristics and fixed effects as in regression (E.9).

Figure E.2: Correlation between Inferred Quality and Product Characteristics

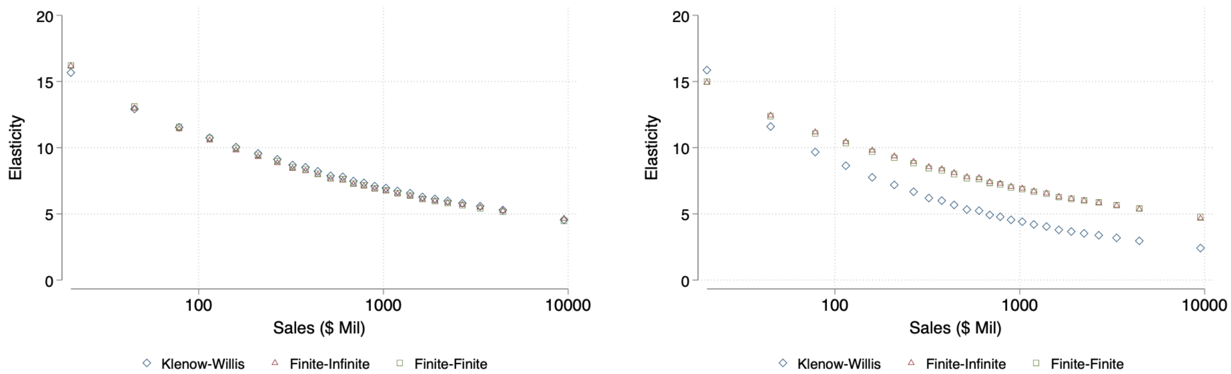


Note: The figure reports the relationship between product characteristics and inferred quality. In the CES DP case, the inferred quality measure follows from Equation (15). For the Kimball specification, inferred quality is obtained inverting demand as in Appendix *B*.2.2. The coefficients referring to the DP approach (CES and Kimball) and the Kimball IV case are obtained from regression in Equation (*E*.9). We consider the following product characteristics: horsepower, footprint, miles-per-dollar and style (suv, truck, van). Continuous variables are in log. The coefficients referring to the OLS and IV estimates of the CES specification are obtained from Equation (29), where product characteristics are used to proxy for quality. All regressions use the entire sample and includes time and product (model) fixed effects. Standard errors are clustered at the producer level, the bands around the estimates show the 95% confidence intervals.

CES and the Kimball specifications. This is a direct consequence of the ability of the DP approach to correctly estimate own-price elasticities, as shown in the previous section. Notice that the correlation between inferred quality and product characteristics differs across model specifications. Even though the correlations exhibit the same qualitative patterns, the magnitude is stronger in the CES specification compared to Kimball. The quantitative difference across models suggests that accounting for heterogeneity in price elasticity has a first order role in quantifying the role of quality.

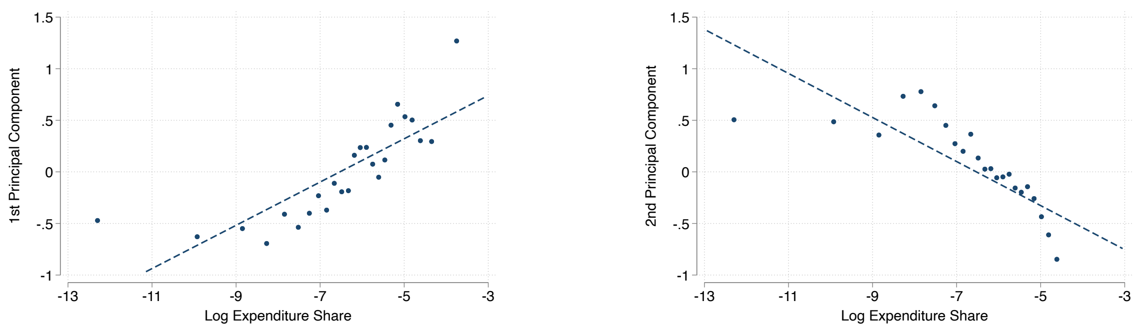
E.7 Additional Tables and Figures

Figure E.3: Comparison across Kimball Specifications



Note: The left panel shows a binscatter representation of the relationship between sales in millions of US dollars and the Kimball own-price elasticities estimated using the DP approach. The right panel shows the relationship between sales in millions of US dollars and Kimball price elasticities estimated using the IV approach. All three Kimball specifications (Finite-Finite, Finite-Infinite, and Klenow-Willis) are considered. Observations with sales less than $10mil are dropped.

Figure E.4: PCA - Correlation Market Shares



Note: The left (right) panel shows the relationship between the (log) expenditure shares and the first (second) principal component scores from a singular value decomposition of product characteristics. The following characteristics are included: footprint, horse-power, miles-per-gallon, curbweight, height, style dummies, luxury dummy, electric dummy, US brand dummy. We standardize all variables before performing the singular value decomposition.

Table E.5: PCA - Correlation Market Shares

	Full Sample	Car	Suv	Truck	Van
1st Principal Component	0.12	0.07	-0.03	0.17	0.15
	(0.01)	(0.01)	(0.01)	(0.02)	(0.03)
2nd Principal Component	-0.31	-0.32	-0.53	-0.90	-0.15
	(0.01)	(0.02)	(0.02)	(0.06)	(0.05)
N	9694	6126	2242	684	642

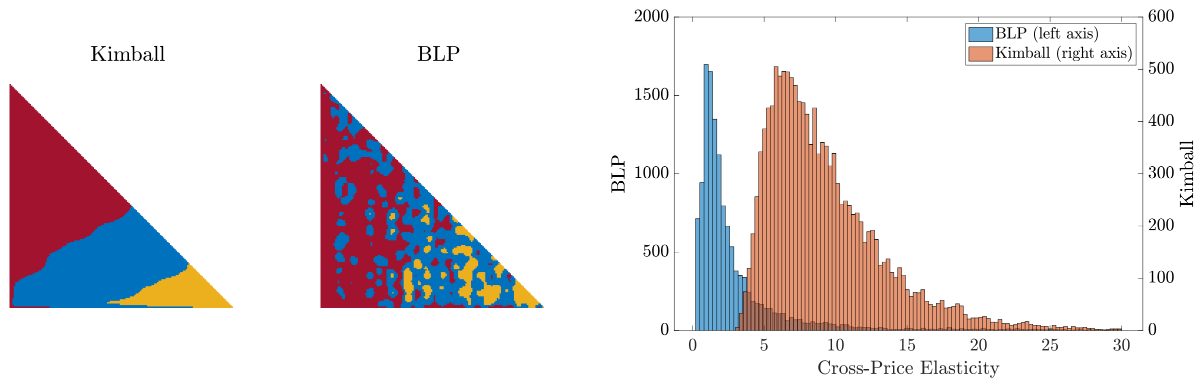
Note: The table reports the correlation coefficients between the log expenditure shares and the first two principal component scores from a singular value decomposition of product characteristics. The following characteristics are included: footprint, horsepower, miles-per-gallon, curbweight, height, style dummies, luxury dummy, electric dummy, US brand dummy. We standardize all variables before performing the singular value decomposition. The first column uses all sample, while all the remaining columns consider a product style (car, SUV, truck, and van) at the time.

Table E.6: PCA - Variance Decomposition

	Full Sample	Car	Suv	Truck	Van
1st Principal Component	33.83	36.99	41.32	55.65	42.06
2nd Principal Component	13.37	14.06	14.32	10.79	14.56

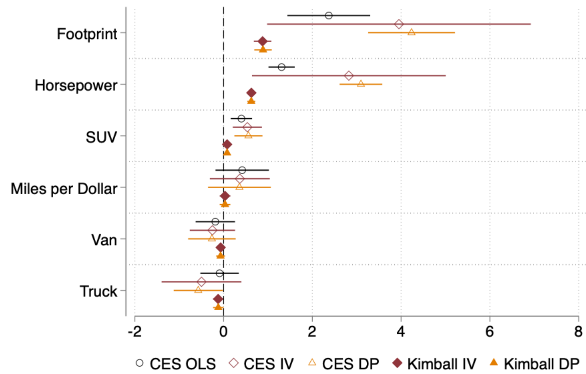
Note: The table reports the variance explained by the first two principal components from a singular value decomposition of product characteristics. The following characteristics are included: footprint, horsepower, miles-per-gallon, curbweight, height, style dummies, luxury dummy, electric dummy, US brand dummy. We standardize all variables before performing the singular value decomposition. The first column uses all sample, while all the remaining columns consider a product style (car, SUV, truck, and van) at the time.

Figure E.5: Cross-price Elasticities



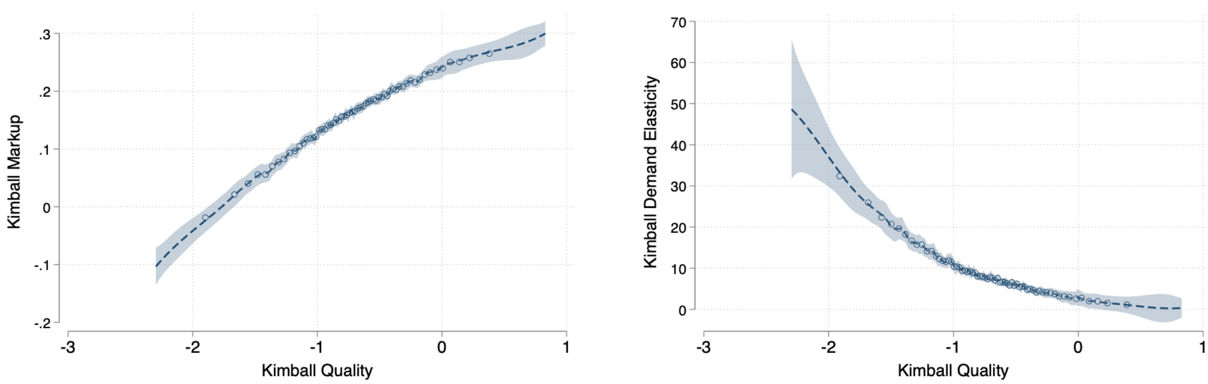
Note: The left panel shows the matrix of cross-price elasticities from the Kimball (left) and the BLP (right) specifications for the year 1981. Cross-price elasticities for the Kimball specification are constructed using Equation (11), while BLP cross-price elasticities are constructed according the correction in footnote **??**. We report the lower half as the matrix is symmetric by construction. The yellow area represents cross-price elasticities above the 90th percentile, the blue area represents cross-price elasticities above the median and below the 90th percentile, the red area represents cross-price elasticities below the median. Products are ordered on the horizontal axis in descending order by market share. The matrix is smoothed with a gaussian kernel and discretized. The right panel shows the histograms of the cross-price elasticities from the Kimball (red, right axis) and the BLP (blue, left axis) specifications for the year 1981.

Figure E.6: Correlation between Inferred Quality and Product Characteristics



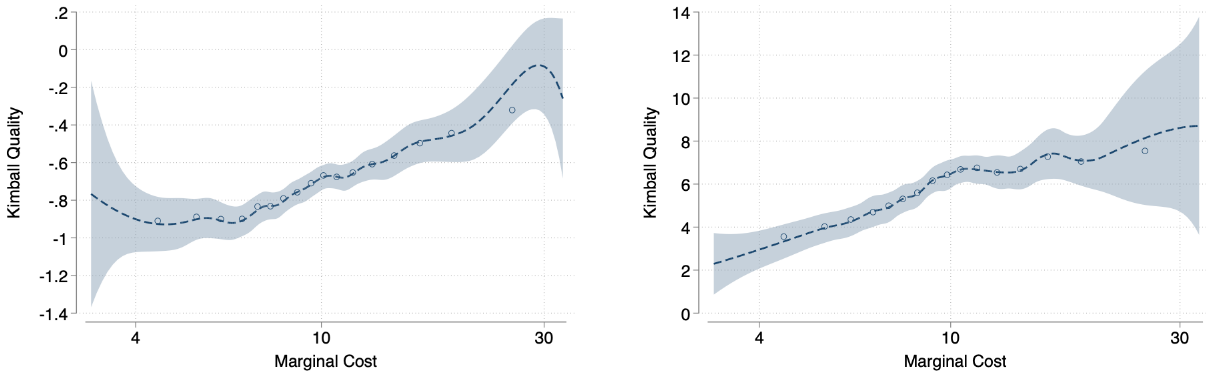
Note: The figure reports the relationship between product characteristics and inferred quality. In the CES DP case, the inferred quality measure follows from Equation (15). For the Kimball specification, inferred quality is obtained inverting demand as in Appendix *B*.2.2. The coefficients referring to the DP approach (CES and Kimball) and the Kimball IV case are obtained from regression in Equation (*E*.9). We consider the following product characteristics: horsepower, footprint, miles-per-dollar and style (suv, truck, van). Continuous variables are in log. The coefficients referring to the OLS and IV estimates of the CES specification are obtained from Equation (29), where product characteristics are used to proxy for quality. All regressions use the entire sample and includes time and producer fixed effects. Standard errors are clustered at the producer level, the bands around the estimates show the 95% confidence intervals.

Figure E.7: Quality, Own-price Elasticity and Markups



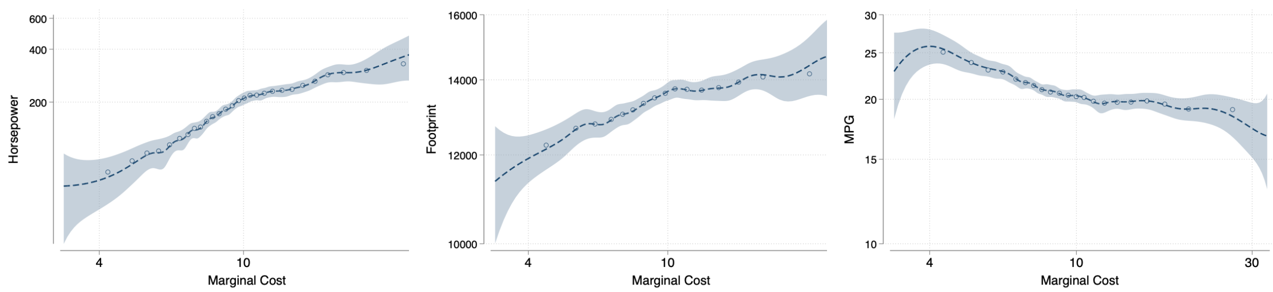
Note: The left panel shows the relationship between the measure of inferred quality and the markups. The right panel shows the relationship between the measure of inferred quality and the demand elasticity estimated from the Finite-Finite Kimball specification using the DP approach. Markups are computed under the assumption of monopolistic competition, $\mu_{it} = \frac{1}{\sigma_{it}-1}$, where σ_{it} is the estimated own-price elasticity. We approximate the relationship using a polynomial of degree 3 after absorbing model fixed effects. We report 95% confidence intervals. Standard errors are clustered at the model level.

Figure E.8: Marginal Cost and Quality



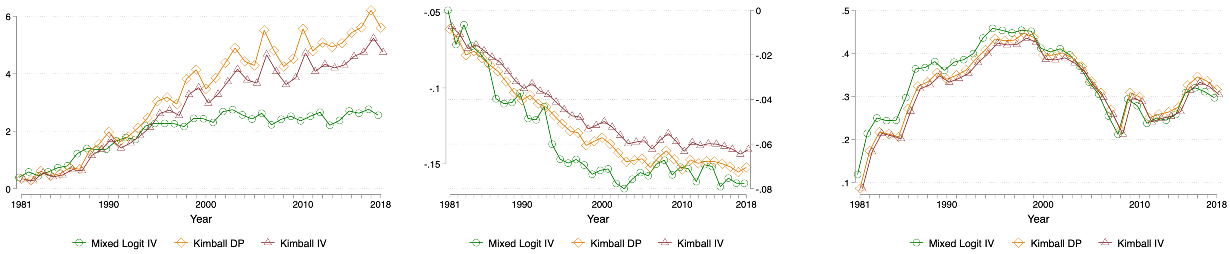
Note: The right panel shows the relationship between the implied marginal cost and a proxy of input costs, namely the price of steel multiplied to the weight of each vehicle. The left panel shows the relationship between the implied marginal cost and the measure of inferred quality estimated from the Finite-Finite Kimball specification using the DP approach. The marginal cost of each model is inferred as follow: $mc_{it} = \frac{p_{it}}{1+\mu_{it}}$, where $\mu_{it} = \frac{1}{\sigma_{it}-1}$ is the markup computed under the assumption of monopolistic competition and σ_{it} is the estimated own-price elasticity. We approximate the relationship using a polynomial of degree 3 after absorbing model fixed effects. We report 95% confidence intervals. Standard errors are clustered at the model level.

Figure E.9: Marginal Cost and Product Characteristics



Note: Each panel shows the relationship between the inferred marginal cost and a product characteristic. We consider horsepower (left), footprint (center) and miles-per-gallon (right). Marginal cost is inferred from $mc_{it} = \frac{p_{it}}{1+\mu_{it}}$, where μ_{it} is the markup computed under the assumption of monopolistic competition using the own-price elasticities estimated from the Finite-Finite Kimball specification using the DP approach. All variables are in level.

Figure E.10: Trends in Markups and Marginal Cost



Note: The left panel shows the estimated trend in the own-price elasticity over the period 1980-2018. The center panel shows the estimated trend in the markup over the period 1980-2018. Markups are computed under the assumption of monopolistic competition, $\mu_{it} = \frac{1}{\sigma_{it}-1}$, where σ_{it} is the estimated own-price elasticity. The right panel shows the estimated trend in the real marginal cost. The real marginal cost is computed from $mc_{it} = \frac{p_{it}}{1+\mu_{it}}$ using prices in 1980 US dollars. Trends are obtained regressing the outcome variables at the product-year level on a time trend, controlling for product fixed effects. In the marginal cost case, we also control for product characteristics such as horsepower, miles-per-dollar, footprint, curbweight, years since design, style, electric and luxury dummies. The BLP and Finite-Finite Kimball specifications (both IV and DP) are considered.