Measuring the Gains from Creative Destruction*

David Argente†
Pennsylvania State University

Chang-Tai Hsieh‡
University of Chicago

Munseob Lee§
University of California San Diego

July 2022

Abstract

We measure the aggregate price index when innovation takes the form of creative destruction. When new or improved products do not completely replace products that are its close substitutes, the standard price index formula by Feenstra (1994) *overstates* the contribution of new products but *understates* the effect of the improvement of existing products. The former shows up empirically as a decline, and the latter as an increase, in the residual revenue share of incumbent products (after controlling for the price change of the product). We use barcode level data on consumer packaged goods to show that these two groups of products account for 58% of the barcodes, where the share of products whose residual revenue share falls is significantly larger than the share of products whose residual revenue share increases. We use data on prices and expenditures of products that are not likely to have close substitutes to measure the effect of *all* types of innovation, including creative destruction. Our estimate of the aggregate price of consumer packaged goods exceeds the price calculated from the Feenstra (1994) formula by 0.74 percentage points per year between 2004 and 2019.

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 central challenge in measuring price indices is that the mix of products changes over time. For example, in the most detailed data of consumer packaged goods at the 12-digit barcode level, new barcodes account for 16.8 percent of total sales.[1] An influential paper by Feenstra (1994) shows that if the new products are new horizontal varieties as in Romer (1990), then their effect on welfare is captured by two statistics, the revenue share of the new products and the elasticity of substitution between the horizontally differentiated varieties. Many authors, including Broda and Weinstein (2010), Handbury and Weinstein (2015), Argente and Lee (2021), and Argente, Hsieh and Lee (forthcoming), assume that new products are horizontally differentiated varieties and use the expenditure share of the new varieties as a summary statistic of their contribution on the aggregate price index.

A recent paper by Aghion, Bergeaud, Boppart, Klenow and Li (2019) shows that the Feenstra (1994) summary statistic approach can still be used when new products are not new horizontal varieties but instead are perfect substitutes for previously existing products. In this case, when new varieties are the result of creative destruction as in Aghion and Howitt (1992) and Grossman and Helpman (1991), the summary statistics are the revenue share of new products *net of the revenue share of exiting products* and the same elasticity of substitution. The intuition is that the difference between the revenue share of the new products and the products that exit captures the quality improvement due to creative destruction.

However, new products may also lower the market share of existing competing products but not entirely replace them. Figure 1 (right panel) plots the change in the revenue share of incumbent products against the revenue share of new goods in a product category (e.g. soap) in barcode level data of consumer packaged goods.[2] The figure shows clearly the negative relationship between the expenditure on entering products in a product category and the decline in expenditure on incumbent products in the same product category. For comparison, the left panel shows the relationship between the expenditure share of exiting products and the share of new barcodes in the product category.[3] So more entry is not only associated with more exit but also with a larger decline in the market share of surviving incumbent products.

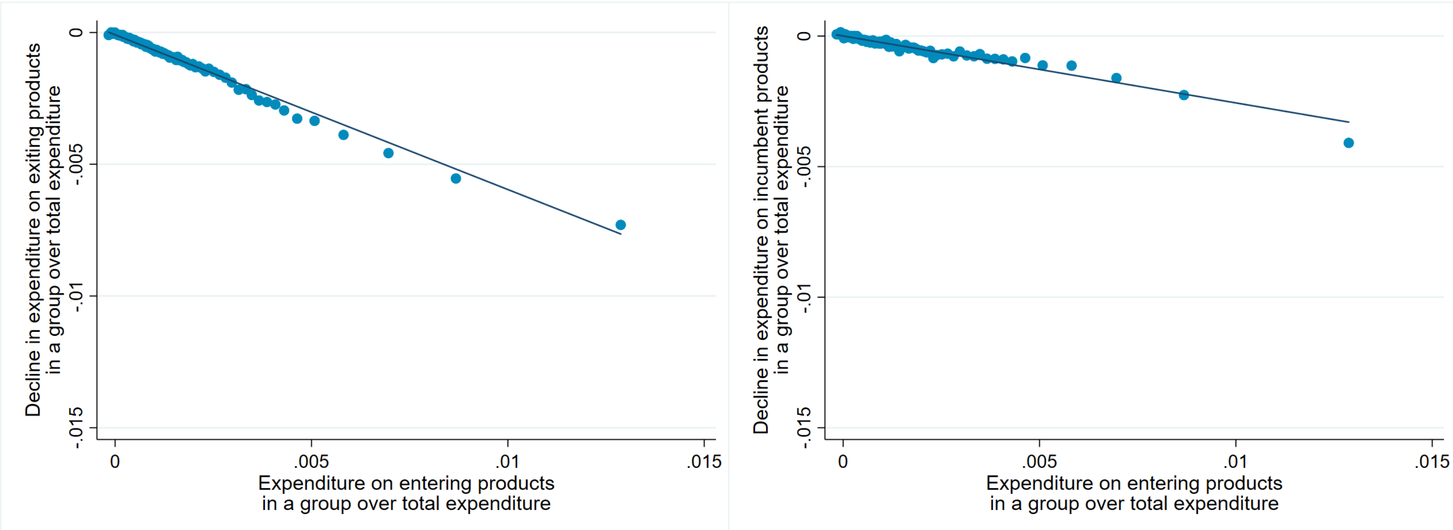
In principle, we can use the same Feenstra (1994) summary statistic when new goods also lower the market share of incumbent firms but does not completely replace them. This requires that we identify the goods that are adversely affected for each new product. Although Figure 1 shows that entry in a product category in the barcode data is associated with a

[1]The expenditure share of new barcodes is calculated by comparing year-on-year differences in a quarterly data. For example, in 2005Q1, new barcodes are the set of barcodes available in 2005Q1 that didn't exist in 2004Q1.

[2]We describe this data in more detail later. The data is from 2004 to 2019.

[3]The regression coefficients are -0.59 for the left panel and -0.26 for the right panel.

Figure 1: Entry, Exit, and Incumbents



Entry and Exit | Entry and Incumbent Share

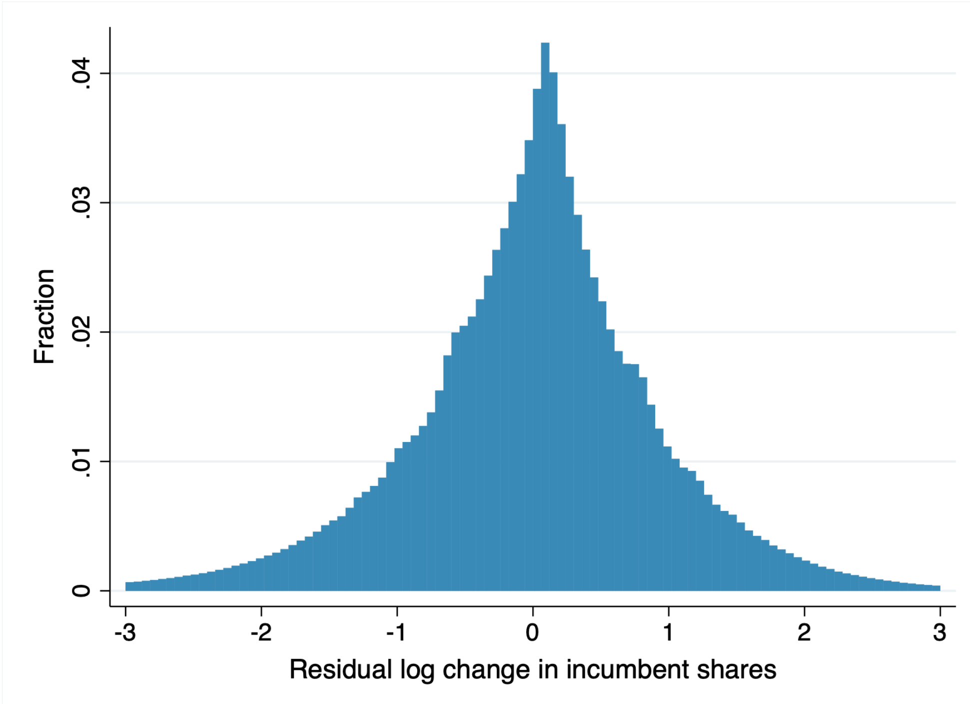
Notes: Left panel shows scatterplot of revenue share of exiting products (multiplied by -1) and the expenditure share of entering products in each product category (117 in total) in the barcode level data. x-axis variables are grouped into 100 equal-sized bin. Right panel shows scatterplot of the change in the revenue share of incumbent products in a product category and the revenue share of entering products in the same category. Data are barcode level data on consumer packaged goods from 2004 to 2019.

decline in the share of incumbent products in the same category, it seems clear that the product category is much too coarse. Figure 2 shows the distribution of residual log change in incumbent shares after controlling for log price change and product category fixed effects. It is clear that there are incumbent products that are close substitutes for new products in the group, and these are the ones whose residual market shares fall, and other incumbent products that are not close substitutes, and these are ones where the residual market shares fall by less or even increase.

One solution is to match new and improved products with incumbent or exiting products with similar characteristics, which of course is the long standing practice of many statistical agencies (such as the U.S. BEA). However, this exercise is likely to be time consuming and error prone. In addition, products with different physical characteristics may in fact be close substitutes. For example, Adam Smith famously argued that linen shirts in 18th century Europe appear to be quite different from togas used in ancient Rome and Greece but that the two types of clothing are in fact close substitutes.[4]

[4] "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

Figure 2: Distribution of Change in Residual Incumbent Shares



Notes: Figure shows the histogram of change in the residual of log incumbent shares at the barcode level after controlling for log price change and product category fixed effects. The data contains 117 product categories.

This paper proposes a new methodology to measure the welfare gains from creative destruction that circumvents the problem of matching products that are close substitutes. This method requires that we identify a set of products that do *not* have close substitutes, or put differently, a set of products that are horizontally differentiated. Empirically, we show that this is equivalent to finding products where changes in expenditure shares are fully explained by prices. The main advantage of our procedure is that we do not need to know anything about the products – new, exiting, or surviving – for which there are close substitutes.

The key is that we identify a set of products where the change in its expenditure share is fully explained by its price. Once we have identified this bundle of "no residual expenditure" 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 "no residual expenditure" bundle and; 2) the change in the market share of the "no residual expenditure" bundle, adjusted by a function of the price elasticity of demand. This is exactly Feenstra (1994)'s formula, with one key difference. Namely, the "new goods" term is not the welfare gain from

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, Wealth of Nations.

new goods because it also includes the effect of price changes among incumbent products with close substitutes. However, the sum of the two terms is the sum of price changes among all incumbent products and the welfare gain from new products, even if we cannot separately identify these two forces.

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. We then choose the products where the regression coefficient are negatively and significantly correlated for our set of "no residual expenditure" products.

We then calculate the change in the price index using our procedure and also using a standard Feenstra (1994) price index formula. 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. Empirically this suggests that innovation among products for which there are close substitutes primarily takes the form of new barcodes instead of improvements in existing barcodes.

Our work builds on the large body of work that measures the welfare effect of new varieties, including Broda and Weinstein (2010), Broda and Weinstein (2006), Broda and Weinstein (2010)), Aghion, Bergeaud, Boppart, Klenow and Li (2019), and Argente and Lee (2021). We have two messages relative to this literature. First, the estimates of the welfare gain of new varieties in these papers are likely to overstate the real gains if some of the products that are close substitutes lost market share but did not entirely exit. Second, we cannot separately measure the effect of the welfare gain from new products from the price change of incumbent varieties. What we can do instead is to measure the *net* effect of these two forces by identifying a set of "no residual expenditure" products and measuring the price change and the change in the expenditure share of these group of products.

The rest of the paper is organized as follows. Section 1 develops a conceptual framework, derives the aggregate price index, and shows that that the aggregate price index can be measured by identifying a set of products where the residual expenditure share is zero. The next section shows how we identify products with "zero residual expenditures." In Section 3, we show that the change in the residual expenditure share of 58% of products is non-zero, and describe our procedure to select products where the change in the residual expenditure share is likely to be zero. Section 4 estimates parameters, presents our results, and compares the changes of our price index to other indexes in the literature. The next section studies residual market shares at the product level further. The last section concludes.

1 Measuring Welfare Gain from Creative Destruction

The utility function is:

$$\mathbb{U} = \left[\sum_{i \in \mathbb{H}} C_i^{\frac{\sigma-1}{\sigma}} + \sum_g \left(\sum_{i \in g, \mathbb{V}} C_i^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1} \frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}$$

where $\mathbb{V}$ denotes the set of products for which there are close substitutes, with an elasticity of substitution given by ρ, and $\mathbb{H}$ denotes the products that are less substitutable, with an elasticity of substitution $\sigma << \rho$. For products $i \in \mathbb{V}$, g denotes the set of products that are its close substitutes.

The change in the aggregate price index P is:

$$\begin{aligned} d\ln P = & \sum_{i \in \mathbb{I}} \omega_i \, d\ln\, P_i + \frac{1}{\sigma-1} d\ln\left(1 - \text{New} + \text{Exit}\right) \\ & - \frac{\rho-\sigma}{(\sigma-1)(\rho-1)} \sum_g \omega_g \, d\ln\left(1 - \text{New}_g + \text{Exit}_g\right) \end{aligned} \tag{1}$$

where $\mathbb{I}$ denotes the set of incumbent products, P_i is the price of product i, ω_i is the Sato-Vartia weight of product i as a share of all incumbent products, "New" and "Exit" are the revenue shares of *all* new and exiting products in total expenditures, "New_g" and "Exit_g" are the revenue shares of new and exiting products of group g in total expenditures of group g, and $\omega_g \equiv \sum_{i \in \mathbb{V}, \mathbb{I}, g} \omega_i$ is the Sato-Vartia share of group g.[5]

The first line in equation 1 is the standard Feenstra (1994) price index formula, which is the sum of the weighted average of the price change of incumbent products and the product of $1/(\sigma-1)$ and the change in the market share of the incumbent products. What is different in our case is the second line in equation 1, which is the product of $\frac{\rho-\sigma}{(\sigma-1)(\rho-1)}$ and the weighted average of the change in the share of incumbent products $i \in \mathbb{V}$. This second term is zero when $\rho = \sigma$ or when the share of incumbent products $i \in \mathbb{V}$ is constant.

It useful to see how innovation maps into the aggregate price index in equation 1. Many authors, including Broda and Weinstein (2010), Broda and Weinstein (2006), and Argente and Lee (2021), assume that new products are new horizontal varieties as in Romer (1990). In this case, the last term in equation 1 is zero and aggregate effect of new varieties is captured by their market shares, which is the "New" term in the first line in equation 1.

[5] Refer to Appendix B for derivation of this equation.

Now suppose that new products are the result of creative destruction as in Aghion and Howitt (1992) and Grossman and Helpman (1991). In addition, suppose the new products are perfect substitutes of incumbent products such that with limit pricing new products entirely replaces the incumbent products. This formulation thus assumes $\rho = \sigma$, in which case the second line in equation 1 is also zero. This is the case considered by Aghion, Bergeaud, Boppart, Klenow and Li (2019). Compared to the papers that assume new products are horizontally differentiated varieties, what is different when new products are the result of Schumpeterian creative destruction is that after the introduction of new products the competing incumbent products exit. In this case, the welfare effect of creative destruction is given by the difference between "New" and "Exit" in the first line in equation 1. Intuitively, the improvement in quality from creative destruction is given by the difference in the market shares between the new products and the products it replaces.

In the empirical section, we will show that the empirically relevant case is when new products are close substitutes for some incumbent products but does not entirely replace all the incumbent products. In this case, the products adversely affected by the new products are the products that exit and the incumbent products that lose market share but do not exit. The Feenstra (1994) formula captures the former but not the latter, where the effect of the loss in market share of incumbent products is given by the second line in equation 1. So when creative destruction results in new varieties, the standard Feenstra (1994) price index understates the change in the aggregate price index, and the magnitude of the understatement depends on $\rho - \sigma$ and the decline in the market share of incumbent products $i \in \mathbb{V}$

It is also possible that the Feenstra (1994) price index formula *overstates* the price index. This happens when creative destruction improves the quality of an incumbent product, and this quality improvement results in the exit of incumbent products. In this case, the share of the remaining incumbent products rises, so the standard Feenstra (1994) formula *overstates* the change in the aggregate price index.

To summarize, the bias in the aggregate price index calculated from the Feenstra (1994) formula when $\rho > \sigma$ and when innovation takes the form of creative destruction depends on the extent to which creative destruction takes the form of new varieties on the one hand, or quality improvements in existing products, on the other hand. The net effect of these two forces is summarized by the change in the share of incumbent products $i \in \mathbb{V}$.

Note that when creative destruction creates new varieties, the price index is biased when incumbent products lose market share but do not exit. In contrast, when creative destruction improves the quality of incumbent products, the price index is biased only when incumbent products exit. If the competing incumbent products lose market share but do not exit, then there is no measurement bias. Empirically there is no change in the total market share of incumbent products in the group. Also, note that if quality improves for product $i \in \mathbb{H}$, there is also no bias as the market share of incumbent products $i \in \mathbb{V}$ are unaffected.

There are two empirical challenges with calculating the aggregate price index following equation 1. First, we need to measure "New$_g$" and "Exit$_g$", which are the market shares of new and exiting products for each group g. That is, we need to match new and exiting products with the incumbent products for which they are close substitutes. One can attempt to do this by grouping products with similar physical attributes, it is possible that products that appear to be quite different may in fact be close substitutes. More importantly, matching products with similar attributes is labor-intensive and the result likely riddled with error. The second issue is that we also need an estimate of the elasticity of substitution between products in the same group, ρ.

Now suppose that over a given time period, we can identify a set of incumbent products $\mathbb{C} \in \mathbb{H}$ that are horizontally differentiated. We can then measure the aggregate price index with data on prices and expenditures only for the products in the set $\mathbb{C}$:

$$d \ln P = \sum_{i \in \mathbb{C}} \widetilde{\omega}_i \, d \ln P_i \; + \; \frac{1}{\sigma - 1} d \ln \left(1 - \text{New} - S_{\mathbb{H} \notin \mathbb{C}, \mathbb{I}} - S_{\mathbb{V}, \mathbb{I}} + \text{Exit}\right) \tag{2}$$

where $S_{\mathbb{V},\mathbb{I}}$ is the revenue share of the incumbent products in set $\mathbb{V}$, $S_{\mathbb{H}\notin\mathbb{C},\mathbb{I}}$ is the revenue share of the incumbent products in set $\mathbb{H}$ not in the chosen set $\mathbb{C}$, and $\widetilde{\omega}_i$ is the Sato-Vartia revenue share of product i as a share of all products in $\mathbb{C}$. The first term in equation 2 is the weighted average of price changes among the products in $\mathbb{C}$, and the second term is the change in the expenditure share of the same set of products. Equation 2 is essentially the Feenstra (1994) formula, with the difference that we only need information on the price and revenue share of the products in $\mathbb{C}$, and not for all incumbent products.

The key advantage of equation 2 is that we do not need to know anything about the products with close substitutes, nor do we need to know the elasticity of substitution ρ between the products with close substitutes. The reason is because everything that goes on among the products not in the chosen set $\mathbb{C}$ is captured in the second term in equation 2. To see this, note that the difference between the first new variety term in equation 1 and the new variety term in equation 2 is $S_{\mathbb{H}\notin\mathbb{C},\mathbb{I}} + S_{\mathbb{V},\mathbb{I}}$. In turn, the change in these shares are equal to:

$$d \ln S_{\mathbb{V},\mathbb{I}} \; = \sum_{i \in \mathbb{V},\mathbb{I}} \omega_i \, d \ln P_i \; + \; \frac{\rho - \sigma}{\rho - 1} \sum_{g} \omega_g \, d \ln \left(1 - \text{New}_g + \text{Exit}_g\right) \tag{3}$$

$$d \ln S_{\mathbb{H}\notin\mathbb{C},\mathbb{I}} \; = \sum_{i \in \mathbb{H}\notin\mathbb{C},\mathbb{I}} \omega_i \, d \ln P_i \tag{4}$$

The first term in equation 3 is the average change in prices among incumbent products in $\mathbb{V}$, the second term in equation 3 is the bias in the Feenstra price index when innovation takes the form of creative destruction and $\rho \neq \sigma$, and equation 4 is the average change in prices of incumbent products in $\mathbb{H}$ that are not in the chosen set $\mathbb{C}$.

In sum, the "new variety" term in equation 2 measures the welfare effect of three things: (1) welfare effect of *all* new varieties; (2) price changes among incumbent products in set $\mathbb{V}$ and; (3) price changes among incumbent products in set $\mathbb{H}$ that are not in the chosen set $\mathbb{C}$. So while we can no longer interpret the "new variety" term in equation 2 as only measuring the effect of new varieties, the sum of this term and the average change in the prices of products in the chosen set $\mathbb{C}$ measures the net effect of entry and exit and price changes in all incumbent products. The key empirical challenge then is to identify products for which there are no close substitutes, which we tackle in the next section.

2 Identifying Products with No Close Substitutes

In this section, we develop a method to choose a set of horizontally differentiated products $\mathbb{C}$. To begin, it is useful to see the telltale signs of a product with close substitutes. The expenditure share of such a product is given by:

$$S_i = \left(\frac{P_g}{P_i}\right)^{\rho-1} \left(\frac{P}{P_g}\right)^{\sigma-1} \tag{5}$$

where P_g is the aggregate price index for the group of products in g defined as

$$P_g = \frac{1}{\rho-1} d\ln\left(1 - \text{New}_g + \text{Exit}_g\right) + \sum_{i\in\mathbb{I},g} \widehat{\omega}_i d\ln P_i \tag{6}$$

which is decreasing in the market share of new products (net of the market share of exiting products) and the weighted average in the price of the incumbent products.[6]

The residual change in the share of the product $\tilde{S}_i$ is then:

$$d\ln\tilde{S}_i \;\equiv\; d\ln\frac{S_i}{(P_i/P)^{\sigma-1}} \;=\; \frac{\rho-\sigma}{\sigma-1}\, d\ln\left(P_g/P_i\right)$$

Now consider how two types of innovation affect the residual market share. First, suppose that innovation primarily takes the form of creation of new products that are close substitutes for some incumbent products. That is, suppose that $d\ln P_g < 0$ and $d\ln P_g - d\ln P_i < 0$ for an incumbent product i. In this case, the residual change in the price-adjusted expenditure share of the incumbent product *decreases.* So an empirical sign of the case when the Feenstra

[6] $\widehat{\omega}_i$ is the Sato-Vartia weight as a share of all incumbent products in g.

(1994) formula understates the price index is a decline in the residual share of incumbent products.

The second case is when innovation takes the form of quality improvements in the incumbent products. When this happens to products that have close substitutes, then $d \ln P_g - d \ln P_i > 0$ for these products. In this case, the price-adjusted share of these products *increases*. So the sign of the case where the Feenstra (1994) formula overstates the price index is an increase in the residual share of incumbent products.

Now consider the same price-adjusted market share of a horizontally differentiated product. The share of the product is:

$$S_i = \left(\frac{P}{P_i}\right)^{\sigma-1} \tag{7}$$

The change in the residual market share for this type product is zero.

So the telltale sign of a horizontally differentiated product is that the change in expenditures after controlling for the change in its relative price is zero. On the other hand, the residual change in its expenditure share of a product with close substitutes is either negative, if new products are introduced in the group, or positive, if the price of a product with close substitutes has declined significantly.

We implement this intuition with regressions at the product level. From equations 5 and 7, the change in the expenditure share on a product is:

$$d \ln S_i = -(\sigma - 1) d \ln \left(\frac{P_i}{P}\right) + \lambda_i \, d \ln \tilde{S}_i \tag{8}$$

where λ_i is an indicator variable for a product with close substitutes ($i \in \mathbb{V}$). We can implement this equation at the product level using time-series variation as follows:

$$d \ln S_i = \beta_i \, d \ln \left(\frac{P_i}{P}\right) + \epsilon_i. \tag{9}$$

Note that when $\lambda_i = 0$, the coefficient β_i is unambiguously negative given our assumption that $\sigma > 1$. However, given that some products have close substitutes, β_i will be biased toward zero, and possibly even positive, as follows:

$$\beta_i = -(\sigma - 1) + \lambda_i \, \frac{\mathrm{Cov}(d \ln P_i, \, d \ln \tilde{S}_i)}{\mathrm{Var}(d \ln P_i)}.$$

The bias in the regression coefficient for products $i \in \mathbb{V}$ depends on probability the product has close substitutes λ_i and the correlation of the price and the residual change in the product's market share.

In sum, the coefficient from a regression of the change in the expenditure share of a product on the change in the relative product of the product indicates whether the product is likely to have close substitutes. Horizontally differentiated products are those with an unambiguously negative coefficient for β_i. Given that our methodology can be applied, even if the set of horizontally differentiated products $\mathbb{C}$ is arbitrarily small, we select as horizontally differentiated products whose estimated β_i is negative and statistically different than zero at the 1% level. The last restriction also allows use to include in $\mathbb{C}$ only products with small variation in their residual change in expenditure share, which further ensures that products in this set are less likely to have close substitutes. The next subsection provides more details of our methodology to choose products in the set $\mathbb{C}$.

3 Product Level Regressions

In this section, we implement equation 9 with the barcode level data in the Nielsen Consumer Panel. 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. 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. Each barcode is classified into one of 104 product categories.[7] For example, a 31 oz bag of Tide Pods (UPC 037000930389) is mapped to product category "Detergent." Nielsen constructs projection weights that make the sample representative of the US urban population. We use these weights to calculate total expenditures of each barcode.

Our sample covers all 16 years available of the Nielsen Consumer Panel Data (2004-2019). 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. 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_i is estimated separately for each of them.

To estimate β_i we need to calculate the aggregate price index, which we calculate through an iterative procedure. First, we calculate the conventional price index of each of the 104 product categories in the Nielsen data using the entire set of barcodes available in each

[7]Barcodes in the Nielsen Consumer Panel are organized into a hierarchical structure. Each UPC 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. Throughout the paper we refer to product groups as product categories.

product category. This step requires an estimate of the elasticity of substitution across barcodes within product categories, which we obtain following the procedure developed by Feenstra (1994) and extended by Broda and Weinstein (2006) and Broda and Weinstein (2010).[8] 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 category is our initial guess for the aggregate price of the product category. Second, we estimate equation 9 for each barcode in the Nielsen data using a linear regression and select barcodes using our selection criteria (i.e. those whose estimated β_i is negative and statistically different than zero at the 1% level). Third, we calculate a new product category 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.[9] The outcome of this procedure is both, a set of barcodes where the unit price is likely to fully capture changes in market share and an unbiased product category price index that is calculated using these set of *chosen* barcodes, which we discuss in detail in the 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

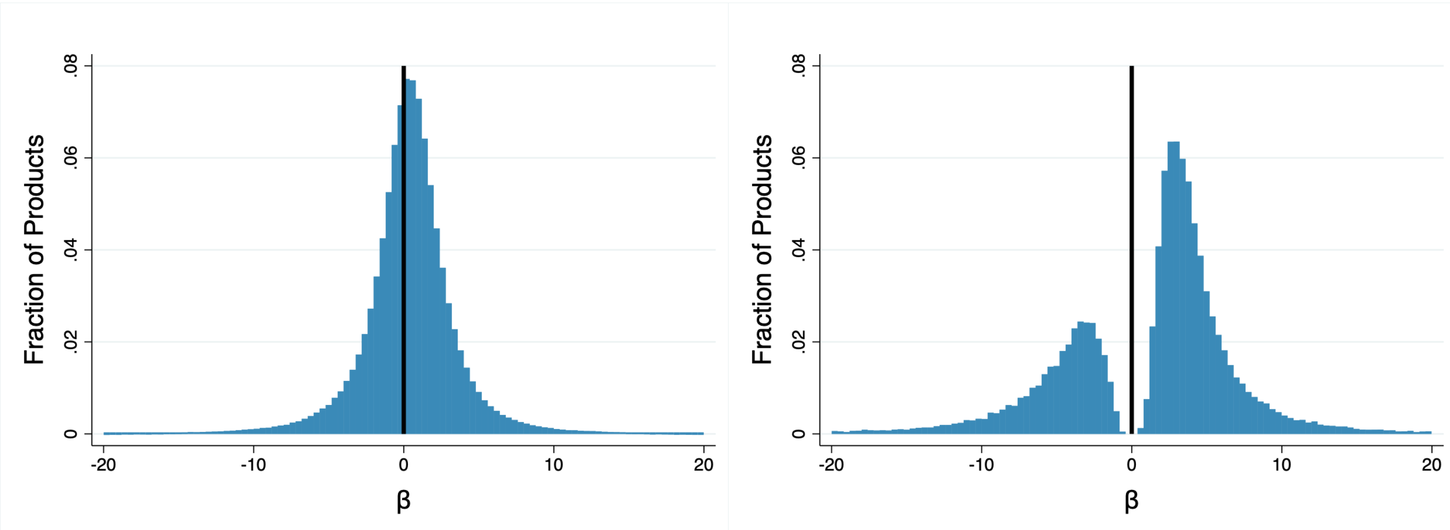
Notes: The table reports descriptive statistics of estimated elasticities of substitution for each product category in the scanner data. We use moment conditions of the double-differenced residuals in demand and supply with the GMM estimation approach.

Panel (a) of Figure 3 shows the distribution of coefficients, β_i. The average estimate for β_i 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 that the change in the residual market share of these products is non-zero.

[8]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 C.

[9]Our convergence criteria is that the difference between the guess and the product category 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).

Figure 3: Price Elasticity of Demand in Nielsen Data



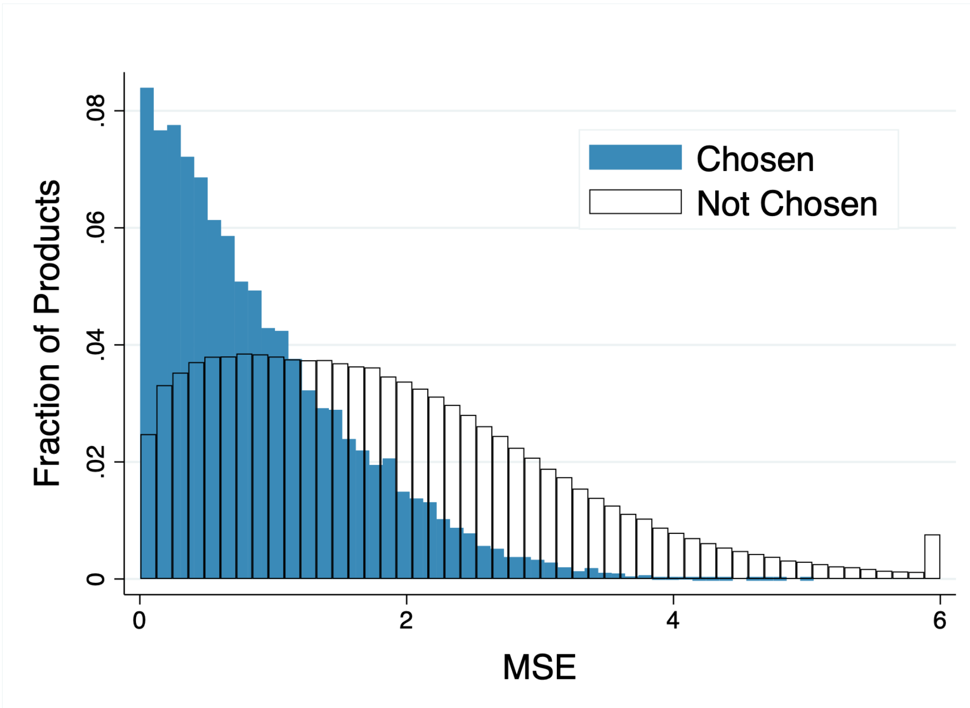
(a) All Barcodes (b) Significant at 1%

Notes: The figure shows the distribution of the coefficients β_i estimated using equation 9 for barcodes in the Nielsen data (2004-2019). Panel (a) includes all barcodes (650,021 barcodes). Panel (b) includes barcodes whose coefficient is significantly different than zero at 1 percent confidence level (18,994 barcodes with negative coefficient and 38,020 barcodes with positive coefficient). The black vertical line marks products whose coefficient is zero.

Panel (b) in Figure 3 shows the distribution of coefficients that are significant at 1%. We choose the barcodes where the coefficients are to the left of the vertical black line; the set of products $\mathbb{C}$, where the unit price is more likely to fully incorporate changes in expenditures. Figure 4 shows the distribution of the mean squared errors for regressions involving products $\mathbb{C}$ and the rest of products. Recall that the telltale sign of horizontally differentiated product is that the change in residual expenditure share is close to zero. Figure 4 shows this is more likely the case for products $\mathbb{C}$ relative to the rest of the products. Products that contribute more barcodes to the chosen sample are less likely to subject to creative destruction. 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.[10]

[10] The product category 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). Table A.I reports the shares of chosen barcodes by department. Less barcodes are chosen from more durable categories, such as "Health and Beauty Aids" and "General Merchandise."

Figure 4: Mean Squared Errors - Product Level Regressions



Notes: Figure shows the histogram of the mean squared errors estimated after implementing equation 9 for barcodes in the Nielsen data (2004-2019). Chosen products are those in the set $\mathbb{C}$ and include barcodes whose coefficient is negative and significantly different than zero at 1 percent confidence level (18,994 barcodes).

4 Estimate of Aggregate Price Index

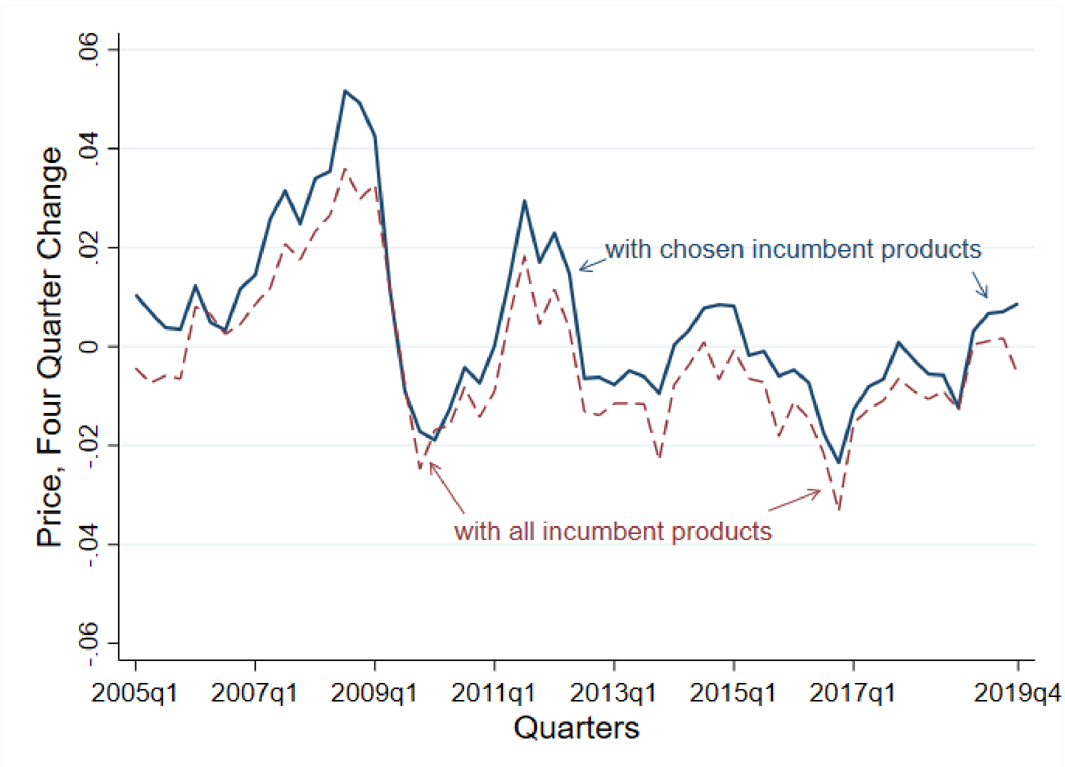
We now have all the necessary ingredients to calculate the aggregate price index. For each of the 104 Nielsen product categories, we have estimates of the elasticity of substitution within the product category, the choice of barcodes in the set $\mathbb{C}$, the price indices for the barcodes in this set, and the expenditure share of the barcodes in the set $\mathbb{C}$.

Figure 5 shows the price indexes calculated using data on prices and expenditure shares of the set $\mathbb{C}$ from equation 2. For comparison, we also estimate the price index using the Feenstra price index formula with data on prices and expenditures of all incumbent products:

$$d\ln P = \sum_{i\in\mathbb{I}} \omega_i \, d\ln\, P_i \, + \, \frac{1}{\sigma-1} d\ln\left(1 \, - \, \text{New} \, + \text{Exit}\right) \tag{10}$$

Figure 5 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 incumbent products. The mean percentage points difference is 0.74.[11]

Figure 5: Price Index with Chosen vs. All Incumbent Products



Notes: Figure shows the one year changes in the price index calculated with chosen (equation 2) vs. all incumbent products (equation 10).

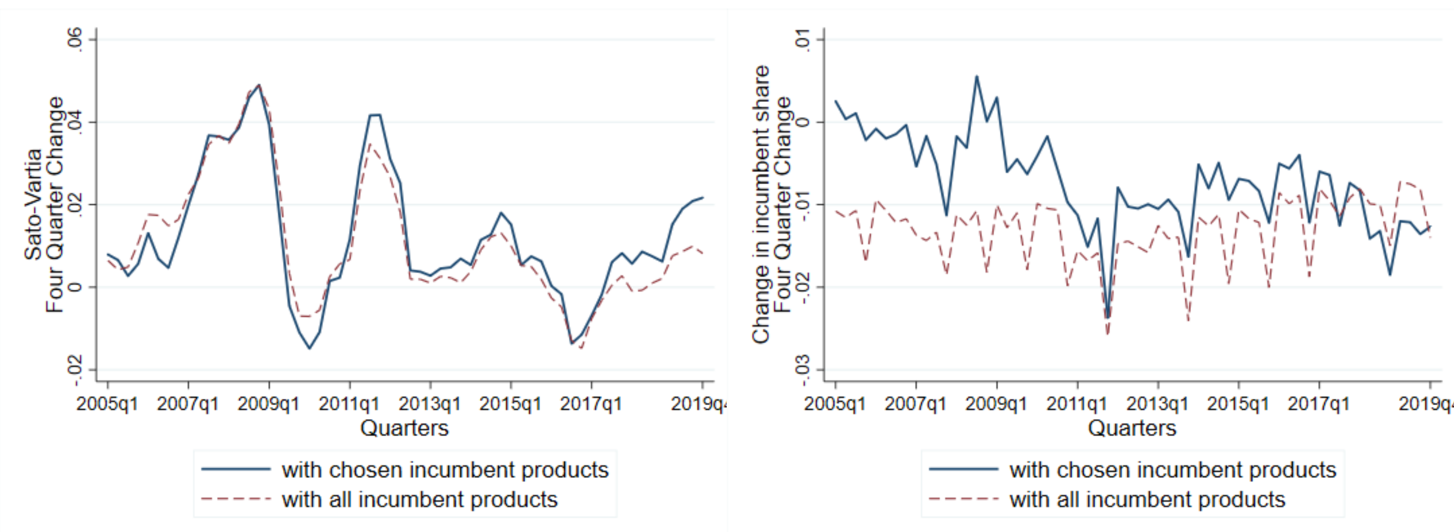
The gap between the two price indices in Figure 5 comes from the barcodes in the Nielsen data, where the regression of changes in expenditures on prices yields a positive coefficient. At the barcode level, a positive regression coefficient indicates either that the residual share of the barcode increased or that the residual share share fell. The latter indicates that the Feenstra (1994) price index formula does not fully account for the effect of the product's price change, and the latter suggests that the Feenstra (1994) formula overstates the contribution of new products. The bias in conventional estimates of the price index depends on the mass of barcodes where positive residual share growth vs. barcodes with negative growth in the residual share. If the former exceeds the latter, the price indices calculated from the unit price of all incumbent products will overstate the true inflation rate, and vice versa.

The change in the two price indices shown in Figure 5 can be decomposed into two

[11]Figure A.1 shows that our results are very similar if we focus on barcodes that are present in the data at least four years to select the set of chosen products.

components: (i) the weighted average of the change in prices of incumbent products and; (ii) the change in incumbent shares adjusted by the elasticity of substitution. Figure 6 shows the one year changes in the two components separately. Note that the second component in equation 10, when estimated using the set of all incumbent products, is commonly used as a measure for the welfare gains from new varieties/innovation. We find that most of the gap between the two price indices in Figure 5 comes the change in the incumbent share adjusted by the elasticity of substitution. Specifically, the difference in the "new variety" terms accounts for almost 80% of the 0.74 percentage point gap in the growth rate of the two price indices.

Figure 6: Two Components in the Price Index with Chosen vs. All Incumbent Products



(a) The first component (Sato-Varta)

(b) The second component (change in incumbent share)

Notes: Figure shows the one year changes in two components in the price index calculated with chosen (equation 2) vs. all incumbent products (equation 10).

Table 2 reports the percentage points differences between price with chosen vs. all incumbent products by department. The conventional price index computed using all incumbent products is downward biased in all 10 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 new or improved ones do not completely replace products that are its close substitute more.

Table 3 shows the weighted average of the change in residual market share and the Sato-Vartia weight of these two types of barcodes in the Nielsen data. The top panel shows these

Table 2: Percentage Points Difference in Price with Chosen vs. All Incumbent Products by Department

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

Notes: Table reports the gap in one year changes in the price index calculated with chosen (equation 2) vs. all incumbent products (equation 10).

two statistics for products where the regression coefficient is positive because both the change in prices and the change in residual expenditure shares are positive. These are the products where the Feenstra (1994) formula does not fully account for the aggregate effect of price changes of the incumbent products. The first row in the top panel shows that the growth in the residual share averages 0.45 percentage points per year.

The bottom panel in Table 3 shows the products where the regression coefficient is positive because the change in prices and residual expenditure shares are both *negative*. These are products that have lost market share because of the entry of new products that are its close substitutes. The first row shows that average growth in the residual market share is -5.24 percent per year in the Nielsen data. So the magnitude of the residual change in the market share is larger for the products where the standard price index formula (equation 10 overstates the contribution of new varieties compared to those where the standard price indices does not fully account for price changes among the incumbent products.

The last row in two panels shows the Sato-Vartia share of these two groups of products. Comparing these two panels, we see that the products where we overstate the contribution of new products also account for a larger share of expenditures compared to the products in the top panel where we understate the contribution of price changes at the product level. The share of this second group is 26% compared to 24.9% of the products in the top panel.

Table 3: Products with Non-Zero Residual Market Shares Growth

Δ Prices > 0, Δ Share > 0	
Average Δ Expenditure Share	0.45%
Share of Products	24.9%
Δ Prices < 0, Δ Share < 0	
Average Δ Expenditure Share	-5.24%
Share of Products	26.0%

Note: Table reports average change in $\ln \tilde{S}$ (in percentage points per year) and the Sato-Vartia share of products where a regression of the change in relative quantity on relative prices yields a positive coefficient. See section 2 for details on the regression. Top panel reports average growth in the residual share and share of products where, on average, Δ prices and Δ residual share are both positive. Bottom panel reports the same for products where Δ prices and Δ residual share are both negative.

The gap in the share of the products and mismeasurement in the aggregate price in top panel vs. the bottom panel is behind the difference in our estimate of the aggregate inflation rate and estimates based on the Feenstra (1994) price index formula in equation 10. When the product of the two rows in the top panel is larger than the product of the two rows in the bottom panel, then the effect of overstating the contribution of new products exceeds the effect of understating the contribution of price changes among incumbent products. In our case, the share of products in the bottom panel is larger than in the top panel. In addition, conditional on mismeasuring aggregate prices, average mismeasurement in the bottom panel is larger than in the top panel. These two forces thus explain why on aggregate we find that true inflation rate is likely to be higher than the inflation rate calculated from equation 10.

5 Residual Market Shares at the Product Level

In this section, we analyze price mismeasurement for individual products to provide useful intuition for the aggregate bias shown in the previous section. We average residual market shares across all periods where the barcode registered positive sales. Table 4 shows the growth in residual expenditure shares (column 1) and unit price (relative to the aggregate price) growth (column 2) grouped by the duration of the barcodes in the Nielsen data. Specifically, our Nielsen data covers the 2004-2019 period, so we group barcodes into those that are present at the beginning of our sample in 2004 and exited prior to the end of the sample ("left censored"), entered after 2004 and exited prior to 2019 ("not censored"), entered after 2004 and are present at the end of the sample ("right censored"), and present in the sample in all periods ("right and left censored").

Table 4: Residual Expenditure Growth at the Product Level

	Residual Market Shares Growth	Relative Unit Price Growth
All (650,021 barcodes)	-1.94	0.47
Not censored (20% of exp.)	-4.85	-0.28
Right censored (32% of exp.)	0.90	0.52
Left censored (19% of exp.)	-4.43	0.53
Right and Left censored (29% of exp.)	0.32	1.07

Notes: Table shows annual average residual market shares growth and relative unit price growth.

Table 4 shows that the change in residual market shares is *negative* among exiting barcodes. Residual expenditure growth averages -4.85 percent per year among barcodes that are not censored and -4.4 percent per year among left censored barcodes. Furthermore, these two groups of barcodes collectively account for almost 40% of the share of expenditures in an average year. This indicates that, among these groups of products, innovation primarily takes the form of creation of new products that are close substitutes for some incumbent products. Thus, conventional prices indices formulas understate the price index particularly for these products. Changes in residual market shares are positive among the two other groups of products (right censored and right and left censored), but the magnitudes are smaller compared to the decline in the residual shares of exiting barcodes.

Figure 7 shows the relationship between residual change in the share and the growth in the relative unit price across all barcodes. We show this relationship separately for each of the four groups of products in Table 4. The difference in the residual change in the share among the left censored and not-censored barcodes and the other two groups seen in Table 4 can be also be seen in this Figure. More importantly, the figure shows that, within each of these groups, residual change in the price-adjusted expenditure share is increasing in the growth of the unit price.

Large representation of products with negative price growth and negative residual market share growth indicates that large amount of incumbent products are negatively affected by entry of new products. When they are partially, not completely, substituted by new products, the change of their unit prices and market shares biases aggregate price measurement. This partial substitution between new products and incumbent products is consistent with the

Figure 7: Residual Market Share Growth vs. Relative Price Growth

Notes: The figure shows the relationship between average residual market share growth and changes in relative price in the Nielsen data for four types of products: left censored, right censored, censored on both sides, and not censored. Average residual market share growth is calculated using equation 8. Change in relative prices are grouped into 100 equal-sized bins for each of the four types of products, and we show the average residual market share growth of each bin.

work by Argente, Lee and Moreira (2021), who find that the sales of individual products decline at a steady pace throughout most of the product life cycle. They find that the decline in the market share of incumbent products can be explained by the introduction of new products that, either through business stealing or cannibalization, *partially* substitute existing products.

Several recent papers have interpret the "new variety" term in Feenstra (1994)'s formula as the measuring the welfare effect of new varieties. Recent example of papers following these interpretation are Broda and Weinstein (2010), who compute welfare gains of new product varieties in the U.S retail sector, and Aghion, Bergeaud, Boppart, Klenow and Li (2019), who estimate it for the entire private non-farm sector. This interpretation is problematic since the new variety term in Feenstra (1994)'s formula does not capture the full effect on the expenditure share of incumbent products. And, for similar reasons, the other term in Feenstra (1994)'s formula, the weighted average of the price change of incumbent products, is also biased.

An alternative method to quantify the gains from new products is to accurately match them to the incumbent varieties affected by their entrance. This is the path attempted by Nordhaus (1997) for a single product, lightning. However, conducting this sort of matching involves, undoubtedly, difficult judgement calls and could bias price indices in particular ways (e.g. residual expenditure share of lightning is negative). Our methodology allows us to bypass these difficulties. And, although it does not allow us to separate the welfare gain from new goods from the effect of price changes among incumbent products with close substitutes, the sum of the two is an unbiased estimate of living standards in the presence of creative destruction.

Our analysis is also related to Bils and Klenow (2001) and Bils (2009) that assess the unmeasured quality growth for durable goods. Quality growth is unmeasured when the BLS tracks prices of incumbent products. One can arrive at vastly different measures of price inflation and real growth under different assumptions and practices on how to deal with product replacement. Bils and Klenow (2001) use the US Consumer Expenditure surveys and estimate "quality Engel curves." Bils (2009) uses CPI microdata to decompose the observed price increases into quality changes and true inflation. In particular, Bils (2009) compares prices and market shares across model changes in vehicles and consumer electronics. In our context, those are sectors where it is less burdensome to identify goods that are adversely affected for each new product. Still, those grouping may not be perfect, and more importantly, not be plausible for other sectors.

6 Conclusion

Standard price index formulas, such as the one developed by Feenstra (1994), account for the introduction of new products that are new horizontal varieties or for new varieties that are perfect substitutes for previously existing products. However, when new or improved products do not completely replace products that are its close substitutes, the standard price index formula exaggerates the contribution of new products and understates the effect of the improvement of existing products.

We proposed a new methodology to measure the price index when innovation creates new products or improves existing products that have close but not perfect substitutes. Our method does not require that we match new products with the exiting products or products with shrinking market share that are its close substitutes, but it does require that we identify a set of products where changes in their expenditure share is fully explained by the change in their prices (i.e. products that are not likely to be close substitutes of new or improved products).

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 change in the market share of this bundle, adjusted by a function of the price elasticity of demand. The sum of the two terms is an unbiased estimate of the *net* effect of price changes of all incumbent products and the welfare gain from new products, even if we can not separately identify these two forces. We use barcode level data on prices and expenditures from 2004-2019 to implement our methodology. We find that the aggregate price of consumer packaged goods exceeds the price calculated from the standard Feenstra (1994) formula by 0.74 percentage points per year.

References

Aghion, Philippe and Peter Howitt, "A Model of Growth Through Creative Destruction," *Econometrica*, 1992, pp. 323–351.

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

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

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

_ , Munseob Lee, and Sara Moreira, "The Life Cycle of Products: Evidence and Implications," *mimeo*, 2021.

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," *American Economic Review*, 1994, pp. 157–177.

Grossman, Gene M and Elhanan Helpman, "Quality Ladders in the Theory of Growth," *The Review of Economic Studies*, 1991, *58* (1), 43–61.

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.

Romer, Paul M, "Endogenous technological change," *Journal of Political Economy*, 1990, *98* (5, Part 2), S71–S102.

APPENDIX

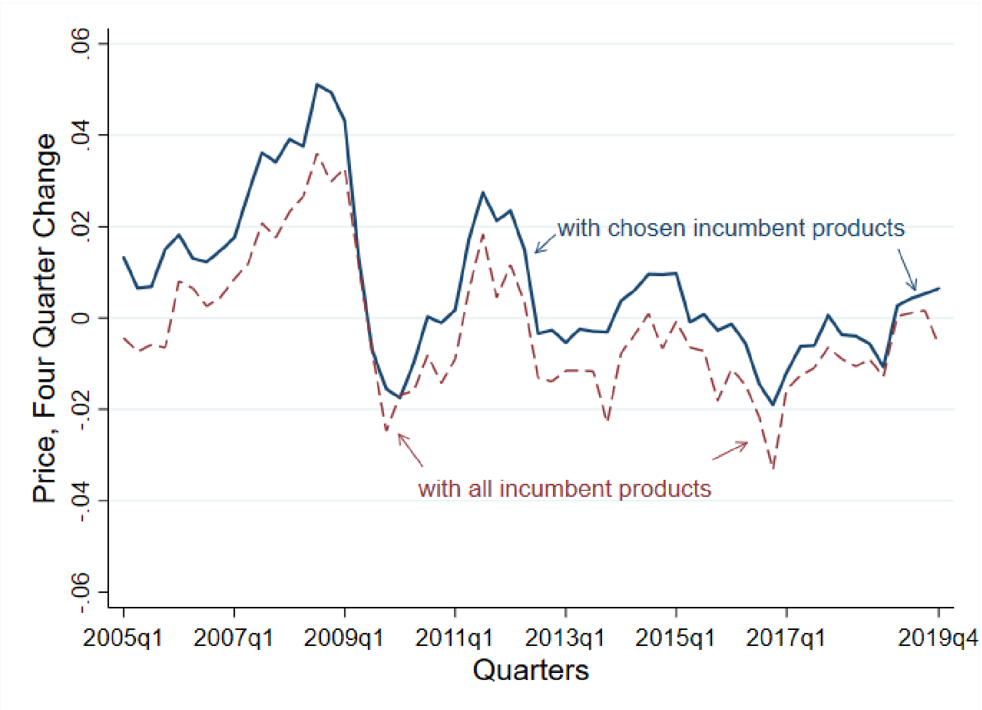
A Additional Figures and Tables

Table A.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.

Figure A.1: Price Index with Chosen vs. All Incumbent Products - Long Duration



Notes: Figure shows the one year changes in the price index calculated with chosen (equation 2) vs. all incumbent products (equation 10). The chosen set of products is drawn from those that lasted at least four years in the data.

B Proof

For incumbent $\mathbb{H}$ goods, the expenditure share of good i over total expenditure is defined from the CES structure as follows:

$$s_i \equiv \frac{\text{Expenditure}_i}{\text{Total Expenditure}} = \left(\frac{P}{P_i}\right)^{\sigma-1} \tag{11}$$

For incumbent $\mathbb{V}$ goods, the expenditure share of good i over total expenditure is defined from the CES structure as follows:

$$s_i \equiv \frac{\text{Expenditure}_i}{\text{Total Expenditure}} = \left(\frac{P}{P_g}\right)^{\sigma-1} \left(\frac{P_g}{P_i}\right)^{\rho-1} \quad \text{for } i \in g \tag{12}$$

Define $\widetilde{P}_i$ as an ideal price index of good i such that

$$s_i = \left(\frac{P}{\widetilde{P}_i}\right)^{\sigma-1} \tag{13}$$

Then, $\widetilde{P}_i$ becomes

$$\widetilde{P}_i = P_i \tag{14}$$

for $\mathbb{H}$ goods and

$$\widetilde{P}_i = \left(\frac{P_i^{\rho-1}}{P_g^{\rho-\sigma}}\right)^{\frac{1}{\sigma-1}} \tag{15}$$

for $\mathbb{V}$ goods.

Define s_I as the expenditure share of all incumbent goods over total expenditure:

$$s_I \equiv \frac{\text{Expenditure}_I}{\text{Total Expenditure}} = \left(\frac{P}{Q}\right)^{\sigma-1} \tag{16}$$

where Expenditure_I is an expenditure on all incumbent goods and Q is a CES price of representative incumbent product.

By taking logs, we get

$$d\ln P = \frac{1}{\sigma-1} d\ln s_I + d\ln Q \tag{17}$$

Define $\widetilde{s}_i$ as the expenditure share of good i over the expenditure on all incumbent goods. Then, following the CES structure, we get

$$\widetilde{s}_i \equiv \frac{\text{Expenditure}_i}{\text{Expenditure}_I} = \left(\frac{Q}{\widetilde{P}_i}\right)^{\sigma-1} \tag{18}$$

By taking logs, we get

$$d\ln Q = \frac{1}{\sigma-1} d\ln \widetilde{s}_i + d\ln \widetilde{P}_i \tag{19}$$

By applying Sato-Vartia weights and summing across all incumbents, we get

$$\begin{aligned} d\ln Q &= \sum_{i\in\text{incumbents}} \omega_i d\ln \widetilde{P}_i \\ &= \sum_{i\in\text{incumbents in } \mathbb{H}} \omega_i d\ln \widetilde{P}_i + \sum_g \sum_{i\in\text{incumbents in } \mathbb{V}_g} \omega_i d\ln \widetilde{P}_i \end{aligned} \tag{20}$$

where the Sato-Vartia weights within incumbent products are given as

$$\omega_i \equiv \frac{\frac{d\widetilde{s}_i}{d\ln \widetilde{s}_i}}{\sum_{i\in \text{incumbents}} \left[\frac{d\widetilde{s}_i}{d\ln \widetilde{s}_i}\right]} \tag{21}$$

By combining equations 20 with 14 for $\mathbb{H}$ goods, and 15 for $\mathbb{V}$ goods, we get

$$d\ln Q = \sum_{i\in \text{incumbents in } \mathbb{H}} \omega_i d\ln P_i + \sum_g \sum_{i\in \text{incumbents in } \mathbb{V}_g} \omega_i \frac{1}{\sigma - 1} \left[(\rho - 1) d\ln P_i - (\rho - \sigma) d\ln P_g\right] \tag{22}$$

From the CES structure, we know

$$d\ln P_g = \frac{1}{\rho - 1} d\ln s_{I,g} + d\ln P_i \tag{23}$$

where the expenditure share of incumbent goods within g is defined as

$$s_{I,g} = \frac{\text{Expenditure}_{I,g}}{\text{Expenditure}_g} \tag{24}$$

where $\text{Expenditure}_{I,g}$ is expenditure on incumbent goods in g and Expenditure_g is total expenditure in g.

By combining equations 22 and 23, we get

$$d\ln Q = \sum_{i\in \text{incumbents in } \mathbb{H}} \omega_i d\ln P_i + \sum_g \sum_{i\in \text{incumbents in } \mathbb{V}_g} \omega_i \left[d\ln P_i - \frac{\rho - \sigma}{(\sigma - 1)(\rho - 1)} d\ln s_{I,g}\right] \tag{25}$$

By combining equations 17 and 25, we get

$$d\ln P = \frac{1}{\sigma - 1} d\ln s_I + \sum_{i\in \text{incumbents}} \omega_i d\ln P_i - \frac{\rho - \sigma}{(\sigma - 1)(\rho - 1)} \sum_g \omega_g d\ln s_{I,g} \tag{26}$$

Recall that s_I is the share of all incumbents with respect to total expenditure and $s_{I,g}$ is the share of incumbents within g with respect to expenditure on g. The first two components are the same as the standard price index (Feenstra, 1994; Broda and Weinstein, 2010), and the last component represents the bias. Note that when $\rho = \sigma$ the bias becomes zero.

C 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):

$$\Delta^{\underline{u},t}\ln S_{ig} = (1-\sigma_g)\Delta^{\underline{u},t}\ln P_{ig} + \iota_{ig} \tag{27}$$

$$\Delta^{\underline{u},t}\ln P_{ig} = \frac{\delta_g}{1+\delta_g}\Delta^{\underline{u},t}\ln S_{ig} + \kappa_{ig} \tag{28}$$

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 27 and 28 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:

$$G(\beta_g) = E_T[\nu_{ig}(\beta_g)] = 0 \tag{29}$$

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:

$$\hat{\beta}_g = \arg\min_{\beta_g \in B} G^*(\beta_g)' W G^*(\beta_g) \quad \forall\, i \in \omega_b \tag{30}$$

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