Adam Smith's Linen Shirt*

David Argente†
Pennsylvania State University

Chang-Tai Hsieh‡
University of Chicago

Munseob Lee§
University of California San Diego

March 2022

Abstract

We show that the perceived quality of products, even when defined at the finest level of aggregation, changes over time. We develop a methodology to choose products where quality is likely to be constant. We show that this sample of "constant quality" products can be used to construct an aggregate price index that accounts for changes in quality. This index is the product of the conventional price index of the "constant quality" products and the expenditure share of these products. A similar procedure can also be applied to the price index of more aggregated product categories, where the price index may not fully incorporate new variety growth and quality changes within the product category. We apply this procedure to barcode level data of consumer packaged goods and the BEA's official data on the 212 product categories in the CPI. Our procedure generates annual inflation rates that are 0.74 percentage points higher for consumer packaged goods between 2004 and 2019 and 0.36 percentage points higher for the entire CPI between 1959 and 2019.

JEL Classification Numbers: D11, D12, E01, E31
Keywords: Price Index, Cost of Living, Inflation

*We thank Alberto Cavallo and Marc-Andreas Muendler 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 Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

†Email: dargente@psu.edu. Address: 403 Kern Building, University Park, PA 16801.
‡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.

> A linen shirt, for example, is, strictly speaking, not a necessary of life. The Greeks and Romans lived, I suppose, very comfortably though they had no linen. But in the present times, through the greater part of Europe, a creditable day-labourer would be ashamed to appear in public without a linen shirt, the want of which would be supposed to denote that disgraceful degree of poverty which, it is presumed, nobody can well fall into without extreme bad conduct.
>
> *Adam Smith*

The central challenge in measuring price indices is that the mix and quality of products changes 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. The official US consumer price index (CPI) incorporates the effect of quality change for some products (such as computers) using hedonic techniques, but does not make this adjustment for most of the products in the CPI. The Bureau of Economic Analysis (BEA) also does not account for the creation of new varieties in the CPI, except for the occasional creation of new product categories. Out of the 212 product categories in the CPI between 1959 and 2019, only 28 of these product categories were introduced after 1959. For example, the BEA introduced a new product category for "audio streaming and radio services (including satellite radio)" in January 2007, long after these services have been available.[1]

To address this challenge, researchers have turned to data at finer and finer level of detail to distinguish products that are identical over time from those that are not. For example, the recent body of work using detailed data of consumer packaged goods at the 12-digit barcode level assumes that products with the same barcode are identical over time or across space whereas a new barcode represents a "new" product. See, for example, Broda and Weinstein (2010), Handbury and Weinstein (2015), Argente and Lee (2021), and Argente, Hsieh and Lee (forthcoming).

However, Adam Smith's example of the linen shirt suggests that using finely disaggregated data does not solve the measurement problem. First, the perceived quality of a product can change even when its physical characteristics is the same. To paraphrase Adam Smith, no self-respecting worker in 18th century Europe would wear a tunic, even if the tunic is, in physical terms, exactly the same as those worn in ancient Greece and Rome. The change in the price of a tunic does not measure the change in its effective price for the simple reason that the same tunic provides very different utility in 18th century Europe compared to Greece and Rome.

Second, the example of the linen shirt also suggests that products that appears to be "new" may not actually be new. In Smith's example, a European linen shirt did not exist

[1]Sirius XM satellite radio has been available since September 2001.

in Greece and Rome, and appears to be quite different from a tunic, but provides exactly the same utility to 18th century Europeans as a tunic in ancient Greece and Rome. The implication for price measurement is that what we measure as the gain from new varieties may instead reflect the change in quality and the unit price of an existing product that was mistakenly identified as a new good.

This paper proposes a new methodology to measure the price index when we do not know for sure whether a new good is in fact new or whether an existing good delivers the same utility over time. This method does not require that we isolate new goods from old goods, but it does require that we identify a set of products where, on average, quality is constant. There are two key advantages of our procedure. First, we do not need to identify all the products that are comparable over time nor which products are new varieties. Second, we do not need to know the change in quality or prices for the group of products that we are not sure are comparable over time.

The key is that we identify a set of products where average quality change is zero. Once we have identified this bundle of "constant-quality" products, the change in the aggregate price index is given by the product of two terms: 1) the weighted average of the change in prices of the products in this "constant-quality" bundle and; 2) the the change in the market share of the "constant-quality" bundle, adjusted by a function of the price elasticity of demand. This is exactly Feenstra (1994)'s formula, with two differences. First, the weighted average of the price change of the "constant-quality" bundle is not the price change of *all* products with constant quality, simply because many products that are comparable over time are likely not included in the bundle. Second, the "new goods" term is not the welfare gain from new goods because it also includes the effect of prices with constant quality not included in the "constant-quality" bundle. However, the product of the two terms is an unbiased estimate of the *net* effect of prices changes of products with constant quality and the welfare gain from new products, even if we can not separately identify these two effects.[2]

We implement our methodology with data on prices and expenditures at the barcode level, the finest level of disaggregation possible, for a large sample of consumer packaged goods. Using 16 years of data at quarterly frequency and defining products as barcodes, we estimate the correlation between expenditure shares and prices for each barcode in the data. We find that even at the barcode level, changes in prices and expenditures are positively and significantly correlated for 58% of all barcodes, which suggests that quality change is likely to be present for these barcodes.

[2]Under the CES unified price index by Redding and Weinstein (2020), the price index is given by the product of two terms: 1) a combination of geometric average of the change in prices and the change in shares weighted by the elasticity of substitution and; 2) the same variety correction term. The same argument holds if some of products that are classified as common do not actually have the same quality over time.

We then develop a strategy to select products where quality is more likely to be constant. In particular, we focus on products that have a negative and significant relationship between expenditure shares and prices over time. We then calculate the aggregate price index as the product of the price index of the "constant quality" products and the expenditure share of these products adjusted by the demand elasticity. Under our assumption that average quality is constant among the products where the covariance between price and expenditure shares is negative, the product of these two terms yields an estimate of the aggregate price index that accounts for quality change for all the products in the data.

We then calculate the change in the price index using our procedure and also using a conventional procedure where we assume in the scanner data that goods with common barcodes are identical over time and goods with new barcodes are new products. Our estimate of the inflation rate of consumer packaged goods using this procedure are, on average, 0.74 percentage points *higher* over the 2004-2019 period than the inflation rate estimated following conventional methods.

We also show how a similar procedure can be used with more aggregated price data, where the issue is whether the price index fully accounts for variety growth as well as quality change within the product category. We apply this procedure to the BEA's data on prices and expenditures on the roughly 200 product categories in the consumer price index. We show that changes in prices and expenditure shares are positively and significantly correlated for 68% of the products in the CPI. This suggests that the BEA's price indices for these product categories does not fully capture the effect of quality changes or variety growth in these product categories.

We then show that the key is to identify a set of product categories where the BEA's price index is likely to fully account for variety and quality growth within the product category. The aggregate price index, one that properly adjusts for variety and quality growth in *all* the product categories, can then be measured as the product of the price index of the bundle of product groups where the price index incorporates variety and quality growth and the expenditure share of this bundle. We implement this procedure by choosing product categories in the BEA's data where the correlation between price and expenditure changes are negative. Our estimates suggest that the official CPI *understates* the inflation rate by 0.36 percentage points per year between 1959 and 2019.

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 because they do not take variety growth into account. A second body of work by Nordhaus (1997), Bils and Klenow (2001), and Bils (2009) measure the bias because standard prices indices do not adjust for quality growth. We have two messages relative to this literature. First, and this is presumably what Adam Smith would say, when there is quality growth, it is not possible to separately identify variety growth from quality change. Second, what we can do instead is to measure the *net* effect of these two forces by identifying a set of products where quality change and variety growth (if the products are aggregated) are less likely to be an issue.

We highlight two papers in particular that we build on. First, we build on Bils (2009)'s idea that the change in a product's expenditure share can be used to infer its true price. In our case, we use this idea to to infer the combination of the new variety growth and quality change for *all* products in the economy. Second, Redding and Weinstein (2020) also show that quality is not likely to be constant for given barcode. They develop an alternative price index that assumes that the geometric mean of quality change in the incumbent barcodes is zero. Although our evidence suggests that their assumption is not likely to hold for all products with a constant barcodes over a period of time, our basic idea is that we can measure the aggregate price index once we identify a set of products for which their assumption of no quality growth is more likely to hold.

The rest of the paper is organized as follows. Section 1 describes the data. Section 2 develops a simple conceptual framework under homothetic CES preferences. In Section 3 we show that the quality of most products is not constant over time and describe our procedure to select products likely to have that property. Section 4 estimates parameters, presents our results, and compares the changes of our price index to other indexes in the literature. The last section concludes.

1 Data

We use two datasets. The first dataset is the U.S Bureau of Economic Analysis (BEA)'s official statistics of prices and expenditures of 212 product categories in personal consumption expenditures. We use two NIPA tables with data on prices and expenditures of these 212 product categories: "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)."[3] These product categories account for about two-thirds of domestic final spending.

Our sample consists of monthly data on prices and expenditures of these product categories from 1959 to 2019, and we compute changes in prices and expenditures of each product category over 12 months. Out of the 212 products, 182 products are covered contin-

[3] Details on the methodology used by the BEA to prepare estimates of PCE can be found in Chapter 5 of the NIPA handbook.

uously 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 28 categories were introduced after 1959 and have stayed until December 2019. For example, 3 products were introduced in January 1977: (i) "Video discs, tapes, and permanent digital downloads," (ii) "Personal computers/tablets and peripheral equipment," and (iii) "Computer software and accessories." Most recently, 2 products were introduced in January 2007: (i) "Video streaming and rental," and (ii) "Audio streaming and radio services (including satellite radio)."

The second data we use is Nielsen's Consumer Panel data. This data tracks the shopping behavior of 40,000 to 60,000 households in the US. Each household uses in-home scanners to record their purchases. The data contains slightly under one million distinct 12-digit barcodes. A barcode is uniquely assigned to each specific good available in stores. Barcodes were created so retail outlets could determine prices and inventory accurately and to improve transactions along the supply chain (Basker and Simcoe, 2021). For this reason, barcodes are by design unique to every variety: changes in any attribute of a good (e.g., form, size, package, formula) result in a new barcode. Producers have a strong incentive to purchase barcodes for all products that have more than a trivial amount of sales because the codes are inexpensive, and they allow sellers to access stores with scanners.

Each barcode is classified into one of 104 product groups. For example, a 31 oz bag of Tide Pods (UPC 037000930389) is mapped to product group "Detergent." The product groups in the Nielsen data are roughly at the same level of aggregation as the product groups in the BEA's data. Nielsen constructs projection weights that make the sample representative of the US urban population. We use these weights to calculate total aggregate expenditures on each barcode. We follow Redding and Weinstein (2020) and aggregate prices and expenditures of each barcode to a quarterly frequency. Four-quarter differences are then computed by comparing values for a quarter of each year relative to the same quarter of the previous year.

2 Conceptual Framework

In this section, we derive a price index for homothetic CES preferences with two nests (i.e. across barcodes and across products).[4] The utility function is:

$$\mathbb{U} = \left(\sum_{g} \left(\sum_{i \in g} \left(\phi_{ig} C_{ig} \right)^{\frac{\sigma_g - 1}{\sigma_g}} \right)^{\frac{\sigma_g}{\sigma_g - 1} \frac{\theta - 1}{\theta}} \right)^{\frac{\theta}{\theta - 1}}$$

where g denotes a product group, i denotes products within a product group, θ and σ_g denote the elasticity of substitution across and within product groups, and ϕ_{ig} is the quality of barcode i in product group g. We assume i corresponds to a barcode in the Nielsen data and g corresponds to a product group in the Nielsen and a product category in the BEA data.[5]

The change in the aggregate price index is given by:

$$d \ln P = \sum_{g} \omega_g \left(\frac{1}{\sigma_g - 1} d \ln S_{\mathbb{I}_g} + \sum_{i \in \mathbb{I}_g} \omega_{ig} \, d \ln \frac{P_{ig}}{\phi_{ig}} \right) \tag{1}$$

where $\mathbb{I}_g$ denotes the set of incumbent barcodes in a product group g, ω_g is the Sato-Vartia weight of a product group g and ω_{ig} is the Sato-Vartia weight of barcode i within the set of incumbent barcodes in a product group g, $S_{\mathbb{I}_g}$ is the revenue share of the incumbent barcodes of a product group g, and P_{ig} is the unit price of barcode i in a product group g. Equation 1 is the well-known formula for the price index by Feenstra (1994). The first term is the change in the price index that comes from the net introduction of new barcodes. The second term is the Sato-Vartia weighted average of the change in the *quality-adjusted* price P_{ig}/ϕ_{ig} of the set of incumbent barcodes.

We now discuss how to measure the aggregate price index in equation 1. First, we discuss the measurement exercise with barcode level data. Second, we discuss how to measure the aggregate price index with the aggregated product-group level data provided by the BEA.

[4] Homothetic CES preferences are prominent across several fields including macroeconomics and international trade. Furthermore, Argente and Lee (2021) and Redding and Weinstein (2020) show using barcode data that the Sato-Vartia index derived using CES preferences generates similar changes in the cost of living as superlative indexes that are exact for flexible functional forms, such as the Fisher and Tornqvist indexes.

[5] We assume stable preferences over time so there is a well defined price index over time. Baqaee and Burstein (2021) compute the price index holding preferences fixed at those for consumers at a given point in time.

2.1 Barcode Level Data

The main empirical problem with implementing equation 1 with barcode level data is that what we observe is the unit price of the barcode P_{ig} and not the quality-adjusted price P_{ig}/ϕ_{ig}. In Adam Smith's example, we do not directly observe the change in the quality of a roman tunic over time. The bias in the price index when we use the data on unit prices at the barcode level is

$$-\sum_g \omega_g \, d\ln\phi_g.$$

where $d\ln\phi_g \equiv \sum_{i\in\mathbb{I}_g}\omega_{ig}\, d\ln\phi_{ig}$ is the Sato-Vartia weighted average of quality growth among the set of incumbent barcodes in product group g.

Now suppose that for each product group g there is a subset of barcodes $\mathbb{C}_g$ within the group of incumbent barcodes where the Sato-Vartia weighted average of the quality change is zero:

$$\sum_{i\in\mathbb{C}_g}\omega_{ig}\, d\ln\phi_{ig} = 0.$$

The change in the aggregate price index is then given by:

$$d\ln P = \sum_g \omega_g \left(\frac{1}{\sigma_g - 1} d\ln S_{\mathbb{C}_g} + \sum_{i\in\mathbb{C}_g}\omega_{ig}\, d\ln P_{ig} \right). \tag{2}$$

Comparing with equation 1, the first term in equation 2 is the change in the revenue share of the set $\mathbb{C}_g$, and not the set of all incumbent barcodes $\mathbb{I}_g$, and the second term is the Sato-Vartia weighted average of the *unit price* of the set of barcodes in $\mathbb{C}_g$, and not the average of the *quality-adjusted* price of all incumbent barcodes.

We do not need to know the change in quality of the set $\mathbb{C}_g$ for the simple reason that they aggregate to zero. We also do not need to observe prices or qualities of the other incumbent barcodes because the effect of the change in the quality-adjusted price of these barcodes is captured in the first term in equation 2. To see this, note that $d\ln S_{\mathbb{C}_g} = d\ln S_{\mathbb{I}_g} + d\ln S_{\mathbb{D}_g}$ where $\mathbb{D}_g$ denotes the set of incumbent barcodes not in $\mathbb{C}_g$. Furthermore, it is easy to show that $\frac{1}{\sigma_g-1}\, d\ln S_{\mathbb{D}_g}$ is equal to the Sato-Vartia average of the change in the quality-adjusted price of the barcodes not in set $\mathbb{C}_g$. The "new variety" term in equation 2 thus captures the change in the price index from barcodes where the change in average quality is likely to be non-zero. Of course the effect on this term depends on whether the change in the quality-adjusted price is positive or negative.

It is also possible that many of the barcodes we deem as new could be the same as another seemingly unrelated barcode in the past. Harkening back to Adam Smith's example, suppose that Roman tunics disappeared when Europeans started to wear linen shirts, and tunics provide the same utility in Roman days as linen shirts in 19th century Europe. In this case, the difference between the revenue share of linen shirts and the revenue share of Roman tunics reflects the difference in the quality-adjusted price of a linen shirt vs. a tunic, where the two should be regarded as the same good. That is, the new variety term also captures the effect of changes in the quality-adjusted price of comparable barcodes that are misidentified as new barcodes.

In sum, the "new variety" term in equation 2 captures three things: (1) the change in the quality-adjusted price of incumbent barcodes where we are not sure that the average quality change is constant; (2) the change in the quality-adjusted price of barcodes that we did not identify as the same barcode and; (3) entirely brand new barcodes. So while the new variety term can not be interpreted as the welfare gain from new barcodes, the sum of the "new variety" and the "incumbent product" price terms in equation 2 yields an unbiased estimate of the aggregate price index, including the effect of varieties that are actually new and quality change for all the products.

2.2 BEA product-category level data

Now suppose each of the 212 products in the CPI corresponds to a product category indexed by g, and all we observe is the BEA's price index and expenditure of the product category g. The key question is whether the BEA's price index captures variety and quality growth within the product category. Specifically, suppose that the BEA's price index P_g is given by:

$$d\ln P_g \equiv \sum_{i\in\mathbb{I}_g} \omega_{ig}\, d\ln P_{ig} \;+\; (1-\lambda_g)\left(\frac{1}{\sigma_g - 1} d\ln S_{\mathbb{I}_g} + d\ln\phi_g\right) \tag{3}$$

The first term is the Sato-Vartia average of the change in unit prices of the incumbent products in product-category g; the second term measures the extent to which the BEA's price index captures variety and quality change within the product category. Equation 3 assumes the BEA accurately measures the unit price of incumbent varieties in each product category, but possibly understates the effect of new varieties and quality changes within the product category. $\lambda_g = 0$ denotes no mismeasurement; $\lambda_g = 1$ denotes the other extreme where the BEA makes no adjustment whatsoever for quality change and new varieties.

The resulting bias in the CPI is thus given by

$$-\sum_{g}\omega_g\,\lambda_g\,\left(\frac{1}{\sigma_g-1}d\ln S_{\mathbb{I}_g}+d\ln\phi_g\right).$$

The first term is the bias from missing variety growth. The second term is the bias from not fully capturing changes in quality among the set of incumbent products. Clearly, if $\lambda_g = 0$ then there is no bias in the aggregate price index computed from the BEA's prices at the product category level.

The issue is that we do not know whether the BEA properly accounts for new varieties and quality changes. For some product categories, such as computers, it employs hedonic techniques but for most product categories it does not. And aside from occasionally splitting product categories and creating new product categories, the BEA does not account for new varieties when computing price indices.

However, suppose that we can identify a set $\mathbb{C}$ of product categories where the CPI fully captures the effect of variety and quality growth, or $\lambda_g = 0$.[6] Given the set $\mathbb{C}$, the change in the aggregate price index can be measured as

$$d\ln P=\frac{1}{\theta-1}d\ln S_{\mathbb{C}}\;+\sum_{g\in\mathbb{C}}\omega_g P_g\,. \tag{4}$$

where P_g is the price index at the product-category level given by equation 3 and $S_{\mathbb{C}}$ is the revenue share of the product categories in $\mathbb{C}$. The price index, one that properly accounts for quality change and variety growth in all the product categories, is the Sato-Vartia average of the price index at the product category level of the set of product categories in the set $\mathbb{C}$ and the change in the expenditure share of these product categories, where the latter is adjusted by the elasticity of substitution between product categories θ.

The intuition behind equation 4 is as follows. The second term measures the effect of price changes among the product categories in the set $\mathbb{C}$. We do not need to measure quality change and new varieties in these product group because the BEA's price index already incorporates their effect. The first term captures the net effect of new products, price changes, and quality changes in the product categories where the CPI does not fully capture the effect of variety and quality growth, or $\lambda_g > 0$. The effect of new varieties in these product categories shows up in this term, but it should be clear that one should not interpret the "new variety" term as only the effect of new varieties as it also includes the effect of price and quality changes. In addition, by assumption, the effect of variety growth in the set of product groups in $\mathbb{C}$

[6] Or where the net effect of new varieties and quality change is zero, or $\frac{1}{\sigma_g-1}d\ln S_{\mathbb{I}_g}\;+\;d\ln\phi_g\;\approx\;0$.

shows up in the second term in equation 4. However, as in equation 2 in the case of the barcode level data, the sum of the two terms in equation 4 yields an accurate estimate of the quality change and variety growth within *all* product groups in the CPI.

In sum, to measure the price index using the detailed data at the barcode level, we need identify the set of products $\mathbb{C}_g$ where the average quality change is zero. With these sets of products, we then measure the aggregate price index from equation 2. Likewise, to measure the price index from the BEA's product-category level data, we need to identify a set of product categories where the BEA's price index fully captures the effect of variety and quality growth. With this set of product category, we then measure the aggregate price index from equation 4.

3 Quality of a Variety Over Time

In this section, we explore whether the quality of barcode is constant over time, and whether the price index of a BEA product category incorporates the effect of quality changes and net products within the product. We first derive a product-level regression equation from the conceptual framework in Section 2 and implement it in the two datasets. We then use the results from this regression to choose a set of products $\mathbb{C}_g$ in the barcode level data and a set of product-categories $\mathbb{C}$ in the BEA's data.

3.1 Product-level Regression Equation

We start with regressions at the barcode level. From equation 2, the change in the expenditure share on a barcode is

$$d \ln S_{ig} = -(\sigma_g - 1) d \ln \tilde{P}_{ig} + (\sigma_g - 1)\, d \ln \phi_{ig}$$

where S_{ig} is the expenditure share of barcode i in product group g, $\tilde{P}_{ig}$ is the unit price of barcode i relative to the aggregate price of product group g. The expenditure share of a barcode is decreasing in its relative price and increasing in its quality.

We can implement this equation in the barcode level data barcode-by-barcode using time-series variation as follows:

$$d \ln S_{ig} = \beta_{ig}\, d \ln \tilde{P}_{ig} + \epsilon_{ig}. \tag{5}$$

Note that if $d \ln \phi_{ig} \approx 0$, the coefficient β_{ig} is unambiguously negative given our assumption that $\sigma_g > 1$. In this case β_{ig} will be biased toward zero, and possibly even positive, as follows:

$$\beta_{ig} = -(\sigma_g - 1) + \frac{\mathrm{Cov}(d\ln \tilde{P}_{ig}, (\sigma_g - 1) d\ln \phi_{ig})}{\mathrm{Var}(d\ln \tilde{P}_{ig})}.$$

The bias in the regression coefficient depends on the correlation of the price and the quality change at the barcode level.

Turning to the BEA's product-category level data, the change in the expenditure share of product-category g is:

$$d\ln S_g = -(\theta - 1) d\ln \tilde{P}_g + (\theta - 1)\,\lambda_g \left(d\ln \phi_g + \frac{1}{\sigma_g - 1} d\ln S_{\mathbb{I}_g} \right)$$

where $\tilde{P}_g \equiv P_g/P$ is the price of a product g relative to the aggregate price index and $d\ln \phi_g$ is the average quality change among the incumbent products in product category g. The expenditure share of a product category is decreasing in the relative price of the product category and in λ_g.

We can implement this equation in the BEA's data product-by-product using time-series variation as follows:

$$d\ln S_g = \beta_g \, d\ln \tilde{P}_g + \epsilon_g. \tag{6}$$

Note that if $\lambda_g = 0$ or if the sum of the new variety and quality change terms is zero, the coefficient β_g is unambiguously negative given our assumption that $\theta > 1$. On the other hand, if $\lambda_g > 0$ then β_g will be biased toward zero or could even be positive as follows:

$$\beta_g = -(\theta - 1) + \frac{\mathrm{Cov}\left(d\ln \tilde{P}_g, (\theta - 1)\,\lambda_g \left[d\ln \phi_g + \frac{1}{\sigma_g - 1} d\ln S_{\mathbb{I}_g} \right] \right)}{\mathrm{Var}(d\ln \tilde{P}_g)}$$

Note that the effect of unmeasured quality growth $\lambda_g d\ln \phi_g$ on the regression coefficient β_g is the same as that of unmeasured variety growth $\lambda_g \frac{1}{\sigma_g - 1} d\ln S_{\mathbb{I}_g}$. At the product-group level, the effect of unmeasured quality change on the price mismeasurement is isomorphic to that of unmeasured new variety growth.

In sum, the coefficient from a regression of the change in expenditure share of a product category on the change in the BEA's price index of the product category tells us whether the unmeasured quality change and new varieties are likely to be be a problem for a given price index. Similarly, the coefficient from a regression of the change in the expenditure share of a barcode on the change in the unit price of the barcode indicates whether the quality of the barcode is likely to have changed. In the next subsection we will use this idea to choose the product categories $\mathbb{C}$ from the BEA data where the BEA's prices are likely to be accurate and the barcodes $\mathbb{C}_g$ in the Nielsen data where quality is likely to have remained constant.

3.2 Results

We begin by implementing equation 5 with the Nielsen's barcode data. Our sample covers all 15 years available of the Nielsen Consumer Panel Data (2004-2019). We use quarterly level data and take four-quarter differences in order to control for seasonality. We focus on barcodes that are present in the data at least two years, in order to have at least 8 observations in each regression. Our total sample consists of 686,282 barcodes; β_{ig} is estimated separately for each of them.

In order to select barcodes less likely to be affected by quality changes and introduction of new varieties, we focus on those whose estimated β_{ig} is negative and statistically different than zero at the 1% level. This is, barcodes whose changes in prices and quantities are more likely to be driven by supply rather than demand changes.

Notice that in order to estimate β_{ig} we need to calculate $\tilde{P}_{ig}$, which depends on the product group price index, P_g. In order to implement equation 5 using an unbiased price index of the product group, we estimate β_{ig} through an iterative procedure. First, we calculate the conventional price index of the product group using the entire set of barcodes available in each product group of the Nielsen data. This step requires an estimate of the elasticity of substitution across barcodes within product groups, which we obtain following the procedure developed by Feenstra (1994) and extended by Broda and Weinstein (2006) and Broda and Weinstein (2010).[7] The average of the elasticity of substitution that we obtain is 6.42. The distribution of elasticities can be found in Table 1. The price index obtained using the full set of barcodes in each product group is our initial guess for P_g. Second, we estimate equation 5 for each barcode in the Nielsen data using a linear regression and select barcodes using our selection criteria (i.e. those whose estimated β_{ig} is negative and statistically different than zero at the 1% level). Third, we calculate a new product group price index using only the set of *chosen* barcodes. And, lastly, we check whether this index is the same as our guess and, otherwise, we use it as a new guess and continue the iterative process until convergence.[8] The outcome of this procedure is both, a set of barcodes that are less likely to be affected by quality changes and introduction of new varieties and an unbiased product group price index that is calculated using these set of *chosen* barcodes, which we discuss in detail in the

[7]The procedure consists of estimating a demand and a supply equation for each barcode in data using information of prices and quantities. Details on the estimation procedure can be found in Appendix A.

[8]Our convergence criteria is that the difference between the guess and the product group price index estimated using only the set of chosen barcodes is less than 0.001 percentage points. The average number of iterations across products is 4.77 (median 5).

next section.

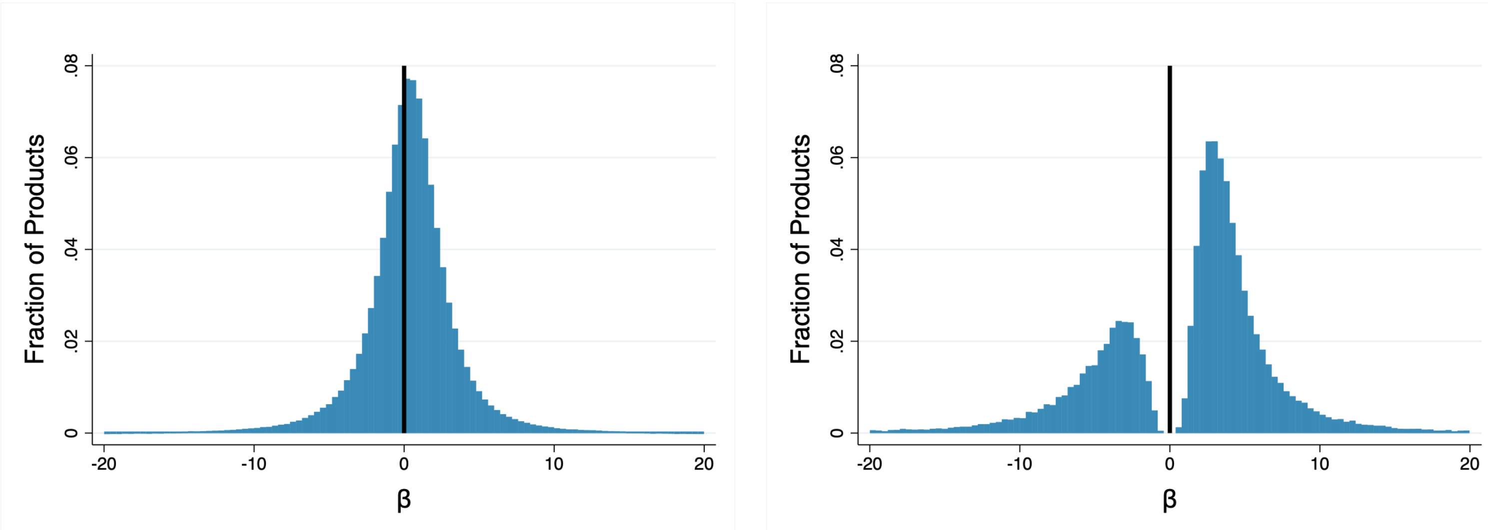
Table 1: Estimated Elasticities of Substitution

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

Panel (a) of Figure 1 shows the distribution of coefficients, β_{ig}. The average estimate for β_{ig} for barcodes that lasted at least 8 quarters in the data is 0.37 on average (median 0.41). In fact, approximately 58% of barcodes have a positive coefficient indicating again the presence of quality changes even at the barcode level.

Figure 1: Product-level Regression Results Using Scanner Data



(a) All Barcodes

(b) Significant at 1%

Notes: The figure shows the distribution of the coefficients β_{ig} estimated using equation 5 for barcodes in the Nielsen Consumer Panel 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). The black vertical line marks products whose coefficient is zero.

Panel (b) in Figure 1 shows the distribution of coefficients that are significant at 1%. For the set of products $\mathbb{C}_g$ where demand shifts are less likely to be prevalent, we choose the barcodes where the coefficients are to the left of the vertical black line. Products that contribute more barcodes to the chosen sample are less likely to be subject to changes in quality. Some examples are: sugar sweeteners, butter and margarine, canned fruit, canned seafood, frozen juices, eggs, milk, cheese, canned drinks, and frozen items such as vegetables, meat, and baked goods. [9]

We implement the same procedure to estimate equation 6 with the BEA's product category data. Our sample covers 61 years available of the BEA data. We use monthly level data and take twelve-month differences to control for seasonality. We focus on products that are present in the data at least two years in order to have at least 24 observations in each regression. Importantly, our iterative procedure is implemented across product categories. In other words, our initial guess is the conventional aggregate price index, P. For this step,

[9] The product groups that contribute a higher share of barcodes to the sample of chosen barcodes are butter and margarine, seafood (canned), sugar sweeteners, eggs, fresh meat, meal/poultry/seafood frozen, paper products, breakfast foods (frozen), milk, pizza/snacks (frozen), detergents, tobacco and accesories, juices (frozen), dough products, juice (canned), prepared foods (frozen), nuts, charcoal/logs accesories, baked good (frozen). Appendix B reports the shares of chosen barcodes in each department.

we again need an estimate of the elasticity of substitution across product categories, which we obtain following a similar procedure as the one used for the Nielsen data. The elasticity of substitution across product categories we obtain is 2.88.

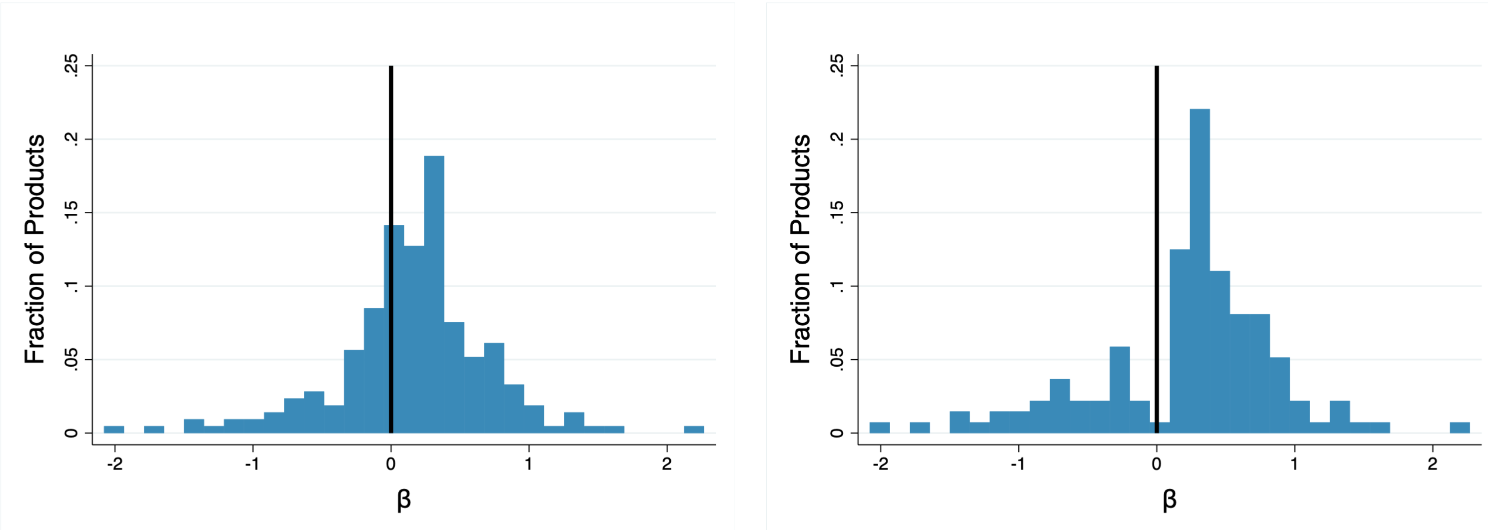
Given an initial guess of P, we estimate equation 6 for each product category. We then select product categories using our selection criteria and, using only the set of *chosen* product categories, we calculate a new aggregate price index. As before, we iterate until convergence using the same convergence criteria used for the Nielsen data.[10]

Panel (a) of Figure 2 shows the distribution of the estimates of β_g. The average estimate for β_g is positive and approximately 0.15 on average (median 0.19) In fact, approximately 68% of products have a positive coefficient indicating the presence of quality shocks and/or new varieties within these products. Panel (b) shows the distribution of coefficients that are significant at 1%. The figure shows that there are more products whose estimated elasticity of demand is positive, again suggestive that quality changes and new varieties for many product categories are not measured in the BEA data. Our sample of *chosen* products $\mathbb{C}$ includes only those to the left of the vertical black line. Examples of *chosen* products are durables such as household cleaning products, personal computers/tablets, and computer software and accesories. These are products whose price index is regularly adjusted for quality changes by the BLS. Other chosen products are child care services, dental services, individual and family services, religious organizations' services to households, and vocational rehabilitation services.[11]

[10]The iterative procedure using the BEA data converges after 5 iterations.

[11]Appendix B reports the list of the 33 products chosen out of 212 products.

Figure 2: Product-level Regression Results Using BEA Data



(a) All Products

(b) Significant at 1%

Notes: The figure shows the distribution of the coefficients β_g estimated using equation 6 for products in the BEA data (1959-2019). Panel (a) includes all products that lasted at least 2 years in the data (212 products). Panel (b) shows the coefficients of products that lasted at least 2 years and whose whose coefficient is significantly different than zero at 1 percent confidence level (34 products). The black vertical line marks products whose coefficient is zero.

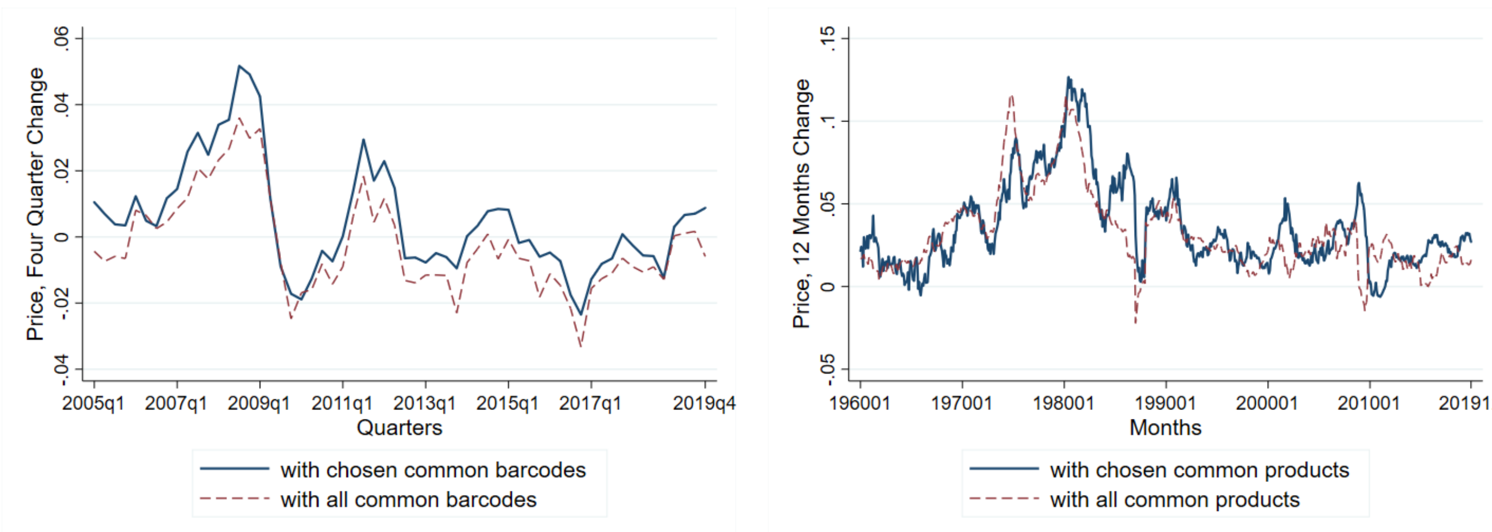
4 Price Index

We now have all the necessary ingredients to estimate the aggregate price index. For the Nielsen data, these are estimates of the elasticity of substitution within product groups σ_g, the choice of barcodes in the set $\mathbb{C}_g$, the price indices for the barcodes in this set, and the expenditure share of the barcodes in the set $\mathbb{C}_g$. For the BEA data, these are estimates of the elasticity of substitution across product categories θ, the choice of product categories in the set $\mathbb{C}$, the BEA's price indices for product categories in this set, and the expenditure share of the products in the set $\mathbb{C}$.

Figure 3 shows the price indexes for the scanner and BEA data sets calculated using equations 2 and 4, respectively. For comparison, we also estimate the price index using all the common products (equation 1), which corresponds to the standard Feenstra-CES index. Panel (a) of Figure 3 shows the price indices calculated from the scanner data from the first quarter of 2004 to the last quarter of 2019. The price index calculated using our procedure from data on prices and the expenditure share of the set of chosen products (the set $\mathbb{C}_g$ is on average higher than the price index with all common barcodes. The mean percentage points difference is 0.74.

Panel (b) shows the price indices calculated from the BEA data from January 1959 to December 2019 (12-month changes). The price index calculated from the price and expenditure share of the chosen common products (the set $\mathbb{C}$) is on average higher than the price index computed using all common products. The mean percentage points difference is 0.36.

Figure 3: Price Index with Chosen vs. All Common barcodes/product categories



(a) Nielsen

(b) BEA

Notes: Panel (a) and (b) show the one year changes in Feenstra-CES price index calculated with chosen (equation 2) vs. all common products (equation 1). Panel (a) uses the scanner data from the first quarter of 2004 to the last quarter of 2019. Panel (b) uses the BEA data from January 1959 to December 2019.

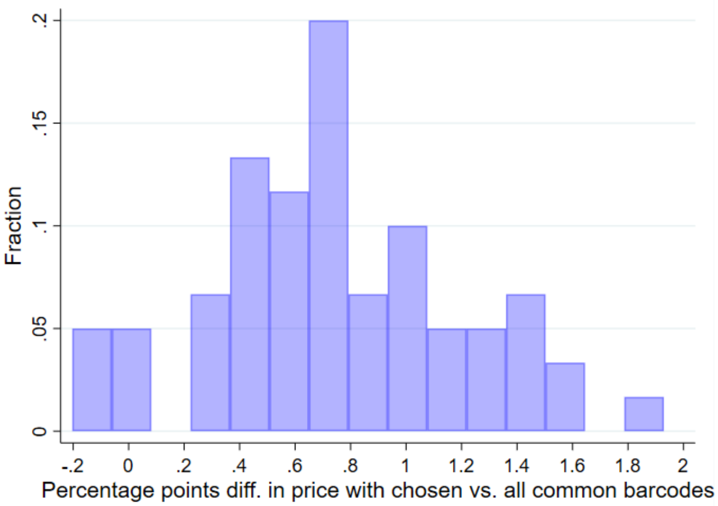
Figure 4 shows histograms of the percentage points difference in the Feenstra-CES price index with chosen vs. all common barcodes/product categories. Panel (a) of Figure 4 uses the scanner data from the first quarter of 2004 to the last quarter of 2019 and plots the gap between the indexes depicted in Panel (a) of Figure 3. The mean percentage points difference is 0.74 with a standard deviation of 0.45. There are only 4 out of 60 quarters when the difference is negative. Overall, the conventional price index with all barcodes underestimates the inflation rate in the consumer-packaged goods sector.[12]

Panel (b) uses the BEA data from January 1959 to December 2021 and plots the gap between prices with chosen vs. all common products depicted in Panel (a) of Figure 3. The mean percentage points difference is 0.36 with a standard deviation of 1.69. There are certain

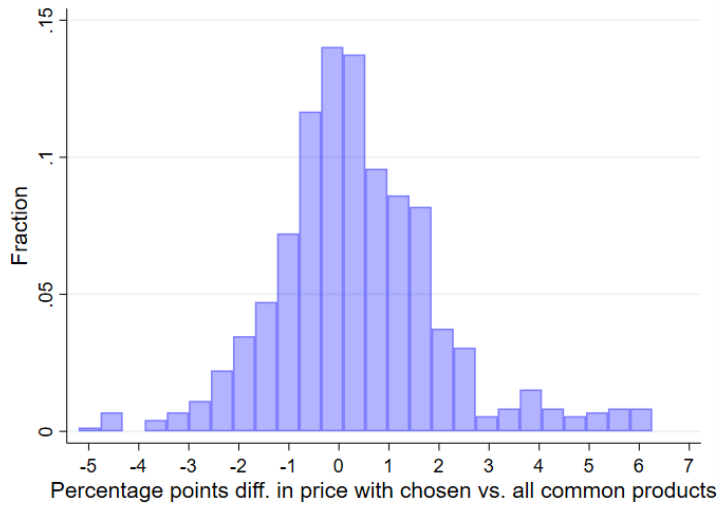
[12]We find no correlation between the percentage points difference and quarterly unemployment rate in this sample.

months when the difference is negative but, on average, the conventional price index with all products underestimates the increase in the cost of living in the US from 1959 to 2019.[13]

Figure 4: Histogram of the Percentage Points Difference: Chosen vs. All Common barcodes/product categories



(a) Nielsen



(b) BEA

Notes: Panel (a) and (b) show the percentage points difference in Feenstra-CES price index with chosen (equation 2) vs. all common products (equation 1). Panel (a) uses the scanner data from the first quarter of 2004 to the last quarter of 2019. Panel (b) uses the BEA data from January 1959 to December 2019.

Recall that the products we excluded from the sets $\mathbb{C}$ and $\mathbb{C}_g$ are those where the regressions of changes in expenditures on prices yields a positive coefficient. At the barcode level, a positive regression coefficient could indicates that quality of the barcode improved or because the quality of the barcode fell, and the quality change was not reflected in the unit price. The bias in conventional estimates of the price index using the barcode level (which ignores the quality change) depends on the prevalence of barcodes where quality improved vs. barcodes where quality fell. If the former exceeds the latter, then standard price indices will overstate the true inflation rate, and vice versa.

Similarly, at the product-group level, a positive regression coefficient could be because the BEA's price index either overstates or understates the sum of quality change and new variety growth. And again, the bias in the BEA's consumer price index depends on the prevalence of product-groups where quality change and new variety growth is understated

[13]The percentage points difference is correlated with monthly unemployment rate with a correlation of 0.1896, which is statistically significant at one percent level. Thus, the conventional price index underestimate the cost of living more during recessions.

vs. product-groups where they are overstated.

Table 2: Share of Products with Positive Regression Coefficients

	Barcode Data	BEA Data
Δ Prices > 0, Δ Share > 0	24.9%	25.4%
Δ Prices < 0, Δ Share < 0	26.0%	35.4%

Note: The table reports the Sato-Vartia share of products where, on average, Δ prices and Δ expenditure share are both positive (row 1) and where Δ prices and Δ expenditure share are both negative (row 2).

Table 2 shows the Sato-Vartia weight of these two types of products in the barcode data (column 1) and in the BEA product-category data (column 2). The first row shows the Sato-Vartia weight of products where the coefficient is positive because both the change in prices and the change in expenditure shares are positive. For the barcode level data, these are the products where unmeasured quality change is likely to be positive. For the BEA product-group level data, these are the product groups with unmeasured variety and and/or quality growth. The prices of these products thus *overstate* the inflation rate because they ignore quality improvement (barcode data) or variety growth and quality improvement (BEA data). The first row shows that these products account for 24.9% of the products in the Nielsen data and 25.4% in the BEA data.

The second row in Table 2 shows the products where the regression coefficient is positive because the change in prices and expenditure shares are both *negative*. These are products where quality change is likely to be *negative* for the barcode data or BEA product groups where varieties are exiting and where product quality falls. These products, where prices *understate* the true inflation rate, account for a larger share of expenditures compared to the products where prices overstate the inflation rate.

For the barcode level data, the share of this second group is 26% compared to 24.9% of the products in the first row. This gap is what is behind the 0.74 percentage point difference between our estimate of the inflation rate of consumer package goods (based on barcode data) and estimates that do not take quality change into account. For the BEA data, the gap is larger. The share of this second group, where inflation is understated, is 35.4% compared to 25.4% of the products in the first row, where inflation is overstated. The result is that our estimate of the CPI that exceeds the official CPI by 0.36 percentage points per year.

5 Conclusion

Measuring changes in the cost of living faces two main challenges: i) the perceived quality of a product can change even when its physical characteristics is the same, and ii) a product that appears to be "new" may not actually be new. These two challenges were perfectly epitomized in Adam Smith's linen shirt example. This paper proposes a new methodology to measure the price index that addresses the issues raised by Adam Smith. Our method does not require that we isolate new goods from old goods, but it does require that we identify a set of products where, on *average*, the quality is constant. The key advantage of our procedure is that this set of products does not need to include *all* the products that are comparable over time. With this set of products, the price index is then given by the product of two terms: 1) the weighted average of the change in prices of the products in this bundle and; 2) the product of the function of the price elasticity of demand and the change in the market share of this bundle. The product of the two terms is an unbiased estimate of the *net* effect of prices changes of products with constant quality and the welfare gain from new products, even if we can not separately identify these two effects.

Using both data underlying the scanner and BEA data on prices and quantities, we first show that the perceived quality of most barcodes/product categories is not constant. We then develop a strategy to select barcodes/product categories that are less likely to suffer from quality changes and use this set of barcodes/product categories to calculate an unbiased price index. We find that conventional price indexes substantially underestimate changes in the cost of living because the set of products that are comparable over time have experienced a decline in perceived quality.

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.

Argente, David O, Chang-Tai Hsieh, and Munseob Lee, "Measuring the Cost of Living in Mexico and the US," *American Economic Journal: Macroeconomics*, forthcoming.

Baqaee, David and Ariel Burstein, "Welfare and Output with Income Effects and Taste Shocks," *National Bureau of Economic Research Working Paper Series*, 2021.

Basker, Emek and Timothy Simcoe, "Upstream, Downstream: Diffusion and Impacts of the Universal Product Code," *Journal of Political Economy*, 2021, *129* (4), 1252–1286.

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.

Handbury, Jessie and David E Weinstein, "Goods prices and availability in cities," *The Review of Economic Studies*, 2015, *82* (1), 258–296.

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.

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 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):

$$\Delta^{\underline{u},t}\ln S_{ig} = (1-\sigma_g)\Delta^{\underline{u},t}\ln P_{ig} + \iota_{ig} \tag{7}$$

$$\Delta^{\underline{u},t}\ln P_{ig} = \frac{\delta_g}{1+\delta_g}\Delta^{\underline{u},t}\ln S_{ig} + \kappa_{ig} \tag{8}$$

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 7 and 8 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:

$$G(\beta_g) = E_T[\nu_{ig}(\beta_g)] = 0 \tag{9}$$

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:

$$\hat{\beta}_g = \arg\min_{\beta_g \in B} G^*(\beta_g)' W G^*(\beta_g) \quad \forall\, i \in \omega_b \tag{10}$$

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 of each country's population. 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.

We follow the same procedure using the BEA data from 1959 to 2019 and estimate elasticity of substitution, σ, across product categories.

B List of Chosen barcodes/product categories

B.1 BEA Data

Our of 212 product categories, the 33 product categories we choose are child care, household cleaning products, personal computers/tablets and peripheral equipment, computer software and accessories, cable, satellite, and other live television services, dental services, electric appliances for personal care, individual and family services excluding Household insurance normal losses, employment agency services, clothing repair, rental, and alterations, live entertainment, excluding sports, luggage and similar personal items, medical expenditures of foreigners, residential mental health and substance abuse, miscellaneous personal care services, nonprescription drugs, food products, not elsewhere classified, owner-occupied mobile homes, specialty outpatient care facilities and health and allied services, hair, dental, shaving, and miscellaneous personal care products except electrical products, other road transportation service, owner-occupied stationary homes, religious organizations' services to households, social advocacy and civic and social organizations, repair and hire of footwear, other social assistance, not elsewere classified, tenant landlord durables, tenant-occupied mobile homes, tenant-occupied stationary homes, labor organization dues, vocational rehabilitation services, tenant-occupied, including landlord durables.

B.2 Nielsen Data

Table B.I reports the shares of chosen barcodes by department. Less barcodes are chosen to have a constant quality from more durable categories, such as "Health and Beauty Aids" and "General Merchandise."

Table B.I: 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.