Bias in the PCE Price Index*

David Argente[†] Yale University

Chang-Tai Hsieh[‡] University of Chicago

Munseob Lee[§] UC San Diego

May 2025

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. When we apply this procedure to the BEA data from 1959 to 2019, our ideal inflation rate *exceeds* the official PCE inflation rate by 0.41 percentage points per year on average. The aggregate bias was close to zero before the BLS introduced 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 Numbers: D11, D12, E01, E31
Keywords: Price Index, Inflation

*We thank Alberto Cavallo, Marc-Andreas Muendler, and Dan Sichel for helpful comments. Researchers' own analyses 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.

[†]Email: david.argente@yale.edu. Address: 165 Whitney Ave, New Haven, CT 06511.
[‡]Email: chsieh@chicagobooth.edu. Address: 5807 South Woodlawn Ave., Chicago, IL 60637.
[§]Email: munseoblee@ucsd.edu. Address: 9500 Gilman Drive #0519, La Jolla, CA 92093-0519.

It has long been recognized that the central challenge in measuring price indices lies in the fact that the mix and quality of products change over time. For example, Nordhaus (1997), Bils and Klenow (2001), and Bils (2009) document that quality improvement are not fully taken into account in standard price indices. While the official Personal Consumption Expenditures Price Index (PCE) incorporates the effect of quality change for some products, such as computers and TVs, using hedonic techniques, this adjustment is not made for many products.

The second concern, which has not be as widely recognized, arises from the mis-classification of new varieties within product categories.[1] Even when new varieties are captured, correctly classifying them remains challenging: should a tablet be recorded as a new product within the television category or the computer category, and if so, which one? Or are tablets its own distinct category? This misweighting bias becomes especially significant when new varieties account for a meaningful share of consumer spending or when the affected categories undergo large price changes. For instance, product categories such as TVs or computers are particularly susceptible to misweighting bias because of their steep price declines over time and the frequent introduction of new varieties.

This paper introduces a methodology that addresses the challenge of accounting for missing quality and for new product categories and varieties, while also tackling weighting bias. Leveraging a CES utility function, we demonstrate that each of these sources of bias manifests as non-zero residual demand growth at the product-category level. The telltale sign that the BEA omits quality and variety growth is increasing residual demand in an incumbent category, whereas the hallmark of misweighting is declining residual demand.

Building on this insight, we develop a methodology to calculate an unbiased price index by identifying a subset of product categories in which residual demand growth averages zero. We refer to these as the *chosen* product categories–a set of categories whose observed prices fully incorporate the effects of quality changes and new varieties (i.e., expenditure-share shifts driven solely by price changes). Our approach offers three key advantages: first, it does not require separating new goods from existing ones; second, it does not depend on direct measures of quality or detailed product-attribute data; and third, it avoids the need to classify new varieties across categories.

Once we have identified this bundle of chosen categories, the change in the aggregate price index is given by the product of two terms: (1) the weighted average change in prices of the

[1]Both the CPI and the PCE Price Index use the same price data collected by the Bureau of Labor Statistics (BLS) to measure inflation. The CPI maintains fixed expenditure weights until new consumption data become available. In contrast, the PCE index utilizes survey data to update its expenditure weights, allowing it to adjust for changes in consumer spending patterns; however, this data does not include information on the introduction of new products. Our application uses data from the PCE index, but our findings apply also to the CPI.

products in this chosen bundle, and (2) the change in the bundle's market share, adjusted by a function of the price elasticity of demand. This formulation mirrors that of Feenstra (1994), with two distinctions. First, the weighted average price change of the chosen bundle does not represent all products with zero residual demand growth. This is because many items whose unit prices fully capture quality or variety changes may not be included in this bundle. Second, the "new goods" term is not solely the welfare gain from new products; it also reflects the impact of products excluded from the chosen bundle. However, the product of these two terms yields an unbiased estimate of the combined effect of product price changes and welfare gains from new products, even when those two effects cannot be identified separately.

We implement our methodology using public price and expenditure data from the Bureau of Economic Analysis (BEA) for roughly 200 product categories in the PCE Price Index. For the PCE Price Index, we begin by estimating the correlation between expenditure shares and prices for each category, focusing on those exhibiting a negative and significant relationship over time. Notably, prices and expenditure shares are positively correlated for 67% of products in the PCE index. This finding suggests that the BEA's price indices for these categories do not fully capture the effects of quality changes or variety growth. We then calculate the aggregate price index as the product of the price index of the *chosen* product categories and the expenditure share of these categories, adjusted by the demand elasticity. Under our assumption that the average change in residual demand is zero among categories where the covariance between price and expenditure shares is negative and significant, the product of these two terms yields an estimate of the aggregate price index that accounts for biases. Our estimates suggest that the PCE index understates the annual inflation rate by 0.41 percentage points on average between 1959 and 2019.[2]

To ensure that this trend is not driven by variety growth, we estimate the new-variety bias at the product-category level using measures of job creation and destruction from the Business Dynamics Statistics (BDS) data following Aghion et al. (2019), and we reach the same conclusion. Thus, the BEA's quality adjustments are likely overstating the impact of residual demand growth, primarily because data on expenditure shares are unavailable for calculating quality and variety biases.

Importantly, the bias in the PCE Index is particularly pronounced among product categories in which the Bureau of Economic Analysis (BEA) began adjusting for quality in the late 1990s, following the release of the Boskin Commission's 1996 report. Recall that a positive residual demand indicates that the BEA has overlooked quality growth, while a negative residual demand signals misweighting of varieties. Because these biases pull in opposite

[2]In the NielsenIQ data, prices and expenditure shares are positively correlated for 58% of products. The aggregate price of consumer packaged goods exceeds the price calculated from the Feenstra (1994) formula by 0.74 percentage points per year between 2004 and 2019.

directions, correcting only for missing quality bias can exacerbate the bias arising from misweighting varieties. Our estimates show that categories subject to quality adjustment after the Boskin Commission exhibit more negative residual demand growth.

Our work is related to two bodies of work. First, Broda and Weinstein (2006, 2010), Aghion, Bergeaud, Boppart, Klenow and Li (2019), and Argente and Lee (2021) estimate the bias in price indices from not taking into account variety growth. A second body of work by Nordhaus (1997), Bils and Klenow (2001), and Bils (2009) measures the bias because standard price indices do not adjust for quality growth. We have two messages relative to this literature. First, when there is quality growth, it is not possible to separately identify variety growth from quality change. Second, instead of attempting to separate these effects, we can measure their net effect by identifying a set of products for which quality change and variety growth are less likely to be an issue.

We highlight two papers that serve as the foundation for our work. First, we build upon the idea presented by Bils (2009) that changes in a product's expenditure share can be used to infer its true price. In our case, we employ this concept 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 products with constant barcodes over time, 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.

The rest of the paper is organized as follows. Section 1 develops a simple conceptual framework to compare the official BEA price index with an ideal aggregate index and shows how various biases manifest as nonzero residual demand at the product level. Section 2 describes our iterative 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 scanner data and present our main findings on the bias in the official PCE 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. We demonstrate that these two error sources manifest empirically as 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 Ideal 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 of 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.

[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.

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* varieties in j.[4] 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 PCE 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 assume that the "official" price index is a weighted average of prices of each product category as measured by the BEA:

$$d \ln P^{BEA} \equiv \sum_{j \in \mathbb{I}} \omega_j^{BEA} d \ln P_j^{BEA} \tag{4}$$

where the weight ω_j^{BEA} is measured from the expenditure shares in the BEA S_j^{BEA}:[5]

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

The difference between the BEA's price index (Equation 4) and the "ideal" price index

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

[5] We assume that the weights 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.

(Equation 1) is then:

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

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

The first line captures two upward-bias components arising from missing new products and missing new varieties. The first is the new-product bias: when new product categories enter, incumbents' share falls ($d\ln S_{\mathbb{I}} < 0$), and scaling by $1/(\sigma-1)$ converts this into a positive adjustment to the measured price index. The second is the new-variety bias: as new varieties arrive, incumbents' share drops ($d\ln X_{\mathbb{I}_j} < 0$), and each category's decline is 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

$$
\sum_{j\in\mathbb{I}} \left(\omega_j^{\text{BEA}} - \omega_j\right) d\ln P_j^{\text{BEA}} = \text{Cov}\left(\omega_j^{\text{BEA}} - \omega_j,\; d\ln P_j^{\text{BEA}}\right),
$$

where the weight discrepancy $\omega_j^{\text{BEA}} - \omega_j$ arises from both the omission and misclassification of new varieties and depends on $d\ln S_j^{\text{BEA}} - d\ln S_j$. In particular, the BEA's expenditure share for category j is

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

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

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

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

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

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:[6]

$$d\ln S_j^{\text{BEA}} + (\sigma-1)\, d\ln\!\left(\tfrac{P_j^{\text{BEA}}}{P}\right) = \lambda_j\,(\sigma-1)\, d\ln Q_{\mathbb{I}_j} \;-\; \frac{\sigma-1}{\eta_j-1}\, d\ln X_{\mathbb{I}_j} + \; d\ln S_j^{\text{BEA}} - d\ln S_j. \tag{6}$$

Equation 6 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 6, 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 in Equation 6, $d\ln S_j^{\text{BEA}} - d\ln S_j$, 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.

By 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

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

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

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 6 for the products in $\mathbb{C}$ can then be expressed as:

$$d\ln P = \sum_{j\in\mathbb{C}} \alpha_j \Big(\tfrac{1}{\sigma-1}\, d\ln S_{\mathbb{I}_j} + d\ln P_j^{\text{BEA}}\Big). \tag{7}$$

Equation 7 shows that the change in the aggregate price index can be calculated using only

[6] Details on the derivation can be found in Appendix D.

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 7. 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 6 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:

$$d \ln S_j^{BEA} = \beta_j\ d\ \ln \left(\frac{P_j^{BEA}}{P} \right) + \epsilon_j. \tag{8}$$

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

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

When the residual demand is zero, the coefficient β_j is *unambiguously* negative given our assumption that $\sigma > 1$. On the other hand, if this assumption fails, an omitted-variable bias arises and β_j will be biased toward zero or could even turn positive. Intuitively, the covariance between the relative-price change and the residual demand is positive because each bias term in our two key identities moves both measures in the same direction. For instance, when quality improves (i.e. $d \ln Q_{\mathbb{I}_j} > 0$), the residual-demand term

$$\lambda_j (\sigma - 1)\, d \ln Q_{\mathbb{I}_j} > 0$$

in Equation 6 and the corresponding term

$$\sum_{j \in \mathbb{I}} \omega_j\, \lambda_j\, d \ln Q_{\mathbb{I}_j} > 0$$

in Equation 5 both increase. Because both the residual demand and $d\ln(P_j^{\text{BEA}}/P)$ rise together under quality growth,

$$\text{Cov}\left[d\ln\left(\frac{P_j^{\text{BEA}}}{P}\right), \text{Residual Demand}_j\right] > 0.$$

2.1 Iterative Procedure

Our procedure begins by implementing Equation 8 product-by-product in the data. Estimating β_j requires knowing each product's relative price P_j^{BEA}/P and thus the ideal aggregate index P. Because P is unknown a priori, we adopt an iterative procedure:

1. Take the BEA aggregate price index, given σ, as an initial guess for the ideal aggregate price index.[7]

2. Estimate Equation 8 for each product category in the BEA data and select those categories for which β_j is negative and statistically different from zero at the 1% level.

3. Calculate a new aggregate price index using only the set of *chosen* products via Equation 7.

4. Compare the newly constructed index to the previous guess. If the average absolute log-difference exceeds 0.001, 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 an ideal aggregate price index calculated using these *chosen* categories.[8]

3 Empirical Application

In this section, we apply our iterative procedure to the BEA price and expenditure data to construct an ideal aggregate index that corrects for the quality, variety, and weighting biases described above. To validate our approach, we repeat the exercise using detailed NielsenIQ scanner 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.

[7] Our initial guess is a Sato-Vartia price index using all product categories. We use $\sigma = 4$ for our benchmark case and explore other values in robustness checks.

[8] In our application with the NielsenIQ data, we implement this procedure at the barcode level within each product category. We estimate the elasticity of substitution using the approach of Redding and Weinstein (2020) and take the exact price index from the Feenstra (1994) formula as the initial guess.

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).[9] 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.[10]

3.2 Residual Demand: BEA Data

We implement our iterative procedure for each of the BEA's 212 product categories, using the full 61 years of available data. When estimating Equation 8, we work with monthly data and take twelve-month differences to control for seasonality.

Figure 1 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 0.14, with a median of 0.18), and about 67% of products have a positive coefficient. The right panel restricts attention to coefficients significant at the 1% level, where positive estimates are even more prevalent. Our "chosen" set $\mathbb{C}$ comprises the 20 categories to the left of the vertical black line—those with significantly negative β_j. Examples include child care services, electric personal-care appliances, and small electric household appliances.[11]

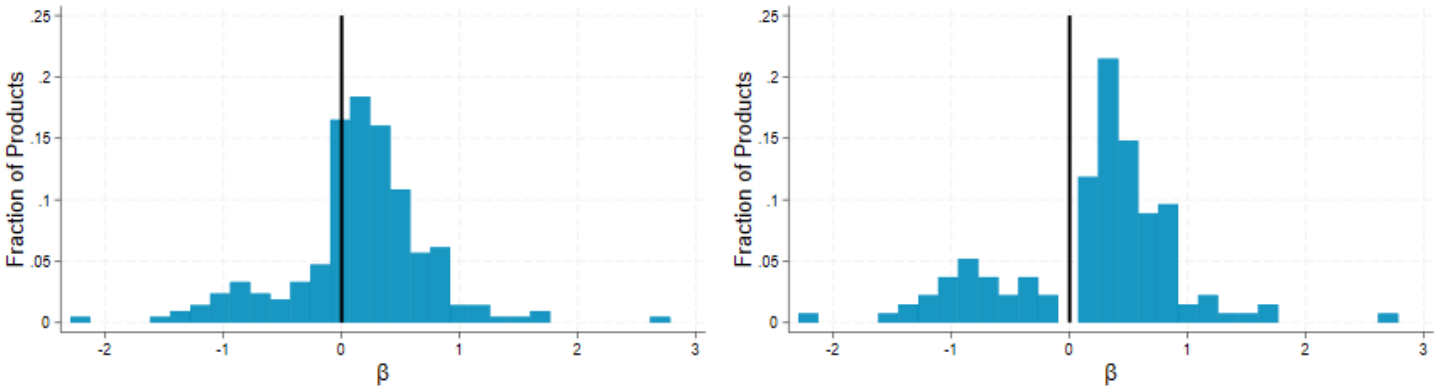
Given the chosen product categories, we calculate the ideal aggregate price index using Equation 7. Table 1 reports the mean and standard deviation of the bias for various values of the elasticity of substitution. Under our benchmark $\sigma = 4$, the average bias is -0.41 percentage points with a standard deviation of 1.29. In other words, ideal aggregate inflation exceeds BEA aggregate inflation by 0.41 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 BEA aggregate inflation generally understates price increases. Under $\sigma = 3$ and $\sigma = 5$, the average biases remain sizable.

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

[10] See Appendix A for further details.

[11] Appendix B lists all 20 chosen categories along with their estimated coefficients and standard errors.

Figure 1: 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 8 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.

Table 1: Bias under Various Elasticity of Substitution Values

	$\sigma = 3$	$\sigma = 4$ (benchmark)	$\sigma = 5$
Average Bias	-0.99	-0.41	-0.38
Standard Deviation of Bias	1.41	1.29	1.17

Notes: The table reports the mean and standard deviation of the bias–measured as the percentage-point difference between the BEA price index (Equation 4) and the one-year change in the ideal price index (Equation 7)–from January 1960 to December 2019, under various values of the elasticity of substitution.

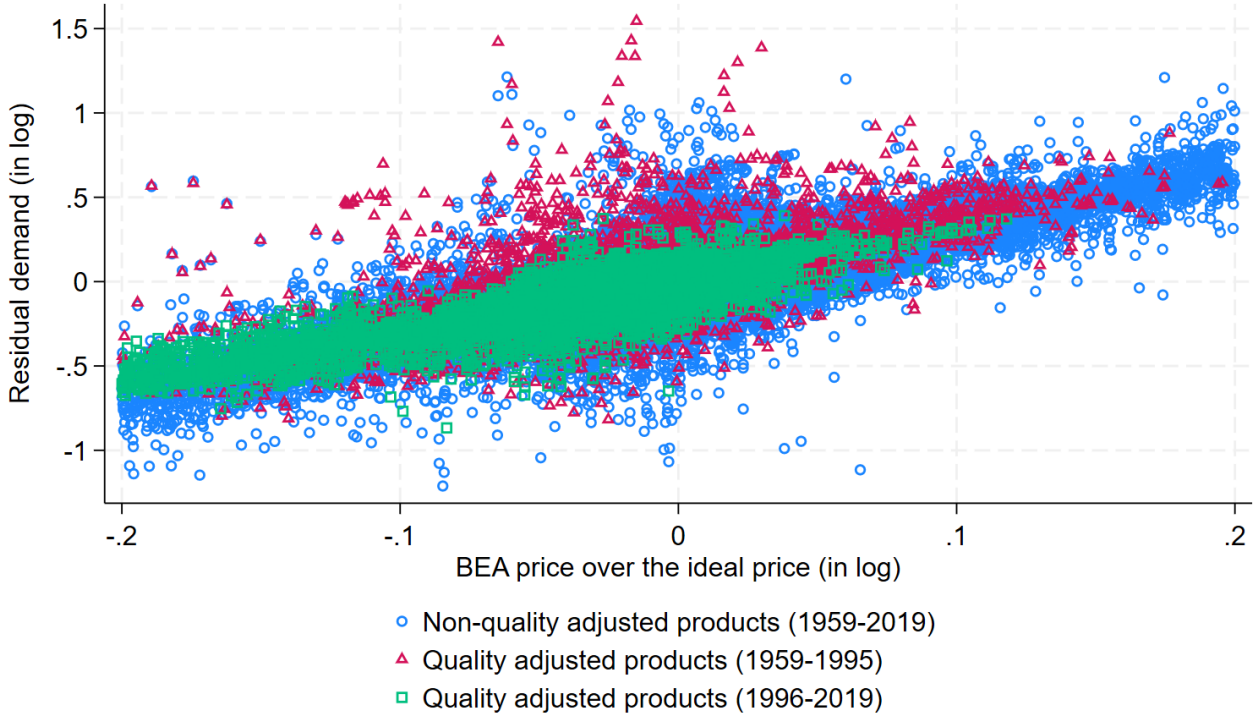
The evidence in Table 1 that the BEA's inflation rate is too *low* may seem surprising. After all, the BEA index omits the direct effects of new varieties and does not fully capture quality improvements–both of which would tend to make its inflation rate too *high*. However, missing new varieties also distorts the expenditure weights used to aggregate product-level prices. The resulting bias depends on the covariance between a category's importance of new varieties and its BEA-measured inflation rate. When new varieties are more prevalent in categories where the BEA's measured inflation is lower, the overall inflation rate implied by the BEA's data will be too low.

Figure 2 provides suggestive evidence that this is indeed the case. The figure plots monthly

observations at the product-category level, with each point showing a category's residual demand (from Equation 6) against its bias in the BEA price index, $d\ln(\frac{P_j^{\text{BEA}}}{P})$. The figure illustrates the positive covariance between these two measures. Furthermore, recall that residual demand is negative when new varieties are omitted from the BEA's expenditure data. The figure shows that observations with more negative residual demand–reflecting omitted new varieties–tend to coincide with larger negative price biases, indicating those categories have the lowest measured inflation relative to the ideal index.

The figure also distinguishes the 28 quality-adjusted categories before and after 1996; after hedonic adjustments begin, these categories shift further into the bottom-left quadrant, suggesting that partial quality corrections have unmasked even larger misweighting errors. In Section 4.2, we discuss this further and examine how adjusting for quality growth within product categories can also exacerbate the bias arising from misweighted expenditure shares.

Figure 2: Residual Demand vs. BEA Price over the Ideal Price



Notes: The figure shows the relationship between residual demand and the BEA price relative to the ideal price. Residual demand is calculated from Equation 6 for each month. The BEA price relative to the ideal price is $d\ln(\frac{P_j^{\text{BEA}}}{P})$. We report non–quality-adjusted products for the entire sample period, and quality-adjusted products separately for the pre-Boskin period (1959–1995) and the post-Boskin period (1996–2019). For graphical clarity, we trim the top and bottom 1% of observations where the logged BEA price relative to the ideal price lies below –0.2 or above 0.2.

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 20 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 scanner data at the barcode level. The larger "chosen" sample in the NielsenIQ data allows us to report the full distributions of R^2 and mean-squared error, and to assess the sensitivity and robustness of our iterative procedure.

3.3 Residual Demand: NielsenIQ Data

In this section, we apply our iterative procedure to barcode-level data from the NielsenIQ Consumer Panel and uncover three key findings. First, as in the BEA data, the vast majority of barcodes yield positive β_j estimates, indicating pervasive residual-demand biases at the product level. Second, consistent with our BEA results, the average aggregate bias is negative—about -0.74 percentage points per year—though it varies considerably across departments (from -0.11 in fresh produce to -2.79 in general merchandise). Third, leveraging the richness of the scanner data, we show that our chosen barcodes deliver substantially smaller mean-squared errors and higher R^2 values in the regressions, demonstrating that they effectively exhibit near-zero average residual demand. To assess robustness, we then reestimate the bias under multiple subsamples—short- and long-duration barcodes, those with particularly low mean-squared errors or high R^2, and even random 50% draws of the chosen set—and find that our bias estimates remain remarkably stable throughout.

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.[12]

We implement our iterative procedure at the barcode level within each product category using all 16 years of the Nielsen Consumer Panel data (2004–2019). Following Redding and Weinstein (2020), we aggregate each barcode's prices and expenditures to a quarterly

[12]Barcodes in the Nielsen Consumer Panel are organized hierarchically: each UPC belongs to one of 1,070 product modules, which are grouped into 104 product groups, which in turn fall under 10 major departments. Throughout the paper, we refer to product groups as product categories.

frequency.[13] 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 8 separately for each one.

As before, estimating Equation 8 requires an initial guess of its product category's price index. Therefore, we construct the exact price index for each of the 104 Nielsen product categories using all available barcodes within that category.[14] 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).[15] 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 104 product categories, using the same selection and convergence criteria as in the BEA implementation.[16] As before, this yields a set of barcodes whose unit prices are likely to fully capture changes in market share, and an unbiased product-category price index computed using these chosen barcodes.

Panel (a) of Figure 3 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 3 shows the distribution of coefficients significant at the 1% level. As before, we select barcodes with coefficients to the left of the vertical black line.[17]

Building on our barcode-level selection, we compute the ideal 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 0.74 percentage points per year. Unlike the BEA index, which covers the

[13] Nielsen 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.

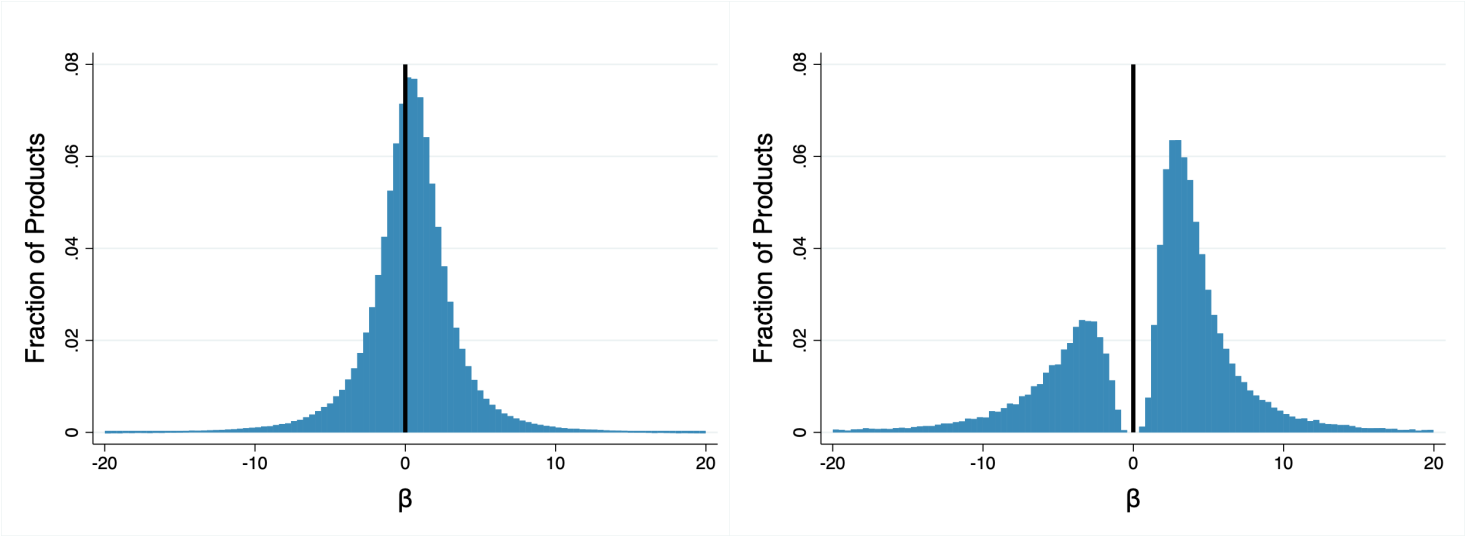
[14] The exact index multiplies a Sato–Vartia price index by a variety-correction term that captures the welfare effect of products entering and exiting.

[15] This procedure involves estimating demand and supply equations for each barcode using observed prices and quantities; see Appendix E.1 for details.

[16] The average number of iterations across categories is 4.77 (median 5).

[17] The product categories that contribute the largest share of chosen barcodes include butter and margarine; canned seafood; sugar sweeteners; eggs; fresh meat; frozen meal/poultry/seafood; paper products; frozen breakfast foods; milk; frozen pizza/snacks; detergents; tobacco and accessories; frozen juices; dough products; canned juice; frozen prepared foods; nuts; charcoal/log accessories; and frozen baked goods. Table E.II reports the shares of chosen barcodes by department, with fewer barcodes selected from more durable categories such as Health and Beauty Aids and General Merchandise.

Figure 3: Distribution of Coefficients β_i in NielsenIQ Data



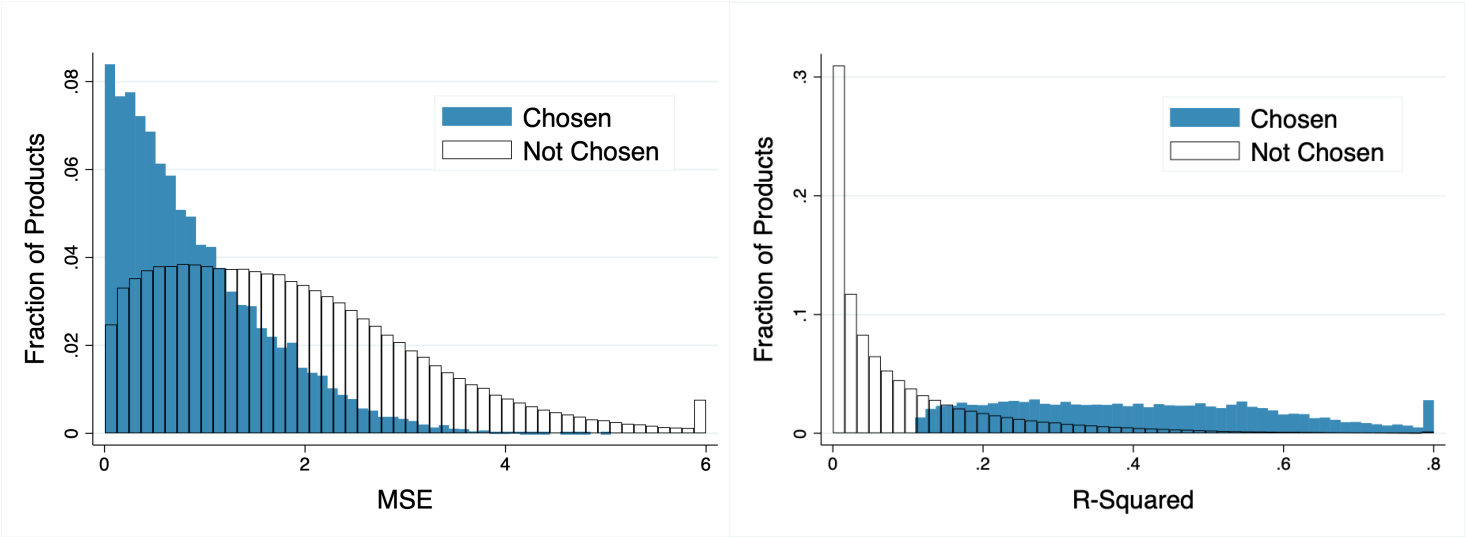
(a) All Barcodes (b) Significant at 1%

Notes: The figure shows the distribution of the coefficients β_i estimated using equation 8 for barcodes in the Nielsen data (2004-2019). Panel (a) includes all barcodes (650,021 barcodes). Panel (b) includes barcodes whose coefficient is significantly different than zero at 1 percent confidence level (18,994 barcodes with negative coefficient and 38,020 barcodes with positive coefficient). The black vertical line marks products whose coefficient is zero.

entire economy, the NielsenIQ index represents only the consumer packaged-goods sector; nevertheless, substantial heterogeneity remains across departments. Table E.III reports the average point-differences between the chosen-barcode index and the incumbent-only index across the 10 major departments. In every department, the conventional index understates inflation: Non-Food Grocery and Frozen Foods show biases of –0.90 and –0.83, respectively, while Health and Beauty Aids and General Merchandise exhibit the largest gaps (–1.01 and –2.79), reflecting especially high rates of entry and turnover in those categories.

The richness of the NielsenIQ data gives us a large "chosen" set—about 2.7% of all barcodes, or roughly 18,994 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 4 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.

Figure 4: MSE and R-Squared - Product Level Regressions



Notes: Panel (a) shows the histogram of mean-squared errors estimated from Equation 8 for barcodes in the Nielsen 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 negative and significantly different from zero at the 1 percent level (18,994 barcodes).

Table 2: Ideal Price Index - Robustness Checks

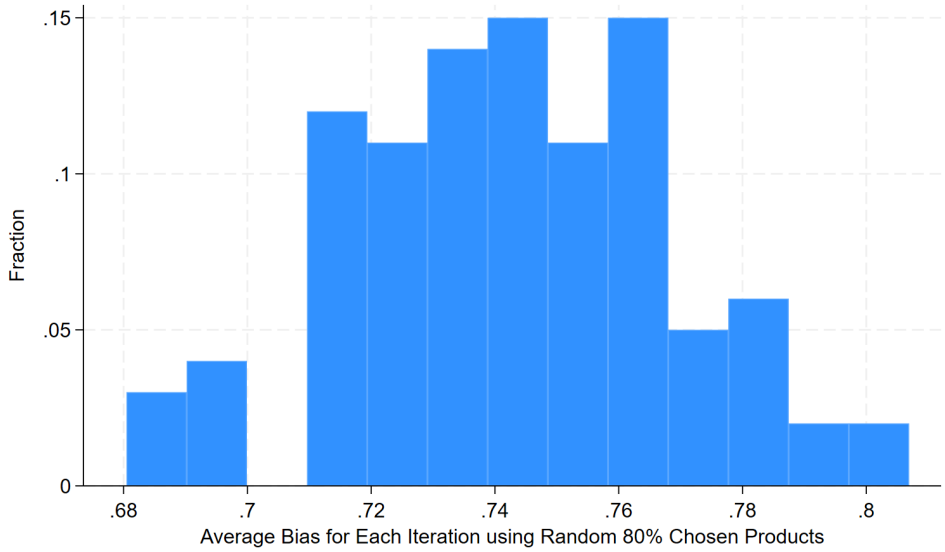
Subsample	Difference (pp)
$\beta < 0$ + significant at 1% + long duration (16 qtrs)	0.98
$\beta < 0$ + significant at 1% + MSE< 1	0.89
$\beta < 0$ + significant at 1% (**baseline**)	**0.74**
$\beta < 0$ + significant at 5%	0.67

Notes: The table reports the average annual point-difference between the ideal price index constructed from chosen barcodes (those with $\beta_i < 0$ and significant at 1%, unless otherwise noted) and the conventional index using all 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% sig." uses a 5% significance cutoff for $\beta_i < 0$.

Lastly, the large size of our chosen-barcode set in the NielsenIQ data enables us to test the robustness of the ideal 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 ideal index under four alternative selection rules, as shown in Table 2. 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 0.74 percentage points, showing that our ideal index is robust to the exact composition of the chosen set and providing additional validation to our procedure to select chosen products.

Figure 5: 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 0.74.

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 5 presents the average bias across 100 iterations using random 80% subsets. The mean bias is 0.74, identical 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 0.74 (p-value=0.3077). Moreover, the average bias across iterations remains tightly concentrated between 0.7 and 0.8, 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 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 BEA's price index, but it simultaneously increases the downward gap arising from misweighting.

To make this concrete, we focus on the BEA's PCE 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 PCE index actually grows. By eliminating the bias that pushed the official index upward, we inadvertently magnify the remaining misweighting bias that pulls it downward, resulting in a larger net error than before.

4.1 Correcting Prices for New Varieties

We begin by isolating the impact of new varieties on the BEA price index, setting aside quality growth for now. To do so, we construct two alternative BEA price indices. The first, P^{NV_1}, corrects only for net entry of new product categories; the second, P^{NV_2}, also adjusts for new varieties within existing categories.

Specifically, define

$$d\ln P^{NV_1} \;\equiv\; \frac{1}{\sigma-1}\,d\ln S_{\mathbb{I}} \;+\; \sum_{j\in\mathbb{I}} \omega_j^{\text{BEA}}\; d\ln P_j^{\text{BEA}},$$

which removes the new-product bias but leaves the within-category variety bias and the expenditure-share bias intact. Comparing P^{NV_1} to the true index P yields

$$d\ln\!\left(\tfrac{P^{NV_1}}{P}\right) = -\sum_{j\in\mathbb{I}} \omega_j^{\text{BEA}} \frac{1}{\eta_j-1}\,d\ln X_{\mathbb{I}_j} \;+\; \sum_{j\in\mathbb{I}} (\omega_j^{\text{BEA}}-\omega_j)\,d\ln P_j^{\text{BEA}}. \tag{9}$$

The new-product term drops out, leaving only the variety-bias and the expenditure-share bias. Alternatively, we can apply the Feenstra correction for within-category variety entry as well. Define

$$d\ln P^{NV_2} \;\equiv\; \frac{1}{\sigma-1}\,d\ln S_{\mathbb{I}} \;+\; \sum_{j\in\mathbb{I}} \omega_j^{\text{BEA}}\Big(d\ln P_j^{\text{BEA}} \;+\; \frac{1}{\eta_j-1}\,d\ln X_{\mathbb{I}_j}\Big).$$

In this case, the only remaining bias is from misweighted shares:

$$d\ln\!\left(\tfrac{P^{NV_2}}{P}\right) = \sum_{j\in\mathbb{I}} (\omega_j^{\text{BEA}}-\omega_j)\,d\ln P_j^{\text{BEA}}. \tag{10}$$

Thus, by approximating $d\ln S_{\mathbb{I}}$ and $d\ln X_{\mathbb{I}_j}$ in the data, we can quantify how each partial correction—first for new product entry, then for new-variety entry—affects the overall bias

in the BEA index.

Although the BEA does not directly measure the contributions of new products or new varieties, we can approximate these biases using establishment-level employment data following Aghion et al. (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. Accordingly, the change in the employment share of incumbent establishments, $d \ln S_{\mathbb{I}}$, measures the evolution of these incumbents.

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

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

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 3 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 BEA index and the ideal index even more negative. The final line shows the aggregate bias after applying the full Feenstra correction (our P^{NV_2} in Equation 10): the variety-adjusted inflation rate is now 0.43 percentage points per year below the ideal rate. Although this gap is larger than the unadjusted BEA bias, it is smaller than the bias that corrects only for within-category varieties.

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.

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

Table 3: Bias after Variety Correction

	$\sigma = 3$	$\sigma = 4$ (benchmark)	$\sigma = 5$
Original bias, $d \ln \left(\frac{P^{BEA}}{P} \right)$	-0.99	-0.41	-0.38
Bias after variety correction of BEA product, $d \ln \left(\frac{P^{NV_1}}{P} \right)$	-1.05	-0.46	-0.41
Bias after variety correction *within* BEA product, $d \ln \left(\frac{P^{NV_2}}{P} \right)$	-1.03	-0.43	-0.38

Notes: The table reports the average bias under three measures—original bias (Equation 5), bias after correcting for new products only (Equation 9), and bias after correcting for new varieties within each category (Equation 10)—over January 1960 to December 2019 for various values of the elasticity of substitution. The original average biases match those in Table 1.

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 4 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 (–0.08 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 4 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 –0.08 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 –0.15 log 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.

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

[19] Appendix Section C lists these quality-adjusted categories.

Table 4: Residual Demand Growth: Quality vs Non-Quality Adjusted Products

	Residual Demand Growth
all: 1959-2019	
All (212 product categories)	-0.02
Quality adjusted (28; 27% of exp.)	-0.08
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.)	-0.04
Non-quality adjusted (179; 70% of exp.)	-0.01
post-Boskin Commission: 1996-2019	
All (212 product categories)	-0.04
Quality adjusted (28; 27% of exp.)	-0.15
Non-quality adjusted (184; 73% of exp.)	-0.02

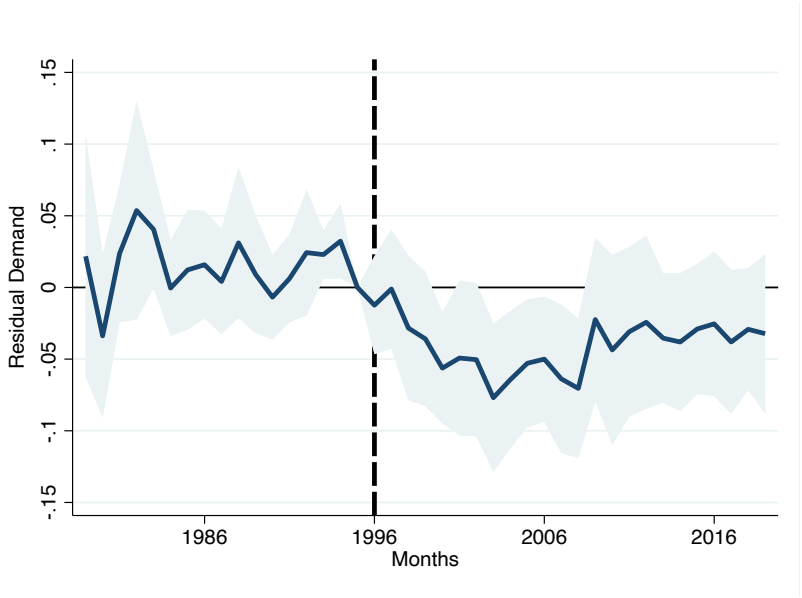
Notes: The table reports average residual demand growth over three periods: (i) 1959–2019, (ii) 1959–1995 (pre-Boskin), and (iii) 1996–2019 (post-Boskin). Numbers in parentheses indicate the number of product categories and their aggregate expenditure shares for quality-adjusted and non-quality-adjusted groups. Residual demand is calculated from Equation 6 for each month, and we report the period average.

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:

$$Y_{it} = \alpha \; + \; \sum_{k=-\infty}^{\infty} \gamma_k \, \mathbb{1}\{K_{it} = k\} \; + \; \theta_i \; + \; \lambda_t \; + \; \zeta \, X_{it} \; + \; \epsilon_{it}, \tag{11}$$

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 6: Event Study: Quality-Adjusted Categories



Notes: The graph shows the evolution of the residual demand growth before and after the Boskin Comission. The panels plot the coefficients of γ_k after estimating Equation 11. 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 6 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.

The combined impact of BEA quality adjustments and the resulting increase in residual demand is that the aggregate bias in the BEA's inflation rate has grown. Specifically, between 1959 and 1995, the BEA's measured inflation was 0.35 percentage points per year below the ideal rate. In the 1996–2019 period, this gap widened to 0.51 percentage points per year.

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-constant residual demand growth, reflecting both quality- and variety-related distortions. By selecting the subset of categories with statistically negative and stable pric-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 "chosen" index reveals that official PCE inflation has, on average, understated the true cost-of-living increase by 0.41 percentage points per year. This understatement is especially pronounced in categories with rapid product turnover and steep price declines, where missing varieties are most important.

We validate these findings using rich NielsenIQ scanner data at the barcode level. There, our procedure selects nearly 19,000 barcodes–about 2.7 percent of all varieties–and delivers similarly large negative biases (–0.74 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 ideal index, confirming the robustness of our approach.

Our results carry two immediate policy lessons. 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 provides 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.

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.

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.

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.

ONLINE APPENDIX

A 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 A.I lists those 27 categories along with year of introduction.

Table A.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 elsewere 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.

B List of Chosen Product Categories

Our of 212 product categories, the 20 chosen product categories are listed below.

Table B.I: List of 20 Chosen Product Categories

Product Category	Estimated Coefficient	Standard Error
Child care	-0.85	0.12
Cable, satellite, and other live television services	-0.33	0.08
Laundry and drycleaning services	-1.36	0.19
Electric appliances for personal care	-1.17	0.08
Repair of household appliances	-1.05	0.16
Employment agency services	-0.82	0.27
Group housing (23)	-0.99	0.20
Package tours	-0.72	0.17
Clothing repair, rental, and alterations	-0.79	0.14
Live entertainment, excluding sports	-0.65	0.10
Luggage and similar personal items (part of 119)	-0.21	0.05
Standard clothing issued to military personnel	-2.24	0.53
Miscellaneous personal care services	-1.05	0.14
Owner-occupied mobile homes	-0.66	0.13
Hair, dental, shaving, and miscellaneous personal care products except electrical products	-0.69	0.07
Other road transportation service	-0.44	0.07
Social advocacy and civic and social organizations	-1.42	0.18
Repair and hire of footwear	-1.02	0.14
Small electric household appliances	-0.39	0.13
Tenant-occupied mobile homes	-1.05	0.14

Notes: The table reports the list of 20 chosen product categories in the BEA data and its estimated coefficient and standard error.

C 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 C.I are identified to be quality adjusted.

Table C.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.

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

D "Residual" Product Demand: Derivation

In this section, we provide a derivation of Equation 6. We begin by writing the expenditure share of product j:

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

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

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

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

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

E NielsenIQ Consumer Panel

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 scanner data. We use moment conditions of the double-differenced residuals in demand and supply with the GMM estimation approach.

Table E.II: Share of Chosen Barcodes by Department

Department	Share of Chosen Barcodes (%)
Health and Beauty Aids	2.01
Dry Grocery	3.10
Frozen Foods	3.43
Dairy	3.81
Deli	3.62
Packaged Meat	3.80
Fresh Produce	3.72
Non-Food Grocery	2.96
Alcohol	2.34
General Merchandise	1.96

Notes: The table reports the shares of chosen barcodes by ten departments in the scanner data.

Table E.III: Percentage Points Difference in Price with Chosen vs. All Incumbent Products by Department

Department	Percentage Points Difference
Health and Beauty Aids	1.01
Dry Grocery	0.49
Frozen Foods	0.83
Dairy	0.06
Deli	0.44
Packaged Meat	0.47
Fresh Produce	0.11
Non-Food Grocery	0.90
Alcohol	0.20
General Merchandise	2.79

Notes: Table reports the gap in one year changes in the price index calculated with chosen vs. all incumbent products.

E.1 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):

$$\Delta^{\underline{u},t}\ln S_{ig} = (1-\sigma_g)\Delta^{\underline{u},t}\ln P_{ig} + \iota_{ig} \tag{12}$$

$$\Delta^{\underline{u},t}\ln P_{ig} = \frac{\delta_g}{1+\delta_g}\Delta^{\underline{u},t}\ln S_{ig} + \kappa_{ig} \tag{13}$$

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 12 and 13 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:

$$G(\beta_g) = E_T[\nu_{ig}(\beta_g)] = 0 \tag{14}$$

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:

$$\hat{\beta}_g = \arg\min_{\beta_g \in B} G^*(\beta_g)' W G^*(\beta_g) \quad \forall\, i \in \omega_b \tag{15}$$

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 Nielsen Homescan 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.