

Adam Smith's Linen Shirt*

David Argente[†]

Pennsylvania State University

Chang-Tai Hsieh[‡]

University of Chicago

Munseob Lee[§]

University of California San Diego

December 2021

Preliminary and Incomplete

Abstract

We show that the perceived quality of products, even when defined at the finest level of aggregation, changes over time. This contrasts with the standard assumptions used in the construction of price indexes, which require the quality of products to stay constant over time, at least on average. We develop a methodology to select products whose price changes are mainly driven by supply changes. We then show theoretically that this sample can be used to construct an unbiased aggregate price index. This index is the product of the conventional price index and the share of common varieties, and both components fully compensate for movements of each other. We find that previous estimates significantly underestimate the cost of living in the US.

JEL Classification Numbers: D11, D12, E01, E31

Keywords: Price Index, Cost of Living, Inflation

*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: munseoble@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. 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 products 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.¹

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 19th century Europe would wear a tunic, even if the tunic is 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 19th century Europe compared to Greece and Rome.

Second, Adam Smith suggests that product that appears to be “new” may not actually be new. In his example, a European linen shirt did not exist in the days of the Roman empire, and appears to be quite different from a tunic, but provides exactly the same utility to 19th 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 that addresses the issues raised by Adam Smith. 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*, the quality is constant for the group. 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.

¹See [Broda and Weinstein \(2010\)](#), [Handbury and Weinstein \(2015\)](#), [Argente and Lee \(2021\)](#), and [Argente et al. \(forthcoming\)](#).

The key is that we identify a set of products where the average quality is constant. Once we have identified this bundle of 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 bundle and; 2) the product of the function of the price elasticity of demand and the change in the market share of this bundle. This is exactly [Feenstra \(1994\)](#)'s formula, with two differences. First, the weighted average of the price change of the “old” 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, for the simple reason that it also includes the effect of prices with constant quality not included in the bundle of old goods. 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 this procedure using scanner level data on prices and quantities at the barcode level for a large sample of nondurable goods at the quarterly level over 15 years. Using 15 years of data at quarterly frequency and defining products as barcodes, we estimate the correlation between expenditure shares and prices for each product in the data. Our results show that for most products, expenditure shares and prices are positively and significantly correlated, an indication of the presence of quality changes over time at the product level. This is, defining products at finer levels of aggregation does not guarantee constant quality over time.

We then develop a strategy to select products that are less likely to suffer from quality changes. In particular, we focus on products that have a negative and significant relationship between expenditure shares and prices over time. Furthermore, we choose products that lasted at least 2 years in the data to reduced measurement error in our estimation procedure. We then calculate the exact price index using only the set of *chosen* products as common products under the procedure described above. In particular, we calculate the Sato-Vartia price index and the variety correction term developed by [Feenstra \(1994\)](#). Under our assumptions that the set of chosen products have constant quality, the product of these two terms yields an unbiased exact price index, regardless of the size of the set of common products or the magnitudes of each of these two terms separately. In fact, according to the theory, using any subset of products within the chosen set should yield identical exact price indexes. This is, despite the fact that the magnitude of the Sato-Vartia index and the variety correction

²Under 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.

term would differ if calculated in a subset of products within the chosen set, in every case they should perfectly compensate to yield identical exact price indexes.

Using this insight from the theory, we calculate the elasticities of substitution within product groups that are consistent with the expected compensation between the Sato-Vartia index and the variety correction term. We then use these elasticities to calculate the change in the price index over the past 15 years using our procedure and also using a conventional procedure where we assume that goods with common barcodes are identical over time and goods with new barcodes are new products. Our estimates show that conventional price indexes significantly underestimate inflation rates in the US. The gap between our estimates of the inflation rate and that estimated following conventional methods is 1.91 percentage points per year over the 2005-2019 period.

Our work is related to the literature estimating CES price indexes, including [Feenstra \(1994\)](#), [Broda and Weinstein \(2006, 2010\)](#), and [Argente and Lee \(2021\)](#). This work implicitly assumes that quality change for products is zero, and we provide evidence that this assumption does not hold in the data. [Redding and Weinstein \(2020\)](#) develop an alternative price index that assumes that the geometric mean of the quality parameters among products with common barcodes is constant. Our evidence also suggests that this assumption is not likely to hold for all products with constant barcodes over a period of time.

Our work is also related to the literature quantifying biases in measures of inflation, in particular quality and variety biases. For example, [Nordhaus \(1997\)](#), [Bils and Klenow \(2001\)](#), and [Bils \(2009\)](#) document the bias in price indexes over time because quality improvements are not fully taken into account. [Broda and Weinstein \(2010\)](#) quantify the size of the variety bias in the US CPI over time. [Argente et al. \(forthcoming\)](#) decompose the gap in the price index of Mexico relative to the US into the biases from imputation, sampling, quality, and variety. Here we argue that, when there is uncertainty about the set of common products across time periods (something likely to happen in the data as highlighted by the linen shirt example), quality and variety biases cannot be separated. The relative magnitude of the variety bias and the quality bias could be affected if we identify too few or too many products in common across periods, particularly if there are a significant number of products not considered identical (i.e. they have different barcodes) yielding the same utility as products considered common across periods. Nonetheless, in our framework, as long as there are a group of products with average constant quality, the aggregate bias is the product of the variety and the quality biases. Therefore, the aggregate bias is not affected by the possibility that some of the products not considered common may in fact be the same product. For sure, the magnitude of the variety and the quality bias, taken individually, could be biased, but the product of the variety and quality terms are not biased.

The rest of the paper is organized as follows. Section 1 develops a simple conceptual framework under homothetic CES preferences. Section 2 describes the data. 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. In Section 4 we develop our price index. Section 5 presents our results and compares the changes of our price index to others indexes in the literature. The last section concludes.

1 Conceptual Framework

We use a price index for homothetic CES preferences with two nests.³ The utility function is:

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

where g denotes a group (e.g beer, soda), i is indexes product within a group, θ and σ_g denote the elasticity of substitution across and within groups, ϕ_{ig} is the quality of product i in group g .

The change in the aggregate price index is given by:

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

where \mathbb{I}_g denotes the set of incumbent products in group g , ω_g is the Sato-Vartia weight of group g and ω_{ig} is the Sato-Vartia weight of product i within the set of incumbent products in group g , $S_{\mathbb{I}_g}$ is the revenue share of the incumbent products of group g , and P_{ig} is the unit price of product i in 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 products; 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 products.

The main problem with implementing this formula is that what we observe is the unit price of the product P_{ig} and not the quality-adjusted price P_{ig}/ϕ_{ig} . In Adam Smith's example,

³Homothetic 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.

we do not directly observe the change in the quality of a roman tunic over time. Therefore, when we use the observed unit price of the set of incumbent products to calculate the price index, the resulting bias is given by the Sato-Vartia weighted average of the quality change:

$$- \sum_g \omega_g \left(\sum_{i \in \mathbb{I}_g} \omega_{ig} d \ln \phi_{ig} \right).$$

The price index is biased downwards when the weighted-average of the quality change is positive, and vice-versa.

Now suppose that there is a subset of products \mathbb{C}_g within the group of incumbent products 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:

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

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 products \mathbb{I}_g , and the second term is the Sato-Vartia weighted average of the *unit price* of the set of products in \mathbb{C}_g , and not the average of the *quality-adjusted* price of all incumbent products. 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 products because the effect of the change in the quality-adjusted price of these products is captured in the first term in equation 2. To see this, it is useful to note that $d \ln S_{\mathbb{C}_s} = d \ln S_{\mathbb{I}_s} + d \ln S_{\mathbb{D}_s}$ where \mathbb{D}_s denotes the set of incumbent products not in \mathbb{C}_s . Furthermore, it is easy to show that $\frac{1}{\sigma_g - 1} d \ln S_{\mathbb{D}_s}$ is equal to the Sato-Vartia average of the change in the quality-adjusted price of the products not in set \mathbb{C}_g . The first term in equation 2 thus captures the change in the price index from products where 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 products we deem as new could be the same as another seemingly unrelated product 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 quality-adjusted price of a linen shirt relative to a tunic. That is, the new variety term also captures the effect of changes in the quality-adjusted price of comparable products that are misidentified as new products.

In sum, the “new variety” term in equation 2 captures the effect of three things: (1) the change in the quality-adjusted price of incumbent products where we are not sure that the average quality change is constant; (2) the change in the quality-adjusted price of products that we did not identify as the same product and; (3) entirely brand new products. So while the new variety term can not be interpreted as the welfare gain from new products, the product of the “new variety” term and the Sato-Vartia weighted average of the change in the unit price of products in the set \mathbb{C}_g captures the change in the aggregate price index, including the effect of varieties that are actually new.

Finally, the set \mathbb{C}_g does not need to include *all* products where on average the quality change is zero. The only requirement is that quality change among the products in the chosen set average to zero. To be sure the relative magnitude of the two terms in equation 2 depends on the products chosen for the set \mathbb{C}_g but the product of the two terms do not. For example, suppose that the average quality of products in any subset of products in the set \mathbb{C}_g average to zero, and we instead use this subset to calculate the price index. If the change in the unit price of this subset of \mathbb{C}_g is lower than in the full set of products, then the second term in equation 2 will be larger (lower average price change of “existing” products) and the second term will be larger because the products dropped from the subset of “common” goods are the ones with higher price increases. If our assumption that the average quality change is zero in the two set of products, these two effects will exactly offset so that the change in the aggregate price index is unchanged.

2 Data

The Nielsen Consumer Panel data for the US tracks the shopping behavior of 40,000 to 60,000 households in 48 contiguous states plus Washington D.C. Each household uses in-home scanners to record their purchases. The US data contain 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 product: 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. For each barcode, the data also contain information on the brand, size, packaging, and other rich sets of product features.

Each barcode is classified into one of the 1,070 product modules that are organized into 104 product groups, that are then grouped into 10 major departments. For example, a 31 oz bag of Tide Pods (UPC 037000930389) is mapped to product module “Detergent-Packaged” in product group “Detergent” that belongs to the “Non-Food Grocery” department. The 5 largest of our 104 product groups in order of expenditure are pet food, carbonated beverages, paper products, bread and baked goods, and candy.

The data also contains information on each purchasing trip the panelist makes, including information on the retailer, the retailer’s location, the date of the transaction, and the expenditures and prices of each barcode purchased in each store. Furthermore, the data have demographic variables such as age, education, annual income, marital status, and employment that are updated annually based on surveys sent to the households. Nielsen constructs projection weights that make the sample representative of the US urban population that we use in our calculations to construct a demographically balanced sample of households.

The choice of data frequency requires a trade-off between choosing a sufficiently high frequency that keeps us from averaging out most of the price variation and a low enough frequency that enables us to be reasonably confident that purchase and consumption quantities are close. We follow [Redding and Weinstein \(2020\)](#) and use a quarterly frequency in our baseline specification. Four-quarter differences were then computed by comparing values for the fourth quarter of each year relative to the fourth quarter of the previous year.

3 Quality of a Variety Over Time

To explore whether the quality of product is constant over time, we map products in our theory to barcodes in the data, the finest aggregation of products available. The change in the expenditure share on a barcode is given by:

$$d \ln S_{ig} = -(\sigma_g - 1)d \ln P_{ig} + (\sigma_g - 1)d \ln \phi_{ig} + (\sigma_g - 1)d \ln P_g$$

where P_g is the aggregate price index of group g as in equation 1. Given that a product i is mapped to a single product group g , we implement the equation using the following specification:

$$d \ln S_{ig} = \beta_{ig} d \ln P_{ig} + \epsilon_{ig}. \quad (3)$$

We implement equation 3 product by product using time-series variation. Note that if $d \ln \phi_{ig} \approx 0$, the coefficient β_{ig} is unambiguously negative given our assumption that $\sigma_g > 1$ and that ϵ_{ig} captures measurement error in expenditure shares or prices. On the other hand, if there are quality changes that are positively correlated with prices, β_{ig} will be biased toward zero or could even be positive as follows:

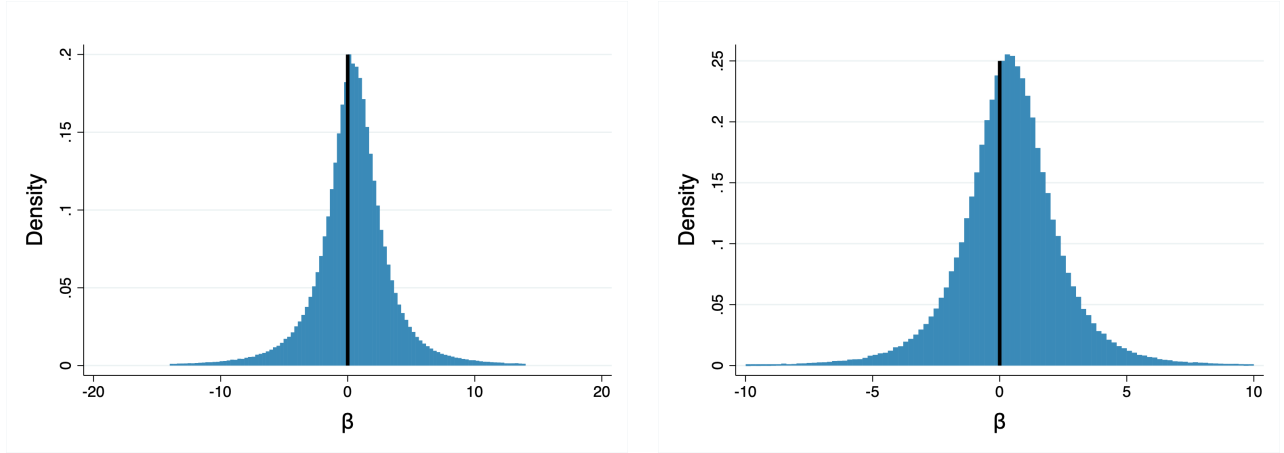
$$\beta_{ig} = -(\sigma_g - 1) + \frac{\text{Cov}(d \ln P_{ig}, (\sigma_g - 1)d \ln \phi_{ig})}{\text{Var}(d \ln P_{ig})}.$$

3.1 Empirical Implementation

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 products 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 products; β_{ig} is estimated separately for each of them.

Figure 1 shows the distribution of coefficients, β_{ig} . The average estimate for β_{ig} , using either all products or only those that lasted in the data for at least five years, is positive and very similar in magnitude, approximately 0.31 on average (median 0.4) In fact, approximately 60% of products have a positive coefficient indicating both the presence of quality shocks and their positive correlation with prices.

Figure 1: Single Barcode Regression: Coefficients



(a) All barcodes

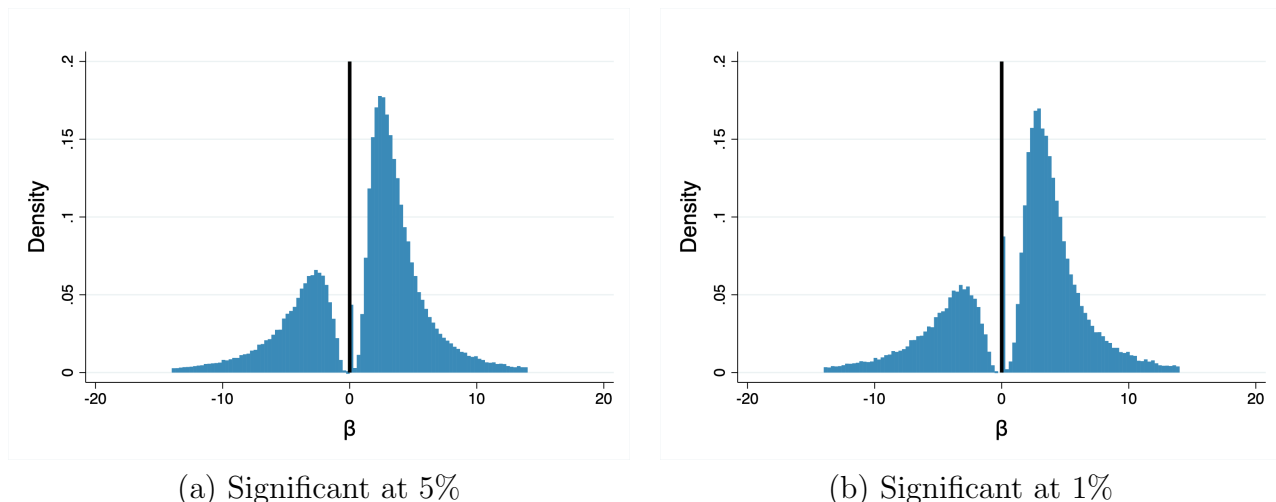
(b) More than 5 years

Notes: The figure shows the distribution of the coefficients β_{ig} estimated using equation 3 for products in the Nielsen Consumer Panel data (2004-2019). Panel (a) includes all products that lasted at least 8 quarters in the data (approximately 686 thousand). Panel (b) shows the coefficients of products that lasted at least 20 quarters in the data (approximately 288 thousand). The black vertical line marks products whose coefficient is zero.

3.2 Chosen Sample

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

Figure 2: Single Barcode Regression: Significant Coefficients



Notes: The figure shows the distribution of the coefficients β_{ig} estimated using equation 3 for products in the Nielsen Consumer Panel data (2004-2019). Panel (a) includes all products whose coefficient is significantly different than zero at 5 percent confidence level (39,644 products). Panel (b) includes all products whose coefficient is significantly different than zero at 1 percent confidence level (18,691 products). The black vertical line marks products whose coefficient is zero.

Panel (a) of Figure 2 shows the distribution of coefficients that are significant at 5% and Panel (b) those that are significant at 1%. The figure shows that there are more products whose estimated elasticity of demand is positive, again consistent with the prevalence of quality changes at the barcode level. Our sample of *chosen* products includes only those to the left of the vertical black line. Some of the product groups that contribute more products to the chosen sample seem to be less likely to be subject to temporary demand fluctuations. 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.

4 Price Index

We now use equation 2 to estimate the aggregate change in prices. As a reminder to the reader, the key is that the set of products for which we measure the change in the expenditure share and the change in prices are products for which the average quality change is zero. To implement equation 2 in the data, for each product group and quarter, we calculate both the Sato-Vartia price index and the variety correction term. For each product group, we have 60 observations, one per quarter. As before, we use four-quarter differences to control

for seasonality. Furthermore, using the methodology developed in the previous section, we identify the set of products whose quality is approximately constant. In total, we use 18,691 products, those whose β_{ig} is negative and significant at 1 percent level in equation 3.

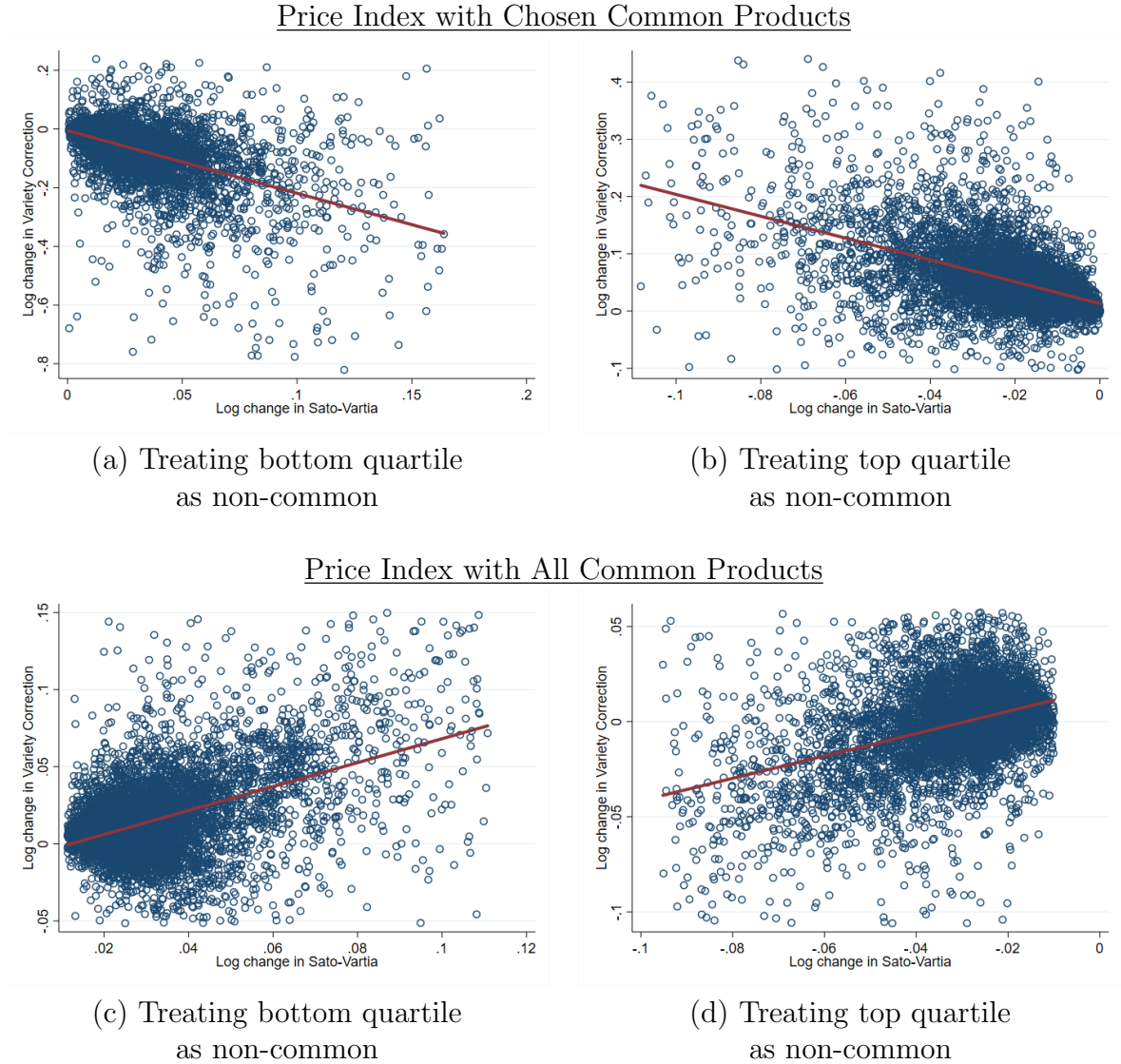
4.1 Net Effect

Our theory predicts that the product of the variety correction and the Sato-Vartia price index is an unbiased estimate of the aggregate price index regardless of the size of the set of chosen products. We first investigate whether the variety correction term compensates when the Sato-Vartia component either increases or decreases, when we adjust the size of the set of chosen products. Suppose that we treat the bottom quartile of year-on-year of price changes as non-common goods. Then, the Sato-Vartia component would naturally increase because the bottom quartile is excluded from the set of chosen common goods. Likewise, if we treat top quartile of year-on-year price changes as non-common goods, the Sato-Vartia component would naturally decrease. Our theory predicts that the variety correction term will move to the opposite direction, fully compensating increase or decrease in the Sato-Vartia term.

Panel (a) and (b) of Figure 3 describes the compensation when the elasticity of substitution is assumed to be 2, which makes $\frac{1}{\sigma_g-1} = 1$ in the variety correction term. The y-axis is log change in the variety correction and the x-axis is log change in the Sato-Vartia component. Panel (a) treats the bottom quartile as non-common and panel (b) treating top quartile as non-common. Both panels show that the variety correction term moves to the opposite direction of the change in Sato-Vartia component as predicted by the theory.

Panel (c) and (d) of Figure 3 show that the compensation is not guaranteed when we use the set of *all* common products. This is because the quality of these products is changing over time, even if we define products at the finest level of aggregation. Panel (c) treats bottom quality of price changes among all common products as non-common. The Sato-Vartia component naturally increases, but the variety correction term *increases*, which biases the price index. We also see that the Sato-Vartia component and the variety correction term move to the same direction in Panel (d) when we treat top quartile of price changes among all common chosen goods as non-common.

Figure 3: Compensation of Variety Correction Term when Sato-Vartia Component Changes



Notes: Panel (a) and (b) are the experiment with the price index with chosen common products. Panel (c) and (d) are the experiment with the price index with all common products. The elasticity of substitution is assumed to be 2, which makes $\frac{1}{\sigma_g - 1} = 1$ in the variety correction term. Each dot represents a combination of product group and a quarter. For visualization, we trim top and bottom 1% of each variables.

4.2 Elasticity of Substitution Estimation

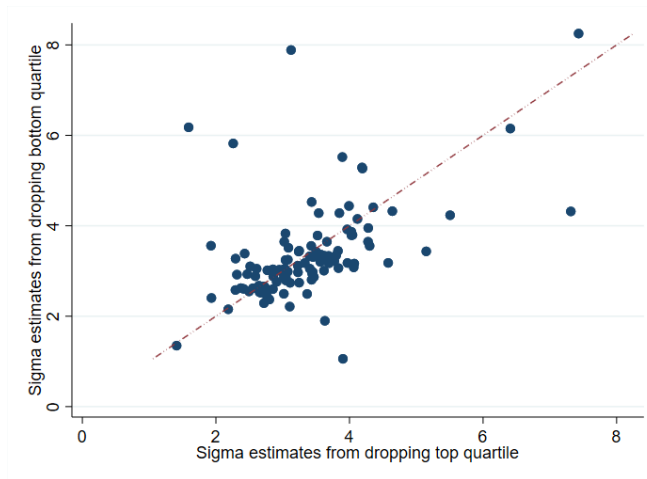
We propose a strategy to estimate elasticity of substitution taking advantage of the fact that the set of chosen goods are likely to have constant quality and by targeting full compensation of the variety correction term when the Sato-Vartia component changes. We find the elasticity

of substitution that makes the coefficient in the following regression exactly one:

$$d \ln S_{C_g}^{\frac{1}{\sigma_g-1}} = \beta_g \times d \sum_{i \in C_g} \omega_{ig} \ln P_{ig} + \epsilon_g \quad (4)$$

where left hand side and right hand side variables come from equation 2 given the elasticity of substitution, σ_g .

Figure 4: Estimated Elasticities of Substitution



Notes: Each dot represents a combination of product group and a quarter. For visualization, we trim top and bottom 1% of each variables. A dotted line is 45 degree reference line. The correlation is 0.50 and significant at 1 percent level.

Figure 4 shows estimated elasticities of substitution from dropping bottom/top quartile of common products’ price changes. Estimated elasticities are correlated. The correlation is 0.50 and significant at 1 percent level. We take the average of two inferred elasticities of substitution to get our preferred estimates.

Column (1) of Table 1 reports descriptive statistics of our estimated elasticities of substitution. The median group has an elasticity of 3.24 and the mean group of 3.28. Column (2) of Table 1 reports descriptive statistics of elasticities of substitution following the methodology developed by Feenstra (1994) (e.g. Broda and Weinstein, 2010; Hottman et al., 2016; Redding and Weinstein, 2020; Argente et al., forthcoming), which is based on a GMM estimation approach using the double-differenced residuals in demand and supply as moment conditions. Section A in the Online Appendix describes in detail this method. Our benchmark estimates are lower than the estimates from the GMM with double-differentiated residuals, the median group has an elasticity of 6.32. The elasticities under both methodologies have an statistically significant positive correlation of 0.44. We report results under both sets of elasticities

Table 1: Estimated Elasticities of Substitution

| Percentile | Elasticities of Substitution | |
|------------|------------------------------|--------------------------------------|
| | (1) VC-SV Compensation | (2) GMM w/ double-diff. residuals |
| 10 | 2.59 | 5.12 |
| 25 | 2.78 | 5.52 |
| median | 3.24 | 6.32 |
| mean | 3.31 | 6.42 |
| 75 | 3.61 | 7.24 |
| 90 | 4.11 | 7.89 |

Notes: The table reports descriptive statistics of estimated elasticities of substitution under three different strategies. Column (1) uses full compensation of the variety correction term when the Sato-Vartia component changes. Column (2) uses moment conditions of the double-differenced residuals in demand and supply with the GMM estimation approach.

in later sections.

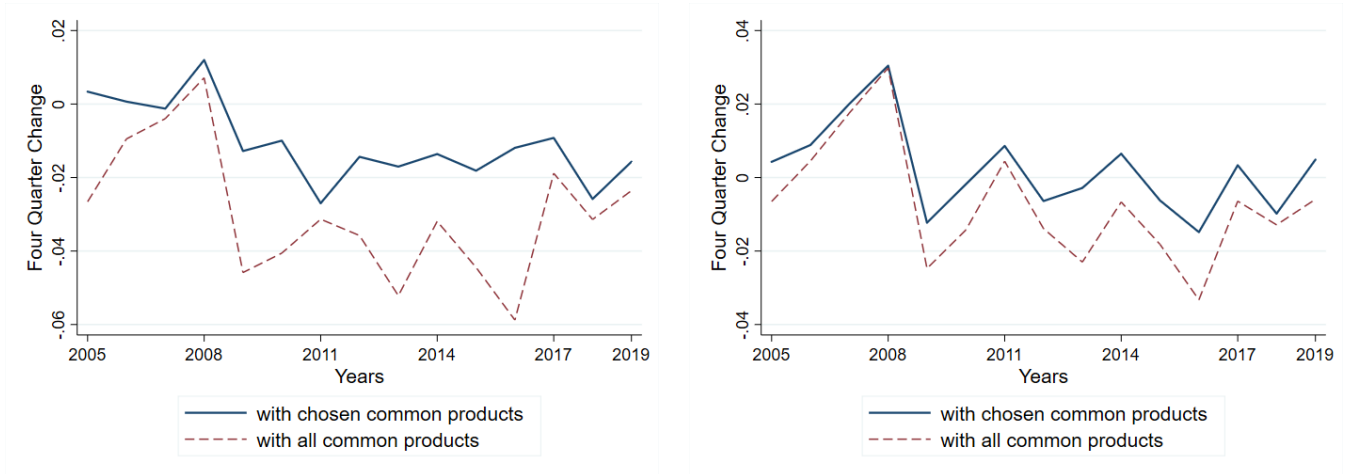
5 Results

Given estimated elasticities of substitution, we calculate the price index with chosen common products (equation 2) and the price index with all common products (equation 1), which corresponds to the standard Feenstra-CES index. Figure 5 shows the four-quarter changes of both of these indexes. Panel (a) uses our benchmark elasticities and Panel (b) uses the GMM estimated elasticities for comparison. Four-quarter changes in the price index calculated with chosen common products are higher than four-quarter changes in the price index calculated with all common products in all periods. On average, four-quarter changes in the aggregate price index calculated with chosen common products are 1.91 (0.94) percentage points higher under our benchmark elasticities (GMM elasticities).

Conventional price indexes underestimate the cost of living in the US because the set of all common products have experienced, on average, a decline in perceived quality over time. They become obsolete. This is consistent with the documented decline of prices and shares over the products' life cycle (Argente et al., 2021). For example, a product's quality could wane as competing firms introduce similar new products and/or as the firm improves upon its own products.

We find very similar results when we examine each of the main departments in our data.

Figure 5: Four-Quarter Changes Chosen vs. All Common Products



(a) VC-SV compensation elasticities

(b) GMM Elasticities

Notes: Panel (a) and (b) show the four-quarter changes in Feenstra-CES price index calculated with chosen (equation 2) vs. all common products (equation 1). Panel (a) uses our benchmark elasticities using full compensation between the variety correction term and the Sato-Vartia component. Panel (b) uses elasticities estimated from moment conditions of the double-differenced residuals in demand and supply with the GMM.

Table 2 reports changes in the aggregate price index using both our index and the standard Feenstra-CES index by department. As before, we find that the conventional price index is downward biased. This is the case in all eight departments. Importantly, the size of bias is large in “Health and Beauty Aids” and “General Merchandise.” Both departments have a larger share of semi-durable consumer goods (e.g., razors, lamps), which are products that become obsolete more rapidly; the sales and prices of these products decline fast shortly after introduction.

Table 2: Four-Quarter Changes Chosen vs. All Common Products (Department)

| Department | Percentage Points Diff. b/w Chosen vs. All Common Products | |
|------------------------|--|--------------------------------------|
| | (1) VC-SV Compensation | (2) GMM w/ double-diff. residuals |
| Health and Beauty Aids | 2.22 | 1.17 |
| Dry Grocery | 0.90 | 0.37 |
| Frozen Foods | 1.05 | 0.50 |
| Dairy | 0.71 | 0.10 |
| Packaged Meat | 0.52 | 0.15 |
| Non-Food Grocery | 0.88 | 0.55 |
| Alcohol | 0.49 | 0.12 |
| General Merchandise | 4.25 | 2.28 |

Notes: The table reports average percentage points difference in four-quarter changes of the price index calculated from the chosen (equation 2) and all common products (equation 1) by department. Column (1) uses our benchmark elasticities estimated from VC-SV compensation. Column (2) uses uses elasticities estimated from moment conditions of the double-differenced residuals in demand and supply with the GMM estimation approach.

6 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 scanner level data on prices and quantities at the finest level of aggregation, we first show that the perceived quality of most products is not constant. We then develop a strategy to select products that are less likely to suffer from quality changes and use this set of products 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 (at least regarding physical characteristics) have experienced a decline in perceived quality.

References

- Argente, David and Munseob Lee, “Cost of living inequality during the great recession,” *Journal of the European Economic Association*, 2021, 19 (2), 913–952.
- , —, and Sara Moreira, “The Life Cycle of Products: Evidence and Implications,” *mimeo*, 2021.
- Argente, David O, Chang-Tai Hsieh, and Munseob Lee, “Measuring the Cost of Living in Mexico and the US,” *American Economic Journal: Macroeconomics*, forthcoming.
- 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.
- Hottman, Colin J, Stephen J Redding, and David E Weinstein, “Quantifying the sources of firm heterogeneity,” *The Quarterly Journal of Economics*, 2016, 131 (3), 1291–1364.
- 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 Alternative 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 \(2006\)](#):

$$\Delta^{u,t}\ln S_{ig} = (1 - \sigma_g)\Delta^{u,t}\ln P_{ig} + \iota_{ig} \quad (5)$$

$$\Delta^{u,t}\ln P_{ig} = \frac{\delta_g}{1 + \delta_g}\Delta^{u,t}\ln S_{ig} + \kappa_{ig} \quad (6)$$

Note that when the inverse supply elasticity is zero (i.e. $\delta_g=0$), the supply curve is horizontal and there is no simultaneity bias in σ_g . Equations 5 and 6 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:

$$G(\beta_g) = E_T[\nu_{ig}(\beta_g)] = 0 \quad (7)$$

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:

$$\hat{\beta}_g = \arg \min_{\beta_g \in B} G^*(\beta_g)' W G^*(\beta_g) \quad \forall i \in \omega_b \quad (8)$$

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. If the procedure renders imaginary estimates or estimates of the wrong sign, we use a grid search to evaluate the GMM objective function above.