Discussion Paper Series

IZA DP No. 18423 February 2026

Inferring Prices from Quantities

David Argente
Yale University

Chang-Tai Hsieh
University of Chicago

Munseob Lee
University of California, San Diego and IZA@LISER

IZA@LISER NETWORK
ISSN: 2365-9793

Inferring Prices from Quantities*

Abstract

Measuring aggregate inflation is subject to two opposing biases: unobserved quality and variety growth, and the use of incorrect weights when new varieties are misclassified. We show that it is possible to measure an aggregate price index free of these biases when we have a subset of products where these two errors average to zero. This procedure does not require us to distinguish new from existing goods, measure quality attributes directly, or classify new varieties into the appropriate category. We implement this approach using BEA data from 1959 to 2019, approximating the official PCE price index with a CES aggregate of BEA prices at the product level. Our estimate of the inflation rate exceeds the CES aggregate of BEA prices by 0.3 to 1.0 percentage points per year on average. The aggregate bias was close to zero prior to the BLS introducing hedonic adjustments, which suggests that only adjusting for quality bias can lead to an underestimation of overall inflation, particularly in quality-adjusted categories.

JEL classification

D11, D12, E01, E31

Keywords

price index, inflation

Corresponding author

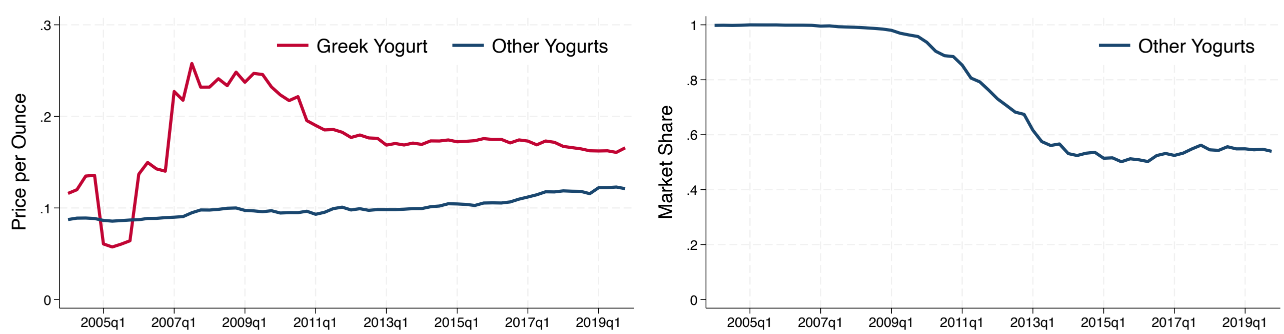
Munseob Lee
munseoblee@ucsd.edu

* We thank Alberto Cavallo, Marc-Andreas Muendler, Matthew Shapiro, Dan Sichel, and Dominic Smith for helpful comments. Researchers' own analyses are calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researchers and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

The central challenge in measuring price indices is that the mix and quality of products change over time. A large empirical literature addresses this problem by collecting more and more granular data on product quality and new products, in what Boskin et al. (1996) call the equivalent of "house-to-house combat." The challenge is not only how to incorporate new products and changing quality into prices of individual products, but also how to classify new products. Should the iPad be classified a new television, a new computer, a new smartphone, or simply as its own product? Although there is no obvious answer these classifications are made all the time.

In this paper we take a different approach. Rather than assembling detailed product-level data, we use demand theory to infer the aggregate price index directly from expenditure shares. Consider quality improvement in Greek yogurt. Over the decade between 2007 and 2017, the quality of Greek yogurt improved, including higher protein content, lower sugar, changes in texture, and the inclusion of probiotics.[1] Although there is no sign of the quality improvement in the price of yogurt (left panel in Figure 1), the effect on expenditure shares is dramatic: the market share of regular yogurt fell by almost 50 percentage points (right panel in Figure 1).

Figure 1: Expenditure Shares and Prices: Yogurt



(a) Unit Price (b) Expenditure Share

Notes: Panel (a) plots the unit-value price per ounce for Greek yogurt and for regular yogurts separately. Panel (b) plots the expenditure share of regular yogurt within the yogurt category over time. We use the NielsenIQ Consumer Panel data for this calculation.

We can formally use this logic to infer the quality-adjusted price of Greek yogurt with a CES utility framework:

$$d \ln \text{share} = -(\sigma - 1) d \ln(p/P)$$

[1] See *"The Greek Yogurt Market,"* DASO (2017), https://www.daso.gr/2017/10/the-greek-yogurt-market/.

where share and p are the market share and price of regular yogurt, P is the weighted average of the prices of regular and Greek yogurt, and σ is the elasticity of substitution. Under one key assumption, that we know the price of regular yogurt p, we can infer the weighted average of the prices of the two types of yogurts P using only data on expenditure shares. Importantly, we do not need direct information on the quality of Greek yogurt. Intuitively, any quality change in Greek yogurt appears in the expenditure share of regular yogurt, after controlling for the price of regular yogurt.

In short, assuming there is a set of products for which we can measure prices accurately, we can infer a weighted average of prices for the other products solely from data on expenditure shares. The empirical challenge is to identify a set of products for which measured prices are reasonably immune to the measurement issues that plague many products. In this paper, we show that measurement error in prices and misclassification of new products manifests empirically as non-zero residual demand growth. Building on this insight, we identify a subset of product categories where residual demand growth averages zero, which we refer to as the *chosen* products.

Once we have identified this bundle of chosen products, the change in the aggregate price index is given by the product of two terms: (1) the weighted average change in prices of the products in the chosen bundle, and (2) the change in the bundle's market share, adjusted by the price elasticity of demand. This formula is essentially Feenstra (1994), except that the "new goods" term in our formula is not solely the welfare gain from new varieties; it also captures the aggregate effect of changes in prices and quality among products excluded from the chosen bundle. However, the product of these two terms yields an unbiased estimate of the combined effects of changes in prices and quality among incumbent products and the welfare gains from (properly classified) new varieties, even when these two effects cannot be identified separately. We refer to this aggregate price index constructed from chosen products as the CES Chosen Price Index (CCPI).

We implement our methodology using publicly available price and expenditure data from the Bureau of Economic Analysis (BEA) for roughly 200 product categories in the PCE Price Index. We approximate the official PCE Price Index with a CES aggregate of prices of 200 product categories in the BEA, which we call the CES-BEA price index.[2] We begin by estimating, for each product category, the correlation between expenditure shares and prices over time. Since the residual demand is likely to be close to zero for product categories that exhibit a negative and statistically significant relationship between expenditures and prices, we define these products as our "chosen" product categories. We then calculate the aggregate price index as the product of the price index of the "chosen" product categories and the ex-

[2] This approximation closely tracks the official PCE price index (see Figure A.1).

penditure share of these categories, adjusted by the demand elasticity. Our estimates suggest that the CES-BEA price index understates the annual inflation rate by 0.3-1.0 percentage points on average between 1959 and 2019.

There are two papers that serve as the foundation for our work. First, Bils (2009) shows that changes in a product's expenditure share can be used to infer its true price. We use the same insight to infer the combined effects of new variety growth and quality change for all products in the economy. Second, Redding and Weinstein (2020) demonstrate that quality is unlikely to remain constant for a given barcode. They introduce an alternative price index by assuming that the geometric mean of quality change among incumbent barcodes is zero. While our evidence suggests that this assumption may not hold for all incumbent barcodes, our point is that we can measure the aggregate price index once we identify a set of products for which the assumption of zero residual demand growth is more likely to hold.

Our work is also related to the large body of work that quantifies the mismeasurement from missing quality and new varieties. Papers that quantify the bias from missing quality growth include Nordhaus (1997), Bils and Klenow (2001), Bils (2009), Argente, Hsieh and Lee (2023), and Ehrlich, Haltiwanger, Jarmin, Johnson, Olivares, Pardue, Shapiro and Zhao (2024), and Broda and Weinstein (2006, 2010), Aghion, Bergeaud, Boppart, Klenow and Li (2019), and Argente and Lee (2021) estimate the bias in price indices from missing new varieties. We don't attempt to separately identify variety growth from quality change. Instead, we measure their net effect by identifying a set of products for which quality change and variety growth are less likely to be an issue.

The paper is organized as follows. Section 1 develops a conceptual framework to compare the CES-BEA index with the CES Chosen Price Index (CCPI) and shows how various biases show up as nonzero residual demand at the product level. Section 2 describes our procedure for identifying the subset of products whose average residual demand is zero. In Section 3, we implement this methodology using both BEA and NielsenIQ Consumer Panel data and present our main findings on the bias in the CES-BEA index. Section 4 quantifies the net errors that arise when individual biases–such as new-variety omissions or incomplete quality adjustments–are corrected in isolation. The final section concludes.

1 Conceptual Framework

This section develops a model in which the measurement problem for the aggregate price index stems from missing quality and new varieties in prices and misclassification of new varieties in expenditures. We demonstrate that the telltale sign of these two errors is a nonzero residual demand growth at the product-category level. We then illustrate how this insight allows us to identify both a set of product categories in the BEA data and a set of barcodes in the NielsenIQ data for which these errors average to zero, and to construct the aggregate price index using changes in expenditures and prices for those items.

1.1 Aggregate Price Index

Utility is a CES aggregate of consumption of BEA product categories (hereafter "products") indexed by j:

$$\mathbb{U} = \left(\sum_j C_j^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$$

where C_j is itself a CES aggregate of varieties within the product indexed by k:

$$C_j = \left(\sum_{k \in j} \left(Q_{jk} \, C_{jk} \right)^{\frac{\eta_j - 1}{\eta_j}} \right)^{\frac{\eta_j}{\eta_j - 1}} .$$

Here, Q_{jk} denotes quality and C_{jk} the units of individual varieties in the category of products j.[3] New products may appear as new product categories (indexed by j) or new varieties within an existing category (indexed by k).

The change in the aggregate price index P is then:

$$d \ln P = \frac{1}{\sigma - 1} \, d \ln S_{\mathbb{I}} + \sum_{j \in \mathbb{I}} \omega_j \, d \ln P_j \tag{1}$$

where $S_{\mathbb{I}}$ denotes the share of incumbent products as a share of total expenditures and P_j is the price of an incumbent product. The first term in Equation 1 captures the net effect of entry and exit of products, measured as the change in the market share of the incumbent products multiplied by $\frac{1}{\sigma-1}$. The second term in Equation 1 captures the effect of the change in price of incumbent products, where the weight of each product ω_j is defined as:

$$\omega_j \equiv \frac{\frac{dS_j}{d \ln S_j}}{\sum_{l \in \mathbb{I}} \frac{dS_l}{d \ln S_l}}$$

Here, S_j is the revenue of j as a share of all incumbent products.[4]

[3] We assume stable preferences over time so there is a well defined price index over time. Baqaee and Burstein (2023) compute the price index holding preferences fixed at a given point in time.

[4] Note that given the CES utility function the change in the log aggregate price in equation 1 is exact. That is, it holds for changes in prices of any magnitude. The key is that we use Sato-Vartia expenditure weights w_j to aggregate the price changes of individual products. See appendix B for the derivation of this result.

Finally, the price of product j in Equation 1 is defined as:

$$d \ln P_j \equiv \frac{1}{\eta_j - 1} d \ln X_{\mathbb{I}_j} + d \ln \left(\frac{P_{\mathbb{I}_j}}{Q_{\mathbb{I}_j}} \right) \tag{2}$$

where $X_{\mathbb{I}_j}$ are revenues of incumbent varieties in product j as a share of total revenues of j and $P_{\mathbb{I}_j}/Q_{\mathbb{I}_j}$ is the quality-adjusted price of the representative *incumbent* variety in j.[5] The first term denotes the effect of net entry of varieties in product j; the second term measures the effect of changes in the quality adjusted price of incumbent varieties in j.

1.2 CES-BEA Price Index

The BEA provides data on prices and expenditures for roughly 200 product categories. Its price index makes no adjustment for new varieties but does attempt to adjust for quality changes. Specifically, suppose the BEA price of a product is given by:

$$d \ln P_j^{BEA} \equiv d \ln P_{\mathbb{I}_j} - (1 - \lambda_j) \, d \ln Q_{\mathbb{I}_j} \tag{3}$$

where $d \ln P_{\mathbb{I}_j}$ is the change in the unit price of incumbent products and $\lambda_j \in [0, 1]$ parameterizes the extent to which the BEA accounts for quality change. When $\lambda_j = 0$ there is no bias from missing quality. The other extreme is $\lambda_j = 1$ when the PCE index misses *all* quality change.

We define the CES-BEA price index as a CES aggregate of prices of each product category as measured by the BEA:

$$d \ln P^{BEA} \equiv \frac{1}{\sigma - 1} d \ln S_{\mathbb{I}} + \sum_{j \in \mathbb{I}} \omega_j^{BEA} \, d \ln P_j^{BEA} \tag{4}$$

where the weight ω_j^{BEA} is measured from the expenditure shares in the BEA S_j^{BEA}:[6]

$$\omega_j^{BEA} \equiv \frac{dS_j^{BEA}}{d \ln S_j^{BEA}} \Bigg/ \sum_l \frac{dS_l^{BEA}}{d \ln S_l^{BEA}}$$

In Appendix A, we compare the CES-BEA price index calculated from equation 4 with the official Personal Consumption Expenditures (PCE) Price Index published by the BEA.

[5] $d \ln \left(\frac{P_{\mathbb{I}_j}}{Q_{\mathbb{I}_j}} \right) \equiv \sum_{k \in \mathbb{I}_j} \omega_{jk} \, d \ln \left(\frac{P_{jk}}{Q_{jk}} \right)$ where ω_{jk} is the Sato-Vartia share of variety k among all *incumbent* varieties in product j.

[6] The official PCE price index reflects the entry of new product lines through changes in expenditure weights over time. Accordingly, we include the expenditure share of incumbent products, $S_{\mathbb{I}}$, to mimic this feature of the PCE (Bureau of Economic Analysis, 2023).

Although the PCE uses a chain-weighted Fisher formula while our approach employs a CES aggregator with an explicit variety correction, the two indexes track each other closely over the entire 1959–2019 sample period.[7]

The difference between the CES-BEA price index in Equation 4 and the price index in Equation 1 is then:

$$\begin{aligned} d\ln\left(\frac{P^{BEA}}{P}\right) = &-\sum_{j\in\mathbb{I}}\omega_j^{BEA}\left(\frac{1}{\eta_j-1}\,d\ln X_{\mathbb{I}_j}\right) \\ &+\sum_{j\in\mathbb{I}}\omega_j\,\lambda_j\;d\ln Q_{\mathbb{I}_j} \\ &+\sum_{j\in\mathbb{I}}\left(\omega_j^{BEA}-\omega_j\right)\,d\ln P_j^{BEA} \end{aligned} \quad (5)$$

Equation 5 highlights three sources of bias from using Equation 4 to measure inflation.

The first line captures an upward-bias component arising from missing new products and varieties. As new varieties arrive, incumbents' expenditure shares fall ($d\ln X_{\mathbb{I}_j} < 0$), with each category's decline scaled by $1/(\eta_j - 1)$ and weighted by ω_j^{BEA}; summing across j yields the aggregate variety-bias correction.

The second line of Equation 5 isolates the quality bias arising from incomplete hedonic adjustments in the BEA's price index. Here, $Q_{\mathbb{I}_j}$ measures the quality of incumbent varieties in category j, and λ_j is the share of quality change in that category reflected in BEA prices. If $\lambda_j = 0$, BEA prices fully adjust for quality and this bias vanishes; $\lambda_j > 0$ implies unaccounted quality growth and thus a strictly positive upward adjustment. Together with the new-product bias, the quality bias means that a straightforward weighted average of BEA product-level prices will overstate the true rate of inflation.

The last line in Equation 5 can be written as

$$\sum_{j\in\mathbb{I}}\bigl(\omega_j^{\mathrm{BEA}}-\omega_j\bigr)\,d\ln P_j^{\mathrm{BEA}} = \mathrm{Cov}\bigl(\omega_j^{\mathrm{BEA}}-\omega_j,\ d\ln P_j^{\mathrm{BEA}}\bigr),$$

where the gap between the two weights $\omega_j^{\mathrm{BEA}} - \omega_j$ arises from both the omission and misclassification of new varieties and depends on $d\ln S_j^{\mathrm{BEA}} - d\ln S_j$. In particular, the BEA's

[7] The weights in our CES aggregator are a Sato-Vartia average of the expenditure shares. The BEA uses a Fisher index, which is an average of the Laspeyres and Paasche indices. Fisher weights are not exactly the same as Sato-Vartia weights, but they are similar in that they are both averages of weights at the beginning and end of the period.

expenditure share for category j is

$$S_j^{\text{BEA}} = \frac{\sum_{k\in\mathbb{I}_j} \text{rev}_{jk} + \sum_{k\in\mathbb{N}_j} \gamma_j \, \text{rev}_{jk} + \delta_j \sum_{i\in\mathbb{I}} \sum_{k\in\mathbb{N}_i} (1-\gamma_i)\, \text{rev}_{ik}}{\sum\limits_{i\in\mathbb{I}} \Big(\sum\limits_{k\in\mathbb{I}_i} \text{rev}_{ik} + \sum\limits_{k\in\mathbb{N}_i} \text{rev}_{ik} \Big)},$$

with $\gamma_j \in [0,1]$ denoting the share of new varieties in j correctly classified to j and $\delta_j \in [0,1]$ denoting the share of all misclassified new-variety revenues assigned to j. It can be shown that

$$S_j^{\text{BEA}} - S_j = S_j \big[(\gamma_j + (1-\gamma_j) X_{\mathbb{I}_j}) - 1 \big] \; + \; \delta_j \sum_{i\in\mathbb{I}} (1-\gamma_i)\, S_i \, (1 - X_{\mathbb{I}_i}), \tag{6}$$

where the first term is negative when uncaptured new varieties reduce $X_{\mathbb{I}_j}$, and the second term is positive when misclassified varieties arrive from other categories. Thus, the difference $d \ln S_j^{\text{BEA}} - d \ln S_j$ decreases with the arrival of new varieties and increases with higher misclassification. Because this difference multiplies $d \ln P_j^{\text{BEA}}$ inside a covariance, the expenditure-share bias can be positive or negative, depending on the joint behavior of weight errors and price changes.

There are two implications of this term. First, if

$$\text{Cov}\big(d \ln S_j^{\text{BEA}} - d \ln S_j, \; d \ln P_j^{\text{BEA}} \big) < 0,$$

so that BEA weights overemphasize categories experiencing price declines, then an inflation rate calculated from a weighted average of BEA prices *understates* the true inflation rate.

Second, correcting for new varieties and quality in prices at the product level may not necessarily reduce the overall bias in the aggregate price index. For example, suppose that we correct for quality growth. This has two effects. First, the second term in Equation 5 goes to zero. Second, this also lowers $d \ln P_j^{\text{BEA}}$ for products with high quality growth. But if these are products where there are many unmeasured new varieties in the BEA's expenditure data, this *increases* the aggregate bias from using the wrong weights. The net effect of adjusting for quality growth thus depends on the relative magnitudes of these two biases.

The same logic holds when we adjust the price index for new varieties in the product category. Suppose we do this perfectly. This adjustment has two effects. First, the direct effect of missing new varieties in the price index disappears—the second term in the first line of Equation 5 becomes zero. Second, $d \ln P_j^{\text{BEA}}$ is lower for products with significant new variety growth. However, if we adjust the price indices for new varieties without updating the weights used to aggregate these prices, this second effect increases the aggregate bias. The net effect hinges on the relative magnitudes of these two biases from new varieties: the bias in the price index at the product level and the bias from using incorrect weights to aggregate these prices.

1.3 "Residual" Product Demand

In this subsection, we show how the biases that affect the aggregation of incumbent product prices manifest in the "residual" demand. We then demonstrate how to measure the aggregate price index—including the contribution of new products—by identifying a set of products in the BEA data for which these biases average to zero.

Specifically, the BEA data also provide nominal expenditures for each product. The change in the share of expenditures on j in total expenditures (as measured by the BEA), after adjusting for the product of $\sigma - 1$ and the change in the BEA price index (relative to the aggregate price), is given by:[8]

$$d\ln S_j^{\text{BEA}} + (\sigma-1)\, d\ln\Big(\tfrac{P_j^{\text{BEA}}}{P}\Big) = \lambda_j\,(\sigma-1)\, d\ln Q_{\mathbb{I}_j} \; - \; \frac{\sigma-1}{\eta_j-1}\, d\ln X_{\mathbb{I}_j} + \; d\ln S_j^{\text{BEA}} - d\ln S_j. \tag{7}$$

Equation 7 points to two reasons why the residual demand can be non-zero. The first is the standard quality and new-variety biases in the price index, captured by the first term Equation 7, which are non-negative. Thus, a telltale sign that the BEA is missing quality and variety growth in the price data is a *positive* residual demand.

The second term, $d\ln S_j^{\text{BEA}} - d\ln S_j$, is given by equation 6 and captures mismeasurement in the BEA's expenditure shares due to two opposing forces. When the BEA omits new varieties in category j ($\gamma_j < 1$), its reported share falls more sharply than the true share yielding a negative residual demand. Thus, the hallmark of misweighting is a declining residual demand. In contrast, when the BEA misclassifies entrants from other categories into j ($\delta_j > 0$), its share falls less—or even rises—relative to the true share yielding a positive residual demand. In the extreme case of perfect capture ($\gamma_j = 1$) with no misclassification ($\delta_j = 0$), these two effects cancel and the residual is zero.

1.4 Products with zero residual demand

Suppose we identify a set of products $\mathbb{C}$ where the residual demand averages zero:

$$\sum_{j\in\mathbb{C}} \alpha_j \left[\lambda_j\,(\sigma-1)\, d\ln Q_{\mathbb{I}_j} \; - \; \frac{\sigma-1}{\eta_j-1}\, d\ln X_{\mathbb{I}_j} + \; d\ln S_j^{\text{BEA}} - d\ln S_j\right] \approx 0$$

[8] Details on the derivation can be found in Appendix C.

where α_j are weights that add up to one among the products in $\mathbb{C}$. The weighted average of the residual demand in Equation 7 for the products in $\mathbb{C}$ can then be expressed as:

$$d \ln P = \sum_{j \in \mathbb{C}} \alpha_j \Big(\tfrac{1}{\sigma - 1} \, d \ln S_{\mathbb{I}_j} + d \ln P_j^{\text{BEA}} \Big). \tag{8}$$

We use a Sato-Vartia weight for α_j and refer to this price index as the CES Chosen Price Index (CCPI). Equation 8 shows that the change in the aggregate price index can be calculated using only the expenditure shares and BEA price indices of products in $\mathbb{C}$. We do not require data on new product categories, nor on incumbent products with mismeasured quality or varieties, nor on any classification of products into groups. This is because the effects of all excluded products are captured by the first term in Equation 8. Therefore, the key is to identify products whose residual demand likely averages zero.

2 Identifying Products with Zero Residual Demand

To identify the set of products where the residual demand is likely to average to zero, we rewrite Equation 7 as follows:

$$d \ln S_j^{BEA} = -\,(\sigma - 1)\; d\,\ln \left(\frac{P_j^{BEA}}{P} \right) + \text{Residual Demand}_j$$

Then, we can implement this equation product-by-product using time-series variation as follows:

$$d \ln S_j^{BEA} = \beta_j \; d\,\ln \left(\frac{P_j^{BEA}}{P} \right) + \epsilon_j. \tag{9}$$

The estimate of β_j from an OLS regression is given by:

$$\beta_j = -(\sigma - 1) \, + \, \frac{\text{Cov} \left[d\,\ln \left(\frac{P_j^{BEA}}{P} \right), \, \text{Residual Demand}_j \right]}{\text{Var} \left[d\,\ln \left(\frac{P_j^{BEA}}{P} \right) \right]} \tag{10}$$

When the residual demand is zero, the coefficient β_j is *unambiguously* negative given our assumption that $\sigma > 1$. However, when the residual demand is not zero, then we get an omitted variable bias, where the direction of the bias depends on whether the residual demand is positive or negative.

On the one hand, if the residual demand is positive, either because official prices do not completely capture quality change or more varieties, then the covariance term in equation 10 is positive and β_j is biased towards zero or could even turn positive. On the other hand, if the residual demand is negative because official data on *expenditure shares* of a product does not capture all the new varieties of the product. Furthermore, if this is more prevalent in products where the official inflation rate of the product category is low, then the covariance term in equation 10 is negative and $\beta_j < -(\sigma - 1)$. Based on this logic, we will identify products where the residual demand is likely to average to zero as those for which $\beta_j \approx -(\sigma - 1)$.

2.1 Iterative Procedure

Our procedure begins by implementing Equation 9 product-by-product in the data. Estimating β_j requires knowing each product's relative price P_j^{BEA}/P and thus the aggregate index P. Because P is unknown a priori, we adopt an iterative procedure:

1. Take the CES-BEA aggregate price index, given σ, as an initial guess for the aggregate price index.

2. Estimate Equation 9 for each product category in the BEA data and select those categories for which β_j is close to $-(\sigma - 1)$ and statistically different from zero at the 1% level. In the benchmark case, we allow $0.5 \times \sigma$ deviation in β_j's.

3. Calculate a new aggregate price index using only the set of *chosen* products via Equation 8.

4. Compare the newly constructed index to the previous guess. If the average absolute log-difference exceeds 0.01, replace the guess with the new index and repeat steps 2–4; otherwise, stop.

The outcome of this iterative procedure is both a subset of products whose residual demand is likely to average zero and the CES Chosen Price Index (CCPI) calculated using these *chosen* categories.[9]

2.2 Monte Carlo Validation

In this subsection, we validate our iterative selection procedure through a Monte Carlo exercise that incorporates both sources of measurement error identified in Section 1. We simulate an economy with 100 products observed over 60 time periods. Half of the products are "well-measured," meaning their observed prices and expenditure shares accurately reflect true values. The other half are "mis-measured," exhibiting both types of errors described above.

[9] In our application with the NielsenIQ data, we implement this procedure at the barcode level within each product category.

For each product j in period t, we generate a common aggregate inflation component drawn from $N(0.02, 0.01)$ and idiosyncratic price shocks drawn from $N(0, 0.05)$. The true price change for product j is the sum of these two components. For mis-measured products, we introduce two types of errors corresponding to the biases identified in Section 1.

The first error source is price measurement error. The observed price change for mis-measured products deviates systematically from the true price change by δ^P, so that $d\ln p_{jt}^{\text{obs}} = d\ln p_{jt}^{\text{true}} + \delta^P$. This mimics incomplete quality adjustments or missing variety growth. We consider $\delta^P \in \{0.01, 0.02\}$, representing 1–2 percentage point annual bias in observed prices.[10]

The second error source is expenditure weight distortion. For mis-measured products, the observed expenditure share responds incorrectly to price changes according to:

$$d\ln s_{jt}^{\text{obs}} = -(\sigma - 1) \cdot \text{price\_shock}_{jt} + \epsilon_{jt} + \kappa \cdot \text{price\_shock}_{jt},$$

where $\epsilon_{jt} \sim N(0, 0.05)$ and κ captures the strength of weight distortion. This parameter corresponds to the expenditure-share bias term in Equation 5. When $\kappa < 0$, expenditure shares under-respond to price changes, mimicking the omission of new varieties. When $\kappa > 0$, expenditure shares over-respond, mimicking misclassification. We consider $\kappa \in \{-2.0, 2.0\}$.

This 2×2 design yields four scenarios. For each scenario, we run 25 Monte Carlo replications. In each replication, we construct the CES price index using all products and implement our iterative procedure from Section 2.1 to identify chosen products. We then calculate the CES Chosen Price Index (CCPI) using only chosen products and measure bias as the difference between each index and true aggregate inflation.

Table 1 presents the average bias across replications for each scenario. The CES index shows significant bias ranging from 0.5 to 1.5 percentage points per year, increasing with the magnitude of both error sources as predicted by Equation 5. In contrast, the CCPI reduces bias by 75–95% across all scenarios. On average, 80–90% of products selected as chosen are truly well-measured. The small contamination by mis-measured products explains the residual bias in the CCPI.

[10] Imposing $\delta^P \in \{-0.01, -0.02\}$ flips the direction of the bias. The bias reduction in this case is comparable to the reported case.

Table 1: Monte Carlo Simulation Results

P	W	CES	Bias (pp) CES Chosen Price Index (CCPI)
0.01	−2.0	0.99	0.20
0.01	2.0	0.52	0.07
0.02	−2.0	1.52	0.15
0.02	2.0	0.98	0.06

Notes: P denotes systematic price measurement error added to mis-measured products. W denotes expenditure weight distortion, where negative values represent omission of new varieties and positive values represent misclassification. CES is the conventional CES price index, and CCPI is the CES Chosen Price Index. Results based on 25 Monte Carlo replications per scenario.

Figure 2 displays the probability density functions of residual demand for chosen versus non-chosen products. Residual demand is calculated from Equation 7 as the component of expenditure share changes not explained by relative price movements. Chosen products exhibit distributions tightly centered at zero (solid blue lines), while non-chosen products show systematic deviations (dashed red lines) from zero.

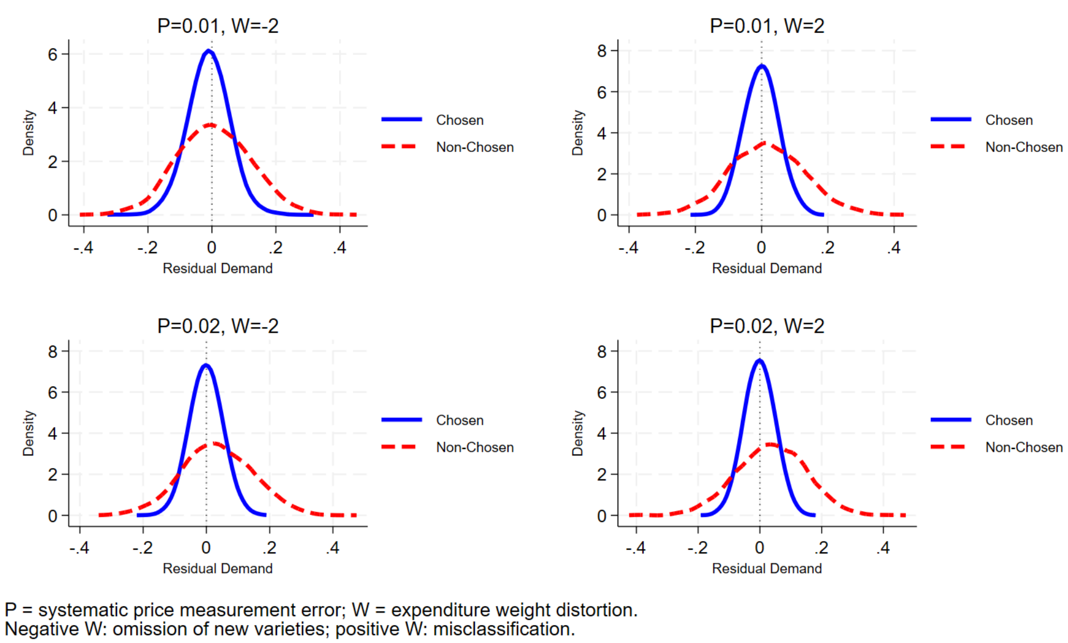


Figure 2: Residual Demand Distributions: Chosen vs. Non-Chosen Products

Notes: Kernel density estimates of residual demand for chosen products (solid blue) and non-chosen products (dashed red) across all four scenarios. Chosen products exhibit residual demand centered at zero, validating that the selection procedure identifies products with no systematic measurement bias.

3 Empirical Application

In this section, we apply our iterative procedure to the BEA price and expenditure data to construct the CES Chosen Price Index (CCPI) that corrects for the quality, variety, and weighting biases described above. To validate our approach, we repeat the exercise using detailed NielsenIQ Consumer Panel data, whose granular coverage enables us to perform several robustness checks. First, we present the BEA implementation; then we describe how we adapt the method to the NielsenIQ dataset.

3.1 BEA Data

We use the U.S. Bureau of Economic Analysis (BEA)'s official statistics on prices and expenditures for 212 product categories in Personal Consumption Expenditures. Specifically, we draw on two NIPA tables: "Price Indexes for Personal Consumption Expenditures by Type of Product" (Table 2.4.4U) and "Personal Consumption Expenditures by Type of Product" (Table 2.4.5U).[11] These categories account for about two-thirds of domestic final spending.

Our sample consists of monthly price and expenditure data for these categories from January 1959 to December 2019, and we compute 12-month changes for each category. Of the 212 products, 182 are covered continuously over this period; two products were combined into one in January 2002 ("Tenant-occupied stationary homes" and "Tenant landlord durables" became "Tenant-occupied, including landlord durables"); and 27 categories were introduced after 1959 and remain through December 2019.[12]

3.2 Residual Demand: BEA Data

We begin by estimating the elasticity of substitution, σ, following Broda and Weinstein (2006).[13] Our estimate yields $\sigma = 2.91$. Then, we implement our iterative procedure for each of the BEA's 212 product categories, using the full 61 years of data. When estimating Equation 9, we work with monthly data and take twelve-month differences to control for seasonality.

Figure 3 presents the distribution of the estimated β_j. The left panel covers all 212 BEA product categories: the mean estimate of β_j is positive (approximately 0.10, with a median of 0.11), and about 68% of products have a positive coefficient. The right panel restricts attention to coefficients significant at the 1% level, where positive estimates are even more

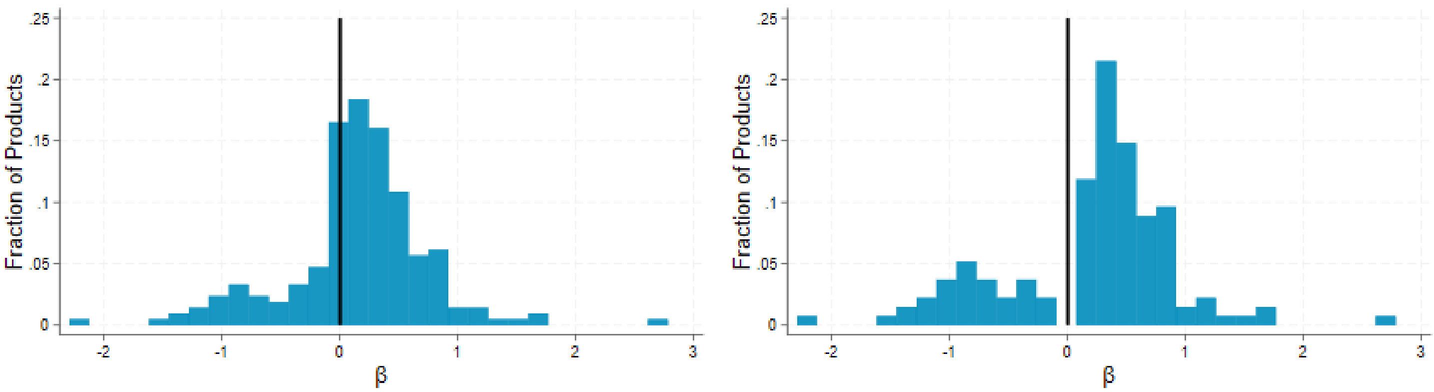
[11] Details on the BEA's methodology for preparing PCE estimates can be found in Chapter 5 of the NIPA Handbook.

[12] See Appendix D for further details.

[13] See Appendix E for details.

prevalent. Our "chosen" set $\mathbb{C}$ consists of the 11 categories with β_j values within the interval $[-(\sigma-1)-0.5\sigma,\ -(\sigma-1)+0.5\sigma]$, that is, from -3.36 to -0.45. Examples include tax preparation and other related services and small electric household appliances.[14]

Figure 3: Distribution of Coefficients β_j in BEA Data



(a) All Products (b) Significant at 1%

Notes: The figure shows the distribution of the coefficients β_j estimated using Equation 9 for BEA product categories (1959–2019). Panel (a) includes all 212 products. Panel (b) displays only those coefficients that are significantly different from zero at the 1 percent significance level. The vertical black line marks the zero coefficient.

Using the chosen product categories, we compute the CES Chosen Price Index (CCPI) following Equation 8. Our benchmark set retains β values within $\pm\ 0.5\sigma$ of the theoretical value $-(\sigma-1)$. As a robustness check, we also consider wider tolerances of $\pm\ \sigma$ and $\pm\ 1.5\sigma$. Table 2 reports the mean and standard deviation of the bias across these ranges. Under the benchmark window ($\pm\ 0.5\sigma$), the average bias is –1.08 percentage points with a standard deviation of 1.96. In other words, the CES Chosen Price Index (CCPI) inflation exceeds the CES-BEA inflation by 1.08 percentage points on average. Although the magnitude and direction of the bias fluctuate between January 1960 and December 2019, its overall sign is negative—that is, the CES-BEA aggregate inflation generally understates price increases. Under $\pm\ \sigma$ and $\pm\ 1.5\sigma$ ranges, the average biases remain sizable.

The evidence in Table 2 that the CES-BEA's inflation rate is too *low* may seem surprising. But remember that the bias can be either positive or negative, depending on the magnitude of missing quality and new varieties vs. the importance of misclassification of new varieties. Furthermore, the former shows up empirically as a positive residual demand whereas the latter shows up as negative residual demand. Figure 4 reports a histogram of residual demand.

[14] Appendix F lists all 11 chosen categories along with their estimated coefficients and standard errors.

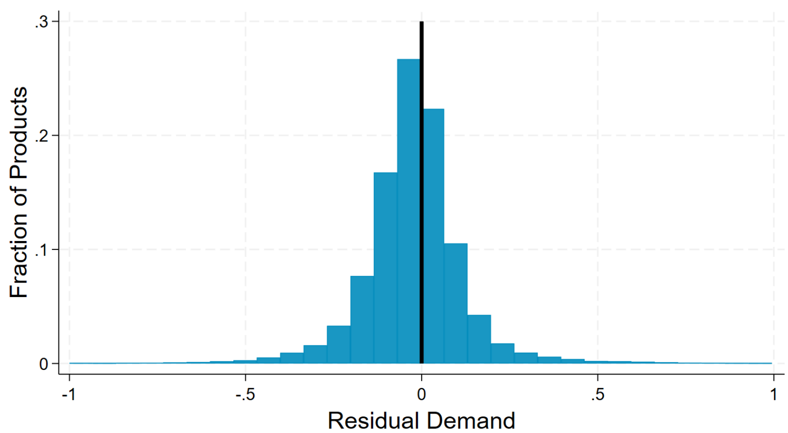
Table 2: **Bias under Various β Selection Windows**

	$\pm\, 0.5\sigma$ (benchmark)	$\pm\, \sigma$	$\pm\, 1.5\sigma$
Average Bias	-1.08	-0.34	-0.31
Standard Deviation of Bias	1.96	0.69	0.54

Notes: The table reports the mean and standard deviation of the bias, measured as the percentage-point difference between the CES-BEA price index (Equation 4) and the one-year change in the CES Chosen Price Index (CCPI) (Equation 8), from January 1960 to December 2019, under various β selection windows. We retain βs within $\pm\, 0.5\sigma$ of the theoretical value $-(\sigma-1)$ as our benchmark set of chosen products.

As can be seen, the residual demands are large, and the fraction of negative residual demand is larger than the fraction of positive residual demand.

Figure 4: Histogram of Residual Demand



Notes: The figure shows the histogram of residual demand. Residual demand is calculated from Equation 7 for each month. For graphical clarity, we trim residual demands below −1 or above 1.

Appendix G provides direct empirical evidence of expenditure weight discrepancies between aggregate and scanner-based data sources. We manually match 21 product categories observed in both the NielsenIQ Consumer Panel data and the Personal Consumption Expenditures (PCE) accounts of the U.S. Bureau of Economic Analysis (BEA) over 2004Q1–2019Q4. For categories such as milk and cereal, BEA-measured expenditures grow substantially faster than their NielsenIQ counterparts, indicating systematic divergence consistent with the omission or misclassification of new varieties in aggregate accounts. Across all matched categories and quarters, the distribution of quarterly differences in log expenditure growth rates exhibits substantial dispersion around zero. These patterns confirm that the expenditure weight distortions are empirically relevant features of the data.

3.3 Residual Demand: NielsenIQ Data

Our iterative procedure, implemented here for the BEA data, focuses on identifying products whose average residual demand is zero in order to construct an unbiased index. In principle, this property should translate into higher R^2 values and lower mean-squared errors in the β_j regressions for the chosen set. Moreover, any subset of these chosen goods–so long as its average residual demand remains zero–should also deliver an unbiased aggregate price index. However, with only 11 chosen categories in the BEA data, it is difficult to rigorously test these statistical properties.

To overcome this limitation, we apply our methodology to the much richer NielsenIQ Consumer Panel data at the barcode level.[15] We implement our iterative procedure at the barcode level within each product category using all 16 years of the NielsenIQ Consumer Panel data (2004–2019). Following Redding and Weinstein (2020), we aggregate each barcode's prices and expenditures to a quarterly frequency.[16] We then compute four-quarter differences by comparing each quarter's value to the same quarter in the previous year. To ensure sufficient observations, we focus on barcodes present in the data for at least two years, yielding a total sample of 686,282 barcodes; we estimate Equation 9 separately for each one.

As before, estimating Equation 9 requires an initial guess of its product category's price index. This step hinges on an estimate of the elasticity of substitution across barcodes, which we obtain following the method of Feenstra (1994) as extended by Broda and Weinstein (2006) and Broda and Weinstein (2010).[17] The average elasticity of substitution is 6.42, and its full distribution is reported in Table E.I. The resulting category-level price index serves as our initial guess for the aggregate price of each product category. We then implement our iterative procedure for each of the 100 product categories, using the same selection and convergence criteria as in the BEA implementation. As before, this yields a set of barcodes whose unit

[15] The NielsenIQ Consumer Panel tracks the shopping behavior of 40,000 to 60,000 U.S. households, each using in-home scanners to record their purchases. The data contain just under one million distinct 12-digit barcodes, with each barcode uniquely assigned to a specific good available in stores. By design, each barcode corresponds to a single variety: any change in a product's attributes (e.g., form, size, packaging, or formula) generates a new barcode. Each barcode is classified into one of 104 product categories. Barcodes in the NielsenIQ Consumer Panel are organized hierarchically: each UPC belongs to one of 1,070 product modules, which are grouped into 100 product groups, which in turn fall under 10 major departments. Throughout the paper, we refer to product groups as product categories.

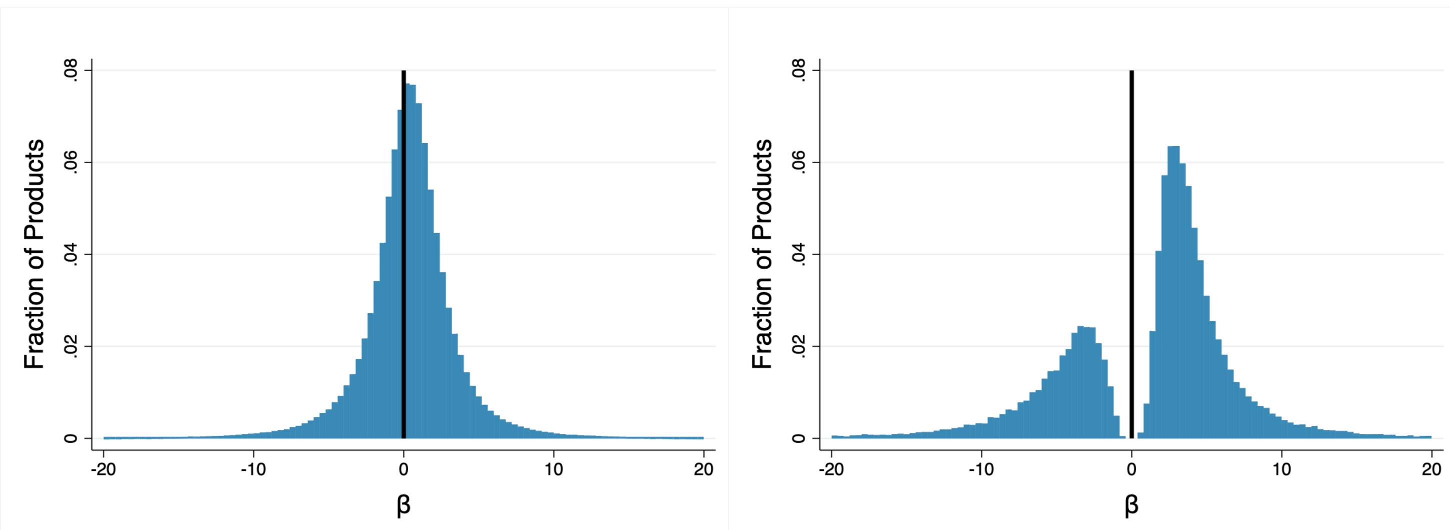
[16] NielsenIQ constructs projection weights to make the sample representative of the U.S. urban population. We use these weights to calculate total expenditures for each barcode.

[17] This procedure involves estimating demand and supply equations for each barcode using observed prices and quantities; see Appendix E for details.

prices are likely to fully capture changes in market share, and an unbiased product-category price index computed using these chosen barcodes, which we refer to as the CES Chosen Price Index (CCPI) for each product category.

Panel (a) of Figure 5 shows the distribution of coefficients β_i, where i indexes barcodes within product categories. For barcodes present in the data for at least eight quarters, the mean estimate of β_i is 0.37 (median 0.41). As in the BEA data, many barcodes—approximately 58%—have positive coefficients, indicating that their residual market-share changes are nonzero. Panel (b) of Figure 5 shows the distribution of coefficients significant at the 1% level. As before, we retain β's within $\pm$ 0.5σ of the theoretical value $-(\sigma-1)$ as our benchmark set of chosen products.

Figure 5: Distribution of Coefficients β_i in NielsenIQ Data



(a) All Barcodes

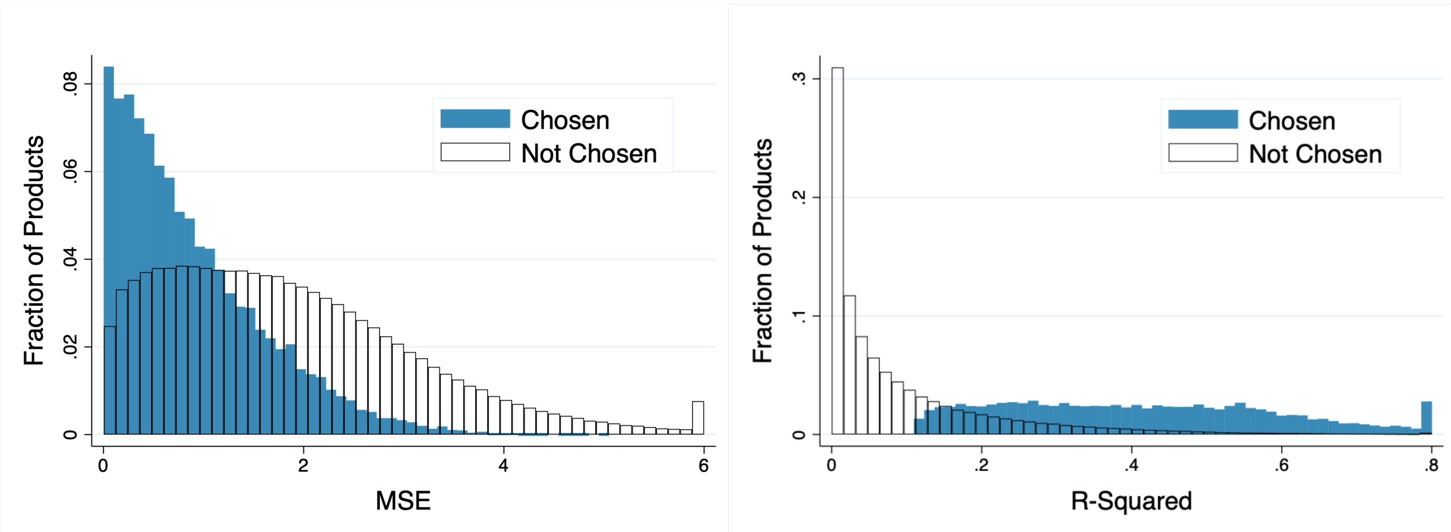
(b) Significant at 1%

Notes: The figure shows the distribution of the coefficients β_i estimated using equation 9 for barcodes in the NielsenIQ data (2004-2019). Panel (a) includes all barcodes (648,944 barcodes). Panel (b) includes barcodes whose coefficient is significantly different from zero at a 1 percent confidence level (18,649 barcodes with negative coefficient and 38,551 barcodes with positive coefficient). The black vertical line marks products whose coefficient is zero.

Building on our barcode-level selection, we compute the aggregate price index using only the chosen barcodes in each category. As with the BEA data, this "chosen-barcode" index consistently exceeds the conventional index that aggregates all incumbent varieties–on average by 0.75 percentage points per year.

The richness of the NielsenIQ data gives us a large "chosen" set—about 2.0% of all barcodes, or roughly 13,045 individual products—allowing us to rigorously describe their statistical properties. We begin by examining their performance in the β_i regressions. Panel (a) of

Figure 6: MSE and R-Squared - Product Level Regressions



Notes: Panel (a) shows the histogram of mean-squared errors estimated from Equation 9 for barcodes in the NielsenIQ data (2004–2019), while Panel (b) displays the corresponding R^2 distribution. Chosen products belong to the set $\mathbb{C}$ and comprise barcodes whose estimated coefficient is within $\pm$ 0.5σ of the theoretical value $-(\sigma-1)$ and significantly different from zero at the 1 percent level (13,045 barcodes). MSE is winsorized at 6 and R-squared is winsorized at 0.8 for visualization purposes.

Figure 6 compares the distribution of mean-squared errors (MSE) for chosen barcodes versus the remainder. Chosen barcodes are far more likely to exhibit low MSE values, indicating that their price movements explain market-share changes much more precisely than for other products. Panel (b) shows the corresponding R^2 distributions: chosen barcodes achieve substantially higher R^2 on average, confirming that our selection isolates items whose residual demand is nearly zero.

Table 3: Bias under Alternative Specifications

Subsample	Difference (pp)
$\beta < 0$ + significant at 1% + long duration (16 qtrs)	-0.97
$\beta < 0$ + significant at 1% + MSE < 1	-0.89
$\beta < 0$ + significant at 1% (**baseline**)	**-0.75**
$\beta < 0$ + significant at 5%	-0.67

Notes: The table reports the average bias, measured as the percentage-point difference between the conventional index using all barcodes (as in Broda and Weinstein (2010)) and the CES Chosen Price Index (CCPI) constructed from chosen barcodes in the NielsenIQ dataset. Subsamples are defined as follows: "long duration" requires barcodes to appear for at least 16 quarters; "MSE < 1" restrict to barcodes whose regression meet this threshold; "5%" uses a 5% significance cutoff for $\beta_i < 0$. We retain β's within $\pm\ 0.5\sigma$ of the theoretical value $-(\sigma-1)$ in all cases.

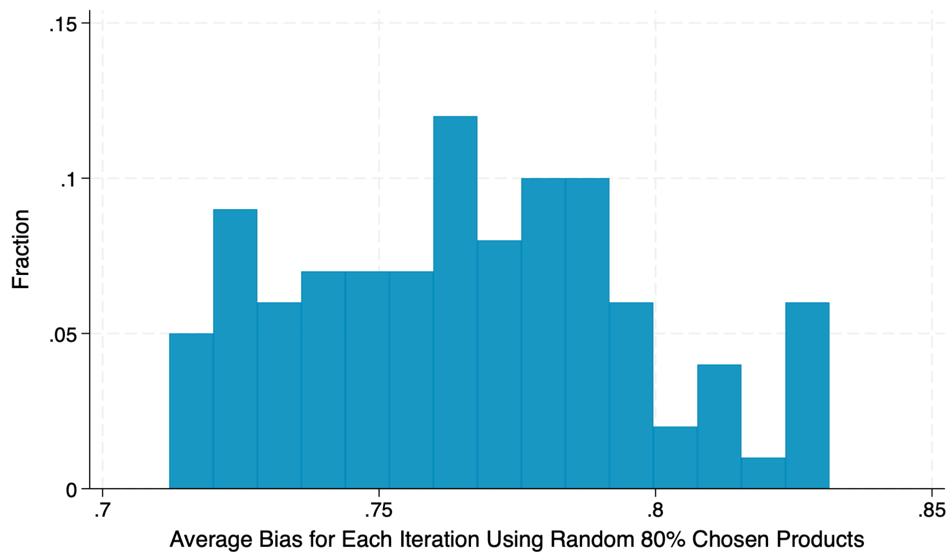
Lastly, the large size of our chosen-barcode set in the NielsenIQ data enables us to test the robustness of the aggregate price index across multiple subsamples. Theory predicts that any subset of chosen products with average residual demand near zero should reproduce the same aggregate index. To assess this, we reestimate the aggregate price index under four alternative selection rules, as shown in Table 3. These rules include: barcodes with $\beta < 0$ and significant at 1% that last at least 16 quarters; those with $\beta < 0$, significant at 1%, and mean-squared error below 1; our baseline ($\beta < 0$, significant at 1%); and barcodes significant at 5%. In each case, the sign of the bias and the average point-difference between the chosen-barcode index and the incumbent-only index remain close to our benchmark of 0.75 percentage points, showing that our aggregate price index is robust to the exact composition of the chosen set and providing additional validation to our procedure to select chosen products.

We conduct an additional robustness check by randomly selecting 80% of the baseline set of chosen products in each iteration. If all chosen products truly exhibit zero residual demand, the resulting price index should remain unchanged when using any subset of them. Figure 7 presents the average bias across 100 iterations using random 80% subsets. The mean bias is 0.76, very close to the bias obtained using the full set of chosen products. A t-test fails to reject the null hypothesis that the mean of these 100 observations equals 0.75 (p-value=0.3077). Moreover, the average bias across iterations remains tightly concentrated between 0.71 and 0.83, indicating a narrow range.

4 Partial Bias Corrections

In this section we examine the pitfalls of addressing individual biases in isolation. Recall that a positive residual demand indicates overlooked quality and variety growth, while a negative

Figure 7: Average Bias by Iteration Using 80% Random Subsamples



Notes: The figure shows the histogram of 100 iterations using random 80% chosen products. The bias with all chosen products is 0.75.

residual demand signals misweighting of varieties. Because these forces act in opposite directions, fixing only one can amplify the other: correcting for missing quality growth removes the upward gap in the CES-BEA price index, but it simultaneously increases the downward gap arising from misweighting.

To make this concrete, we focus on the CES-BEA price index, for which we can separately approximate both the new-variety bias and the quality bias. When we adjust the index perfectly for quality improvements and new-variety entry–yet leave the misweighting in place–the net bias in the CES-BEA index actually grows. By eliminating the bias that pushed the official index upward, we inadvertently magnify the remaining misweighting bias, pulling it downward and resulting in a larger net error.

4.1 Correcting Prices for New Varieties

We begin by isolating the impact of new varieties on the CES-BEA price index, abstracting from quality growth. To this end, we construct an alternative index, P^{NV}, which adjusts for the entry of new varieties within existing categories.

Specifically, define

$$d\ln P^{NV} \equiv \frac{1}{\sigma - 1}\, d\ln S_{\mathbb{I}} + \sum_{j\in\mathbb{I}} \omega_j^{\text{BEA}} \Big(d\ln P_j^{\text{BEA}} + \frac{1}{\eta_j - 1}\, d\ln X_{\mathbb{I}_j} \Big).$$

In this case, the variety-bias term drop out, leaving misweighted expenditure shares as the only remaining source of bias:

$$d\ln\left(\frac{P^{NV}}{P}\right) = \sum_{j\in\mathbb{I}} (\omega_j^{\text{BEA}} - \omega_j)\, d\ln P_j^{\text{BEA}}. \tag{11}$$

Thus, by approximating $d\ln X_{\mathbb{I}_j}$ in the data, we can quantify how a partial correction for new-variety entry affects the overall bias in the BEA index.

Although the BEA does not directly measure the contributions of new varieties within each product category, we can approximate these biases using establishment-level employment data following Aghion, Bergeaud, Boppart, Klenow and Li (2019). We employ the Business Dynamics Statistics (BDS), which reports annual changes in the employment share of incumbent establishments by 4-digit NAICS code.[18] Under the assumption that introducing a new product requires opening a new establishment, continuing plants serve as a proxy for incumbent products.

In particular, we construct the new-variety bias as follows. For each product category j, the fraction of employment retained by incumbents between $t-1$ and t is

$$\frac{X_{\mathbb{I}_{j,t}}}{X_{\mathbb{I}_{j,t-1}}} = \frac{\frac{\text{Employment}_{j,t} - \text{Job Births}_{j,t}}{\text{Employment}_{j,t}}}{\frac{\text{Employment}_{j,t-1} - \text{Job Deaths}_{j,t}}{\text{Employment}_{j,t-1}}},$$

so that $-\frac{1}{\eta_j - 1}\, d\ln X_{\mathbb{I}_j}$ captures the effect of new-variety entry on the price index. We obtain each η_j from the import-demand estimates of Broda and Weinstein (2006), aggregating HS-10 import data from the Census Bureau to the product-category level using the HS–to–NAICS concordance of Pierce and Schott (2012) and weighting observations by import shares.

Table 4 reports the impact of the variety correction on the BEA inflation rate. As expected, adjusting for new varieties within categories lowers the BEA's measured inflation, making the gap between the variety-adjusted CES-BEA index and the CES Chosen Price Index (CCPI) even more negative. The final line shows the aggregate bias after applying the full Feenstra correction (our P^{NV} in Equation 11): the variety-adjusted inflation rate is now 1.11 percentage points per year below the CES Chosen Price Index (CCPI).

[18] We map PCE categories to NAICS codes using the BEA's concordance, relying on 2007 definitions and adjusting for secondary-output reclassifications.

Table 4: Bias after Variety Correction

	$\pm$ 0.5σ (benchmark)	$\pm$ σ	$\pm$ 1.5σ
Original bias, $d\ln\left(\frac{P^{BEA}}{P}\right)$	-1.08	-0.34	-0.38
Bias after variety correction *within* BEA product, $d\ln\left(\frac{P^{NV}}{P}\right)$	-1.11	-0.36	-0.40

Notes: The table reports the average bias under two measures—original bias (Equation 5) and bias after correcting for new varieties within each category (Equation 11)—over January 1960 to December 2019 for various values of β selection windows. The original average biases match those in Table 2.

4.2 Correcting Prices for Quality Growth

Adjusting for quality growth within product categories can also exacerbate the bias arising from misweighted expenditure shares. By removing the portion of measured price increases that reflects improved features or performance, quality corrections lower the observed price change $d\ln P_j^{BEA}$ for those products–and if those same products suffer from unmeasured entry of new varieties, their underweighted shares become even more consequential.

Both the BLS and the BEA recognize the importance of quality adjustments. The BLS uses hedonic regressions to isolate the quality-driven component of price change for rapidly evolving goods. Similarly, following the Boskin Commission's 1996 recommendations, the BEA began applying quality adjustments in the late 1990s to 28 of its 212 PCE product categories, including autos, trucks, appliances, TVs, clothing, homes, and phones.[19]

To understand the impact of these corrections, Table 5 compares average residual demand growth in quality-adjusted versus non-adjusted categories over two subperiods: 1959–1995 (pre-Boskin) and 1996–2019 (post-Boskin). Across the full sample, quality-adjusted categories exhibit significantly negative residual demand (–0.05 log points per year), while non-adjusted categories hover near zero (–0.01). Because negative residual demand signals omitted new varieties–and hence underweighted expenditure shares–this pattern indicates that the BEA's quality corrections have widened the misweighting bias.

Table 5 also shows that the gap in average residual demand growth between quality-adjusted and non-adjusted categories widened after the BEA began its adjustments. In the pre-Boskin era (1959–1995), quality-adjusted categories averaged –0.02 log points of residual demand growth per year, compared to –0.01 for non-adjusted categories. In the post-Boskin era (1996–2019), the residual demand in quality-adjusted categories deepened to –0.10 log

[19] Appendix H lists these quality-adjusted categories.

points per year, while non-adjusted categories only declined to –0.02. This pattern suggests that isolated quality corrections have amplified the remaining misweighting bias.

Table 5: Residual Demand Growth: Quality vs Non-Quality Adjusted Products

	Δ Residual Demand
all: 1959-2019	
All (212 product categories)	-0.02
Quality adjusted (28; 27% of exp.)	-0.05
Non-quality adjusted (184; 73% of exp.)	-0.01
pre-Boskin Commission: 1959-1995	
All (206 product categories)	-0.01
Quality adjusted (27; 30% of exp.)	-0.02
Non-quality adjusted (179; 70% of exp.)	-0.01
post-Boskin Commission: 1996-2019	
All (212 product categories)	-0.03
Quality adjusted (28; 27% of exp.)	-0.10
Non-quality adjusted (184; 73% of exp.)	-0.02

Notes: The table reports average residual demand growth over three periods: (i) 1959–2019, (ii) 1959–1995 (pre-Boskin), and (iii) 1996–2019 (post-Boskin). Numbers in parentheses indicate the number of product categories and their aggregate expenditure shares for quality-adjusted and non-quality-adjusted groups. Residual demand is calculated from Equation 7 for each month, and we report the period average.

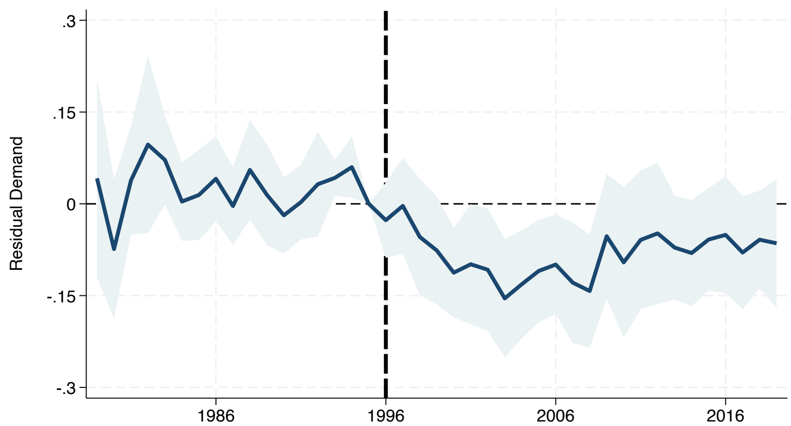
To further illustrate how residual demand evolves after quality adjustments, we employ an event-study approach comparing quality-adjusted and non-adjusted product categories before and after the Boskin Commission's recommendations. Our sample spans 1978–2018 and includes all BEA product categories, of which 28 received hedonic quality adjustments. By including a pre-Boskin period, we can test for any pre-existing trends. Categories that never received quality adjustments serve as the comparison group. Let Y_{it} denote residual demand growth for category i at time t. We estimate:

$$Y_{it} = \alpha + \sum_{k=-\infty}^{\infty} \gamma_k \mathbb{1}\{K_{it} = k\} + \theta_i + \lambda_t + \zeta X_{it} + \epsilon_{it}, \tag{12}$$

where θ_i are product-category fixed effects and λ_t are year fixed effects. K_{it} measures years relative to 1996–the year of the Boskin Commission report and the start of BEA's hedonic adjustments–so γ_k for $k < 0$ captures pre-trends and γ_k for $k \geq 0$ captures the dynamic effect

k years after implementation. X_{it} is a category-specific variety-correction term that controls for the impact of entry and exit on residual demand growth. Standard errors are clustered at the product-category level.

Figure 8: Event Study: Quality-Adjusted Categories



Notes: The graph shows the evolution of the residual demand growth before and after the Boskin Commission. The panels plot the coefficients of γ_k after estimating Equation 12. The vertical line marks the year the BEA started adjusting for quality 28 product categories. The gray area depicts the 95% confidence interval computed after clustering the standard errors at the product category level.

Figure 8 plots residual demand growth before and after the Boskin Commission. Conditional on category and time fixed effects, there is no evidence of pre-trends, consistent with the timing of the report's publication being as good as random. After 1996, residual demand growth declines sharply for the quality-adjusted categories. Under the assumption that the BLS's hedonic adjustment fully removes the quality bias (i.e. $\lambda_j = 0$) and after accounting for new-variety bias, this decline measures the bias arising from unmeasured new varieties in the BEA's nominal expenditures. The effect is especially large for quality-adjusted categories, which tend to have higher rates of product entry.

5 Conclusion

This paper introduces a new methodology for constructing an aggregate price index that jointly corrects for two of the most pervasive measurement problems in inflation statistics: unobserved quality and variety growth within product categories, and misweighting due to the omission or misclassification of new varieties in expenditure shares. Central to our approach is the identification of a "chosen" set of products—those whose residual demand growth averages zero—so that the remaining biases cancel out at the aggregate level.

Applying our method to the BEA's PCE data from 1959 to 2019, we show that most product categories exhibit non-zero residual demand growth, reflecting both quality- and variety-related distortions. By selecting the subset of categories with statistically negative and stable price-share elasticities, we construct an unbiased index without having to separate new goods from incumbents, directly measure quality attributes, or classify new varieties across groups. Using publicly available PCE data on prices and expenditure shares, our CES Chosen Price Index (CCPI) reveals that a CES aggregate of BEA prices has, on average, understated the true cost-of-living increase by 0.3-1.0 percentage points per year.

We validate these findings using rich NielsenIQ Consumer Panel data at the barcode level. There, our procedure selects 13,045 barcodes, about 2.0 percent of all varieties, and delivers similarly large negative biases (–0.75 percentage points) in the consumer packaged-goods sector. The large sample of "chosen" barcodes allows us to demonstrate that any sufficiently large subsample with average zero residual demand replicates the CES Chosen Price Index (CCPI), confirming the robustness of our approach.

Our results carry two immediate lessons for policy. First, partial bias corrections, whether for quality or for new varieties alone, can inadvertently worsen the net error if the remaining biases are left unaddressed. Second, because these two sources of bias pull the aggregate index in opposite directions, they must be corrected jointly. More broadly, our methodology offers a practical roadmap for statistical agencies to improve inflation measurement using readily available price and expenditure data, even in the absence of detailed product-level attributes or exhaustive variety tracking.

References

Aghion, Philippe, Antonin Bergeaud, Timo Boppart, Peter J Klenow, and Huiyu Li, "Missing growth from creative destruction," *American Economic Review*, 2019, *109* (8), 2795–2822.

Argente, David and Munseob Lee, "Cost of living inequality during the great recession," *Journal of the European Economic Association*, 2021, *19* (2), 913–952.

__ , Chang-Tai Hsieh, and Munseob Lee, "Measuring the cost of living in mexico and the united states," *American Economic Journal: Macroeconomics*, 2023, *15* (3), 43–63.

Baqaee, David R and Ariel Burstein, "Welfare and output with income effects and taste shocks," *The Quarterly Journal of Economics*, 2023, *138* (2), 769–834.

Bils, Mark, "Do higher prices for new goods reflect quality growth or inflation?," *The Quarterly Journal of Economics*, 2009, *124* (2), 637–675.

__ and Peter J Klenow, "Quantifying Quality Growth," *American Economic Review*, 2001, pp. 1006–1030.

Boskin, Michael J., Ellen R. Dulberger, Robert J. Gordon, Zvi Griliches, and Dale. Jorgenson, "Toward a More Accurate Measure of the Cost Of Living," *Final Report to the Senate Finance Committee*, 1996.

Broda, Christian and David E Weinstein, "Globalization and the Gains from Variety," *The Quarterly Journal of Economics*, 2006, *121* (2), 541–585.

__ and __ , "Product creation and destruction: Evidence and price implications," *American Economic Review*, 2010, *100* (3), 691–723.

Bureau of Economic Analysis, *Concepts and Methods of the U.S. National Income and Product Accounts* U.S. Department of Commerce 2023. Chapters 4–5.

Ehrlich, Gabriel, John C Haltiwanger, Ron S Jarmin, David Johnson, Ed Olivares, Luke W Pardue, Matthew D Shapiro, and Laura Zhao, "Quality adjustment at scale: Hedonic vs. exact demand-based price indices," *NBER Working Paper No. 31309*, 2024.

Feenstra, Robert C, "New product varieties and the measurement of international prices," *The American Economic Review*, 1994, pp. 157–177.

Hansen, Lars Peter, "Large sample properties of generalized method of moments estimators," *Econometrica*, 1982, pp. 1029–1054.

Nordhaus, William D, “Do Real-Output and Real-Wage Measures Capture Reality? The History of Lighting Suggests Not.,” in Timothy F. Bresnahan and Robert J. Gordon, eds., *The Economics of New Goods*, University of Chicago Press, 1997, pp. 29–66.

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.

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.

Sato, Kazuo, “The ideal log-change index number,” *The Review of Economics and Statistics*, 1976, pp. 223–228.

Vartia, Yrjö O, “Ideal log-change index numbers,” *Scandinavian Journal of Statistics*, 1976, pp. 121–126.

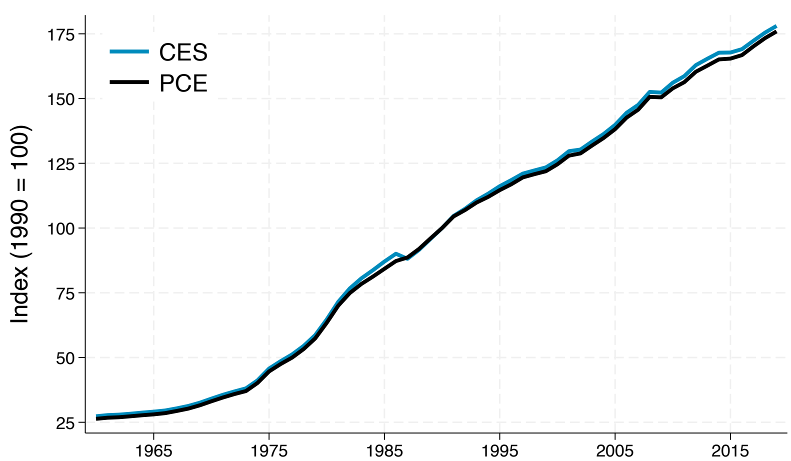
ONLINE APPENDIX

A Comparing PCE vs. CES

The official Personal Consumption Expenditures (PCE) Price Index is constructed using a chain-weighted Fisher index. A key feature of this methodology is that expenditure weights are regularly updated using observed consumption expenditures. As a result, once new goods or product categories enter measured consumption with positive expenditures, they receive nonzero weights in the aggregate price index. This feature allows the PCE to flexibly incorporate changes in the composition of consumption over time.

Our dataset uses the most detailed publicly available product-line definitions underlying the PCE and accounts for approximately 96% of total PCE spending. Using these data, we construct a CES price index based on BEA prices and expenditure shares, explicitly incorporating a correction for changes in product variety. Despite differences in aggregation theory and weighting schemes, the resulting CES price index aligns closely with the official PCE Price Index over the entire sample period. Figure A.1 illustrates the comparison between cumulative year-to-year inflation implied by the two indexes from 1959 to 2019.

Figure A.1: CES Price Index vs Official PCE



Notes: The figure plots cumulative price indexes normalized to 1990 = 100. "CES" denotes the CES price index with a variety correction, constructed using an elasticity of substitution of 2.91. "PCE" is the official Personal Consumption Expenditures (PCE) price index published by the Bureau of Economic Analysis (BEA).

B Derivation of CES price index

Below, we clarify that our formulas are exact for CES, not first-order approximations. Start from CES utility $U = \left(\sum_j c_j^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$. Minimizing expenditure $\sum_j p_j c_j$ subject to a target utility level yields the expenditure function $E(\mathbf{p}, U) = U \cdot e(\mathbf{p})$, where the unit expenditure function is:

$$e(\mathbf{p}) = \left(\sum_k p_k^{1-\sigma}\right)^{\frac{1}{1-\sigma}} \equiv \Phi^{\frac{1}{1-\sigma}}.$$

The exact cost-of-living index answers, how much more must a consumer spend to achieve the same utility under new prices? For CES, this is $P = E(\mathbf{p}_{t'}, U)/E(\mathbf{p}_t, U) = e(\mathbf{p}_{t'})/e(\mathbf{p}_t)$, so $\Delta \ln P = \Delta \ln e(\mathbf{p})$ is the exact price index.

The CES expenditure share is:

$$s_j = \frac{p_j^{1-\sigma}}{\sum_k p_k^{1-\sigma}} = \frac{p_j^{1-\sigma}}{\Phi}.$$

Taking logs:

$$\ln s_j = (1-\sigma)\ln p_j - \ln \Phi.$$

Now consider discrete changes between any two periods t and t':

$$d \ln s_j = (1-\sigma) d \ln p_j - d \ln \Phi.$$

Since $e(\mathbf{p}) = \Phi^{1/(1-\sigma)}$, we have $\ln \Phi = (1-\sigma)\ln e(\mathbf{p})$, so $d \ln \Phi = (1-\sigma)\Delta \ln P$. Substituting:

$$d \ln s_j = (1-\sigma) d \ln p_j - (1-\sigma) d \ln P = (1-\sigma)(d \ln p_j - d \ln P).$$

First, divide both sides by $1-\sigma$. Then multiply by ds_j and divide by $d \ln s_j$:

$$\frac{1}{1-\sigma} ds_j = \frac{ds_j}{d \ln s_j} d \ln p_j - \frac{ds_j}{d \ln s_j} d \ln P$$

Then, after we take the sum of both sides over all products j, and rearranging:

$$d \ln P = \sum_j \omega_j \, d \ln p_j$$

where ω_i is the Sato-Vartia weight defined as

$$\omega_i \equiv \frac{\frac{ds_i}{d\ln s_i}}{\sum_j \frac{ds_j}{d\ln s_j}}$$

This expression is exact and not an approximation. This is precisely the insight underlying Sato (1976) and Vartia (1976). When we observe non-zero residual demand in the data, this reflects measurement problems (quality bias, variety bias, or mismeasured shares), not approximation error.

C "Residual" Product Demand: Derivation

In this section, we provide a derivation of Equation 7. We begin by writing the expenditure share of product j:

$$d\ln S_j = -\left(\sigma - 1\right)\, d\,\ln\left(\frac{P_j}{P}\right).$$

Adding $(\sigma-1)\,d\,\ln\left(\frac{P_j^{BEA}}{P}\right)$ on both sides

$$\begin{aligned} d\ln S_j + (\sigma-1)\,d\,\ln\left(\frac{P_j^{BEA}}{P}\right) &= (\sigma-1)\,d\,\ln\left(\frac{P_j^{BEA}}{P}\right) - \left(\sigma-1\right)\, d\,\ln\left(\frac{P_j}{P}\right) \\ &= (\sigma-1)d\,\ln P_j^{BEA} - (\sigma-1)d\,\ln P_j. \end{aligned}$$

To obtain an expression for the right hand side, recall Equation 2

$$d\ln P_j = \frac{1}{\eta_j - 1}\,d\,\ln X_{\mathbb{I}_j} + d\ln P_{\mathbb{I}_j} - d\ln Q_{\mathbb{I}_j}$$

and Equation 3

$$d\ln P_j^{BEA} = d\ln P_{\mathbb{I}_j} - d\ln Q_{\mathbb{I}_j} + \lambda_j d\ln Q_{\mathbb{I}_j}.$$

Combining these three equations, we get:

$$d\ln S_j + (\sigma-1)\,d\,\ln\left(\frac{P_j^{BEA}}{P}\right) = -\frac{\sigma-1}{\eta_j-1}\,d\,\ln X_{\mathbb{I}_j} + (\sigma-1)\lambda_j d\ln Q_{\mathbb{I}_j}.$$

Adding $d\ln S_j^{BEA}$ to both sides, and rearranging the previous expression, we obtain the BEA's residual demand

$$d\ln S_j^{BEA} + (\sigma-1)\,d\,\ln\left(\frac{P_j^{BEA}}{P}\right) = -\frac{\sigma-1}{\eta_j-1}\,d\,\ln X_{\mathbb{I}_j} + (\sigma-1)\lambda_j d\ln Q_{\mathbb{I}_j} + d\ln S_j^{BEA} - d\ln S_j.$$

D Entry and Exit of Product Categories in BEA Data

Out of the 212 products, 182 products are covered continuously from January 1959 to December 2019, 2 products were combined into one in January 2002 ("Tenant-occupied stationary homes" and "Tenant landlord durables" became "Tenant-occupied, including landlord durables.'), and 27 categories were introduced after 1959 and have stayed until December 2019. Table D.I lists those 27 categories along with year of introduction.

Table D.I: List of 27 New Product Categories in BEA Data

Product Name	Series ID	Intro. Year
Video discs, tapes, and permanent digital downloads	DOVERC	1977
Personal computers/tablets and peripheral equipment	DCPPRC	1977
Computer software and accessories	DCPSRC	1977
Medical expenditures of foreigners	DMEFRC	1981
Cellular telephone services	DCELRC	1985
Home health care	DHHCRC	1987
Medical laboratories	DMLBRC	1987
Specialty outpatient care facilities and health and allied services	DOMSRC	1987
All other professional medical services	DOMORC	1987
Auto leasing	DALERC	1987
Truck leasing	DTLERC	1987
Meals at limited service eating places	DMLSRC	1987
Meals at other eating places	DMOERC	1987
Meals at drinking places	DMDPRC	1987
Mutual fund sales charges	DMUTRC	1987
Portfolio management and investment advice services	DPMIRC	1987
Homes for the elderly	DELDRC	1987
Residential mental health and substance abuse	DMENRC	1987
Individual and family services	DFAMRC	1987
Vocational rehabilitation services	DVOCRC	1987
Community food and housing / emergency / other relief services	DCFORC	1987
Other social assistance, not elsewhere classified	DSIARC	1987
Exchange-listed equities	DDCERC	1997
Over-the-counter equity securities	DICVRC	1997
Other imputed commissions	DICORC	1997
Video streaming and rental	LA000233	2007
Audio streaming and radio services (including satellite radio)	LA000232	2007

Notes: The table reports the list of 27 new product categories in the BEA data along with year of introduction.

E Elasticity of Substitution Estimation Strategy

In order to obtain the elasticity of substitution, σ_g, for each item, we rely on the method developed by Feenstra (1994) and extended by Broda and Weinstein (2006) and Broda and Weinstein (2010). The procedure consists of estimating a demand and supply equation for each barcode by using only the information on prices and quantities. For this estimation, we face the standard endogeneity problem for a given barcode. Although we cannot identify supply and demand, the data do provide information about the joint distribution of supply and demand parameters.

We first model the supply and demand conditions for each barcode within an item. Specifically, we estimate the demand elasticities by using the following system of differenced demand and supply equations as in Broda and Weinstein (2006):

$$\Delta^{\underline{u},t}\ln S_{ig} = (1-\sigma_g)\Delta^{\underline{u},t}\ln P_{ig} + \iota_{ig} \tag{13}$$

$$\Delta^{\underline{u},t}\ln P_{ig} = \frac{\delta_g}{1+\delta_g}\Delta^{\underline{u},t}\ln S_{ig} + \kappa_{ig} \tag{14}$$

Note that when the inverse supply elasticity is zero (i.e. δ_g=0), the supply curve is horizontal and there is no simultaneity bias in σ_g. Equations 13 and 14 are the demand and supply equations of barcode k in an item i differenced with respect to a benchmark barcode in the same item. The k^{th} good corresponds to the largest selling barcode in each item. The k-differencing removes any item level shocks from the data.

The identification strategy relies on two important assumptions. First, we assume that ι_{ig} and κ_{ig}, the double-differenced demand and supply shocks, are uncorrelated (i.e., $\mathbb{E}_t(\iota_{ig}\kappa_{ig}) = 0$). This expectation defines a rectangular hyperbola in (δ_g, σ_g) space for each barcode within an item, which places bounds on the demand and supply elasticities. Because we already removed any item level shocks, we are left with within item variation that is likely to render independence of the barcode-level demand and supply shocks within an item. Second, we assume that σ_g and ω_g are restricted to be the same over time and for all barcodes in a given item.

To take advantage of these assumptions, we define a set of moment conditions for each item i in a basic heading b as below:

$$G(\beta_g) = E_T[\nu_{ig}(\beta_g)] = 0 \tag{15}$$

where $\beta_g = [\sigma_g, \delta_g]'$ and $\nu_{ig} = \iota_{ig}\kappa_{ig}$.

For each item i, all the moment conditions that enter the GMM objective function can be combined to obtain Hansen (1982)'s estimator:

$$\hat{\beta}_g = \arg\min_{\beta_g \in B} G^*(\beta_g)' W G^*(\beta_g) \quad \forall\, i \in \omega_b \tag{16}$$

where $G^*(\beta_g)$ is the sample analog of $G(\beta_g)$, W is a positive definite weighting matrix, and B is the set of economically feasible β_g (i.e., $\sigma_g > 0$). Our estimation procedure follows Redding and Weinstein (2020) using the NielsenIQ Consumer Panel data from 2004-2019. The elasticities are estimated using data at the quarterly frequency. Households are aggregated using sampling weights to make the sample representative. We weight the data for each barcode by the number of raw buyers to ensure that our objective function is more sensitive to barcodes purchased by larger numbers of consumers. We consider barcodes with more 10 or more observations during the estimation.

Table E.I: Estimated Elasticities of Substitution

Percentile	Elasticities of Substitution
10	5.12
25	5.52
median	6.32
mean	6.42
75	7.24
90	7.89

Notes: The table reports descriptive statistics of estimated elasticities of substitution for each product category in the Consumer Panel data. We use moment conditions of the double-differenced residuals in demand and supply with the GMM estimation approach.

F List of Chosen Product Categories

Our of 212 product categories, the 11 chosen product categories are listed below.

Table F.I: List of 11 Chosen Product Categories

Product Category	Estimated Coefficient	Standard Error
Repair of household appliances	-0.77	0.15
Group housing (23)	-0.59	0.13
Package tours	-0.50	0.14
Lotteries	-3.14	0.78
Motorcycles	-1.85	0.15
Other household services	-0.46	0.14
Owner-occupied mobile homes	-1.65	0.07
Pari-mutuel net receipts	-0.50	0.09
Small electric household appliances	-0.71	0.10
Tax preparation and other related services	-0.67	0.12
Tenant-occupied mobile homes	-1.44	0.10

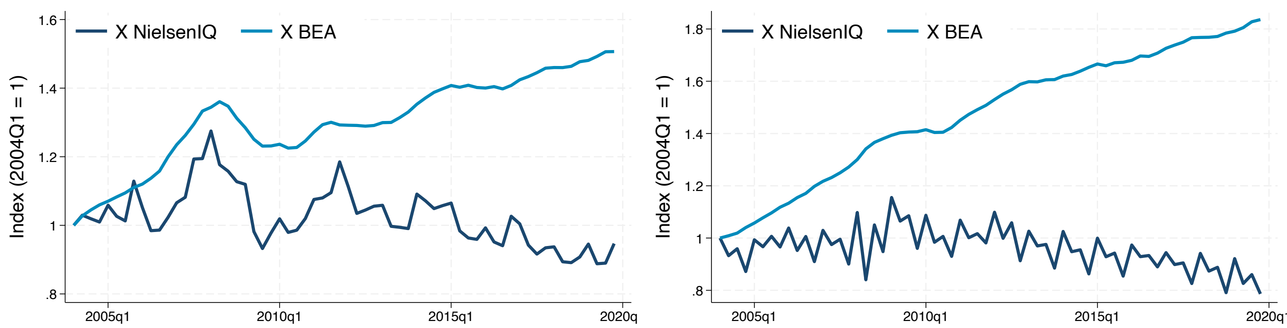
Notes: The table reports the list of 11 chosen product categories in the BEA data and its estimated coefficient and standard error.

G Comparing BEA vs. NielsenIQ Expenditure Shares

We manually match product categories in the NielsenIQ data to those in the Personal Consumption Expenditures (PCE) accounts of the U.S. Bureau of Economic Analysis (BEA). The matching prioritizes accuracy of coverage: several BEA categories aggregate a broader set of products than their NielsenIQ counterparts (and vice versa). We therefore retain only categories for which product coverage is most likely to coincide across the two data sources.

Using this procedure, we successfully match 21 product categories observed in both datasets. The resulting sample spans 64 quarters, from 2004Q1 to 2019Q4. To compare expenditure dynamics across data sources, we normalize expenditures in each category to one in 2004Q1.

Figure G.1: Expenditures Over Time: Milk and Cereal



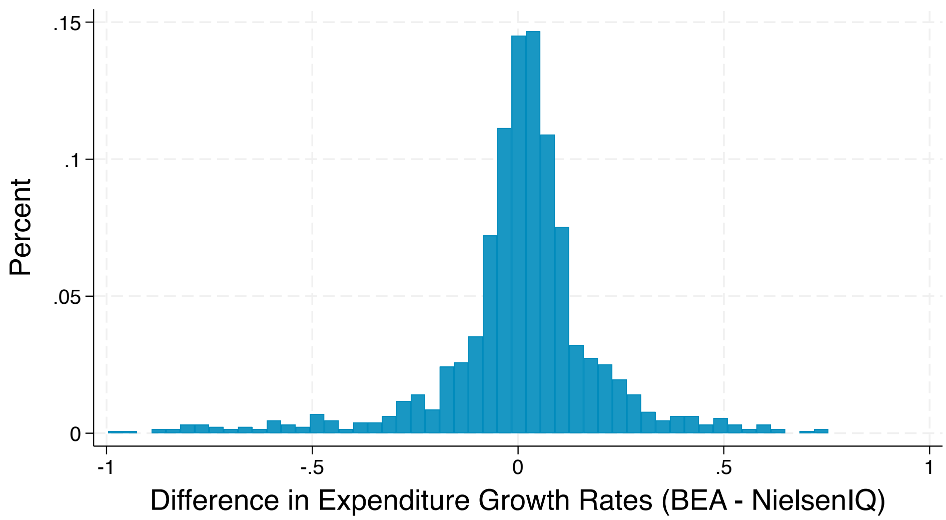
(a) Milk (b) Cereal

Notes: The figure plots the evolution of expenditures over time. Each panel plots indexed expenditures from BEA and NielsenIQ for the indicated category. Expenditures are normalized to 1 in 2004Q1.

Figure G.1 illustrates the evolution of expenditures for two categories: milk and cereal. In both cases, expenditures measured in the BEA grow substantially faster over time than those measured in NielsenIQ, indicating a widening divergence between aggregate and scanner-based expenditure measures.

To summarize these discrepancies more broadly, Figure G.2 plots the quarterly difference in log expenditure growth rates, $\Delta \ln(\text{BEA}) - \Delta \ln(\text{NielsenIQ})$, across all matched categories and quarters. The distribution is centered at zero but exhibits substantial dispersion, indicating sizable and persistent category-level differences in measured expenditure growth between the two data sources.

Figure G.2: Quarterly Difference in Expenditure Growth Rates (BEA – NielsenIQ)



Notes: Histogram of quarterly differences in log expenditure growth rates between BEA and NielsenIQ across matched product categories.

H Identifying Quality Adjusted Product Categories

The BLS provides a list of quality adjusted Entry Level Items (ELI).[20] Out of 273 ELIs, 36 items (13.2%) are quality adjusted. We manually matched ELIs to 212 NIPA product categories we use. 28 (13.2%) product categories in Table H.I are identified to be quality adjusted.

[20]https://www.bls.gov/cpi/quality-adjustment/home.htm

Table H.I: List of 28 Quality Adjusted NIPA Product Categories

Product Name	Series ID	Type
New domestic autos	DNDCRC	Durable
New foreign autos	DNFCRC	Durable
New light trucks	DNWTRC	Durable
Net purchases of used motor vehicles	DNPVRC	Durable
Used autos	DNPURC	Durable
Used light trucks	DUTRRC	Durable
Major household appliances	DMHARC	Durable
Small electric household appliances	DSEARC	Durable
Televisions	DTVSRC	Durable
Other video equipment	DOVARC	Durable
Photographic equipment	DCAMRC	Durable
Personal computers/tablets and peripheral equipment	DCPPRC	Durable
Telephone and related communication equipment	DTCERC	Durable
Women's and girls' clothing	DWGCRC	Nondurable
Men's and boys' clothing	DMBCRC	Nondurable
Shoes and other footwear	DSHURC	Nondurable
Tenant-occupied mobile homes	DTMHRC	Service
Tenant-occupied stationary homes	DTSPRC	Service
Tenant landlord durables	DTLDRC	Service
Tenant-occupied, including landlord durables	LA000630	Service
Owner-occupied mobile homes	DOMHRC	Service
Owner-occupied stationary homes	DOSTRC	Service
Auto leasing	DALERC	Service
Truck leasing	DTLERC	Service
Cable, satellite, and other live television services	DCTVRC	Service
Land-line telephone services, local charges	DLOCRC	Service
Land-line telephone services, long-distance charges	DLDTRC	Service
Cellular telephone services	DCELRC	Service

Notes: The table reports the list of 28 quality adjusted NIPA categories out of 212 NIPA categories we use in the paper.