COST OF LIVING INEQUALITY DURING THE GREAT RECESSION

David Argente
Pennsylvania State University

Munseob Lee
University of California San Diego

Abstract
We construct income-specific price indexes for the period from 2004 to 2016. We find substantial differences across income groups that arise during the Great Recession. The difference in annual inflation between the lowest quartile of the income distribution and the highest quartile was 0.22 percentage points for 2004–2007, 0.85 percentage points for 2008–2013, and 0.02 percentage points for 2014–2016. We find that product quality substitution and changes in the shopping behavior, margins mostly available to richer households, explain around 40% of the gap. Our evidence shows that not accounting for these differences in price indexes could lead to significant biases in the calculation of consumption and income inequality. (JEL: D12, E31, E32, I30)

1. Introduction

For decades, economists have tried to study the patterns of poverty and inequality by focusing only on the disparities in nominal income or nominal wages. The measures of real income inequality that account for differences in the cost of living are arguably better measures of the differences in the standard of living. However, the literature largely ignores this dimension due to the difficulty in calculating price indexes at the income-group level or at finer levels of disaggregation. If price indexes differ across income groups, then estimates of poverty and inequality could be misleading. This

The editor in charge of this paper was Claudio Michelacci.

Acknowledgments: We thank Fernando Alvarez, Erik Hurst, Anil Kashyap, Robert Lucas, and Joseph Vavra for their advice and support. This paper benefited from comments and suggestions by many others, including Bong Geun Choi, Jorge L. Garcia, Jessie Handbury, Chang-Tai Hsieh, Francesco Lippi, Claudio Michelacci, Sara Moreira, Emi Nakamura, Roberto Robatto, Robert Shimer, Jón Steinsson, Nancy Stokey, Chen Yeh, as well as numerous seminar participants. We thank Jacob Orchard for providing exceptional research assistance.

E-mail: dargente@psu.edu (Argente); munseoblee@ucsd.edu (Lee)

Journal of the European Economic Association 2020 00(0):1–40 DOI: 10.1093/jeea/jvaa018

problem could be worse during downturns when the well-known upward bias in the consumer price index (CPI) is larger.[1]

In this paper, we use detailed scanner data collected at the household level to calculate income-specific price indexes for the period 2004–2016. Relying on the method developed by Feenstra (1994), we calculate exact price indexes (EPIs) for different income groups. We find substantial differences across income groups. The difference in annual inflation between the lowest quartile of the income distribution and the highest quartile was 0.22 percentage points for 2004–2007, 0.85 percentage points for 2008–2013, and 0.02 percentage points for 2014–2016. That is, the gap arises at the onset of the Great Recession, grows moderately during the subsequent recovery, and stabilizes after 2014, precisely when variables such as output, household income, and employment reached their prerecession levels.[2] The gap persists after correcting for changes in product variety by households, and it is robust to using a wider set of price indexes.[3]

To understand the source of the gap, we decompose the EPI into four components: (i) change in product prices, (ii) product substitution, (iii) shopping behavior, and (iv) new goods bias adjustment. The first component captures changes in EPI when the expenditure weights are fixed to the level of the previous period. This term resembles the geometric mean formula used by statistical agencies because, due to data limitations, they are unable to update the expenditure weights of products every period and captures changes in the price index attributable to changes in average prices. The second term captures the extent to which, given average prices, the EPI changes due to changes in expenditure weights that result from consumers substituting across varieties. The third term captures differences in shopping behavior, the extent to which consumers can adjust the prices they pay for an item by adjusting the outlet where they purchase, searching for discounts, or using coupons. The last component is the conventional new goods bias, referred to interchangeably as the variety correction term, developed by Feenstra (1994) and Broda and Weinstein (2010). This component captures the importance of quality shifts due to the creation and destruction of varieties.

The magnitude of these four components varies across households because consumers purchase different bundles of goods, differ in the way they adopt new varieties, and respond differently to price changes by substituting across varieties and adjusting their shopping behavior. In fact, when we decompose the gap between the

1. The conventional price indexes do not capture changes in shopping behavior (e.g. store switching, coupon usage, purchases of items on sales, buying in bulk, product quality substitution, etc.), which are prevalent during economic downturns (see, e.g., Chevalier and Kashyap 2019; Coibion, Gorodnichenko, and Hong 2015; Kryvtsov and Vincent 2019; Handbury, Watanabe, and Weinstein 2015). Furthermore, the BLS collects information only on a limited variety of products within a category; therefore, the CPI does not accurately reflect substitution within a category.

2. Pistaferri (2016) documents, for example, that nondurable consumption stopped growing altogether during the recession, and it is only since 2014 that signs of recovery have started appearing.

3. If we use a Törnqvist index, gap is approximately 0.31 percentage points for 2004–2007, 0.63 percentage points for 2008–2013, and −0.53 percentage points for 2014–2016. This index does not take into account changes in product variety over time.

EPI of the highest income quartile and that of the lowest quartile into these components during the recession period, we find that 46% of the gap can be attributable to changes in product prices, 29% to differences in within-category product substitution, 13% to changes in the adoption of product varieties, and 12% to differences in shopping behavior. The decomposition shows that a large part of the gap, particularly during the Great Recession, can be explained by the fact that high-income households were able to cope with the recession by substituting across varieties and adjusting their shopping behavior to a larger extent than low-income households.

We then focus on documenting the extent to which households of different income groups take advantage of these margins in response to changes in aggregate economic conditions. We show that the quality substitution margin is available mostly for high-income households by showing that the relationship between quality (measured by unit price paid) and income is flat for households below the median income. During the recession, this relationship flattens for households above the median income, indicating that they were able to reduce the prices they paid within a category of products by trading down in quality. Although this relationship, also known as the quality Engel curve, has gotten steeper for low-income households recently, by 2016 its slope did not reach its prerecession level, indicating that in recent years the ability of households to take advantage of the quality substitution margin has increased with their income.

High-income households also changed their shopping behavior more relative to low-income households during the Great Recession. For example, from 2008 to 2013, higher income households drastically increased the fraction of coupons they use and the fraction of sale purchases relative to low-income households. In order to summarize the impact of these and other shopping activities on the prices paid by each income group for an identical product, we calculate the price index developed by Aguiar and Hurst (2007). Our estimates indicate that before the Great Recession low-income households paid on average 2% lower prices for the same barcoded item than richer households. This gap completely disappeared by 2013.

We then provide direct evidence that the responses of the high- and low-income households differ when they suffer an income decline by taking advantage of the longitudinal dimension of our data. We investigate three major ways consumers can adjust when they suffer income declines. The first margin is product quality substitution: the reallocation of expenditures from high- to low-quality brands within the same product category. By combining retail-level scanner data and household-level panel data, we are able to measure the degree of product quality substitution for each household. The second is store switching: the reallocation of expenditures towards lower priced retailers. And the third is the frequency at which consumers purchase and the types of purchases they make, such as purchases of items on sale. We explore how households adjust along these margins when local economic conditions deteriorate. We find that households with higher income are better able than lower income households to reduce the average prices they pay after an increase in the unemployment rate of the county where they live. Even though high-income households respond along other margins, such as store switching, the reallocation of expenditures within product categories is the most effective margin to reduce the prices they pay in a given category. This result also holds when we use within-household variation in income.

These findings have important implications for the measurement of real variables such as income and wealth as well as other economic indicators such as the poverty rate and inequality, particularly during economic downturns. They point towards the need for statistical agencies to produce income-specific price indexes. Although the production of these indexes is likely to be costly, an accurate measurement of these indexes will have, undoubtedly, important implications for the measurement of other real variables and could affect the interpretation of the studies that estimate both consumption and income inequality. In particular, given that both the real income and the capacity to spend of low-income households eroded more drastically, official measures could be understating the increases in inequality during the Great Recession.

The rest of this paper is organized as follows: We discuss the related literature in Section 2. Section 3 presents the data used for the empirical analysis. In Section 4, we construct the income-specific price indexes and present evidence that the gap in the price indexes of the top and bottom income groups is robust to using a variety of price indexes. In Section 5, we decompose the gap and discuss the contribution of each component over three subperiods: prerecession, recession, and postrecession. Section 6 presents both cross sectional and longitudinal evidence of the changes in shopping behavior and product substitution across income groups. In Section 7, we discuss the potential implications of our findings and conclude.

2. Related Literature

This paper contributes to various strands of the literature. Our work relates to the literature that focuses on the differences in the cost of living across households. The studies by Michael (1979), Hagemann (1982), and Garner, Johnson, and Kokoski (1996) focus on computing group-specific inflation rates. More recently, Hobijn and Lagakos (2005) explore inflation inequality across households in the United States and find that the average difference in inflation between poor and nonpoor people is less than 0.1%. These studies use published indexes of average prices for broad categories and allow only the shares of expenditures of these categories to vary across groups. As a result, they assume that all households pay the same price for the same product and that all households purchase the same mix of products within a category.

In this paper, we take advantage of detailed scanner data to allow both the share of expenditures of different products within categories and the prices paid for each product to vary across income groups, as in Broda and Romalis (2009). They find that from 1994 to 2005, the inflation for poorer consumers was lower than the inflation for richer consumers and argue that half of the increase in conventional inequality measures was due to a bias caused by ignoring the variation in consumption behavior across income groups. Our paper updates their estimates and builds on their work by computing income-specific elasticities of substitution within product categories to allow for nonhomotheticities across groups. Our findings show that in the recent periods, the inflation for richer households has been lower than that of poorer households. Given the similarity of our methodologies, the differences in our results are likely due to the sample periods examined and the fact that the more recent data are richer. Nielsen no longer supplies the data from 1994 to 2003 so we are unable to replicate their exercise. Our data set is much richer and allows us to compute both the new goods bias correction term and the expenditure weights for each income group every quarter. Broda and Romalis (2009) are unable to compute an EPI for most of the periods in their sample due to lack of household-level information. They examine two different baskets of goods: food and nondurables. For food, which represents 50%–70% of expenditures in their sample, they are unable to compute the EPI for half of the periods and construct a Paasche index instead. Due to the same limitation, they use extrapolations of the new goods bias correction term for that part of the sample. For nondurables, they have household-level information of only one quarter (2003:4). As a result, they construct a Paasche index for most of their sample periods and use the same new goods bias term for all groups.

In subsequent work, Kaplan and Schulhofer-Wohl (2017) and Jaravel (2019) confirm the inflation disparities we find but focus on different aspects of them. Kaplan and Schulhofer-Wohl (2017) focus on understanding the overall distribution of inflation rates across households and find that low-income households experience higher inflation, even after controlling for all other demographics. They estimate that if the differences in inflation rates for goods in our data set extended to the universe of goods and services, they would imply that the difference in real incomes between the top and bottom groups was growing at a rate of nearly 1 percentage point per year faster than the difference in nominal incomes. Jaravel (2019) focuses on the causes of the disparities and argues that increases in the relative demand for products consumed by high-income households led firms to introduce more new products to cater such households. As a result, continuing products decreased their prices due to an increased competitive pressure. His mechanism refers to a long-term trend, whereas we focus on documenting the sources of the difference between rich and poor and how this difference changes over time and in response to the variation in economic conditions. His mechanism is only part of the explanation behind the gap we find given that, during the periods in which we see the inflation disparities rising, the economy experienced a dramatic drop in demand, particularly for goods purchased by high-income households, and a significant decrease in the product entry rate, which is highly cyclical.[4] In fact, we estimate that the adoption of new varieties can explain at most 15% of the inflation disparities we observe.

Our study is also related to the extensive literature that examines the heterogeneity of shopping behavior across income groups. Aguiar and Hurst (2007) find that low-income households shop more intensively and typically pay lower prices for identical products. Griffith et al. (2009) use a national representative sample of UK households

4. See Broda and Weinstein (2010) and Argente, Lee, and Moreira (2018) for evidence on the procyclicality of product entry.

in 2006 and find that the savings from sales and buying in bulk are of a similar order of magnitude to those from purchasing generics and store switching. Our findings show the importance of quality substitution within product categories as a margin of adjustment, particularly for high-income households. Furthermore, the cyclical aspect of the inflation disparities, those explained by adjustments in the shopping behavior of households, is related to the growing literature on how shopping behavior changes over the business cycle. The studies by Nevo and Wong (2019) and Stroebel and Vavra (2019) show that households lower their grocery bill during economic downturns by increasing their shopping effort. McKenzie and Schargrodsky (2011) show that an increase in the activity of searching for shopping was one of the most prevalent adjustment mechanisms used by consumers to cope with the 2002 Argentinean financial crisis. Coibion, Gorodnichenko, and Hong (2015) find that the inflation in effective prices declines significantly with higher unemployment, whereas little change occurs in the posted price's inflation. This difference, they argue, reflects the reallocation of household expenditures across retailers, particularly by households at the highest income quintile. Although store switching was prevalent during the Great Recession, we find that product quality substitution contributed more to the decline of effective prices during this period, particularly for richer households. This finding is consistent with the work by Burstein, Eichenbaum, and Rebelo (2005), who show that a large measurement error emerged in the CPI after households substituted to low-quality goods in the aftermath of Argentina's devaluation, and with the work by Jaimovich, Rebelo, and Wong (2019), who show that higher quality goods lost market share during the Great Recession due to consumers trading down in quality.

3. Data

To construct income-group-specific price indexes, we need to observe not only the households' demographic information but also the details on their purchases. The Nielsen Consumer Panel data set, which is provided by the Kilts-Nielsen Data Center at the University of Chicago Booth School of Business, is the best available source for this information. We use this data set to construct our price indexes. One limitation of the Nielsen Consumer Panel data set is that it does not contain all of the transactions that take place in the market. Given that we can observe only transactions made by households in the sample, in order to directly measure the degree of product quality substitution and store switching, we combine the Nielsen Consumer Panel data set with the Nielsen Retail Scanner data set. Nielsen collects its scanner data at the store level. Thus, these data contain all of the transactions made in a given store-week for each product category covered. We use the Consumer Panel data set from 2004 to 2016 and the Nielsen Retail Scanner data set from 2006 to 2016, which are all the years available for both data sets. Furthermore, in order to approximate the changes in local economic conditions, we use the Bureau of Labor Statistics's (BLS) monthly unemployment data available at the county level. In the following, we discuss each data source in detail and provide information on how we assembled the data.

3.1. Nielsen Consumer Panel Data Set

The Nielsen Consumer Panel data set tracks the shopping behavior of 40,000–60,000 households every year covering 49 states and 2,967 counties in the United States. Each panelist uses in-home scanners to record their purchases.[5] A twelve-digit universal product code (UPC) identifies the items the panelists purchase. The data contain around 3.27 million distinct UPCs grouped into 1,235 product modules that range from food to beauty aids to computer software. Throughout this paper, we refer to product modules as product categories. Our data cover around 40% of all of the expenditures on goods in the CPI.[6]

To check the consistency of the data, Figure A.1 in Appendix A plots a price index constructed using the Nielsen Consumer Panel data set and the BLS's food-at-home CPI for all urban consumers. We use a procedure that mimics the construction of the CPI, and we pick product groups in our data that match those used in the construction of the CPI-U food-at-home. As shown in the figure, our index closely matches the overall patterns of that index.

For each UPC, the data contain information on the brand, size, and packaging, as well as a rich set of other product features. If the panelist purchases the good at a store covered by Nielsen, the price is automatically set to the average price of the good at the store during the week when the purchase was made. If not, the panelist directly enters the price. Nielsen reports detailed transaction information for each product purchased (e.g. UPC code, quantity, price, deals, and coupons). We combine this information with the weight and volume of the product to compute unit values. Because our data lack other measures of quality, we follow the industrial organization and the international trade literature and approximate the quality of a product with its unit value.[7]

The data also contain information about each purchasing trip the panelist makes, such as the retailer, the location, and the date of the transaction. Furthermore, the data have demographic variables such as age, education, annual income, marital status, and employment, which are updated annually based on surveys sent to the panelists. The surveys are sent in the last quarter of each year and the variables are implemented in the first week of January of the following year. Nielsen provided 16 income bins top-coded at $100,000 up to 2005. Since 2006, it has provided 20 income bins top-coded at $200,000. Nielsen asks panelists to report their combined total household annual

5. Nielsen offers a variety of incentives to join and stay active such as monthly prize drawings, gift points, and regular sweepstakes. The incentives are designed to be nonbiasing (i.e. Nielsen does not provide account-specific coupons out of concern for the potential effect on the natural purchase selection of outlets and products).

6. Table A.1 in Appendix A depicts the distribution of UPCs and expenditures across different product groups, as defined by Nielsen. The table shows that our sample includes a wide set of goods, including a few durable goods (e.g. cameras, flashlights, and cookware). Unfortunately, our data do not include other important durables. Our results should be interpreted accordingly.

7. Nielsen provides information on the amount of each product and the unit of measure (e.g. 16 ounces). It also provides information on whether the product is part of a multipack and the units in the multipack. We calculate the total units of a product multiplying this information with the size of the product.

income as of the year-end of the previous calendar year. Nielsen believes panelists are actually reporting their "annualized" estimated income as of the time of the survey and not referring to the previous year's tax returns. Self-reported annual income is likely to be the total labor income. Nielsen constructs projection weights that make the sample representative of the US population that we use in all our calculations.[8]

We restrict our sample to households whose head is between 25 and 64 years of age because the reported annual income might not accurately represent the households' income sources after retirement. Furthermore, a retiree's low opportunity cost of time has a direct bearing on the total cost of consumption, which makes market expenditures a poor proxy for actual consumption (Aguiar and Hurst 2005, 2007). Figure A.2 in Appendix A shows the age profile of the households' average annual income in the Nielsen Consumer Panel data. As the figure shows, it mimics the standard life cycle of earnings.

Each year we divide the households into four groups according to their annual income: (i) less than \$25,000; (ii) \$25,000–\$50,000; (iii) \$50,000–\$100,000; and (iv) more than \$100,000. These groups roughly approximate the quartiles of the cross-sectional distribution of household income. The median household income increased from \$57,674 in 2004 to \$59,534 in 2007, and then dropped to \$54,569 in 2012 during the Great Recession. It recovered to \$60,309 in 2016.[9] On average, approximately 85% of the households remain in the same income bin the following year.

3.2. Nielsen Retail Scanner Data Set

The Retail Scanner data consist of more than 100 billion observations of weekly prices, quantities, and stores, which cover approximately \$2 trillion in sales. This volume represents about 53% of all sales in grocery stores, 55% in drug stores, 32% in mass merchandisers, 2% in convenience stores, and 1% in liquor stores. The collection points include more than 40,000 distinct stores from around 90 retail chains in 49 states and 2,500 counties. The data consist of approximately 3.34 million distinct products, each classified following the same structure as those covered by the Nielsen Consumer Panel data. In comparison to the Nielsen Consumer Panel, the Retail Scanner data cover a wider range of products because it reflects the universe of transactions for the categories it covers, as opposed to the purchases of a sample of households. Overall, the data represent roughly 30% of the total US expenditures on food and beverages and roughly 2% of the total household consumption (Beraja, Hurst, and Ospina 2019).

8. Nielsen has a comprehensive program of dropping and replacing panelists that do not perform to minimum reporting standards. Currently, Nielsen retains about 80% of its active panel each year. Nielsen uses a stratified sampling design to ensure that the panel is demographically balanced.

9. Einav, Leibtag, and Nevo (2010) find that income is not systematically correlated with the quality of recording in the Nielsen Consumer Panel data. Furthermore, to check how representative the households in our data are, we compare the expenditures in the Nielsen Consumer Panel with those of the food-at-home category in the Consumer Expenditure Survey (CEX). For each income group, the expenditure-to-income ratios in both data sets are remarkably similar. For the poorest households, the ratio is 0.25 in the Nielsen Consumer Panel and 0.22 in the CEX. For the highest income households, the ratio is 0.04 in both data sets.

We use the unique store identifier, the barcode of each product, and the date of the purchase to match the information on household purchases contained in the Nielsen Consumer Panel data with transactions at the store level in the Retail Scanner data.

4. Income-Specific Price Indexes

Our analysis begins by calculating the income-specific price indexes with the Nielsen Consumer Panel data. We document a widening gap between the price index of the lowest income group and that of the highest income group. The gap emerges after 2008, stabilizes around 2013, and remains constant until 2016, which is the last year the Nielsen Consumer Panel data are available. Our benchmark specification uses the EPI developed by Feenstra (1994) and Broda and Weinstein (2010) because it is designed to account for the substitution by consumers away from goods whose price is increasing as well as to capture the impact of new and disappearing products. Our findings are, nonetheless, robust to using other commonly used price indexes, including superlative indexes, which do not account for changes in product variety but account for substitution effects (e.g. Fisher, Törnqvist), and fixed-weighted indexes, which do not account for changes in product variety or substitution effects, but that are commonly used by statistical agencies.

4.1. Income-Specific Exact Price Index

Price indexes can vary across households because households buy different bundles of goods at different prices due to differences in tastes, differences in price elasticities, or because households pay different prices for the same goods. To account for differences in tastes, we follow Broda and Romalis (2009). They calculate the price index developed by Feenstra (1994), which allows the set of UPCs and the expenditure shares to differ across income groups and which assumes that each group faces the same elasticity of substitution. This specification reflects differences in tastes across groups but does not reflect differences in price elasticities. In order to account for these differences, we estimate income-specific elasticities of substitution within product categories. This approach implicitly allows for nonhomotheticities, because the elasticities of substitution are allowed to vary across groups. Lastly, to account for differences in the prices paid for the same product, we use the actual price paid by the household, as opposed to the posted price, to calculate the price indexes. Due to data limitations, most of the previous research on the heterogeneity in price indexes has focused exclusively on the variation in consumption bundles and has assumed that all households pay the average price for each category of good. By considering a variation in the prices paid within product categories, our price indexes also capture differences in the consumers' shopping behavior.[10]

10. This approach is also used by Kaplan and Schulhofer-Wohl (2017). As Chevalier and Kashyap (2019) show, some consumers chase discounts, and, thus, the actual prices paid are substantially lower than the

We use a two-tiered Constant Elasticity of Substitution (CES) aggregator to specify the utility function of income group I. The upper level is

$$U_t^I = \left(\sum_{g \in G} (C_{gt}^I)^{\frac{\sigma_I - 1}{\sigma_I}} \right)^{\frac{\sigma_I}{\sigma_I - 1}}, \tag{1}$$

where the product categories are indexed by g, σ_I is the elasticity of substitution across categories by income group I, and G is the set of all product categories. The set G is fixed over time and is equal across income groups. We model the lower tier as

$$C_{gt}^I = \left(\sum_{u \in U_{gt}^I} (d_{ugt}^I C_{ugt}^I)^{\frac{\sigma_g^I - 1}{\sigma_g^I}} \right)^{\frac{\sigma_g^I}{\sigma_g^I - 1}}, \tag{2}$$

where C_{ugt}^I is the total quantity of UPC u that is consumed in product category g by income-group I at time t, σ_g^I is the income-specific elasticity of substitution within product category g, and U_g^I is the set of all possible UPCs within a product category g that group I consumes. The set of existing UPCs in period t is a subset of this set (i.e. $U_{gt}^I \subset U_g^I$) and can vary over time. The parameter d_{ugt}^I captures the perceived quality of each UPC at the income-group level. We assume that it is constant over time (i.e. $d_{ug,t-1}^I = d_{ugt}^I$) just as in Feenstra (1994).

If the set of UPCs available for each group is fixed over time, Sato (1976) and Vartia (1976) derive the EPI in the case of any multilevel CES utility function as

$$\prod_{g \in G} \left\{ \prod_{u \in U_g^I} \left(\frac{p_{ugt}^I}{p_{ugt-1}^I} \right)^{\omega_{ugt}^I} \right\}^{\omega_{gt}^I}. \tag{3}$$

This is the geometric mean of the price changes in each UPC u for income group I that belong to the set U_g^I, where the weights are ideal log-change weights, which are computed as follows:

$$\omega_{ugt}^I = \frac{\dfrac{s_{ugt}^I - s_{ug,t-1}^I}{\ln s_{ugt}^I - \ln s_{ug,t-1}^I}}{\sum_{u \in U_g^I} \dfrac{s_{ugt}^I - s_{ug,t-1}^I}{\ln s_{ugt}^I - \ln s_{ug,t-1}^I}},$$

where

$$s_{ugt}^I = \frac{p_{ugt}^I C_{ugt}^I}{\sum_{u \in U_g^I} p_{ugt}^I C_{ugt}^I}.$$

posted prices. They argue that price indexes that rely on posted prices overstate the price level experienced by consumers.

The weights capture how likely consumers are to substitute within and across categories, and they are income-group specific. The products that are highly substitutable can receive a much smaller weight than the products whose share barely moves after a price change; these products get a weight close to their average expenditure share.[11]

Because this index ignores new and disappearing product varieties, it is often referred to as a "conventional" EPI (CEPI). In order to take into account increases in quality due to the entry and exit of UPCs, Feenstra (1994) generalize the EPI to allow for different, but overlapping, sets of goods in the two periods. If there is a set of UPCs available in both periods, and their quality parameters are constant, then we can derive the following price index by income group, which allows for product creation and destruction as follows:

$$\prod_{g\in G}\left\{\prod_{u\in U_g^I}\left(\frac{p_{ugt}^I}{p_{ugt-1}^I}\right)^{\omega_{ugt}^I}\times\left(\frac{\lambda_{gt}^I}{\lambda_{gt-1}^I}\right)^{\frac{1}{\sigma_g^I-1}}\right\}^{\omega_{gt}^I}, \tag{4}$$

where

$$\lambda_{gt}^I=\frac{\sum_{u\in U_g^I}p_{ugt}^IC_{ugt}^I}{\sum_{u\in U_{gt}^I}p_{ugt}^IC_{ugt}^I}.$$

The cost of living index for income group I is now adjusted for a new goods bias between periods t and $t-1$. As Broda and Weinstein (2010) show, this new goods bias is important because of its cyclical nature. The magnitude of the bias depends on the ratio of the share of common goods and the elasticity of substitution. The ratio of the share of common products within a category in period t relative to the share of common products in period $t-1$ within category g is $\lambda_{gt}^I/\lambda_{gt-1}^I$. This term captures the importance of quality shifts due to the creation and destruction of UPCs in category g for income group I. The within-category elasticity of substitution σ_g^I indicates the effect that these shifts have on the price index and varies by income. As the elasticity of substitution rises (i.e. highly substitutable UPCs), a given movement in the share of common goods over time has a smaller effect on the price index.[12]

4.2. *Estimating Elasticities by Income Group*

In order to estimate the income-specific elasticity of substitution within a product category σ_g^I, we rely on the method developed by Feenstra (1994) and extended by

11. The limit of $\frac{s_{ugt}^I-s_{ug,t-1}^I}{\ln s_{ugt}^I-\ln s_{ug,t-1}^I}$ when $s_{ugt}^I\to s_{ug,t-1}^I$ is $s_{ug,t-1}^I$.

12. We do not allow the entry and exit of a product category. Because the definition of a product category in the Nielsen data is disaggregated, we find some product categories in which no single transaction is made in a quarter. We keep balanced product categories over time that require more than 10 UPCs covered by each income group every quarter. Our baseline results are robust to various cutoffs. In addition, we exclude the cigarette category because its price depends mainly on state taxes instead of a market mechanism.

Broda and Weinstein (2006, 2010). The procedure consists of estimating a demand and supply equation for each UPC by using only the information on prices and quantities. For this estimation, we face the standard endogeneity problem for a given UPC. 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 UPC within a product category. Specifically, we estimate the demand elasticities by using the following system of differenced demand and supply equations, as in Feenstra (1994)[13]:

$$\Delta_g^k \ln s_{ugt} = -(\sigma_g - 1)\Delta_g^k \ln p_{ugt} + \varepsilon_{ugt}^k, \tag{5}$$

$$(1-\rho_g)\Delta_g^k \ln p_{ugt} = \frac{\rho_g}{(\sigma_g - 1)}\Delta_g^k \ln s_{ugt} + \delta_{ugt}^k, \tag{6}$$

where

$$\rho_g \equiv \frac{\omega_g(\sigma_g - 1)}{(1+\omega_g\sigma_g)} \quad \text{and} \quad \omega_g$$

is the inverse supply elasticity of product category g. Note that when the inverse supply elasticity is zero (i.e. ω_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 UPC u in product category g differenced with respect to a benchmark UPC in the same product category. The kth good corresponds to the largest selling UPC in each product category. The k-differencing removes any product category level shocks from the data.

The identification strategy relies on two important assumptions. First, we assume that ε_{ugt}^k and δ_{ugt}^k are uncorrelated (i.e. $\mathbb{E}_t(\varepsilon_{ugt}^k\delta_{ugt}^k) = 0$). Because we already removed any product category level shocks, we are left with a within-product-category variation, which is likely to render independence of the UPC-level demand and supply shocks within a product category. Second, we assume that σ_g and ω_g are restricted to be the same over time and for all UPCs in a given product category.

To take advantage of these assumptions, we multiply equations (5) and (6) to obtain

$$(\Delta_g^k \ln p_{ugt})^2 = \frac{\rho_g}{(\sigma_g - 1)^2(1-\rho)}\left(\Delta_g^k \ln s_{ugt}\right)^2 + \frac{2\rho_g - 1}{(\sigma_g - 1)(1-\rho)}\left(\Delta_g^k \ln s_{ugt}\right)\left(\Delta_g^k \ln p_{ugt}\right). \tag{7}$$

Given equation (7), we define a set of moment conditions for each product category g as follows:

$$G(\beta_g) = \mathbb{E}_t(\nu_{ugt}(\beta_g)) = 0 \quad \forall\, u, g, \tag{8}$$

where $\nu_{ugt} = \varepsilon_{ugt}\delta_{ugt}$ and β_g is a vector that contains σ_g and ω_g.

13. To simplify the notation, in what follows, we omit the superscript *I*.

TABLE 1. Estimated elasticities of substitution by income group.

Percentile	<25k	25k–50k	50k–100k	>100k
5	6.27	5.38	5.55	5.79
25	11.77	10.73	10.67	10.74
Median	20.31	19.14	18.21	19.33
75	44.14	39.18	37.79	39.41
95	149.57	143.44	113.70	137.37

Notes: The table shows the distribution of the within-category elasticity of substitution for each product category and income group. The estimation procedure follows Feenstra (1994) and Broda and Weinstein (2010). The details are discussed in Section 4.2.

For each product category g, all the moment conditions that enter the Generalized Method of Moments (GMM) objective function can be combined to obtain Hansen (1982)'s estimator:

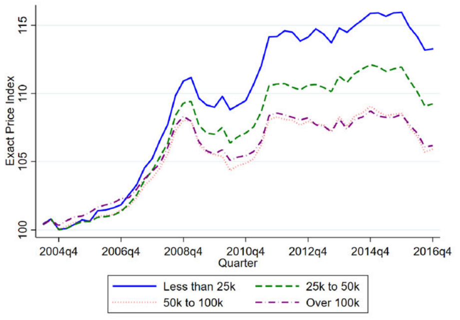
$$\hat{\beta_g} = \arg\min_{\beta_g \in B} G^*(\beta_g)' W G^*(\beta_g) \quad \forall\, g, \tag{9}$$

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. $\sigma_g > 0$). If the procedure renders imaginary estimates or estimates of the wrong sign, we use a grid search to evaluate the aforementioned GMM objective function. Table 1 shows the distribution of the within-category elasticity of substitution by income group for the balanced product category. The median within-category elasticity is between 18 and 20 for all income groups. This number is unsurprisingly high given that it describes the substitutability of a UPC within a finely defined product category. Because we do not find substantial differences in our estimated elasticities across income groups, any difference in the price index across groups is unlikely to come from this source. In fact, when we assume that all income groups have the same elasticity of substitution for a given group g, our results are virtually unchanged.

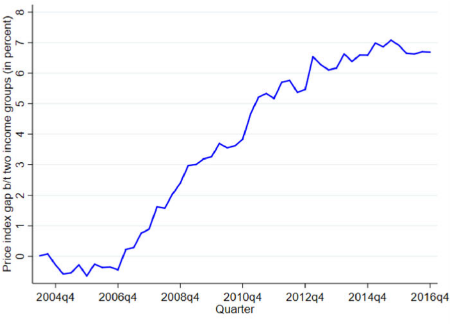
4.3. Results

Panel (a) in Figure 1 compares the estimated EPI of each quartile of the cross-sectional distribution of household income in the United States. The figure shows the quarterly EPI from 2004:1 to 2016:4, normalized to 100 in 2004:1 for all income groups. Panel (a) shows that the indexes for all income groups track each other closely until the end of 2007, when the price indexes of the bottom quartiles start growing significantly faster than those of top quartiles. Panel (b) summarizes the gap (in %) between the EPI of the lowest (less than 25k) and the highest (over 100k) income groups. The gap is zero in 2004:1 by construction. It then increases rapidly between the end of 2007 and the end 2013. By the end of 2016, the last year covered in our data, the EPI of the lowest income group is 6%–7% higher than that of the highest income group.

We summarize the information depicted in Figure 1 by dividing the data into three subperiods: (i) prerecession (2004–2007), (ii) recession (2008–2013), and (iii)



(a) EPI by income groups



(b) Gap between lowest and highest groups

FIGURE 1. EPI by income group and the gap between lowest and highest groups. Panel (a) shows the quarterly EPI, developed in Section 4.1 from 2004:1 to 2016:4. Each line represents the EPI of one of the quartiles of the income distribution: (i) less than \$25,000, (ii) \$25,000–\$50,000, (iii) \$50,000–\$100,000, and (iv) over \$100,000. The first quarter of 2004 is normalized to 100 for all income groups. Panel (b) summarizes the percent gap between the EPI of the lowest (less than 25k) and the highest (over 100k) income groups. By construction, the gap is zero in 2004:1.

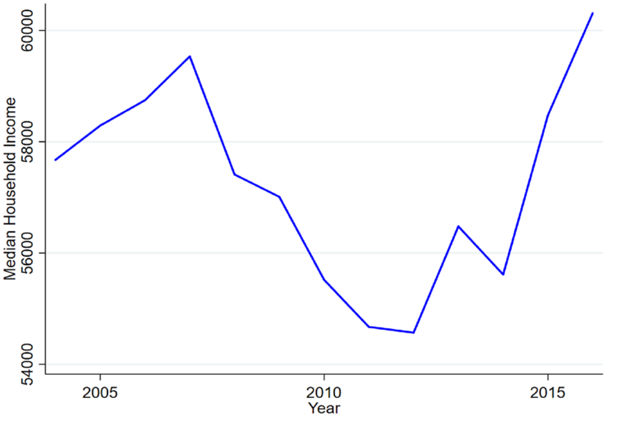


FIGURE 2. Real Income from 2004 to 2016. The figure shows the real median household income for the United States from 2004 to 2016. The series is adjusted using the 2017 CPI-U-RS. Source: U.S. Bureau of the Census.

postrecession (2014–2016). The subperiods reflect the evolution of the real median household income over this period and show that the consequences of the Great Recession lasted for several years. Figure 2 shows that the real median household income increased during 2004–2007, decreased during 2008–2013, and then recovered to the prerecession level during 2014–2016. Table 2 presents the average annual inflation—the average annual change in the price index—for each of these subperiods. Households in the lowest quartile experienced a 1.27% annual inflation before the recession, 1.45% annual inflation during the recession, and −0.36% annual inflation during the postrecession period. The highest income households experienced less

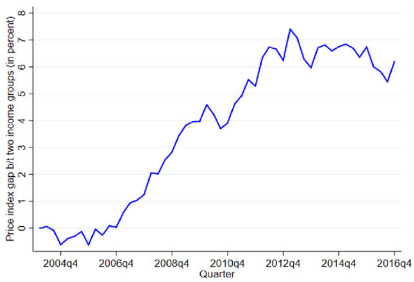
TABLE 2. Average annual inflation rates by subperiod.

	<25k	25k–50k	50k–100k	>100k	Difference
			2004–2007		
EPI	1.27	1.06	0.93	1.05	0.22
Laspeyres	2.70	2.61	2.41	2.39	0.31
Paasche	2.12	2.14	2.01	1.99	0.13
Törnqvist	2.51	2.42	2.21	2.20	0.31
			2008–2013		
EPI	1.45	1.07	0.71	0.60	0.85
Laspeyres	2.52	2.39	2.16	1.76	0.76
Paasche	2.10	2.06	1.87	1.70	0.39
Törnqvist	2.55	2.43	2.25	1.92	0.63
			2014–2016		
EPI	−0.36	−0.48	−0.52	−0.38	0.02
Laspeyres	0.24	0.22	0.21	0.40	−0.16
Paasche	−0.22	−0.25	−0.28	0.15	−0.37
Törnqvist	0.03	0.20	0.21	0.56	−0.53

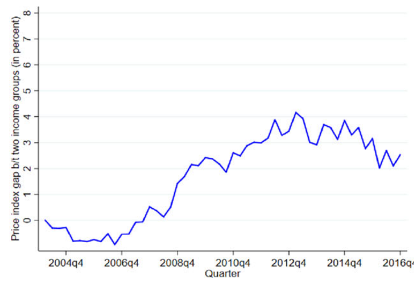
Notes: The table shows the income-specific average annual inflation rates for three subperiods: prerecession, recession, and postrecession. It presents three different specifications: (i) EPI, (ii) Laspeyres, (iii) Paasche, and (iv) Törnqvist. The last column indicates the difference between the average annual inflation of the lowest income group (<25k) and the highest income group (>100k) in percentage points.

pronounced annual inflation rates: 1.05% in the prerecession period, 0.60% during the recession, and −0.38% in the postrecession period. The last column of Table 2 shows that the annual inflation of the top quartile is 0.22 percentage points lower than that of the bottom quartile during the prerecession subperiod, but this difference increases to 0.85 during the recession period. This difference increases by only 0.02 percentage points from 2014 to 2016.

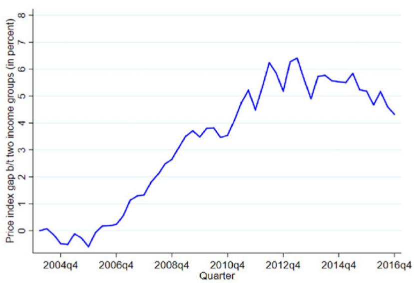
The inflation inequality that emerges over our sample period has been confirmed in subsequent work by Kaplan and Schulhofer-Wohl (2017) and Jaravel (2019), who both find gaps of similar magnitudes across income groups. We further check the robustness of our results by showing that the time-series patterns that we find persist after using other homothetic price indexes that (i) handle substitution effects differently, (ii) omit the new goods bias term, or (iii) do not depend on the estimated elasticity of substitution. For instance, we check the robustness of our results with fixed-weighted indexes that either assume no substitution (Laspeyres) or full substitution (Paasche), indexes with lower substitution bias (geometric Laspeyres and geometric Paasche), superlative indexes (Walsh, Fisher, and Törnqvist), as well as other symmetrically weighted price indexes (Marshall–Edgeworth). Although the magnitudes differ across price indexes, Figure 3 shows that the existence of the gap and its overall time-series patterns are similar to those found using the EPI. Table 2 reports the magnitude of the average annual inflation and the size of the gaps for the Laspeyres, Paasche, and Törnqvist indexes. In all three cases, the gap increases most rapidly from 2008 to 2013. As a result, the inflation inequality that we observe cannot be entirely explained by substitution effects. Furthermore, Figure 3 also shows that the inflation inequality



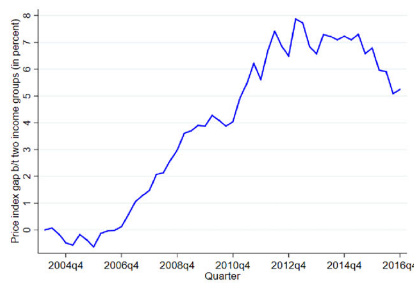
(a) Laspeyres



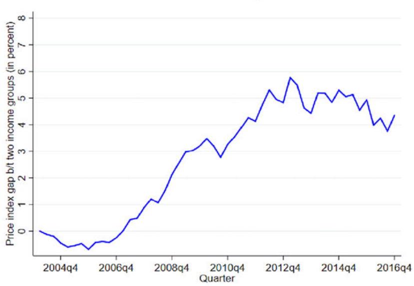
(b) Paasche



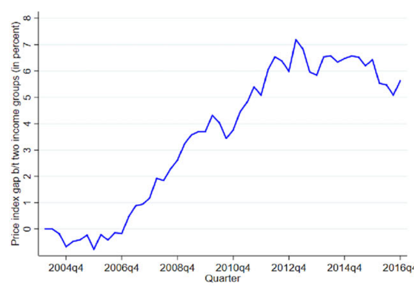
(c) Marshall-Edgeworth



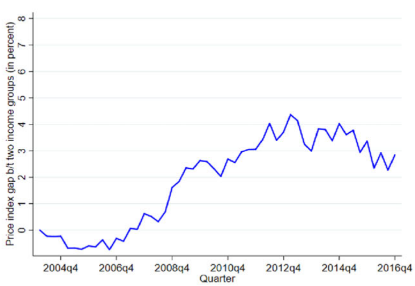
(d) Walsh



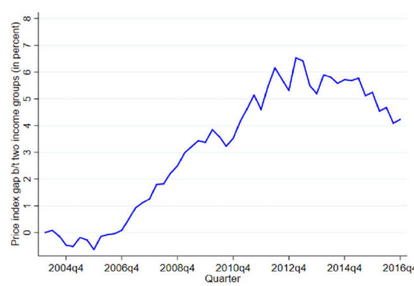
(e) Fisher



(f) Geometric Laspeyres



(g) Geometric Paasche



(h) Törnqvist

FIGURE 3. Gaps between lowest and highest groups under various price indexes. The figure shows the percent gap between the price index of the lowest (less than 25k) and the highest (over 100k) income groups. Each panel calculates the gap under a different price index. By construction, the gap is zero in 2004:1.

cannot be entirely explained by changes in the availability of product varieties favoring the top income groups over this period. This is because the gap exists even when the new goods bias term of the EPI is omitted. Nonetheless, the figure shows that the new goods bias is relevant for the postrecession gap given that, when it is omitted as in all the indexes depicted in Figure 3, the inflation gap decreases. In Appendix B, we compute various homothetic price indexes assuming that all income groups pay the same average price for a given product. Consistent with our findings, there is a gap between the price index of the top and bottom income groups, which grows during the recession and stabilizes and/or decreases in the postrecession period.

5. Decomposition of Price Indexes

What are the sources behind the inflation gap we find and how do they vary over the business cycle? In this section, we decompose the EPI into four components in order to uncover the potential sources. Let $p^I_{ugt} \equiv \bar{p}_{ugt}\theta^I_{ugt}$, where θ^I_{ugt} is the deviation in the price paid by a household in income group I for product u from the average price of product u across income groups, $\bar{p}_{ugt}$. Then, the natural logarithm of the EPI for product category g, income group I, and time t can be written as

$$\ln EPI^I_{g,t} = \underbrace{\sum_{u\in U^I_g} s^I_{ugt-1}\ln\left(\frac{\bar{p}_{ugt}}{\bar{p}_{ugt-1}}\right)}_{\text{change in product prices}} + \underbrace{\sum_{u\in U^I_g}(\omega^I_{ugt} - s^I_{ugt-1})\ln\left(\frac{\bar{p}_{ugt}}{\bar{p}_{ugt-1}}\right)}_{\text{product subsitution}}$$

$$+ \underbrace{\sum_{u\in U^I_g}\omega^I_{ugt}\ln\left(\frac{\theta^I_{ugt}}{\theta^I_{ugt-1}}\right)}_{\text{change in shopping behavior}} + \underbrace{\frac{1}{\sigma^I_g - 1}\ln\left(\frac{\lambda^I_{gt}}{\lambda^I_{gt-1}}\right)}_{\text{new good bias adjustment}}, \tag{10}$$

where the EPI for income group I can be defined as

$$\ln EPI^I_t = \sum_{g\in G}\omega^I_{gt}\ln EPI^I_{g,t}. \tag{11}$$

The first term in equation (10) reflects the change in the average prices of products when their expenditure weights are fixed to the level of the previous period. This term resembles the geometric mean formula used by statistical agencies because, due to data limitations, they are unable to update the expenditure weights of products every period and its magnitude depends on the consumption baskets of each income group.[14] The second term captures the extent to which, given average prices, the EPI changes due to changes in expenditure weights that result from consumers substituting across product

14. In this case, the expenditure weights used are those of the previous quarter, whereas the expenditure weights used in the calculation of the CPI have at least a two-year gap. For instance, CPI data in 2016 and 2017 were based on data collected from the Consumer Expenditure Surveys for 2013 and 2014.

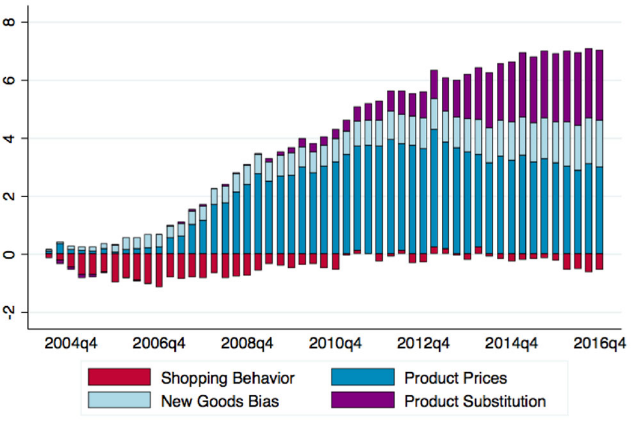


FIGURE 4. Decomposition of the gap in EPI between the lowest and highest income groups. The figure shows the decomposition of the gap in EPI between the lowest and highest income groups. The lowest income group is the set of households earning less than \$25,000 in a year, and the highest income group is the set of households earning more than \$100,000 in a year. Equation (10) decomposes the gap exactly into four components: (i) change in product prices, (ii) new good bias adjustment, (iii) product substitution, and (iv) shopping behavior.

varieties within product categories. The third term estimates the average deviation in the prices paid by a household in income group I from their average price. It captures how the shopping behavior of households changes over time and how it differs across income groups, that is, the extent to which consumers can adjust the prices they pay for an item by adjusting the outlet where they purchase, searching for discounts, or using coupons. Lastly, the fourth component is the conventional new goods bias developed by Feenstra (1994) and Broda and Weinstein (2010). This term captures the importance of quality shifts due to the creation and destruction of varieties. The higher the expenditure share of new products and the lower the expenditure share of existing products, the smaller the EPI. Moreover, the strength of this term depends on the estimated elasticity of substitution between varieties σ_g^I. The importance of this term decreases as varieties become more substitutable.

We decompose the gap in the EPI between the lowest income households ($I = l$) and the highest income households ($I = h$), $\ln epi_t^{I=l} - \ln epi_t^{I=h}$, into these four components. Figure 4 shows the contribution of each of them to the EPI gap. The figure shows that all four components have a positive contribution from 2004 to 2016. Over this period, the difference between the EPI of the bottom quartile and the top quartile is around 7%. Approximately 46% of the gap can be attributable to changes in product prices, 29% to differences in within-category product substitution, 13% to changes in the adoption of product varieties, and 12% to differences in shopping behavior. Table 3 shows the contribution of each component to the annual inflation of each income group over the prerecession, recession, and postrecession subperiods in percent. The last column in the table shows the difference between the annual inflation of the low-income households relative to that of high-income households in percentage points.

TABLE 3. Average annual inflation rates by component and subperiod.

	<25k	25k–50k	50k–100k	>100k	Difference
		2004–2007			
Product prices	1.76	1.73	1.49	1.47	0.29
New goods bias	− 0.26	− 0.30	− 0.35	− 0.38	0.12
Product substitution	− 0.09	− 0.25	− 0.15	− 0.11	0.01
Shopping behavior	− 0.14	− 0.11	− 0.07	0.07	− 0.21
		2008–2013			
Product prices	1.89	1.87	1.65	1.50	0.39
New goods bias	− 0.20	− 0.27	− 0.33	− 0.31	0.11
Product substitution	− 0.20	− 0.42	− 0.47	− 0.45	0.24
Shopping behavior	− 0.03	− 0.11	− 0.14	− 0.14	0.10
		2014–2016			
Product prices	− 0.19	− 0.14	− 0.19	− 0.12	− 0.07
New goods bias	− 0.16	− 0.23	− 0.30	− 0.30	0.14
Product substitution	0.13	− 0.08	− 0.06	− 0.07	0.20
Shopping behavior	− 0.14	− 0.04	0.03	0.12	− 0.26

Notes: The table shows the income-specific average annual inflation rates for three subperiods: prerecession, recession, and postrecession. It presents the four different components defined in Section 5: (i) product price, (ii) new goods bias, (iii) product substitution, and (iv) shopping behavior. The last column indicates the difference between the average annual inflation of the lowest income group (<25k) and the highest income group (>100k) in percentage points.

Change in product prices. During the prerecession subperiod, the gap is explained almost entirely by changes in product prices (70%) and by the new goods bias term (26%). In fact, the contributions of shopping behavior and product substitution were basically zero until the end of 2008. In the recession subperiod, 46% of the gap is explained by differences in the changes in average prices faced by households of different income groups, which explains an approximately 0.39 percentage point difference in the annual inflation of the bottom and top income groups. In the postrecession period, the product prices component contributed to decreasing inflation inequality.

Product Substitution. The relevance of this component starts in the recession and continues, although slightly attenuated, in the postrecession period. It accounts for 29% of the gap during the recession, which is an approximately 0.24 percentage point difference in the annual inflation of the bottom and top income groups. High-income households, in response to a price change, were better able to pay lower prices for the same category of goods by shifting their expenditures to less expensive products. This suggests that low-income households have fewer options for quality substitution during downturns.[15] The contribution of product substitution to the differences in average annual inflation rates between the top and bottom quartiles decreases in the

15. The quality substitution captured by the EPI is along the intensive margin, among varieties already consumed by the households in the previous period. We discuss in detail the implications of the extensive margin of quality substitution in Section 5.1.

postrecession relative to the recession subperiod. Nonetheless, it is still positive and relevant. We provide further evidence of the potential causes in Section 6.

Shopping Behavior. This component contributes only to the EPI gap during the recession. It accounts for 13% of the gap during this subperiod, suggesting that high-income households were able to cope with the recession by adjusting their shopping behavior unlike low-income households. Just as in the prerecession subperiod, in the postrecession the shopping behavior component contributes to significantly reducing the gap by decreasing the difference in average annual inflation rates by 0.26 percentage points.

New Goods Adjustment. This component captures the changes in the availability of product varieties that favor the top income groups. Interestingly, the contribution of this component to the difference in average annual inflation rates between the top and bottom quartiles remained remarkably stable over the three subperiods. It was only during the recession that it had a small decline, which is potentially related to the fact that this period was characterized by a sharp decline in product entry, which is a sign of economic activity by firms and consumers. Michelacci, Paciello, and Pozzi (2019) and Argente, Lee, and Moreira (2018) document a sharp decline in the number of varieties available as well as a sharp decrease in the product entry rate after 2007; both measures only went back to their prerecession levels after 2013. Although in our benchmark specification we assume that each income group has different elasticities of substitution, our results are virtually unchanged if we assume they have the same elasticity, given the small differences in σ^I across groups.[16]

5.1. Discussion

Although the EPI has several advantages, it is important to discuss its limitations in order to clarify the interpretation of our price indexes. First, the calculation of each income-specific EPI uses only the prices paid and the varieties purchased by households in each income category to reflect the prices and varieties available to these households. As a result, despite the fact that the Nielsen Consumer Panel is a very rich data set, our sample could underrepresent the choice sets of households. This is particularly relevant for households that substitute towards products of different qualities along the extensive margin. If high-income households substitute along the extensive margin towards products of lower quality (but not new) during the recession, our decomposition would underestimate the importance of the product substitution component and overestimate the relevance of the new goods bias for inflation inequality, because these purchases

16. Because large elasticities of substitution decrease the relevance of new goods bias, for robustness we recalculate the estimated bias assuming an elasticity of substitution of 10 for all income groups, which is lower than the median elasticity estimated by Broda and Weinstein (2010). The contribution of the new goods bias to the inflation gap during the recession is very similar under this specification and accounts for 14% of the gap during the recession.

will show up as gains in varieties.[17] In Section 6, we combine the household-level panel data with the retail-level scanner data to capture the universe of products in the consumer packaged goods (CPG) sector and quantify the full extent to which product quality substitution allows consumers to decrease the prices they pay; in particular, their economic conditions deteriorate.

Second, our current setup cannot be used to make welfare comparisons across income groups that result from product substitution in response to income declines. We start from a homothetic utility function for each income group. This allows for nonhomotheticicities across groups (because the elasticity of substitution is allowed to vary across income quartiles), but it does not allow for nonhomotheticicities within income groups, at the individual level. In Appendix C, we use a stylized nonhomothetic utility function to quantify the welfare gains of high-income households relative to low-income households over our sample period.

6. Product Substitution and Shopping Behavior

The impact of the changes in product prices and the new goods bias on inflation inequality has been recently explored by Kaplan and Schulhofer-Wohl (2017) and Jaravel (2019). Kaplan and Schulhofer-Wohl (2017) find that, at the height of the Great Recession, the change in average prices contributed significantly to the variance in the cross-sectional distribution of inflation rates because the shift in the relative price of food-at-home made heterogeneity in consumption bundles more important. Jaravel (2019) argues that, in response to new varieties introduced to cater to high-income households, the prices of continuing products in these market segments fell. In what follows, we focus on providing direct evidence of the existence and relevance of the product substitution and shopping behavior components. Both of these margins became relevant to the gap across income groups at the onset of the recession and the literature has yet to fully explore their contribution to inflation inequality.

6.1. Descriptive Evidence: Cross-Sectional Analysis

Product Substitution. To provide further evidence of that the ability of households to take advantage of the product substitution margin differs by income, we begin by exploring the relationship between prices and income purely in the cross section.[18] If

17. Michelacci, Paciello, and Pozzi (2019) estimate that the contribution to expenditure growth of the extensive margin of product substitution is as large as that of the intensive margin. They find that fluctuations in the net additions of products to the consumption basket of households occur within the set of continuously available varieties.

18. This relationship is also known as the quality Engel curve. The curves are computed under the premise that across households, at a point in time, those paying higher unit prices are buying higher quality goods (typically richer households). See Bils and Klenow (2001) for a detailed description and Broda, Leibtag, and Weinstein (2009), Handbury (2019), and Faber (2014) for other applications.

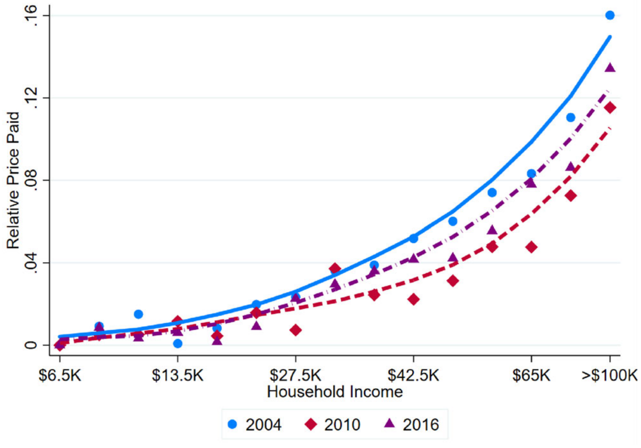
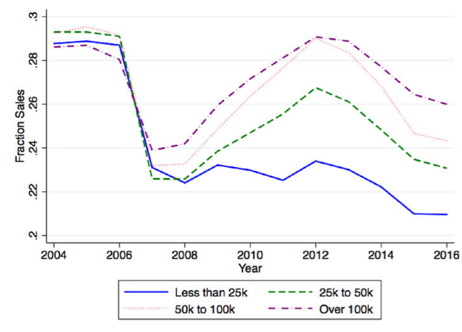


FIGURE 5. Relative price paid by household income. The relative prices are measured in a regression of the log-unit price paid against income category dummies, chain, product category, and county-fixed effects. Each dot represents how much more each income category pays per unit for products with respect to households earning $6,500 per year. The blue dots represent the cross-sectional relation in 2004, the red dots in 2010, and the purple dots in 2016.

quality substitution is a relevant margin of adjustment during the recession, we expect that the relationship between prices and income will flatten. Figure 5 shows how much more or less households pay per unit for products within a category. In the figure, the relative prices are measured in a regression of the log-unit price paid against income category dummies and product category, region, chain, and quarter-fixed effects. Each dot represents how much more households in each income category pay per unit for products within a category than households in the lowest income category of between $5,000 and $8,000. The figure shows that a distinct upward slope exists; high-income households pay around 10%–15% more for products in the same category than low-income households. As income declines, households reduce the relative price paid by switching to cheaper products within a category.

This relation holds when we study specific years within each of the three subperiods. However, the figure shows that in the middle of the recession subperiod (i.e. 2010), the slope distinctively decreases, particularly for households above the median of the income distribution. This decrease indicates that during the recession, these households were able to decrease the prices they paid for a product within a category by substitution towards products of lower quality. Moreover, the fact that the slope is very flat for households below the median of the income distribution indicates that, in the recession, the ability of high-income households to take advantage of the quality substitution margin was significantly greater. Consistent with the results in Section 5, the ability of higher income to take advantage of the product substitution margin remained greater than that of low-income households in the postrecession period. Figure 5 shows that, although the slope of the curve has gotten steeper for

(a) Fraction of Coupons



(b) Fraction of Sales

FIGURE 6. Shopping behavior: coupon usage and purchases on sale. The figure shows the time-series patterns of two shopping activities for each income quartile. Panel (a) shows the coupon usage and panel (b) shows the fraction of expenditures on items on sale. As in Nevo and Wong (2019), an item is defined as being on sale if the household recorded that the item purchased involved a deal. An item is defined as involving coupons if the household recorded that the item purchased involved using either a store coupon or a manufacturer coupon.

low- and middle-income households by 2016, it has not returned its prerecession level.[19]

Shopping Behavior. Households adjust their shopping behavior to pay lower prices for identical goods by shopping more frequently, adjusting the outlet where they purchase, searching for discounts, or using coupons. In this section, we show that both high- and low-income households adjusted their shopping effort differently during the recession subperiod. In fact, in the prerecession subperiod, households in the top quartile of the income distribution paid an average of 2% more than households in the bottom quartile for an identical product. By the end of the recession subperiod, that difference had disappeared. For example, prior to the recession, purchases of items on sale and coupon usage were either stable or declining as a share of total expenditure for all income groups. During the recession, the intensity of these shopping activities increased significantly, especially for high-income groups. Figure 6 shows precisely those patterns across the purchases of sale items and the use of coupons.[20] The figure shows that the increase in the intensity of these activities is more pronounced during the recession and for high-income households. In the postrecession subperiod and consistent with the evidence presented in Section 5, the differences observed across income groups are attenuated.

19. The difference in slopes between 2004 and 2010 for households above the median of the income distribution is statistically significant at the 0.05 level even after allowing the standard errors to be correlated within product categories. We cannot reject that the slope in 2010 is different from that in 2016. These results are shown in Figure A.3.

20. As in Nevo and Wong (2019), an item is defined as being on sale if the household recorded that the item purchased involved a deal. An item is defined as involving a coupon if the household recorded that the item purchased involved using either a store coupon or a manufacturer's coupon.

To be able to capture the impact of all shopping activities (e.g. store switching, coupon usage, etc.) on the prices paid by households, we follow the approach of Aguiar and Hurst (2007). Denote the price of good $i \in I$ purchased by household $j \in J$ in period k. Let X_t^j be the total expenditure during quarter t and Q_t^j the cost of the same bundle if the household paid the average price[21]:

$$X_t^j = \sum_{i \in I, k \in t} p_{i,t}^j q_{i,k}^j,$$

$$Q_t^j = \sum_{i \in I, k \in t} \bar{p}_{i,t} q_{i,k}^j.$$

We then define the price index for the household as the ratio of expenditures at actual prices to the cost of the bundle at the average price:

$$\tilde{p}_t^j \equiv \frac{X_t^j}{Q_t^j}. \tag{12}$$

We normalize the index by dividing through the average price index across households within the quarter, which ensures that for each quarter the index is centered around 1 and allows us to compare the prices paid by the households in the cross section:

$$p_t^j \equiv \frac{\tilde{p}_t^j}{\frac{1}{J}\sum_{j'} \tilde{p}_t^{j'}}.$$

The differences in this price index between households do not reflect differences in the quality of the purchased goods. The price differentials are for the identical goods as measured by UPC codes. All else being equal, households that spend more time searching and shopping would be able to reduce the price paid for a given market good. We use this price index to explore how prices paid for the same goods vary across households of different income groups. Figure 7 plots the difference between the prices paid by households in the top and bottom quartiles. The results indicate that higher income households paid prices that were 2% higher than low-income households in the prerecession period. During the recession, this difference virtually disappeared and remained like that in the postrecession period.

6.2. Direct Evidence: Longitudinal Analysis

In this subsection, we take advantage of the longitudinal dimension of our data and focus on providing a direct measure of product substitution and other shopping

21. We average over households within the quarter to obtain the average price paid, $\bar{p}_{i,t}$, for a given good during that quarter, where the average is weighted by quantity purchased. More specifically,

$$\bar{p}_{i,t} \equiv \sum_{j \in J, k \in t} p_{i,k}^j \left(\frac{q_{i,k}^j}{\bar{q}_{i,t}}\right), \quad \text{where} \quad \bar{q}_{i,t} \equiv \sum_{j \in J, k \in t} q_{i,k}^j.$$

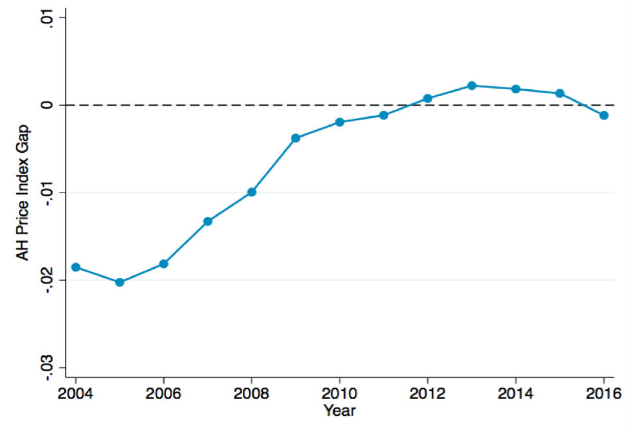


FIGURE 7. Gaps between lowest and highest groups: cross sectional. The figures show the differences between the prices paid by households in the top and bottom quartiles for the same good, which is defined by a UPC code. The price index for each household is constructed following the approach by Aguiar and Hurst (2007). Households are then aggregated to the income group level using the projection factor provided by Nielsen.

activities at the household level. Because we observe the share of expenditures of each household in each product, we construct a household-specific measure of the average quality of goods they purchase and the average quality of the stores they visit. In addition, our data have information on each household's shopping activities, including their number of shopping trips and the fraction of items they purchase on sale. We then explore the sensitivity of these measures to both changes in local economic conditions and changes in reported incomes at the household level. We show that quality substitution is a margin households use not only when aggregate conditions worsen but also when local conditions deteriorate. We also show that high-income households are better able to use this margin due partly to the fact that most of the purchases of low-income households already come from products ranked low in the quality distribution. Our results are also robust after a negative change in the annual income reported by households, which indicates that quality substitution is an important margin of adjustment during economic downturns.

6.2.1. Measurement.. In order to study the extent of quality substitution and store switching, we construct a "ranking" of products and stores based on their relative prices. To do this, we use the Nielsen Retail Scanner data set from 2006 to 2016 to obtain a universe of products sold in the CPG sector. We use information on more than 100 billion transactions at the store level and follow the industrial organization and the international trade literature to calculate the average relative prices of products to approximate their quality. We consider a product with higher relative prices to be higher on the quality ladder. For example, within the milk category, organic whole milk is higher on the quality ladder than regular whole milk given that it is sold, on average, at a higher unit price. Then, we combine this information with the Consumer Panel data to construct a household-specific estimate of product quality in a given period of time. We follow a similar procedure to approximate the quality of the stores. We calculate product quality for 1.76 million distinct barcode-level products and store quality for 67,100 distinct stores in the United States. Because not all products in the Retail Scanner data are consumed by the panelists in the Nielsen Consumer Panel, our final sample includes 0.83 million distinct products and 28,079 distinct stores. We calculate household-level average product and store quality for 114,071 unique households living in 2,833 counties.

Product Quality Substitution. We construct the average relative price at the barcode level and within each product category. First, we measure the log-difference between the price of good j at store s and the median price for product category m at store s in county c as follows:

$$R_{jsct} = \log \frac{P_{jsct}}{\bar{P}_{msct}},$$

where R_{jsct} is the relative unit price and $\bar{P}_{msct}$ is the median unit price of product category m at store s in county c. Therefore, if the price of a high-quality type of milk, say, organic milk, is higher than the median price of milk in that store, then R_{jsct} is positive and high. We compute the average relative unit price for the barcode-level goods across the set of stores within a county c. The average relative unit price of a good is

$$Q_{jct}^{\text{product}} = \sum_{s \in \Omega_c} \omega_{jsct} R_{jsct},$$

where ω_{jsct} is the revenue weight. The Q_{jct}^{product} captures how far a product's average price level is from the median price level of a category in a quarter t within a county c. We construct an average relative price for the products that household h buys in quarter t as follows:

$$Q_{hct}^{\text{product}} = \sum_{j} \psi_{hcjt} Q_{jct}^{\text{product}},$$

where ψ_{hsct} is an expenditure weight. A low (high) value of Q_{hct}^{product} indicates that household h living in a county c buys relatively low (high) quality products in quarter t.

Store Switching. Households may change their shopping behavior by purchasing goods from different stores. To assess the reallocation of expenditures across stores, we first construct the stores' quarterly relative prices as in Coibion, Gorodnichenko, and Hong (2015). First, for each UPC-level good j in store s and county c, we calculate the log-difference between the price of good j in store s and the median price for good

TABLE 4. Descriptive statistics of product and store quality.

	Mean	Standard deviation	10th percentile	90th percentile	Observations
Product quality	2.16	17.79	− 14.85	20.04	1,630,232
Store quality	− 0.13	1.98	− 2.47	2.02	1,630,232

Notes: The table presents descriptive statistics of product and store quality for the households in our data. The product and store rankings are calculated within a county from 0.83 million distinct products and 28,079 distinct stores in the Retail Scanner data and the Consumer Panel data. We use 114,071 unique households living in 2,833 counties from 2006 to 2016.

j across all of the stores in a given county and quarter:

$$R_{jsct} = \log \frac{P_{jsct}}{\bar{P}_{jct}},$$

where R_{jsct} is the relative price and $\bar{P}_{jct}$ is the median price for good j in a county c in a quarter t.

We then compute the average relative price for a store across the set of products available in the county. The average relative price of a store is

$$Q_{sct}^{\text{store}} = \sum_j \omega_{jsct} R_{jsct},$$

where ω_{jsct} is a revenue weight. The Q_{sct}^{store} captures how far a store's average price level is from the median price level in a given county c and quarter t. We construct an average relative price (quality) for the stores at which household h shopped in quarter t as follows:

$$Q_{hct}^{\text{store}} = \sum_s \psi_{hsct} Q_{sct}^{\text{store}},$$

where ψ_{hst} is an expenditure weight. The Q_{hct}^{store} represents the average quality of a store where household h living in county c consumes at. Given this measure, we assess the degree of store switching at the household level. Table 4 reports the descriptive statistics of product quality and store quality at the household level. Both product quality and store quality have means close to zero, and the quality of products is more dispersed across households than the quality of stores.

Shopping Trips And Purchases on Sale. To approximate shopping effort, we use the number of shopping trips and the fraction of items purchased on sale. A shopping trip is defined by the date and location of the transaction (Aguiar and Hurst 2007). That is, transactions at two stores on the same day are counted as two trips. Similarly, two transactions at the same store on two different days are counted as two trips. By making more shopping trips, households can use store and manufacturer discounts more frequently.[22] In addition, the Consumer Panel data define a transaction as being

22. We do not have information on the length of each shopping trip. Aguiar and Hurst (2007) supplements the Consumer Panel data using the American Time Use Survey (ATUS). Because the time spent at shopping

on sale if the household records that the item purchased involved a deal. We measure the fraction of items purchased on sale by a given household in a given quarter.

6.2.2. Impact of an Income Shock on the Households' Shopping Behavior.. We use our household-specific estimates of product and store quality, along with our measures of sale purchases and coupon usage, to assess the changes in product substitution and shopping behavior after an income shock. We represent an income shock by the change in the unemployment rate of the county where a household lives.[23] We use the following empirical specification:

$$Y_{hct} = \lambda + \theta \times \mathrm{UR}_{ct} + \alpha_h + \delta_t + \varepsilon_{ht}, \tag{13}$$

where Y_{hct} represents the variables for the shopping behavior and product substitution of household h living in a county c in quarter t, and UR_{ct} is the unemployment rate of the county c. The α_h and δ_t are household- and time-fixed effects. When all fixed effects are added, the estimates of θ assess the strength of the correlation between the changes in the local unemployment rate and the households' shopping activities and product substitution patterns. Whereas aggregate shocks could lead to simultaneous movements in households' behavior and local economic conditions, we control for time-fixed effects to alleviate this issue. Therefore, by using only the cross-sectional variation, our aim is to show that changes in the shopping effort of households are a relevant margin of adjustment when local unemployment increases.

The results are presented in Table 5. The table shows that the average product quality is negatively correlated with the unemployment rate. A 1 percentage point increase in the unemployment rate is associated with a 0.07 percentage point decrease in the average product quality of the households' consumption bundles. This association indicates that households that face a negative income shock reallocate their expenditures to lower quality products within the same product categories.[24] This finding is consistent with recent work by Jaimovich, Rebelo, and Wong (2019) and Dubé, Hitsch, and Rossi (2018). They find that consumers trade down in the quality of the goods and services they consume during the recession.[25]

The table also shows that households in a region with a higher unemployment rate adjust other margins of their shopping behavior. The coefficients for the variables that measure shopping effort are positive and significant. Households are likely to

in ATUS has remained relatively stable over our sample period (Petev, Pistaferri, and Eksten 2011), we focus only on the number of shopping trips.

23. We check the robustness of results using income drops at the household level in the next section.

24. Our results are very similar if instead of relative unit prices we use the quality measure implied under a nested CES demand structure such as the one used by Hottman, Redding, and Weinstein (2016). The correlation between these two quality proxies is 0.85.

25. Jaimovich, Rebelo, and Wong (2019) also approximates the quality of goods and services by their relative price. Their evidence of quality substitution is from the Yelp! website matched with the U.S. Census of Retail Trade and Compustat. Dubé, Hitsch, and Rossi (2018) finds a negative effect of income on private-label shares.

TABLE 5. Shopping behavior and the local unemployment rate.

	(1) Product quality	(2) Store quality	(3) Trips	(4) Sales
Unemployment rate	−0.0710***	−0.0012	0.0389***	0.0691***
	(0.0145)	(0.0013)	(0.0021)	(0.0098)
Time-fixed effect	Y	Y	Y	Y
Household-fixed effect	Y	Y	Y	Y
Observations	1,630,052	1,630,052	1,630,052	1,630,052
R-squared	0.396	0.622	0.671	0.854

Notes: The table reports the estimates for specification (13). Each observation is at the household × quarter level covering from 2006:1 to 2015:4. The coefficients for product quality and store quality can be interpreted as the percentage change in quality, as represented by the median price, when the local unemployment rate increases by 1 percentage point. The coefficient for the variable trips (sales) is the number of additional shopping trips per store (additional fraction of items purchased on sale) after a 1 percentage point increase in the local unemployment rate. We use the county-level quarterly unemployment rate published by the BLS. Standard errors are presented in parentheses. *Significant at the 10% level; **significant at the 5% level; ***significant at the 1% level.

make more shopping trips and buy more items on sale when the local conditions deteriorate. Households could also reallocate their purchases towards low-price stores when economic conditions worsen. We find a negative but insignificant effect of an increase in the local unemployment rate on the average quality of the stores visited by households. These coefficients could, nonetheless, mask important heterogeneity across households of different income groups. Thus, we examine which types of households are more likely to adjust their shopping behavior after the local economic conditions deteriorate. To do so, we use the following empirical specification:

$$Y_{hct} = \lambda + \theta \times \mathrm{UR}_{ct} + \sum_{I=\{2,3,4\}} \theta_I \times \mathrm{UR}_{ct} \times D_I + \alpha_h + \delta_t + \varepsilon_{ht}, \tag{14}$$

where Y_{hct} and UR_{ct} are specified as before and D_I is a dummy variable that indicates the household's income quartile. The estimates of θ_I indicate whether the response of households to changes in the local unemployment rate varies with their income level.

Table 6 shows that product quality substitution is a relevant margin of adjustment for all income groups except for households at the bottom of the income distribution. Consistent with the findings described in Section 6.1, households belonging to the top income group benefit more from this margin. Higher income households adjust their shopping behavior more with respect to low-income households; they reallocate expenditures across stores and increase more the fraction of items they purchase on sale and the number of shopping trips they make in a given quarter. This adjustment could reflect the fact that high-income households more easily allocate additional resources to shopping.[26]

26. Our results are consistent with the findings of Stroebel and Vavra (2019), who also show heterogeneous responses of different consumers after wealth shocks. Using a similar approach, they

TABLE 6. Shopping behavior and local unemployment rate by income group.

	Product quality (1)	Store quality (2)	Trips (3)	Sales (4)
UR_{ct}	−0.0191	0.0073***	0.0073***	−0.0223*
	(0.0181)	(0.0016)	(0.0026)	(0.0123)
$\text{UR}_{ct} \times D_{\text{income group}=2}$	−0.0676***	−0.0138***	0.0215***	0.0199
	(0.0185)	(0.0016)	(0.0027)	(0.0125)
$\text{UR}_{ct} \times D_{\text{income group}=3}$	−0.0699***	−0.0123***	0.0507***	0.1860***
	(0.0176)	(0.0015)	(0.0025)	(0.0119)
$\text{UR}_{ct} \times D_{\text{income group}=4}$	−0.0882***	−0.0081***	0.0718***	0.1990***
	(0.0208)	(0.0018)	(0.0030)	(0.0141)
Time-fixed effect	Y	Y	Y	Y
Household-fixed effect	Y	Y	Y	Y
Observations	1,630,052	1,630,052	1,630,052	1,630,052
R-squared	0.396	0.622	0.671	0.854

Notes: The table presents the estimates for specification (14). Each observation is at the household × quarter level covering from 2006:1 to 2015:4. The coefficients for product quality and store quality are the percentage change in quality as represented by the median price when the local unemployment rate increases by 1 percentage point. The coefficient for the variable trips (sales) is the number of additional shopping trips per store (additional fraction of items purchased on sale) after a 1 percentage point increase in the local unemployment rate. We use county-level quarterly unemployment rate published by the BLS. Standard errors are presented in parentheses. *Significant at the 10% level; **significant at the 5% level; ***significant at the 1% level.

For the low-income households, only the number of shopping trips is significant. The rest of the margins, including the substitution of product quality, are almost zero or not significant. In other words, the product quality margin is present only for high-income households. A 1 percentage point increase in the unemployment rate is associated with a 0.088 percentage point decrease in the average product quality of households in this group. This is relevant given that on average the unemployment rate increased almost 5 percentage points during the recession. Our results show that, because poor households purchase a larger share of their products at the bottom of the distribution of product quality, they lack the substitution margin when they face an income drop. Higher income households, on the other hand, can do so to mitigate income shocks.

6.2.3. Robustness Checks.. In this subsection, we present two robustness checks of our findings. First, instead of using the self-reported annual income to split households into different income groups, we use the education of the household head to represent lifetime earnings expectations: households with a head that has earned a college degree and those with a head with a lower level of education. Table 7 shows that, after a deterioration of the local economic conditions, households whose head has at least

find that renters are less likely to adjust their shopping behavior after changes in housing wealth than homeowners.

TABLE 7. Shopping behavior and local unemployment rate by education.

	Product quality (1)	Store quality (2)	Trips (3)	Sales (4)
UR_{ct}	− 0.0544***	0.0018	0.0050**	− 0.0231**
	(0.0174)	(0.0015)	(0.0025)	(0.0118)
$\text{UR}_{ct} \times D_{\text{college and above}}$	− 0.0252*	− 0.0038***	0.0503***	0.1370***
	(0.0143)	(0.0013)	(0.0020)	(0.0096)
Time-fixed effect	Y	Y	Y	Y
Household-fixed effect	Y	Y	Y	Y
Observations	1,620,614	1,620,614	1,620,614	1,620,614
R-squared	0.396	0.622	0.671	0.854

Notes: The table presents the estimates for specification (14). Each observation is at the household × quarter level covering from 2006:1 to 2015:4. The coefficients for product quality and store quality are the percentage change in quality as represented by the median price when the local unemployment rate increases by 1 percentage point. The coefficient of the variable trips (sales) is the number of additional shopping trips per store (additional fraction of items purchased on sale) after a 1 percentage point increase in the local unemployment rate. We use the county-level quarterly unemployment rate published by the BLS. Standard errors are presented in parentheses. *Significant at the 10% level; **significant at the 5% level; ***significant at the 1% level.

a college degree substitute more towards lower quality products. These households are also more likely to reallocate their consumption expenditures towards low-price retailers, increase their shopping trips, and increase their purchases of items on sale than those with a head without college education.

Our second robustness check considers a more direct approach to study how households adjust their shopping effort after an income shock. It exploits the panel dimension of our data and uses within-household variation in income. However, the information on income provided by the households in the Consumer Panel has several disadvantages for this type of analysis. First, the income information is top-coded. Thus, we cannot capture income changes for the richest households. Second, the data do not capture many changes in the households' self-reported income for two reasons: (i) households stay an average of only three years in the sample, and (ii) the income information is reported in bins. We observe less than two income changes per household on average. Despite these limitations, Table 8 shows that after controlling for time- and household-fixed effects, product quality substitution remains a relevant margin of adjustment for high-income households after a negative income shock.[27] Although the results presented in the table indicate an increase in the number of shopping trips and the fraction of goods purchased on sale by high-income households relative to low-income households, these estimates are not precise enough under this variation of the data.

27. A negative shock is defined as a decrease of at least one bin in the household's self-reported income relative to the one they reported in the previous year.

TABLE 8. Shopping behavior: negative income shock.

	Product quality (1)	Store quality (2)	Trips (3)	Sales (4)
Negative shock	0.0497	− 0.0045	0.0342	− 0.1807**
	(0.106)	(0.011)	(0.065)	(0.078)
Negative shock $\times D_{\text{income group}=2}$	0.1621	− 0.0005	− 0.1346	0.1446
	(0.139)	(0.015)	(0.085)	(0.102)
Negative shock $\times D_{\text{income group}=3}$	− 0.3958***	0.0193	0.1109	0.2140**
	(0.140)	(0.015)	(0.086)	(0.104)
Negative shock $\times D_{\text{income group}=4}$	− 0.5609***	0.0057	0.1939	0.1335
	(0.210)	(0.023)	(0.129)	(0.155)
Time-fixed effect	Y	Y	Y	Y
Household-fixed effect	Y	Y	Y	Y
Observations	352,555	352,555	352,555	352,555
R-squared	0.626	0.749	0.780	0.928

Notes: The table reports the estimates for specification (14). Each observation is at the household × quarter level covering from 2006:1 to 2015:4. The dependent variable, "negative shock", is defined as a decrease of at least one bin in the household's self-reported income relative to the one they reported in the previous year. The coefficients for product quality and store quality are the percentage change in quality when the household suffers a negative income shock. The coefficient of the variable trips (sales) is the number of additional shopping trips per store (additional fraction of items purchased on sale) when the household experiences a negative income shock. Standard errors are presented in parentheses. *Significant at the 10% level; **significant at the 5% level; ***significant at the 1% level.

7. Conclusion

We find substantial differences in the price index across income groups, which increased during the Great Recession and have remained mostly stable afterwards. In particular, the annual inflation in the cost of living of the highest quartile was on average 0.85 percentage points lower than that of the lowest quartile of the income distribution from 2008 to 2013. We provide evidence that these differences are in part due to the ability of high-income households to adjust their shopping behavior and the quality of products they purchase to mitigate negative income shocks.

Our results have implications for the measurement of consumption and income inequality. Because inflation is lower for high-income households, the inequality in real variables may be rising faster than the inequality in nominal variables. Nonetheless, to fully capture the effect of inflation inequality on consumption and income inequality, it is necessary to extend our analysis to a more comprehensive set of goods than those included in our data. We leave the full economic implications of this analysis for future research.

Appendix A: Figures and Tables

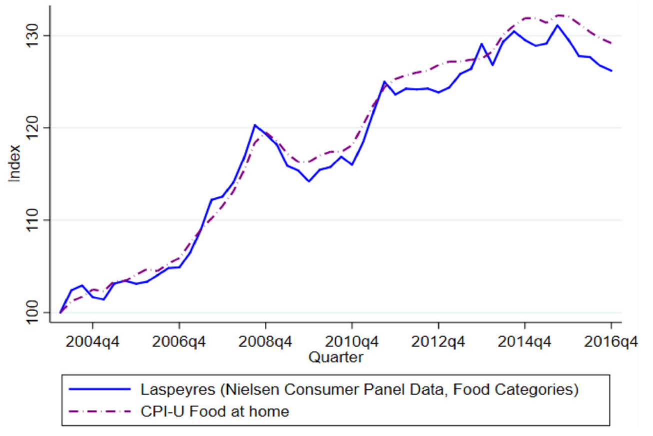


FIGURE A.1. Laspeyres index versus CPI-U food-at-home This graph plots a Laspeyres index from 2004:1 to 2016:4. The graph also plots the CPI for all urban consumers: food-at-home. From the Nielsen Consumer Panel data set, we pick product groups that match the food-at-home category in the CPI-U construction (dry grocery, frozen foods, dairy, deli, packaged meat, and fresh produce).

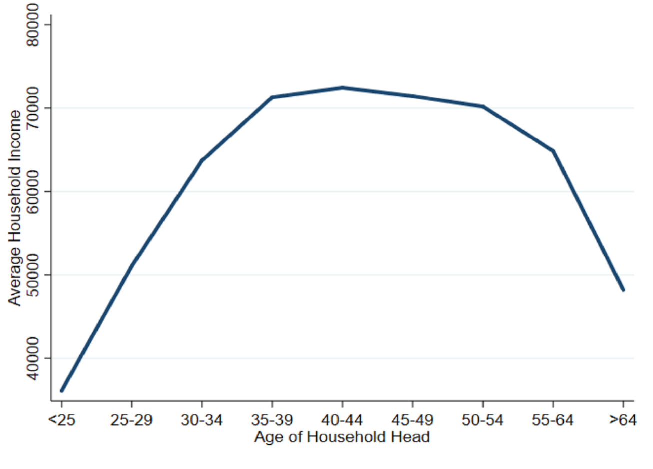
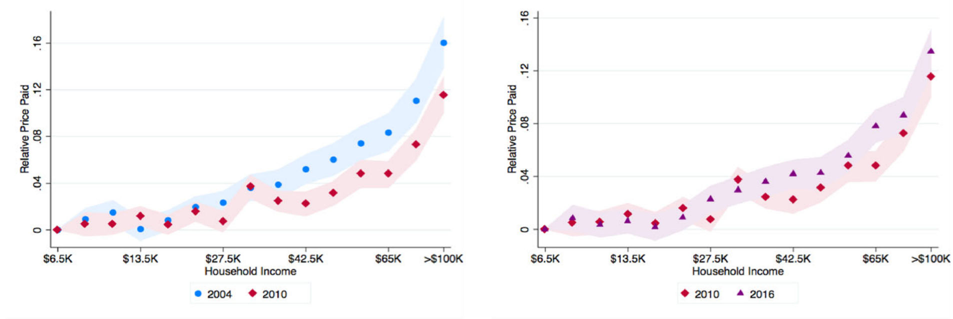


FIGURE A.2. Household income over head age. This graph depicts the annual household income over the age of household head in the Consumer Panel data set from 2004 to 2016. The y-axis depicts the average nominal annual household income and the x-axis depicts the household head's age.



(a) 2004 and 2010

(b) 2010 and 2016

FIGURE A.3. Relative price paid by household income. The relative prices are measured in a regression of the log-unit price paid against income category dummies, chain, product category, and county-fixed effects. Each dot represents how much more each income category pays per unit for products with respect to households earning $6,500 per year. The blue dots represent the cross-sectional relation in 2004, the red dots in 2010, and the purple dots in 2016. The standard errors are clustered at the product-module level. The figures report the 95% confidence interval.

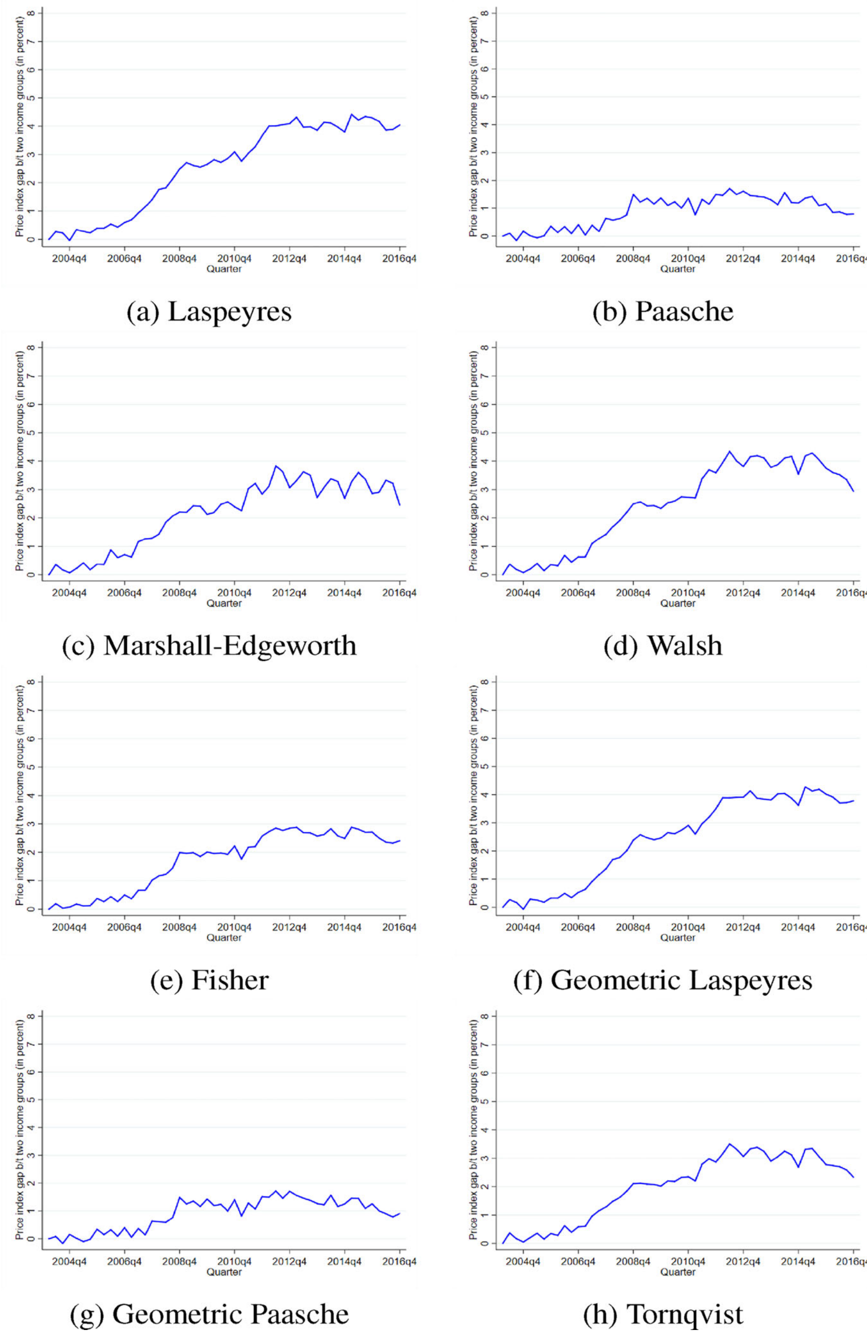
TABLE A.1. Distribution of expenditures over departments.

Code	Description	Product groups	No. of UPCs	Expenditure
0	Health and beauty aids	Baby care, cosmetics, cough & cold remedies, deodorant, hair care, oral hygiene, pain remedies, skin care, shaving	16.3%	10.2%
1	Dry grocery	Baby food, baking mixes, bottled water, candy, carbonated beverages, cereal, coffee, condiments, crackers, pet food, prepared foods, snacks, soup, canned vegetables	31.9%	35.4%
2	Frozen foods	Ice cream, frozen pizza, frozen vegetables	5.2%	9.6%
3	Dairy	Cheese, eggs, yogurt	3.8%	8.3%
4	Deli		1.6%	3.2%
5	Packaged meat		1.4%	3.2%
6	Fresh produce		1.2%	3.5%
7	Nonfood grocery	Detergent, diapers, fresheners/deodorizers, household cleaners, laundry supplies, pet care	12.5%	11.1%
8	Alcohol	Beer, wine, liquor, coolers	3.0%	3.5%
9	General merchandise	Batteries/flashlights, candles, computer/electronic, cookware, film/cameras, insecticides, lawn & garden, motor vehicle, office supplies	22.6%	11.4%

Notes: The Nielsen Consumer Panel data are organized into departments, product groups, product modules, and UPC codes. The departments, product groups, and product modules are all Nielsen-defined codes, while the UPC codes are defined by manufacturers. The table presents the share of UPCs and expenditures in each department in the data from 2004 to 2016.

Appendix B: Price Indexes With Common Prices

In this section, we compute various homothetic price indexes by assuming that all income groups pay the same average price for a given product. This approach is in contrast to the analysis in the main text in which we assume that the prices paid for a product are income-specific. Figure B.1 shows that in all cases and consistent with our previous findings, there is a gap between the price indexes of the top and bottom income groups, which grows during the recession and stabilizes and/or decreases in the postrecession subperiod.



(a) Laspeyres (b) Paasche

(c) Marshall-Edgeworth (d) Walsh

(e) Fisher (f) Geometric Laspeyres

(g) Geometric Paasche (h) Tornqvist

FIGURE B.1. The gaps between lowest and highest groups under various price indexes (common prices). The figure shows the percent gap between the price index of the lowest (less than 25k) income group and that of the highest (over 100k) income group. Each panel calculates the gap under a different price index. By construction, the gap is zero in 2004:1. We use the same prices for a given product across all income groups.

Appendix C: Stylized Nonhomothetic Price Index

We introduce a stylized nonhomothetic price index that builds on the work of Handbury (2019) and Faber and Fally (2019) to quantify the relative welfare gains of different income groups accounting for differences in their opportunities to substitute for quality.

C.1. Framework

We consider a two-tier utility in which the upper tier depends on the utility from consumer packaged goods U_C and the consumption of an outside good z:

$$U = U(U_C^I, z). \tag{C.1}$$

We assume that the outside good is normal. By allowing product-specific demand parameters to be a function of the outside-good consumption, we introduce nonhomotheticity in a reduced-form approach. Utility from consumer packaged goods is defined by

$$U_C^I = \Pi_g [\Sigma_{u \in U_g^I} (C_{ug} \varphi_{ug}^I)^{\frac{\sigma_g^I - 1}{\sigma_g^I}}]^{\alpha_g^I \frac{\sigma_g^I}{\sigma_g^I - 1}}, \tag{C.2}$$

where g is the product category and u is the barcode-level product. The term φ_{ug}^I refers to the perceived quality of product u in product category g at income level I. σ_g^I is the elasticity of substitution between products within each product category g at income level I. We use the same elasticity of substitution as estimated in Section 4.2.

Comparing two goods i and j within the same product category g, expenditures by consumers of income level I are then given by

$$\log \frac{s_{ni}^I}{s_{nj}^I} = (\sigma_g^I - 1)\left(\log \frac{\varphi_{ni}^I}{\varphi_{nj}^I} - \log \frac{p_{ni}^I}{p_{nj}^I}\right). \tag{C.3}$$

Following Faber and Fally (2019), we assume that quality valuations depend on an intrinsic quality term and a multiplicative income-specific term:

$$\log \varphi_{ug}^I = \gamma_g^I \log \phi_{ug}, \tag{C.4}$$

where ϕ_{ug} is the intrinsic quality and γ_g^I is the multiplicative term, which depends on income level I. With the normalization $\frac{1}{N^I} \sum_I \gamma_g^I = 1$, where N^I is the number of income groups, the intrinsic quality term also corresponds to the average quality evaluation across households: $\log \phi_{ug} = \frac{1}{N^I} \sum_I \log \varphi_{ug}^I$.

The price index is income-specific and given by $P_C^I = \Pi_g (P_n^I)^{\alpha_n^I}$, where the price index P_g^I for each category g is defined as

$$P_g^I = \left[\sum_{u \in U_g^I} p_{ug}^{I\ 1-\sigma_n^I} \varphi_{ug}^{I\ \sigma_n^I - 1}\right]^{\frac{1}{1-\sigma_n^I}}. \tag{C.5}$$

As in the construction of our benchmark income-specific indexes, this index allows for both income-specific elasticities of substitution to allow for nonhomotheticities and uses the actual price paid by the household to account for differences in the prices paid for the same product. Furthermore, equation (C.5) incorporates a heterogeneous quality choice by consumers of different incomes. To do so, we estimate unobserved differences in quality evaluation on products across households using observed moments on income-specific product sales and unit values.

C.2. Estimation

Using the elasticities of substitution within a product category, σ_g^I, estimated in Section 4.2, we estimate γ_g^I, which governs the valuation of product quality characteristics across the household income distribution from 2004 to 2016 in the Consumer Panel data set. From equation (C.4), we get the following equation to estimate:

$$\log\left(\varphi_{ugt}^I\right) = \gamma_g^I \log\left(\phi_{ugt}^I\right) + \alpha_t + \varepsilon_{ugt}^I, \tag{C.6}$$

where α_t indicates year effects. To address the concern of correlated measurement errors that appear on both the left- and right-hand sides of the equation, we follow Faber and Fally (2019) and an instrument for $\log(\phi_{ugt}^I)$ with two-year lagged values of product quality. Because we estimate quality evaluation year-on-year, we cannot directly compare price indexes across different years. Instead, within a year or a period, we compare the welfare differences between income groups by computing the gap in their price indexes. Thus, we calculate the average income-specific price indexes for the prerecession, recession, and postrecession subperiods.

The price index of the lowest income group (<25k) is normalized to 100 for all subperiods to compare the welfare gains across income groups within a subperiod. We find that during the prerecession the highest income group paid 12.3% more for their consumed products than the lowest income group. This difference could be interpreted as the welfare gap between groups for goods consumed that belong to the CPG sector. Consistent with our previous findings, during the recession, we find the opposite. The welfare gains of the highest income group relative to the lowest income group were approximately 2.2%. During the postrecession, the gap decreased; the welfare gains of the highest income group relative to the lowest income group were 0.7%.[28]

References

Aguiar, Mark and Erik Hurst (2005). "Consumption Versus Expenditure." *Journal of Political Economy*, 113, 919–948.

28. Our results are qualitatively similar when we use a balanced panel of products across all periods or when we estimate time-invariant quality evaluations to compare the welfare gains across income groups and across subperiods.

Aguiar, Mark and Erik Hurst (2007). "Life-Cycle Prices and Production." *American Economic Review*, 97, 1533–1559.

Argente, David, Munseob Lee, and Sara Moreira (2018). "Innovation and Product Reallocation in the Great Recession." *Journal of Monetary Economics*, 93, 1–20.

Beraja, Martin, Erik Hurst, and Juan Ospina (2019). "The Aggregate Implications of Regional Business Cycles." *Econometrica*, 87, 1789–1833.

Bils, Mark and Peter J. Klenow (2001). "Quantifying Quality Growth." *American Economic Review*, 91, 1006–1030.

Broda, Christian, Ephraim Leibtag, and David E. Weinstein (2009). "The Role of Prices in Measuring the Poor's Living Standards." *The Journal of Economic Perspectives*, 23, 77–97.

Broda, Christian and John Romalis (2009). "The Welfare Implications of Rising Price Dispersion." Working Paper, University of Chicago.

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

Broda, Christian and David E. Weinstein (2010). "Product Creation and Destruction: Evidence and Price Implications." *American Economic Review*, 100, 691–723.

Burstein, Ariel, Martin Eichenbaum, and Sergio Rebelo (2005). "Large Devaluations and the Real Exchange Rate." *Journal of Political Economy*, 113, 742–784.

Chevalier, Judith A. and Anil K. Kashyap (2019). "Best Prices: Price Discrimination and Consumer Substitution." *American Economic Journal: Economic Policy*, 11, 126–159.

Coibion, Olivier, Yuriy Gorodnichenko, and Gee Hee Hong (2015). "The Cyclicality of Sales, Regular and Effective Prices: Business Cycle and Policy Implications." *American Economic Review*, 105, 993–1029.

Dubé, Jean-Pierre, Günter J. Hitsch, and Peter E. Rossi (2018). "Income and Wealth Effects on Private-Label Demand: Evidence From the Great Recession." *Marketing Science*, 37, 22–53.

Einav, Liran, Ephraim Leibtag, and Aviv Nevo (2010). "Recording Discrepancies in Nielsen Homescan Data: Are They Present and Do They Matter?" *Quantitative Marketing and Economics*, 8, 207–239.

Faber, Benjamin (2014). "Trade Liberalization, the Price of Quality, and Inequality: Evidence From Mexican Store Prices." Working Paper, University of California, Berkeley

Faber, Benjamin and Thibault Fally (2019). "Firm Heterogeneity in Consumption Baskets: Evidence from Home and Store Scanner Data." NBER Working Paper No. 23101, University of California, Berkeley.

Feenstra, Robert C. (1994). "New Product Varieties and the Measurement of International Prices." *American Economic Review*, 84, 157–177.

Garner, Thesia I., David S. Johnson, and Mary F. Kokoski (1996). "An Experimental Consumer Price Index for the Poor." *Monthly Labor Review*, 119, 32.

Griffith, Rachel, Ephraim Leibtag, Andrew Leicester, and Aviv Nevo (2009). "Consumer Shopping Behavior: How Much Do Consumers Save?" *Journal of Economic Perspectives*, 23, 99–120.

Hagemann, Robert P. (1982). "The Variability of Inflation Rates Across Household Types." *Journal of Money, Credit and Banking*, 14, 494–510.

Handbury, Jessie (2019). "Are Poor Cities Cheap for Everyone? Non-Homotheticity and the Cost of Living Across U.S. Cities." NBER Working Paper No. 26574, University of Pennsylvania.

Handbury, Jessie, Tsutomu Watanabe, and David E. Weinstein (2015). "How Much Do Official Price Indexes Tell Us About Inflation?" NBER Working Paper No. 19504, University of Pennsylvania.

Hansen, Lars Peter (1982). "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica*, 50, 1029–1054.

Hobijn, Bart and David Lagakos (2005). "Inflation Inequality in the United States." *Review of Income and Wealth*, 51, 581–606.

Hottman, Colin J., Stephen J. Redding, and David E. Weinstein (2016). "Quantifying the Sources of Firm Heterogeneity." *The Quarterly Journal of Economics*, 131, 1291–1364.

Jaimovich, Nir, Sergio Rebelo, and Arlene Wong (2019). "Trading Down and the Business Cycle." *Journal of Monetary Economics*, 102, 96–121.

Jaravel, Xavier (2019). "The Unequal Gains From Product Innovations: Evidence From the US Retail Sector." *The Quarterly Journal of Economics*, 134, 715–783.

Kaplan, Greg and Sam Schulhofer-Wohl (2017). "Inflation at the Household Level." *Journal of Monetary Economics*, 91, 19–38.

Kryvtsov, Oleksiy and Nicolas Vincent (2019). "The Cyclicality of Sales and Aggregate Price Flexibility." Working Paper, Bank of Canada

McKenzie, David and Ernesto Schargrodsky (2011). "Buying Less, but Shopping More: The Use of Nonmarket Labor During a Crisis." *Economia*, 11, 1–43.

Michael, Robert T. (1979). "Variation Across Households in the Rate of Inflation." *Journal of Money, Credit and Banking*, 11, 32–46.

Michelacci, Claudio, Luigi Paciello, and Andrea Pozzi (2019). "The Extensive Margin of Aggregate Consumption Demand." Working Paper, Einaudi Institute for Economics and Finance.

Nevo, Aviv and Arlene Wong (2019). "The Elasticity of Substitution Between Time and Market Goods: Evidence From the Great Recession." *International Economic Review*, 60, 25–51.

Petev, Ivaylo, Luigi Pistaferri, and Itay Saporta Eksten (2011). "Consumption and the Great Recession: An Analysis of Trends, Perceptions, and Distributional Effects." Working Paper, Stanford University

Pistaferri, Luigi (2016). "Why Has Consumption Remained Moderate After the Great Recession." Working Paper, Stanford University

Sato, Kazuo (1976). "The Ideal Log-Change Index Number." *Review of Economics and Statistics*, 58, 223–228.

Stroebel, Johannes and Joseph Vavra (2019). "House Prices, Local Demand, and Retail Prices." *Journal of Political Economy*, 127, 1391–1436.

Vartia, Yrjö O. (1976). "Ideal Log-Change Index Numbers." *Scandinavian Journal of Statistics*, 3, 121–126.

Supplementary Data

Supplementary data are available at *JEEA* online.