# A Distributional PCE Price Index From Aggregate Data

Philipp Hochmuth[†] Markus Pettersson[‡] Christoffer Jessen Weissert[§]

September 21, 2023

**Abstract**

This paper proposes a method to measure individual and aggregate changes in the cost of living when consumer behavior is nonhomothetic and microdata on consumption expenditures are not available. Aggregate prices and expenditure shares together with a single cross-sectional distribution of expenditures are sufficient to create a distribution of nonhomothetic cost-of-living indices with this approach. The cost-of-living indices derive from PIGL preferences, generalize the Törnqvist price index, and only contain two unknown parameters. Because PIGL preferences aggregate consistently, these parameters can be identified from aggregate data. Using US Personal Consumption Expenditure (PCE) data, the method is applied to obtain a nonhomothetic PCE price index covering 72 product groups. This index reveals a 0.5–1.2 percentage point gap in annual inflation rates between the poorest and richest ten percents throughout 2022, and a similar 0.2 percentage point gap on average since 1988, thus suggesting that poorer households are hit harder both in the ongoing inflation surge and in the long run.

**Keywords:** cost of living, inflation inequality, nonhomotheticity, Personal Consumption Expenditures.

**JEL Codes:** C43, D12, D3, E01, E31.

[†] Institute for International Economic Studies, Stockholm University. Email: philipp.hochmuth@iies.su.se.
[‡] Department of Economics, Stockholm University. Email: markus.pettersson@su.se.
[§] Department of Economics and Monetary Policy, Danmarks Nationalbank. Email: cjw@nationalbanken.dk.

# 1 Introduction

Over 160 years of empirical research, following the early work of Engel (1857), suggests that consumption patterns vary systematically with income. Because households consume different bundles, it is widely acknowledged that price movements generate heterogeneous changes in the cost of living at the individual level. Yet, despite a growing interest in distributional questions related to national income and monetary policy, most national accounts-based deflators and inflation measures still rely on aggregate cost-of-living indices that are assumed to apply to everyone. The reason is mainly practical: estimating similar cost-of-living indices for individual households or even subgroups of households generally requires detailed microdata on consumer expenditures that are rarely available to national income accountants. This paper presents a method to overcome this limitation, and uses it to document inflation disparities in US Personal Consumption Expenditure (PCE) data.

The approach considered here is a generalization of the Törnqvist (1936) cost-of-living index that (i) allows consumer behavior to differ with the consumption expenditure level, (ii) can be implemented using publicly available macroeconomic data on prices and expenditure shares only, and (iii) nests the standard Törnqvist index as a limit case. While all other commonly used cost-of-living indices also satisfy (ii), they do so under the assumption of homothetic preferences, in which consumer behavior is independent of the expenditure level. These indices therefore only describe cost-of-living changes for an average household. By contrast, in the framework considered here a single cross-sectional distribution of household consumption expenditures is sufficient to generate a full distribution of cost-of-living indices at the household level in addition to the aggregate-level index.

Like the closely related paper by Hochmuth, Pettersson and Weissert (2023), the cost-of-living index relies on a theoretical foundation with utility-maximizing households whose preferences are of the "price independent generalized linearity" (PIGL) form originally defined by Muellbauer (1975, 1976).[1] These preferences are nonhomothetic, meaning that rich consumers allocate a larger budget share to luxuries than poor consumers, but nevertheless maintain tractable aggregation properties that allow us to consistently estimate any preference parameters from aggregate expenditure data. As shown by Hochmuth, Pettersson and Weissert, PIGL preferences generalize all common homothetic cost-of-living indices, including the Törnqvist index, and allow straightforward decompositions to identify the commodities that drive any overall changes in the cost of living.

The implementation of the PIGL cost-of-living index rests upon a separable preference structure in which commodities are bundled into three intermediate baskets: necessities, luxuries, and homothetic goods. In doing so, the cost-of-living index of any individual or group becomes a function of four components:

(i) their total expenditure share allocated to necessities in some chosen *base period*;
(ii) the aggregate expenditure share allocated to the homothetic bundle in *every period*;
(iii) the prices of each basket; and
(iv) two preference parameters, the elasticity of demand for necessities and a parameter that governs the elasticity of substitution between necessities and luxuries.

[1] These preferences have become popular in the structural change literature; see Boppart (2014), Fan, Peters and Zilibotti (2021), Alder, Boppart and Müller (2022), and Cravino, Levchenko and Rojas (2022).

Behavior within each intermediate basket is homothetic, so the prices of these are captured by standard Törnqvist indices that can be computed from aggregate time series. Thus, given an observed base-period distribution of household expenditures and aggregate time series on prices and expenditure shares, this approach requires the estimation of only two parameters, thereby keeping econometric concerns to a minimum.

The empirical analysis focuses on US PCE data, which are used to construct a nonhomothetic PCE price index over 72 separate commodity groups. The official PCE price index is the consumption deflator in the US national accounts and the main inflation measure for US monetary policy. Yet, the PCE data only covers aggregate consumption expenditures, so distributional inflation analyses with these data have not been possible before. This paper overcomes that limitation through a recent attempt by Garner *et al.* (2022) to distribute PCEs across US households for the year 2019. Since the proposed approach here only needs a single cross-sectional distribution of household expenditures, these estimates for 2019 are sufficient to characterize a full distribution of PCE inflation rates.

The empirical results suggest that consumption-poor households face considerably higher PCE inflation than consumption-rich households, and in particular so during the ongoing inflation surge. Specifically, the bottom decile of the 2019 expenditure distribution faced an annual inflation rate which was 0.8 percentage points larger on average than that of the top decile during 2022, peaking at 1.2 percentage points in June of the same year. A decomposition of this gap identifies price increases for food consumed at home, energy, and motor vehicles as major drivers of the higher inflation of poor households. These are partially offset by increasing costs for restaurant meals and accommodations, transportation services, and financial services, which are consumed proportionately more by the rich.

Similar albeit smaller differences remain when a full 35-year period is considered instead of contemporary short-run developments. Between 1988 and 2023, the gap in annual inflation rates between the top and bottom deciles amounts to 0.2 percentage points on average. This difference adds up to a cumulative 8 percent larger increase in the cost-of-living for the bottom decile over the same period. These long-run findings corroborate much of what Jaravel and Lashkari (2023) find with Consumer Expenditure Survey data, both qualitatively and quantitatively.

Overall, the contribution of this paper is threefold. First, the cost-of-living index presented here extends Hochmuth, Pettersson and Weissert's (2023) nonhomothetic generalization of the Törnqvist price index, allowing commodities to be not only luxuries or necessities but also homothetic. The paper thereby adds to the literature on the economic approach to price index theory following, among many others, Konüs (1939), Samuelson and Swamy (1974), Diewert (1976, 1978), Feenstra (1994), Feenstra and Reinsdorf (2000), and Redding and Weinstein (2020).

Second, to the best of my knowledge this paper obtains the first-ever distribution of PCE inflation rates across households with different levels of consumption expenditures. This adds to the growing efforts by for instance Fixler *et al.* (2017), Piketty, Saez and Zucman (2018), Fixler, Gindelsky and Johnson (2020), and Garner *et al.* (2022) to construct distributional measures of income, consumption, and wealth in the national accounts. Up until now, the primary focus of these efforts has been on nominal variables. Differences in the denominator of real variables, the cost-of-living deflator, has received limited attention thus far.

Lastly, by showing that consumption-poor households in the United States face significantly higher inflation rates than the consumption-rich, both during the current inflation surge and in the long run, this paper contributes to the ever-growing literature on the measurement of inflation inequality that was recently surveyed by Jaravel (2021). Previous studies primarily use detailed microdata such as the Consumer Expenditure Survey or the Kilts-Nielsen Consumer Panel to construct separate homothetic price indices for different consumer groups.[2,3] By contrast, this paper uses a theoretically consistent framework and is the first to consider PCE data.

## 2 PIGL-Törnqvist Model

This section outlines the theoretical model and the corresponding cost-of-living index. Much of the following material draws closely on Hochmuth, Pettersson and Weissert (2023) and readers are referred to that paper for additional details.

### 2.1 Preferences

As in Boppart (2014), suppose a household with consumption expenditure $e$ who is faced with a price vector $\boldsymbol{p}$ has an indirect utility function of Muellbauer's (1975, 1976) PIGL form:

$$V(e, \boldsymbol{p}) = \frac{1}{\varepsilon}\left[\left(\frac{e}{F\big(H(\boldsymbol{p}), B(\boldsymbol{p})\big)}\right)^{\varepsilon} - 1\right] - \frac{\nu}{\gamma}\left[\left(\frac{D(\boldsymbol{p})}{B(\boldsymbol{p})}\right)^{\gamma} - 1\right], \tag{1}$$

where $0 < \varepsilon \leq 1$, $0 < \gamma \leq 1$, and $\nu > 0$. As will become evident below, $\varepsilon$ governs the expenditure elasticity of demand for necessity goods. The parameter $\gamma$ controls the (nonconstant) elasticity of substitution between necessities and luxuries, while $\nu$ is a scale parameter. The functions $B(\boldsymbol{p})$, $D(\boldsymbol{p})$, and $H(\boldsymbol{p})$ are linearly homogenous and are treated throughout as unit cost functions of some intermediate homothetic consumption bundles (which are similarly referred to as the $B$, $D$, and $H$ baskets). The function $F$ is a CES composite of $H(\boldsymbol{p})$ and $B(\boldsymbol{p})$:

$$F\Big(H(\boldsymbol{p}), B(\boldsymbol{p})\Big) = \Big[\theta H(\boldsymbol{p})^{1-\sigma} + (1-\theta) B(\boldsymbol{p})^{1-\sigma}\Big]^{\frac{1}{1-\sigma}}, \tag{2}$$

where $\sigma > 1$ denotes the asymptotic elasticity of substitution between the $B$ and $H$ baskets as $e \to \infty$ and $\theta \in (0, 1)$ is a taste parameter for the $H$ basket. Equation (2) is a generalization of Hochmuth, Pettersson and Weissert's (2023) cost-of-living framework, who consider the special case in which $\theta \to 0$, such that $F\big(H(\boldsymbol{p}), B(\boldsymbol{p})\big) = B(\boldsymbol{p})$.

To add some structure that makes the demand system empirically tractable, a key separability restriction is imposed on preferences:

**Assumption 1.** Preferences are quasi-separable between $B(\boldsymbol{p})$, $D(\boldsymbol{p})$, and $H(\boldsymbol{p})$.[4] ◁

[2] Examples include Hobijn and Lagakos (2005), McGranahan and Paulson (2005), Broda and Romalis (2009), Kaplan and Schulhofer-Wohl (2017), Jaravel (2019), Argente and Lee (2021), Klick and Stockburger (2021), Lauper and Mangiante (2021), and Orchard (2022).

[3] Other recent efforts also focus on nonparametric approaches, see Atkin *et al.* (2020), Baqaee, Burstein and Koike-Mori (2022), and Jaravel and Lashkari (2023).

[4] Quasi-separability groups *prices* of goods in the *expenditure function*, in contrast to direct separability which groups *quantities* of goods in the *utility function*.

Under Assumption 1, the price of an individual commodity occurs in one and only one of the three price functions $B$, $D$, and $H$. This permits two-stage budgeting in which households first allocate expenditures between the three bundles, and subsequently make within-basket decisions conditional on the first-stage allocation. The expenditure share of bundle $C \in \{B, D, H\}$, defined as $w_C \equiv \sum_{j \in J_C} p_j q_j \, / \, e$ with $J_C$ denoting the set of goods in $C$ and $p_j$ and $q_j$ denoting the price and quantity of commodity $j$, is given by Roy's identity as

$$w_D = \nu \left( \frac{F\big(H(\boldsymbol{p}), B(\boldsymbol{p})\big)}{e} \right)^{\varepsilon} \left( \frac{D(\boldsymbol{p})}{B(\boldsymbol{p})} \right)^{\gamma}, \tag{3}$$

$$w_H = \theta \left( \frac{H(\boldsymbol{p})}{F\big(H(\boldsymbol{p}), B(\boldsymbol{p})\big)} \right)^{1-\sigma}, \tag{4}$$

$$w_B = (1-\theta) \left( \frac{B(\boldsymbol{p})}{F\big(H(\boldsymbol{p}), B(\boldsymbol{p})\big)} \right)^{1-\sigma} - \nu \left( \frac{F\big(H(\boldsymbol{p}), B(\boldsymbol{p})\big)}{e} \right)^{\varepsilon} \left( \frac{D(\boldsymbol{p})}{B(\boldsymbol{p})} \right)^{\gamma}. \tag{5}$$

Given that $\varepsilon$ is positive, Equations (3) to (5) highlight that as the expenditure level increases, the budget share for $D$ declines, the share for $H$ remains unchanged, and the share for $B$ increases. Since within-basket behavior is homothetic, it follows that $D$ is a bundle of *necessities*, $H$ is a bundle of *homothetic goods*, and $B$ is a bundle of *luxuries*.[5]

### 2.2 Cost-of-Living Index

To briefly review known results, let the minimum consumption expenditure needed to reach some utility level $u$ when faced by a price vector $\boldsymbol{p}$ be given by the expenditure function $e = c(u, \boldsymbol{p})$. Konüs (1939) defines a cost-of-living index to be the change in minimum expenditures needed to maintain a fixed utility level as prices change from some base-period price vector $\boldsymbol{p}_s$ to a period-$t$ price vector $\boldsymbol{p}_t$:

$$P(u, \boldsymbol{p}_t, \boldsymbol{p}_s) \equiv \frac{c(u, \boldsymbol{p}_t)}{c(u, \boldsymbol{p}_s)}. \tag{6}$$

The Konüs index (6) is independent of the reference utility level if and only if preferences are homothetic, in which case the index becomes a ratio of unit cost functions.[6] The prices of the three intermediate bundles are therefore simply

$$P_{Bt} = \frac{B(\boldsymbol{p}_t)}{B(\boldsymbol{p}_s)}, \qquad P_{Dt} = \frac{D(\boldsymbol{p}_t)}{D(\boldsymbol{p}_s)}, \qquad \text{and} \qquad P_{Ht} = \frac{H(\boldsymbol{p}_t)}{H(\boldsymbol{p}_s)}, \tag{7}$$

where the function arguments on the left-hand sides are left implicit to simplify notation. Moreover, $F\big(H(\boldsymbol{p}), B(\boldsymbol{p})\big)$ is a CES composite. Following Sato (1976) and Vartia (1976), its value in period $t$ relative to any other period $k$ can be written as a function of the expenditure share on $H$ in both periods and the price changes of $B$ and $H$: using Equation (4) it can be

[5] Formally, it is easy to show from Equations (3) to (5) that the expenditure elasticities of demand for $D$, $H$, and $B$ are, respectively, $1 - \varepsilon < 1$, 1, and $1 + \varepsilon \frac{w_D}{w_B} > 1$.

[6] See for instance the *Homogeneity Price Theorem* in Samuelson and Swamy (1974).

shown that

$$\frac{F\big(H(\boldsymbol{p}_t),B(\boldsymbol{p}_t)\big)}{F\big(H(\boldsymbol{p}_k),B(\boldsymbol{p}_k)\big)} \;=\; \left(\frac{P_{Bt}}{P_{Bk}}\right)^{1-\rho_{t,k}} \left(\frac{P_{Ht}}{P_{Hk}}\right)^{\rho_{t,k}}, \tag{8a}$$

where the weight on the $H$ basket is given by

$$\rho_{t,k} \;=\; \frac{L\big(w_{Ht},w_{Hk}\big)}{L\big(w_{Ht},w_{Hk}\big)+L\big(1-w_{Ht},1-w_{Hk}\big)}. \tag{8b}$$

In (8b), $L(\cdot,\cdot)$ denotes the logarithmic mean, defined for positive values $x$ and $y$ as

$$L(x,y) \;=\; \begin{cases} \dfrac{x-y}{\ln x-\ln y} & \text{if } x\neq y,\\ x & \text{if } x=y. \end{cases} \tag{9}$$

Equation (8) is convenient because it implies that we do not need to know the two parameters $\sigma$ and $\theta$ in empirical applications as long as we observe the expenditure shares for the $H$ bundle.

Having established the necessary foundations, we now turn to the cost-of-living index that corresponds to the indirect utility function (1). First, inverting the indirect utility function yields an expenditure function of the form

$$c(u,\boldsymbol{p}) \;=\; \left[\left(1-\frac{\varepsilon\nu}{\gamma}+\varepsilon u\right)B(\boldsymbol{p})^{\gamma} \;+\; \frac{\varepsilon\nu}{\gamma}D(\boldsymbol{p})^{\gamma}\right]^{\frac{1}{\varepsilon}} \frac{F\big(H(\boldsymbol{p}),B(\boldsymbol{p})\big)}{B(\boldsymbol{p})^{\gamma/\varepsilon}}.$$

As shown by Hochmuth, Pettersson and Weissert (2023, Proposition 1), it is possible to pin down the reference utility level $u$ in the expenditure function with the base-period expenditure share on necessities, $w_{Ds}$. Specifically, take the indirect utility function (1) in the base period $s$, substitute for $u$ in any period-$t$ expenditure function, and apply Equation (3). Together with the Konüs definition (6) and the price indices (7) and (8), we then obtain the PIGL cost-of-living index

$$P(u,\boldsymbol{p}_t,\boldsymbol{p}_s) \;=\; \left[\left(1-\frac{\varepsilon w_{Ds}}{\gamma}\right)P_{Bt}^{\gamma} \;+\; \frac{\varepsilon w_{Ds}}{\gamma}P_{Dt}^{\gamma}\right]^{\frac{1}{\varepsilon}} P_{Bt}^{1-\frac{\gamma}{\varepsilon}-\rho_{t,s}}\,P_{Ht}^{\rho_{t,s}}. \tag{10}$$

The cost-of-living index (10) is what is taken to the data and, to reiterate the introduction, it is a function of four components:

(i) the expenditure share $w_D$ allocated to necessities in the *base period* $s$;
(ii) the expenditure share $w_H$ allocated to homothetic goods in *both periods* $s$ and $t$;
(iii) the prices of each basket; and
(iv) the two preference parameters $\varepsilon$ and $\gamma$.

Heterogeneity in the cost of living occurs because the expenditure share for necessities varies with the expenditure level, so richer individuals allocate a lower weight to price changes of necessities, as captured by $P_{Dt}$. Yet, this nonhomotheticity only shows up explicitly in the price index formula through the base-period allocation. Consequently, knowledge about the expenditure distribution for periods other than the base period is not needed. Expenditure shares for the $H$

bundle are required for all periods considered, but these shares are homothetic and thus identical for everyone. They can therefore be obtained from aggregate time series in the data. This makes the PIGL cost-of-living index ideal for the present analysis. Moreover, while Equation (10) is derived for an individual household, an identical formula also holds for any group of households, in which $P_{Dt}$ is weighted by the group's aggregate expenditure share on necessities.

The cost-of-living index (10) gives the total change in the cost of living, but it is also possible to decompose this change into individual contributions of each basket. Following Hochmuth, Pettersson and Weissert (2023, Lemma 1), the period-to-period change in the cost of living for someone with a base-period allocation $w_{Ds}$ can be written as

$$\frac{P(u, \boldsymbol{p}_t, \boldsymbol{p}_s)}{P(u, \boldsymbol{p}_{t-1}, \boldsymbol{p}_s)} = \left(\frac{P_{Dt}}{P_{Dt-1}}\right)^{\frac{\gamma\phi_{t,t-1}(u)}{\varepsilon}} \left(\frac{P_{Bt}}{P_{Bt-1}}\right)^{1-\frac{\gamma\phi_{t,t-1}(u)}{\varepsilon}-\rho_{t,t-1}} \left(\frac{P_{Ht}}{P_{Ht-1}}\right)^{\rho_{t,t-1}}, \tag{11}$$

where the weight $\phi_{t,t-1}(u)$ on necessities, which varies across households, is defined as

$$\phi_{t,t-1}(u) = \frac{L\big(w^h_{Dt}, w^h_{Dt-1}\big)}{L\big(w^h_{Dt}, w^h_{Dt-1}\big) + L\big(\frac{\gamma}{\varepsilon} - w^h_{Dt}, \frac{\gamma}{\varepsilon} - w^h_{Dt-1}\big)}. \tag{12}$$

In Equation (12), $L(\cdot, \cdot)$ is again the logarithmic mean (9) while $w^h_{Dt}$ and $w^h_{Dt-1}$ are *Hicksian* expenditure shares. That is, $w^h_{Dt}$ and $w^h_{Dt-1}$ are the necessity expenditure shares that prevail under period-$t$ and period-$t-1$ prices along the indifference curve associated with the observed base-period allocation $w_{Ds}$. Though not directly observable, Hicksian expenditure shares are straightforward to construct for decomposition purposes: apply Shephard's lemma on the expenditure function in period $k \in \{t, t-1\}$ and use Equations (3) and (10) to get

$$w^h_{Dk} = \frac{w_{Ds} P^\gamma_{Dk}}{(1 - \frac{\varepsilon w_{Ds}}{\gamma}) P^\gamma_{Bk} + \frac{\varepsilon w_{Ds}}{\gamma} P^\gamma_{Dk}}. \tag{13}$$

### 2.3 A Nonhomothetic Törnqvist Index

It remains to parametrize the price indices for the three intermediate baskets. To that end, let $B(\boldsymbol{p})$, $D(\boldsymbol{p})$, and $H(\boldsymbol{p})$ be homogeneous translog expenditure functions. As shown by Diewert (1976), this implies that the corresponding price indices are of the Törnqvist (1936) form. That is, for each bundle $C \in \{B, D, H\}$, we have

$$\frac{P_{Ct}}{P_{Ct-1}} = \prod_{j \in J_C} \left(\frac{p_{jt}}{p_{jt-1}}\right)^{\delta^C_{j,t,t-1}}, \qquad \delta^C_{j,t,t-1} = \frac{w^C_{jt} + w^C_{jt-1}}{2}, \tag{14}$$

where $w^C_j \equiv p_j q_j \,/\, \sum_{j \in J_C} p_j q_j$ is the within-basket expenditure share. Since within-basket behavior is homothetic, these within-shares are identical across agents. The Törnqvist weights can therefore be constructed from aggregate data, just like the weight on the $H$ basket in Equation (10).

The geometric-mean representation of Equations (11) and (14) makes it straightforward to decompose the overall change in the cost of living into contributions of individual commodities, which is useful in applications. While the same is true for any other homothetic geometric-mean parametrization of $P_{Ct}/P_{Ct-1}$, the Törnqvist index is particularly neat because Equations (11) and (14) then nest the standard Törnqvist index as a limit case. We finish the section by stating this as a formal result, since it generalizes Proposition 4 in Hochmuth, Pettersson and Weissert (2023) to include the homothetic bundle $H$.

**Proposition 1.** *Let $B(\boldsymbol{p})$, $D(\boldsymbol{p})$, and $H(\boldsymbol{p})$ be homogeneous translog expenditure functions. If $\varepsilon \to 0$, $\gamma \to 0$, and $\sigma \to 1$, then the PIGL cost-of-living index (11) becomes the standard Törnqvist index:*

$$\frac{P(u, \boldsymbol{p}_t, \boldsymbol{p}_s)}{P(u, \boldsymbol{p}_{t-1}, \boldsymbol{p}_s)} = \prod_{j \in J} \left( \frac{p_{jt}}{p_{jt-1}} \right)^{\delta_{j,t,t-1}}, \qquad \delta_{j,t,t-1} = \frac{w_{jt} + w_{jt-1}}{2},$$

*where $J = J_D \cup J_B \cup J_H$ is the full set of commodities available and $w_j = p_j q_j \,/\, e$ is the expenditure share of commodity $j$.*

*Proof.* If $\varepsilon \to 0$, $\gamma \to 0$, and $\sigma \to 1$, the indirect utility function (1) becomes Cobb-Douglas: $V(e, \boldsymbol{p}) = \ln e - \ln \left[ D(\boldsymbol{p})^{\nu} B(\boldsymbol{p})^{1-\theta-\nu} H(\boldsymbol{p})^{\theta} \right]$. The cost-of-living index between periods $t$ and $t-1$ is then $\left(P_{Dt}/P_{Dt-1}\right)^{\nu} \left(P_{Bt}/P_{Bt-1}\right)^{1-\theta-\nu} \left(P_{Ht}/P_{Ht-1}\right)^{\theta}$, where the weights are time-invariant expenditure shares: $\nu = w_D$, $1 - \theta - \nu = w_B$, and $\theta = w_H$. Substituting in (14), the weight on good $j$ in bundle $C$ becomes $w_C \, \delta^C_{j,t,t-1} = (w_{jt} + w_{jt-1})/2$, since $w_j = w_C \, w^C_j$ holds by definition under Assumption 1. □

## 3 Data and Empirical Implementation

This section applies the theory above to aggregate US data on consumption expenditures and prices. The goal is to construct a nonhomothetic version of the PCE price index, which is the consumption deflator in the US national accounts and the main inflation measure for US monetary policy. All in all, the analysis requires no more than six publicly available tables from the US Bureau of Economic Analysis (BEA), three from the National Income and Product Accounts (NIPA) and another three from the Regional Economic Accounts (REA), which are combined with distributional PCE estimates from Garner *et al.* (2022).

The implementation strategy exploits the separability structure of the theoretical model, in which a good belongs to one and only one of the three commodity bundles. Under such separability, the overall basket expenditure shares ($w_C$) and the individual within-basket expenditure shares ($w^C_j$) can be inferred directly from the observed expenditure shares on individual goods ($w_j$) once we know the sets of goods that belong to each basket. These sets of goods can be identified via the fact that one basket is a bundle of necessities, one is a bundle of luxuries, and one is a bundle of homothetic goods: simply investigate the Engel curves of individual goods and classify them as necessity, luxury, or homothetic based on the slopes of the Engel curves. Following Hochmuth, Pettersson and Weissert (2023), the empirical procedure can then be summarized by the following steps:

(i) classify each individual good considered as "necessity", "luxury", or "homothetic";
(ii) construct the Törnqvist price index (14) for each of the three baskets using within-basket expenditure shares for the sets of commodities identified in (i);
(iii) estimate $\varepsilon$ and $\gamma$ using the expenditure share equation (3); and
(iv) for a given base-period expenditure distribution, construct the corresponding PIGL cost-of-living indices using (ii) and (iii).

The subsections below cover these steps in turn and provide additional detail on the data used in this regard.

### 3.1 Data on US Personal Consumption Expenditures and Prices

The classification of goods relies on annual PCE data by US state, which are available from 1997 onward in REA Table SAPCE3 and include details on over 70 separate consumption categories at the lowest level of product aggregation. The classification exercise also makes adjustments for price differences across states with the regional price parities reported in REA Table SARPP, which are available at an annual frequency starting in 2008 for four broad expenditure groups (goods, housing, utilities, and other services). The cost-of-living index itself is constructed from monthly time series on aggregate US expenditures and prices for each of these consumption categories, which are provided in the underlying detail tables, NIPA Tables 2.4.4U and 2.4.5U. All expenditures in these sources are converted into per-capita terms using the population estimates reported in REA Table SAINC1 and NIPA Table 2.6.

Three consumption categories are excluded from the analysis and subtracted from the expenditure totals: net expenditures abroad by US residents, net foreign travel, and final consumption expenditures of nonprofit institutions serving households. The former two are dropped because their expenditures are not guaranteed to remain positive and because the BEA does not provide any corresponding price indices to use. The latter is excluded from the distributional PCE estimates by Garner *et al.* (2022), and is consequently also dropped here. This restriction results in a data set covering 72 distinct product groups which reflects consumer spending *in* the United States and *by* US *households*.

Lastly, periods that lack data on some of the remaining 72 consumption categories are dropped from the aggregate US time series, as the Törnqvist index does not easily deal with the entry and exit of products. The data used to construct the cost-of-living index subsequently covers the 35-year period starting in January 1988 and ending in January 2023.

### 3.2 Classification of Goods and Basket Price Indices

The classification of a good is implemented by investigating the slope of budget-share Engel curves implied from the cross-sectional variation in consumption expenditures across US states. As in for instance Wachter and Yogo (2010), Orchard (2022), and Hochmuth, Pettersson and Weissert (2023), the classification relies on a simple allocation rule: a good is classified as a necessity if the slope of its estimated Engel curve is negative and statistically significant at the 5 percent level, as a luxury if the slope is positive and similarly significant, and otherwise considered a homothetic good.

The Engel curve of a product $j$ is estimated across states $g$ and years $t$ by regressing the state-level aggregate expenditure share $\overline{w}_{jgt}$ on the corresponding per-capita consumption expenditure $\bar{e}_{gt}$, according to a reduced-form model

$$\overline{w}_{jgt} = \alpha_{jr} + \alpha_{jt} + \beta_{je} \ln \bar{e}_{gt} + \beta_{jp} \ln RPP_{jgt} + u_{jgt}. \tag{15}$$

In this regression, $\alpha_{jr}$ is a dummy for the BEA region in which a state is located, $\alpha_{jt}$ is a good-$j$ time fixed effect, $RPP_{jgt}$ is a price parity adjustment across states, and $u_{jgt}$ is an error term. States are also weighted according to population size in each year.

In Equation (15), the regional fixed effects control for permanent differences in consumption patterns across regions that are unrelated to nonhomotheticity. It would for instance be bold to claim that nonhomothetic preferences alone explains why a landlocked region such as the Rocky Mountain Region exhibits lower expenditure shares on water transportation than, say, the Great Lakes Region, and the regional dummies mitigate these concerns. The time fixed effects controls for aggregate changes in relative prices between goods and for any other common macro shocks, while the regional price parities similarly control for differences in relative prices across states and their evolution over time. These controls are also important, because all else equal we expect the expenditure share of a good to vary across years and states for which its relative price is different, even in the absence of nonhomothetic behavior.

Table 1 presents the estimated slope coefficients $\beta_{je}$ for each good considered. Applying the classification rule above results in 30 necessities, 34 luxuries, and 8 homothetic goods, and most of these are highly intuitive. Goods (in particular nondurable ones) are generally classified as necessities while services are luxuries broadly speaking, which aligns well with the empirical macro evidence on structural transformation in for instance Boppart (2014), Herrendorf, Rogerson and Valentinyi (2014), Comin, Lashkari and Mestieri (2021), and Alder, Boppart and Müller (2022). For comparable individual categories, the results also align with the classifications in similar analyses by Wachter and Yogo (2010), Orchard (2022), and Hochmuth, Pettersson and Weissert (2023), thus suggesting that the approach works well on the state PCE data.

Because the regional price parity data only starts in 2008, Table 1 also shows a version without these controls that utilizes the full state panel starting in 1997. This alternative only changes the classification of six categories compared to the main specification, of which rented and owned housing, now luxuries, are the major ones. That rented housing should be a luxury is particularly unintuitive, and these changes are likely driven by the failure to account for the geographic variation in housing costs, which correlate positively with per-capita expenditures. The remainder of the paper therefore proceeds with the classification from the main estimation.

The classification in Table 1 yields basket expenditure shares at the aggregate US level that are shown in Figure 1a. As expected under sustained economic growth, the expenditure share on homothetic goods remains stable while the necessity share exhibits a noticeable yet small downward shift; over the full sample period, the expenditure share on necessities declines by around 4 percentage points. Dividing individual expenditure shares by these basket shares generates the within-basket shares needed to compute the Törnqvist index (14) for each bundle. The resulting price indices are shown in Figure 1b together with the official PCE price index for reference. These indices reveal that the price of necessities relative to luxuries has increased by almost 40 percent since 1988, which likely explains the small decline in the necessity share.

**Table 1.** Classifying products as necessity, luxury, or homothetic.

| PCE category | (1) | | (2) | |
|---|---|---|---|---|
| | Engel slope | Std. Error | Engel slope | Std. Error |
| **Motor vehicles and parts** | | | | |
| New motor vehicles | −0.0078*** | (0.0012) | −0.0098*** | (0.0011) |
| Net purchases of used motor vehicles | −0.0120*** | (0.0009) | −0.0132*** | (0.0008) |
| Motor vehicle parts and accessories | −0.0046*** | (0.0005) | −0.0056*** | (0.0003) |
| **Furnishings and durable household equipment** | | | | |
| Furniture and furnishings | 0.0059*** | (0.0010) | 0.0047*** | (0.0005) |
| Household appliances | −0.0014*** | (0.0003) | −0.0013*** | (0.0002) |
| Glassware, tableware, and household utensils | 0.0032*** | (0.0003) | 0.0043*** | (0.0002) |
| Tools and equipment for house and garden | −0.0024*** | (0.0001) | −0.0023*** | (0.0001) |
| **Recreational goods and vehicles** | | | | |
| Video, audio, photographic, and IP equipment and media | 0.0318*** | (0.0017) | 0.0298*** | (0.0011) |
| Sporting equipment, supplies, guns, and ammunition | −0.0041*** | (0.0010) | −0.0047*** | (0.0006) |
| Sports and recreational vehicles | 0.0002 | (0.0014) | −0.0014 | (0.0010) |
| Recreational books | 0.0004 | (0.0006) | −0.0004 | (0.0006) |
| Musical instruments | 0.0006*** | (0.0001) | 0.0002*** | (0.0001) |
| **Other durable goods** | | | | |
| Jewelry and watches | 0.0065*** | (0.0007) | 0.0057*** | (0.0004) |
| Therapeutic appliances and equipment | −0.0054*** | (0.0004) | −0.0044*** | (0.0003) |
| Educational books | 0.0012*** | (0.0002) | 0.0008*** | (0.0002) |
| Luggage and similar personal items | 0.0027*** | (0.0004) | 0.0026*** | (0.0002) |
| Telephone and related communication equipment | 0.0008* | (0.0004) | 0.0007*** | (0.0002) |
| **Food and beverages purchased for off-premises consumption** | | | | |
| Food and nonalcoholic beverages | −0.0335*** | (0.0025) | −0.0397*** | (0.0018) |
| Alcoholic beverages | 0.0027*** | (0.0006) | 0.0025*** | (0.0004) |
| Food produced and consumed on farms | −0.0001*** | (0.0000) | −0.0002*** | (0.0000) |
| **Clothing and footwear** | | | | |
| Women's and girls' clothing | −0.0018 | (0.0013) | −0.0004 | (0.0011) |
| Men's and boys' clothing | −0.0040*** | (0.0007) | −0.0027*** | (0.0006) |
| Children's and infants' clothing | −0.0012*** | (0.0001) | −0.0012*** | (0.0001) |
| Other clothing materials and footwear | −0.0027*** | (0.0005) | −0.0015*** | (0.0004) |
| **Gasoline and other energy goods** | | | | |
| Motor vehicle fuels, lubricants, and fluids | −0.0483*** | (0.0031) | −0.0490*** | (0.0019) |
| Fuel oil and other fuels | −0.0027*** | (0.0005) | −0.0042*** | (0.0004) |
| **Other nondurable goods** | | | | |
| Pharmaceutical and other medical products | −0.0471*** | (0.0029) | −0.0359*** | (0.0017) |
| Recreational items | −0.0102*** | (0.0005) | −0.0099*** | (0.0004) |
| Household supplies | −0.0043*** | (0.0004) | −0.0049*** | (0.0003) |
| Personal care products | 0.0044*** | (0.0007) | 0.0091*** | (0.0006) |
| Tobacco | −0.0079*** | (0.0008) | −0.0112*** | (0.0006) |
| Magazines, newspapers, and stationery | 0.0023* | (0.0009) | 0.0033*** | (0.0007) |
| **Housing** | | | | |
| Rental of tenant-occupied nonfarm housing | −0.0342*** | (0.0053) | 0.0387*** | (0.0022) |
| Imputed rental of owner-occupied nonfarm housing | −0.0020 | (0.0106) | 0.0494*** | (0.0053) |
| Rental value of farm dwellings | 0.0014 | (0.0007) | −0.0069*** | (0.0003) |
| Group housing | 0.0008*** | (0.0002) | 0.0001* | (0.0000) |
| **Household utilities** | | | | |
| Water supply and sanitation | −0.0079*** | (0.0005) | −0.0063*** | (0.0004) |
| Electricity | −0.0184*** | (0.0008) | −0.0154*** | (0.0007) |
| Natural gas | −0.0101*** | (0.0008) | −0.0108*** | (0.0008) |

*Continued on the next page*

**Table 1.** Classifying products as necessity, luxury, or homothetic. (Cont.)

| PCE category | (1) Engel slope | (1) Std. Error | (2) Engel slope | (2) Std. Error |
|---|---|---|---|---|
| **Health care** | | | | |
| Physician services | 0.0004 | (0.0021) | −0.0052*** | (0.0014) |
| Dental services | −0.0029*** | (0.0005) | −0.0025*** | (0.0003) |
| Paramedical services | 0.0071*** | (0.0019) | 0.0055*** | (0.0016) |
| Hospitals | −0.0793*** | (0.0046) | −0.0823*** | (0.0029) |
| Nursing homes | −0.0093*** | (0.0010) | −0.0129*** | (0.0008) |
| **Transportation services** | | | | |
| Motor vehicle maintenance and repair | −0.0015 | (0.0011) | 0.0000 | (0.0008) |
| Other motor vehicle services | 0.0048*** | (0.0009) | 0.0061*** | (0.0006) |
| Ground transportation | 0.0038*** | (0.0007) | 0.0054*** | (0.0005) |
| Air transportation | 0.0200*** | (0.0016) | 0.0201*** | (0.0011) |
| Water transportation | −0.0003* | (0.0001) | −0.0001 | (0.0001) |
| **Recreation services** | | | | |
| Membership clubs, sports centers, parks, theaters, museums | 0.0151*** | (0.0019) | 0.0154*** | (0.0013) |
| Audio-video, photographic, and IP equipment services | 0.0085*** | (0.0008) | 0.0080*** | (0.0006) |
| Gambling | 0.0054** | (0.0021) | 0.0056*** | (0.0014) |
| Other recreational services | 0.0054*** | (0.0006) | 0.0046*** | (0.0003) |
| **Food services and accommodations** | | | | |
| Purchased meals and beverages | 0.0090*** | (0.0018) | 0.0078*** | (0.0011) |
| Food furnished to employees (including military) | −0.0011*** | (0.0001) | −0.0013*** | (0.0001) |
| Accommodations | 0.0121*** | (0.0018) | 0.0118*** | (0.0011) |
| **Financial services and insurance** | | | | |
| Financial services furnished without payment | 0.0106*** | (0.0015) | 0.0148*** | (0.0010) |
| Financial service charges, fees, and commissions | 0.0232*** | (0.0026) | 0.0231*** | (0.0016) |
| Life insurance | 0.0026*** | (0.0005) | 0.0050*** | (0.0004) |
| Net household insurance | 0.0005*** | (0.0001) | 0.0003*** | (0.0001) |
| Net health insurance | 0.0141*** | (0.0036) | 0.0060** | (0.0020) |
| Net motor vehicle and other transportation insurance | 0.0053** | (0.0016) | 0.0043*** | (0.0009) |
| **Communication** | | | | |
| Telecommunication services | 0.0098*** | (0.0013) | 0.0115*** | (0.0011) |
| Postal and delivery services | −0.0003*** | (0.0001) | −0.0001 | (0.0001) |
| Internet access | 0.0040*** | (0.0005) | 0.0034*** | (0.0003) |
| **Education services** | | | | |
| Higher education | −0.0100*** | (0.0024) | −0.0071*** | (0.0015) |
| Nursery, elementary, and secondary schools | −0.0004 | (0.0003) | 0.0000 | (0.0002) |
| Commercial and vocational schools | 0.0072*** | (0.0004) | 0.0070*** | (0.0003) |
| **Other services** | | | | |
| Professional and other services | −0.0042* | (0.0018) | −0.0057*** | (0.0012) |
| Personal care and clothing services | 0.0232*** | (0.0022) | 0.0187*** | (0.0014) |
| Social services and religious activities | 0.0221*** | (0.0048) | 0.0182*** | (0.0028) |
| Household maintenance | 0.0047*** | (0.0004) | 0.0061*** | (0.0003) |
| Time and regional fixed effects | ✓ | | ✓ | |
| Controls for regional price parities | ✓ | | | |
| Sample years | 2008–2021 | | 1997–2021 | |
| Observations per good | 714 | | 1,275 | |

*Notes.* Each specification shows the slope coefficients from regressing state-level aggregate expenditure shares on the corresponding level of logarithmized consumption expenditures per capita. States are weighted by their population size. Robust standard errors in parentheses. *, **, and *** denote statistical significance at the 5 percent, 1 percent, and 0.1 percent levels.

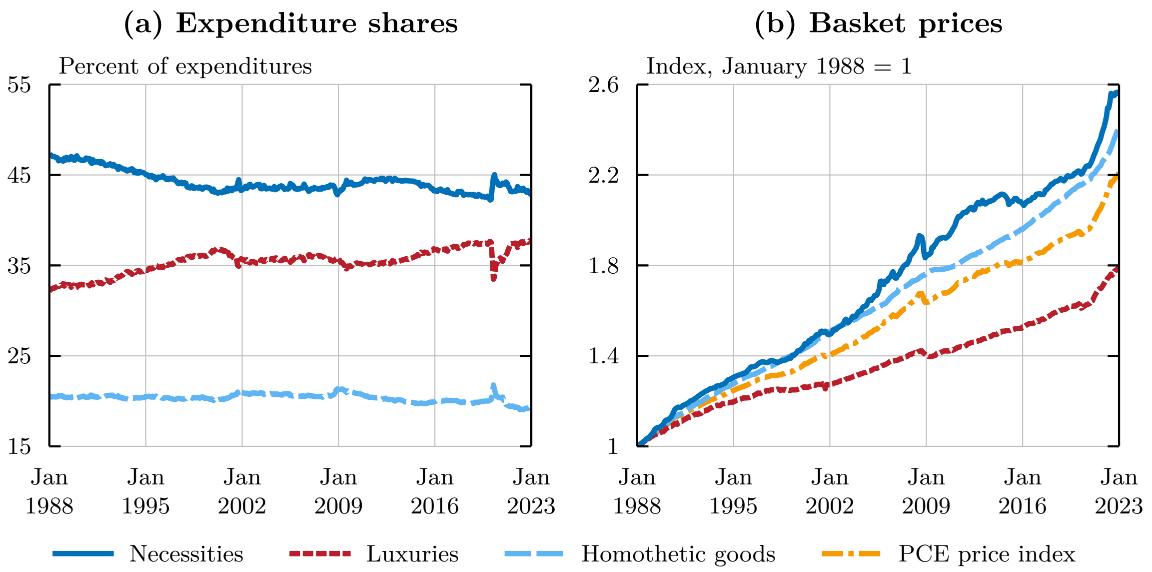


**Figure 1.** Constructed US time series, Jan 1988 to Jan 2023.

### 3.3 Estimation of Preference Parameters

The estimation of the household preference parameters $\varepsilon$ and $\gamma$ relies on the aggregate time series variation displayed in Figure 1. This is possible because PIGL preferences consistently aggregate household-level expenditure shares into market-level expenditure shares as functions of per-capita expenditures. Specifically, for any period $t$ and measure $N$ of households (indexed by $h$ below), the aggregate expenditure share $\overline{w}_{Dt}$ on necessities is the expenditure-weighted average of household-level expenditure shares. Together with Equations (3), (7) and (8), this allows us to write the aggregate expenditure share as

$$\overline{w}_{Dt} \equiv \frac{1}{N}\int_0^N \frac{e_{ht}}{\bar{e}_t}\, w_{Dth}\, dh = \widetilde{\nu}\left(\frac{P_{Bt}^{1-\rho_{t,s}} P_{Ht}^{\rho_{t,s}}}{\bar{e}\kappa_t}\right)^{\varepsilon}\left(\frac{P_{Dt}}{P_{Bt}}\right)^{\gamma}, \tag{16}$$

where $\bar{e}_t$ denotes per-capita expenditures, $\widetilde{\nu} \equiv \nu F\big(H(\boldsymbol{p}_s), B(\boldsymbol{p}_s)\big)^{\varepsilon}\big(D(\boldsymbol{p}_s)/B(\boldsymbol{p}_s)\big)^{\gamma}$ is a scale parameter, and

$$\kappa_t = \left[\frac{1}{N}\int_0^N \left(\frac{e_{ht}}{\bar{e}_t}\right)^{1-\varepsilon} dh\right]^{-\frac{1}{\varepsilon}} \tag{17}$$

is a scale-invariant inequality measure. The parameters $\varepsilon$ and $\gamma$ can therefore be obtained by regressing the aggregate necessity share on per-capita expenditures and the basket price indices according to Equation (16).

Table 2 reports the subsequent estimation results. Note that Equation (16) can be estimated by both linear and nonlinear least squares, the former by taking logs of both sides of (16). The first two columns of Table 2 consequently show linear and nonlinear estimates for the baseline case, which uses the expenditure shares and Törnqvist indices in Figure 1. Both suggest an $\varepsilon$ and a $\gamma$ around 0.275 and 0.354.[7] Reassuringly, these values imply an elasticity of demand for necessities

[7] Obvious caveats here include a potential simultaneous equation bias between expenditure shares and prices,

**Table 2.** Preference parameters in aggregate US data.

| | Baseline | | Fisher | | Using $1 - w_H$ | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $\varepsilon$ | $0.275^{***}$ (0.010) | $0.277^{***}$ (0.010) | $0.275^{***}$ (0.010) | $0.277^{***}$ (0.010) | $0.275^{***}$ (0.008) | $0.277^{***}$ (0.008) |
| $\gamma$ | $0.354^{***}$ (0.018) | $0.356^{***}$ (0.018) | $0.354^{***}$ (0.018) | $0.356^{***}$ (0.018) | $0.385^{***}$ (0.016) | $0.386^{***}$ (0.016) |
| $\nu$ | $2.209^{***}$ (0.135) | $2.244^{***}$ (0.136) | $2.211^{***}$ (0.134) | $2.246^{***}$ (0.136) | $2.660^{***}$ (0.136) | $2.687^{***}$ (0.137) |
| $\sigma$ | | | | | $1.341^{***}$ (0.047) | $1.343^{***}$ (0.047) |
| Method | OLS | NLS | OLS | NLS | OLS | NLS |
| Observations | 421 | 421 | 421 | 421 | 421 | 421 |
| RMSE | 0.013 | 0.006 | 0.013 | 0.006 | 0.012 | 0.005 |
| Adjusted $R^2$ | 0.765 | | 0.766 | | 0.793 | |

*Notes.* RMSE denotes the root mean square error. Robust standard errors in parentheses. $^{*}$, $^{**}$, and $^{***}$ denote statistical significance at the 5 percent, 1 percent, and 0.1 percent levels.

equal to 0.725, which is highly comparable to the elasticities found in richer estimation exercises by for instance Boppart (2014), Comin, Lashkari and Mestieri (2021), and Fan, Peters and Zilibotti (2021).

The remaining columns report additional specifications to show the robustness of these estimates. To assess the sensitivity with respect to the choice of basket price index formula, columns (3) and (4) use basket price indices that are constructed with the Fisher price index formula instead of the Törnqvist index.[8] Columns (5) and (6) substitute the homothetic expenditure share (4) into the necessity expenditure share (16) to obtain an estimating equation of the form $\overline{w}_D = \widetilde{\nu}\left(P_B \,/\, \bar{e}\kappa\right)^{\varepsilon}\left(P_D/P_B\right)^{\gamma}(1 - w_H)^{\varepsilon/(1-\sigma)}$, which does not require a price index construct for the function $F\big(H(\boldsymbol{p}), B(\boldsymbol{p})\big)$. Both alternatives remain very close to the baseline estimates.

## 3.4 Base-Period Distribution of US Personal Consumption Expenditures

The basket prices and aggregate expenditure shares in Figure 1 combined with the parameter values for $\varepsilon$ and $\gamma$ are sufficient to compute the PIGL cost-of-living index at the aggregate level. To obtain household-level cost-of-living indices, however, we additionally need a distribution of expenditure shares on necessities for some chosen base period. Until recently, no such PCE data existed, but a recent paper by Garner *et al.* (2022) bridges this gap by distributing PCE

and potential measurement errors on expenditures and prices. Addressing these issues is somewhat beyond the scope of the paper, since the goal is not necessarily to find *the* correct values for $\varepsilon$ and $\gamma$, but rather to obtain ballpark estimates that aligns with the data such that the PIGL cost-of-living index can be implemented. The final part of the paper includes a sensitivity analysis on the choice of parameter values.

[8] The Fisher index is obtained as the square root of the Laspeyres and Paasche indices: $P_t/P_{t-1} = \sqrt{\sum_j w_{jt-1}\big(p_{jt}/p_{jt-1}\big) \,/\, \sum_j w_{jt}\big(p_{jt-1}/p_{jt}\big)}$.

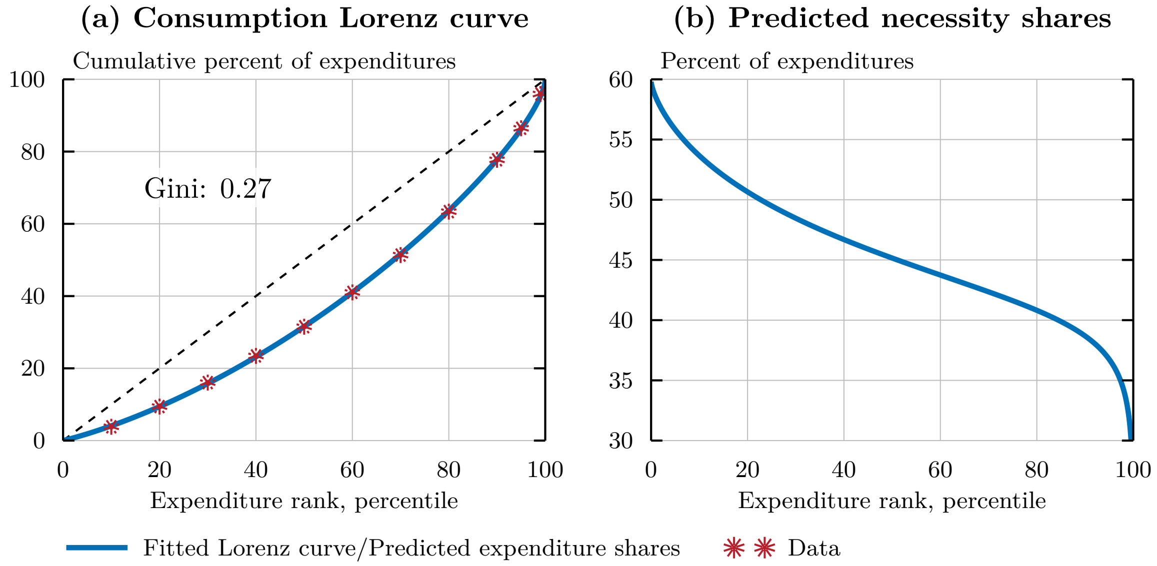


**Figure 2.** Estimated US consumption profiles for base year 2019.

spending in 2019 across US households, which is consequently used as base year here.[9] These estimates divide aggregate spending by decile in the expenditure distribution, both for total expenditures and for 15 broad commodity groups. The former also includes expenditure shares for the top 1 and 5 percents of the distribution.

Garner *et al.*'s (2022) breakdown into deciles and into 15 commodity groups is too coarse to directly infer household-level expenditure shares on the necessity basket identified in Table 1. Fortunately, the theoretical model predicts a direct link between expenditure shares at the household level, the corresponding aggregate expenditure share, and the overall distribution of consumption expenditures. Denote the Lorenz curve associated with the expenditure distribution by $\ell(x)$, where $x$ is the expenditure rank, and its derivative with respect to the rank by $\ell'(x)$. Evaluated at the rank $x_h$ of household $h$, this derivative must satisfy $\ell'(x_h) = e_h \,/\, \bar{e}$. Using the individual expenditure share (3) and the aggregate expenditure share (16) then yields

$$w_{Dh} \;=\; \frac{w_{Dh}}{\overline{w}_D}\,\overline{w}_D \;=\; \left(\frac{e_h}{\bar{e}\kappa}\right)^{-\varepsilon}\overline{w}_D \;=\; \left(\frac{\ell'(x_h)}{\kappa}\right)^{-\varepsilon}\overline{w}_D. \tag{18}$$

Similarly, by Equation (17), the aggregation factor $\kappa$ can be written

$$\kappa \;=\; \left[\int_0^1 \ell'(x)^{1-\varepsilon}\,dx\right]^{-\frac{1}{\varepsilon}}. \tag{19}$$

Thus, the Lorenz curve $\ell(x)$, an empirically observed aggregate expenditure share $\overline{w}_D$ and a parameter value for $\varepsilon$ are sufficient to compute all base-period expenditure shares at the household level.[10]

[9] At the time of writing, similar estimates have been constructed for each year between 2017 and 2020. These estimates can be downloaded from the webpage of the US Bureau of Labor Statistics.

[10] A similar prediction holds for a group, say a decile $d$, over a distribution interval $[x_{d0}, x_{d1}]$. Then $\Delta\ell_d \,/\, \Delta x_d =$

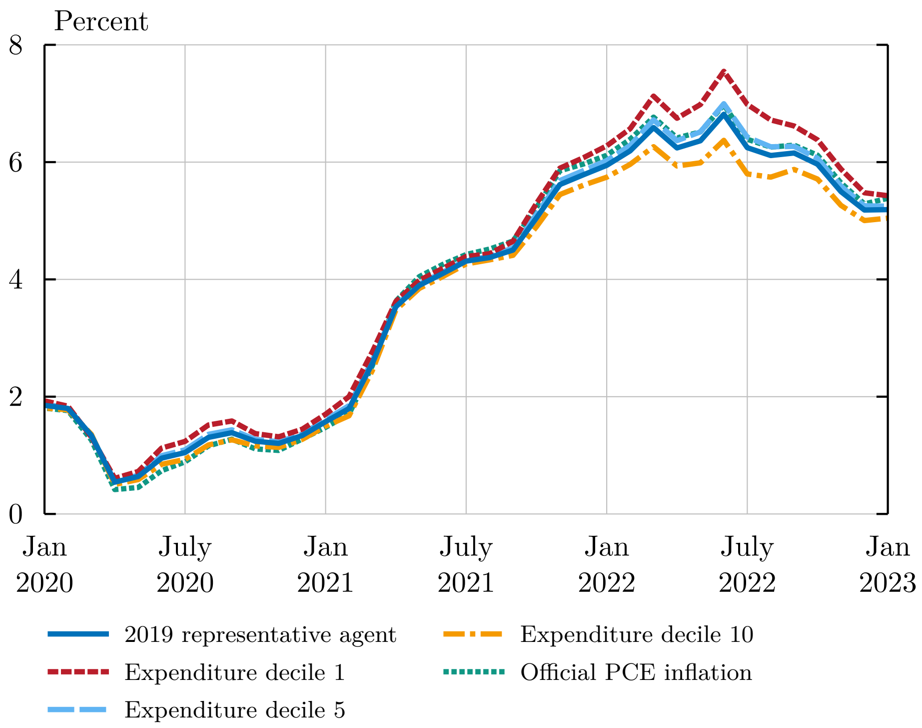


**FIGURE 3.** Annual PCE inflation rates during the inflation surge.

*Notes.* "2019 representative agent" refers to the inflation implied from the cost-of-living index that uses the aggregate expenditure share on necessities in 2019 as weight.

The estimates by Garner *et al.* (2022) are therefore used to construct the Lorenz curve. To that end, this paper follows Sitthiyot and Holasut's (2021) suggestion and parameterizes $\ell(x)$ as a weighted average between an exponential function and the functional form implied by the Pareto distribution: $\ell(x) = (1-\omega)x^{\eta} + \omega\big(1-(1-x)^{1/\eta}\big)$, where $\omega$ and $\eta$ are parameters to estimate. Fitting this function to the distributional PCE data yields the Lorenz curve shown in Figure 2a, which exhibits an $R^2$ of 0.9999. Figure 2b shows the corresponding expenditure share predictions from Equations (18) and (19). Together with the basket prices in Figure 1b and the parameter values in Table 2, these shares yield the nonhomothetic PCE price index that underlies the results explored in the next section.

## 4 PCE Inflation Across the Expenditure Distribution

Unlike most other approaches to measure inflation inequality, a key benefit of the PIGL cost-of-living index is that it can be implemented without requiring detailed micro data for every period of consideration. Once constructed, the nonhomothetic PCE price index therefore easily sheds light on both the very latest inflation developments and on more long-run changes. This section shows the implied distribution of PCE inflation across US households and documents a noticeably higher inflation for consumption-poor households in both cases.

$\bar{e}_d\,/\,\bar{e}$, where $\bar{e}_d$ is group-$d$ average expenditures and $\Delta\ell_d = \ell(x_{d1}) - \ell(x_{d0})$, $\Delta x_d = x_{d1} - x_{do}$. The aggregate expenditure share of the group becomes $\overline{w}_d = \left(\frac{\Delta\ell_d/\Delta x_d}{\kappa/\kappa_d}\right)^{-\varepsilon}\overline{w}_D$, where $\kappa_d$ is a group-specific aggregation factor given by $\kappa_d = \left[\frac{1}{\Delta x_d}\int_{x_{d0}}^{x_{d1}}\left(\frac{\ell'(x)}{\Delta\ell_d/\Delta x_d}\right)^{1-\varepsilon}dx\right]^{-1/\varepsilon}$.

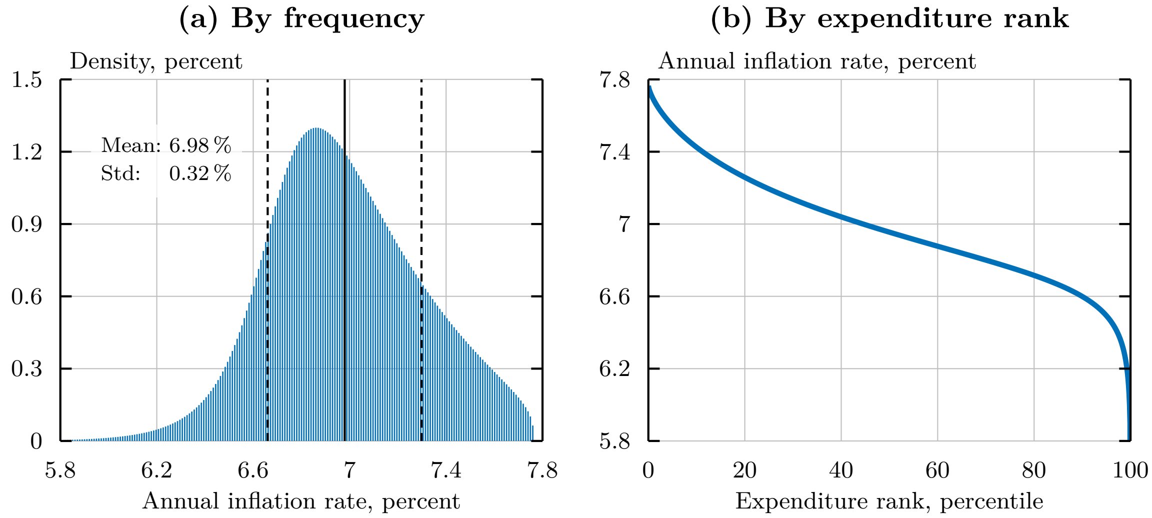


**FIGURE 4.** The cross-sectional distribution of PCE inflation rates, June 2022.

*Notes.* The histogram in Figure 4a is constructed with a bin width of 0.01 percentage points. The solid line represents the mean inflation rate and the dotted lines are the mean plus/minus one standard deviation.

### 4.1 Inflation Inequality in the Ongoing Inflation Surge

Figure 3 shows the annual PCE inflation rates implied by the PIGL cost-of-living index from January 2020 onward for the top, the middle, and the bottom deciles of the cross-sectional expenditure distribution in 2019. Also displayed are the inflation rates for the aggregate expenditure allocation (the representative agent) and, for reference, the official PCE measure. Overall, Figure 3 reveals dynamics of the nonhomothetic PCE indices that closely follow the official PCE inflation measure, with rates remaining between 0 and 2 percent in 2020, then rising sharply during 2021, before reaching the high levels of 2022.

Of more interest is the observation that as the inflation rate levels increase in 2021, then so does the dispersion of rates across groups. In 2020 and the first half of 2021, the inflation rate gap between the top and bottom deciles remains between 0.1 and 0.3 percentage points. After mid-2021, the gap widens, ultimately ranging between 0.5 and 1.2 percentage points throughout 2022, the latter being the peak reached in June 2022. The average inflation rate difference between the top and bottom deciles in 2022 comes in at 0.8 percentage points, thus underlining that consumption-poor households are hit harder by the ongoing inflation surge.

To get an idea of what the entire cross-sectional distribution of inflation rates looks like, Figure 4 plots the density of inflation rates and the variation along the full expenditure distribution in June 2022. The density in Figure 4a exhibits a mean and standard deviation of 6.98 and 0.32 percent, which implies that around 70 percent of inflation rates falls within a narrow band of 6.7 to 7.3 percent.[11] This is also the month with the largest inflation disparities, thus suggesting that the majority of households face similar inflation in most periods. Yet, Figure 4b shows an inflation rate gap of around 1.5 to 2 percentage points between the very poorest and richest.

[11] The distribution of inflation rates is narrower than those obtained by Hobijn and Lagakos (2005) and Kaplan and Schulhofer-Wohl (2017), but the large dispersions in these papers are primarily driven by differences in prices paid and by heterogeneous household characteristics and tastes, which this paper abstracts from.

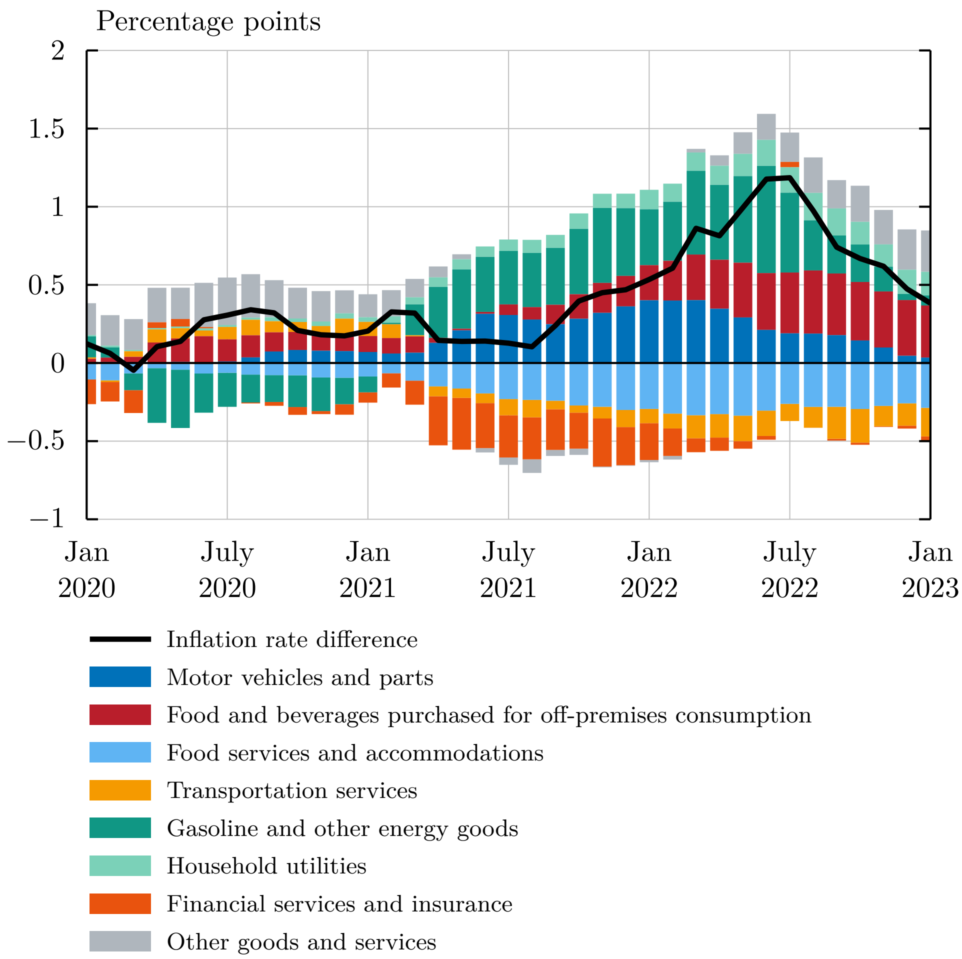


**FIGURE 5.** Inflation rate gap between the top and bottom expenditure deciles.

This difference is considerably larger than in the previous studies by Jaravel (2019), Argente and Lee (2021), and Klick and Stockburger (2021), who typically find a range on the order of 0.3 to 0.4 percentage points between top and bottom groups.

What are the underlying sources of these inflation disparities? The negative relationship between the inflation rate and the consumption expenditure level in Figures 3 and 4b points to the increasing relative price of necessities, since poorer households allocate a higher weight to price changes of these goods than the rich do. The geometric-mean form of the PIGL index allows for a more exact breakdown. Specifically, log differencing the cost-of-living indices of two household groups with each other immediately yields each good's contribution to any overall disparities. Figure 5 shows such a decomposition of the inflation gap between the top and bottom expenditure deciles in order to identify its key drivers. For ease of exposition, this exercise aggregates goods into the 18 broad groups reported in Table 1, with Figure 5 subsequently displaying the seven largest contributors.

From the second half of 2021 onward, the period in which the largest differences occur, Figure 5 identifies increasing costs for food consumed at home, energy, and motor vehicles as the primary sources of the higher inflation faced by the poor. For instance, in June 2022, the peak month with

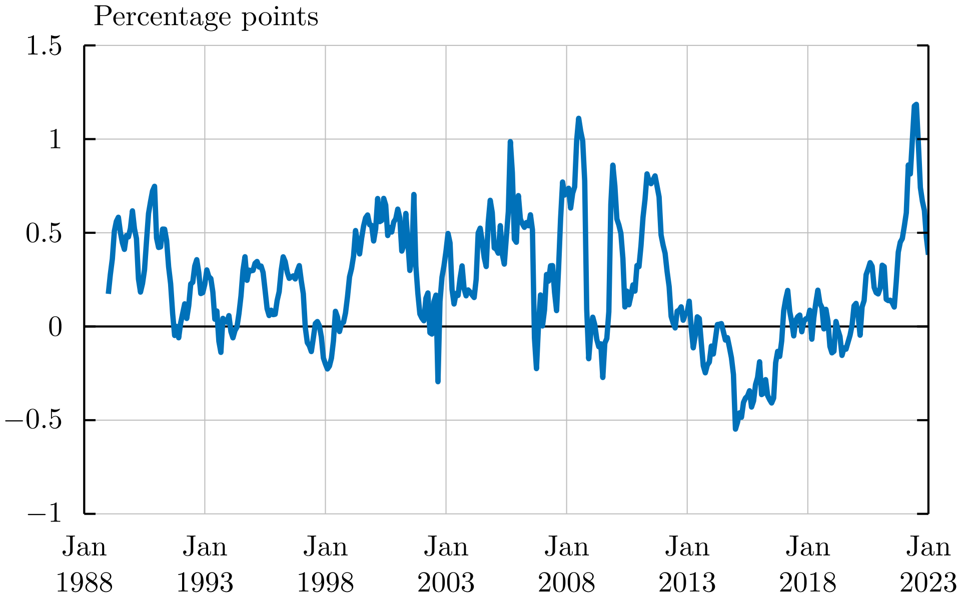


**Figure 6.** Long-run inflation rate gap between the top and bottom expenditure deciles.

respect to inflation inequality, increasing prices for gasoline and other energy products raised the inflation rate for the bottom decile by 0.7 percentage points more than it did for the top decile. The corresponding numbers for food, utilities, and motor vehicles are 0.4, 0.2, and 0.2 percentage points, respectively, for a total which exceeds the overall inflation gap of 1.2 percentage points. Downside contributors are higher prices for food services and accommodations, transportation services, and financial services, which are consumed proportionately more by the rich. The former two in particular are somewhat expected in light of the product groups that raise the inflation gap, since food and energy are vital inputs in these two industries.

In sum, this section shows that the short-run inflation inequality found in Consumer Expenditure Survey data and in US scanner data by for example Kaplan and Schulhofer-Wohl (2017), Jaravel (2019), Argente and Lee (2021), and Klick and Stockburger (2021) also holds up in the PCE data. While corroborating the qualitative conclusions from these studies, the magnitudes found here during the ongoing inflation surge are considerably larger. The reason is much the same as the reason why contemporary inflation is high in general: prices of food and energy are soaring, and these products are consumed proportionately more by the poor.

## 4.2 Inflation Inequality in the Long Run

In the interest of long-run dynamics, Figure 6 extends the inflation rate gap between the top and bottom deciles of the 2019 expenditure distribution to cover the full sample period from January 1988 to January 2023. While 2022 saw the largest gap between these two groups, it is by no means an anomaly. Sizeable differences occur frequently, especially between the end of the 1990s and 2011. These differences also go in both directions: in the beginning of 2015, for instance, the top decile faced a 0.5 percentage point higher inflation rate. For the most part, however, the poor faces a higher inflation rate, with an average gap of around 0.2 percentage points over the full period.

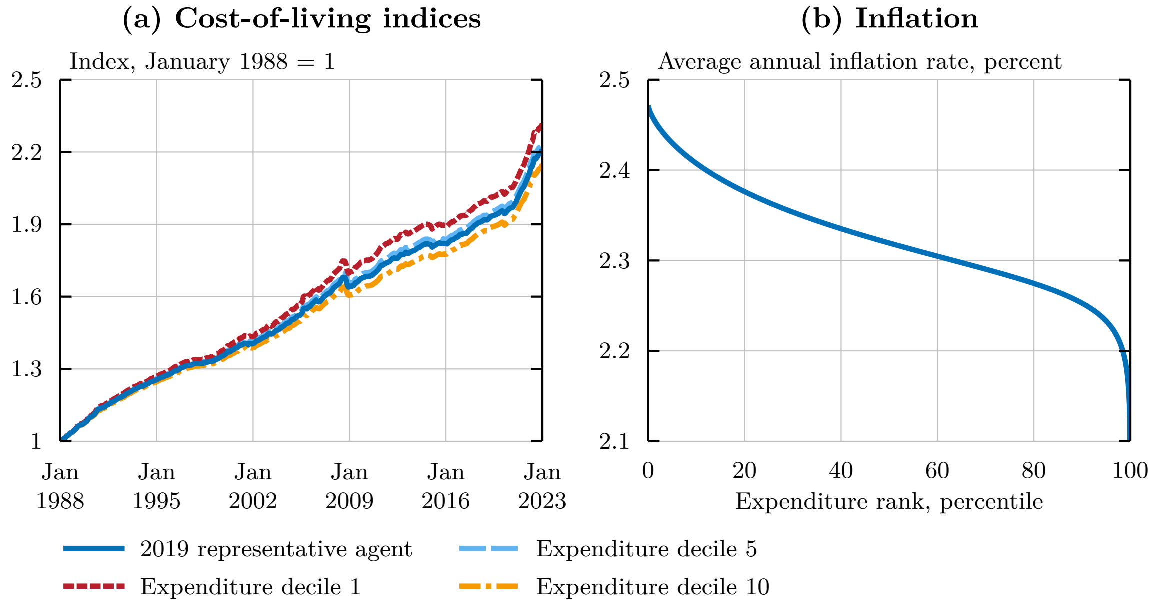


**FIGURE 7.** Long-run inflation inequality.

*Notes.* "2019 representative agent" refers to the cost-of-living index that uses the aggregate expenditure share on necessities in 2019 as weight.

What is the total impact of these developments? Figure 7a plots the cost-of-living indices of the groups considered in Figure 3 for the full sample period, showing that the cost of living has risen more for consumption-poor households also in the long run. Qualitatively, this comes as no surprise in light of the increasing relative price of necessities observed in Figure 1b. Quantitatively, the index of the bottom decile grows by a cumulative 122 percent, compared to 106 percent for the top decile. A back-of-the-envelope calculation with these numbers suggests that cumulative growth in real consumption inequality over this 35-year period, as measured by the change in the 90/10 percentile ratio, has grown by almost 8 percent more than what is implied from a common deflator such as the official PCE price index.

These cost-of-living differences are also present across the whole expenditure distribution. Figure 7b plots the long-run average annual inflation rates throughout the base-year distribution, which reveals a mean inflation rate level of around 2.3 percent per year and a 0.3 percentage point gap between the very top and the very bottom of the distribution. Figure 7b is surprisingly similar to the estimates of Jaravel and Lashkari (2023), who find a nearly identical inflation level and dispersion with Consumer Expenditure Survey data between 1984 and 2019. The nonhomothetic PCE price index consequently reinforces their findings and points to sizeable inflation inequality in the long run as well as in the short run.

## 5 PCE Inflation With Different Parameter Values

The approach in this paper relies on a cost-of-living index that requires the estimation of two exogenous preference parameters. A key concern that this section investigates closer is whether the inflation inequality uncovered above is sensitive to the choice of values for these parameters.

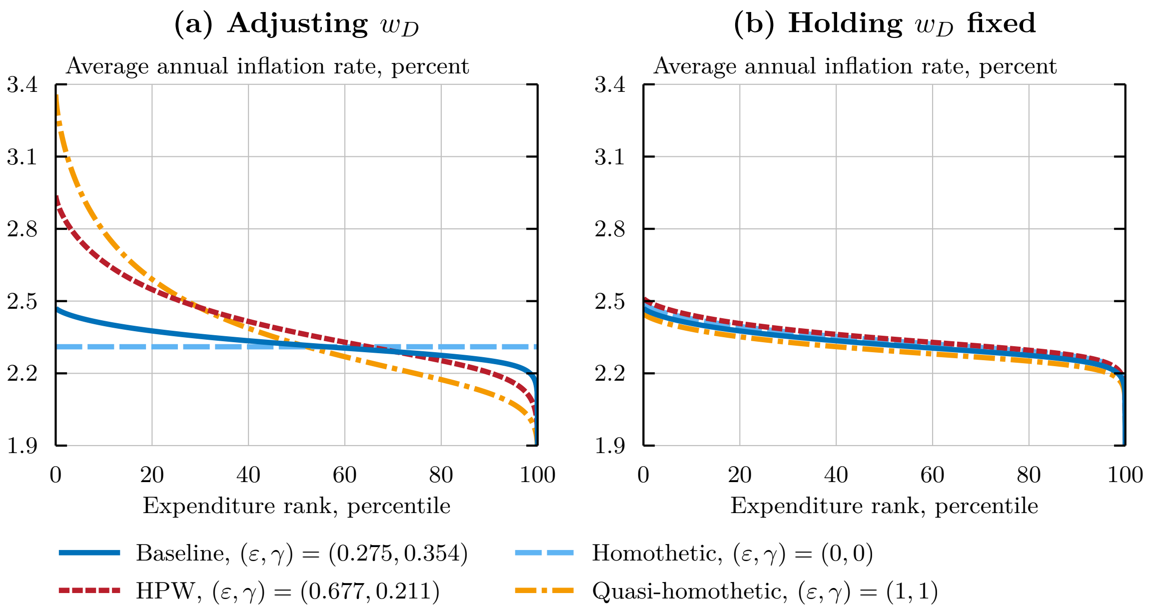


**Figure 8.** Long-run inflation inequality for different parameter values.

As a starting point, suppose we redo the whole exercise for some other feasible parameter choices. Consider for instance the following cases, taken somewhat out of the blue: homothetic Cobb-Douglas preferences, $(\varepsilon, \gamma) \to (0, 0)$; quasi-homothetic preferences, $(\varepsilon, \gamma) = (1, 1)$; and the estimates obtained by Hochmuth, Pettersson and Weissert (2023), $(\varepsilon, \gamma) = (0.677, 0.211)$. Figure 8a compares the average annual inflation rates over the full sample period for these specifications against those of the baseline estimation. At first sight, it seems that the choice of parameter values matters a great deal: the inflation gap between the top and the bottom of the distribution ranges from almost 1.5 percentage points under quasi-homothetic preferences to the zero differences (naturally) obtained with homothetic preferences.

These disparities are driven by two possible channels. First, since we consider predicted necessity expenditure shares $w_{Dh} = \big(\ell'(x_h) \,/\, \kappa\big)^{-\varepsilon}\, \overline{w}_D$, the parameter choice affects the base-period allocation and thereby also the index weight on necessities. Second, $\varepsilon$ and $\gamma$ also affect the predicted substitution behavior of households between the necessity and luxury baskets, as captured by the direct presence of these parameters in the cost-of-living index formula (10). For other applications, it is the sensitivity with respect to the second channel which is of primary interest, because base-period allocations will generally not be predicted from individual expenditure levels and aggregate expenditure shares, but rather taken directly from empirically observed expenditure shares.

A perhaps more interesting analysis is therefore to consider how the results change if we vary the parameter values but keep the base-period consumption patterns fixed according to the baseline estimates. The results from this exercise, shown in Figure 8b, paints a completely different picture regarding the sensitivity of the results to different parameter choices. At any given point in the expenditure distribution, the range of inflation rates between the four cases is never more than 0.06 percentage points. Differences in inflation inequality are even smaller: the inflation rate gap between the top and bottom deciles is always between 0.22 and 0.24 percentage points. Thus, both inflation inequality and the general level of inflation are close to identical across the

**TABLE 3.** Long-run inflation rate gap for different parameter values with $w_D$ fixed.

| $\varepsilon$ | $\gamma$ | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | **0** | **0.1** | **0.2** | **0.3** | **0.4** | **0.5** | **0.6** | **0.7** | **0.8** | **0.9** | **1** |
| **0** | 0.22 | 0.22 | 0.22 | 0.21 | 0.21 | 0.21 | 0.20 | 0.20 | 0.20 | 0.19 | 0.19 |
| **0.1** | 0.23 | 0.22 | 0.22 | 0.22 | 0.21 | 0.21 | 0.21 | 0.20 | 0.20 | 0.20 | 0.19 |
| **0.2** | 0.23 | 0.23 | 0.22 | 0.22 | 0.22 | 0.21 | 0.21 | 0.21 | 0.20 | 0.20 | 0.20 |
| **0.3** | 0.23 | 0.23 | 0.23 | 0.22 | 0.22 | 0.22 | 0.21 | 0.21 | 0.21 | 0.20 | 0.20 |
| **0.4** | 0.24 | 0.23 | 0.23 | 0.23 | 0.22 | 0.22 | 0.22 | 0.21 | 0.21 | 0.20 | 0.20 |
| **0.5** | 0.24 | 0.24 | 0.23 | 0.23 | 0.23 | 0.22 | 0.22 | 0.21 | 0.21 | 0.21 | 0.20 |
| **0.6** | 0.25 | 0.24 | 0.24 | 0.23 | 0.23 | 0.23 | 0.22 | 0.22 | 0.21 | 0.21 | 0.21 |
| **0.7** | 0.25 | 0.25 | 0.24 | 0.24 | 0.23 | 0.23 | 0.22 | 0.22 | 0.22 | 0.21 | 0.21 |
| **0.8** | 0.25 | 0.25 | 0.25 | 0.24 | 0.24 | 0.23 | 0.23 | 0.22 | 0.22 | 0.22 | 0.21 |
| **0.9** | 0.26 | 0.25 | 0.25 | 0.24 | 0.24 | 0.24 | 0.23 | 0.23 | 0.22 | 0.22 | 0.22 |
| **1** | 0.26 | 0.26 | 0.25 | 0.25 | 0.24 | 0.24 | 0.24 | 0.23 | 0.23 | 0.22 | 0.22 |

*Notes.* The table shows the difference in the average annual inflation rate between the top and bottom expenditure deciles over the period January 1988 to January 2023 when the base-period allocation is held fixed at the baseline estimates. All numbers are in percentage points.

different parameter specifications.

This finding is also not limited to the specific cases considered in Figure 8. Holding the base-period allocations fixed, Table 3 presents the difference in average annual inflation rates between the top and bottom deciles of the base-period expenditure distribution for a full grid of $(\varepsilon, \gamma)$ values. Throughout, the inflation gap barely moves, ranging from a low of 0.19 percentage points at $(\varepsilon, \gamma) = (0, 1)$ to a high of 0.26 percentage points at $(\varepsilon, \gamma) = (1, 0)$.

The key takeaway from this section is therefore that the parameters matter only insofar as we need to predict expenditure shares from individual expenditure levels and aggregate expenditure shares. When individual expenditure shares can be inferred directly from the data for some base period, the choice of parameter values seem inconsequential for the inflation inequality measures obtained with the PIGL cost-of-living index. This distinction is important because for most other applications, the latter is arguably the relevant case.

# 6 Conclusion

This paper uses recent estimates of the cross-sectional distribution of US Personal Consumption Expenditures to construct a first-ever distribution of PCE inflation across US households. The underlying cost-of-living index originates from a theoretically sound nonhomothetic utility function, generalizes the Törnqvist price index, and only contains two unknown parameters. A central point in the paper is that the implementation of this index requires no more than a handful of publicly available tables from the BEA's national and regional accounts, provided that a single cross-sectional distribution of consumption expenditures exists.

Using 2019 as base year, the empirical analysis reveals PCE inflation rates that are considerably higher for consumption-poor households than for consumption-rich households. This finding applies for both the short run and the long run. For instance, the annual inflation rate of the bottom expenditure decile is on average 0.8 percentage points higher than that of the top decile during the current inflation surge. A similar 0.2 percentage point difference holds over the 35-year period starting in 1988.

Another key result is that measured inflation levels and inflation disparities are almost completely insensitive to changes in the price index parameters if the base-period consumption allocations are exogenous. Future work will have to evaluate the generality of this finding, but taken at face value, it suggests that statistical agencies should aim to construct distributional consumption data at increasingly detailed levels of product aggregation. Practitioners could then infer the expenditure share on necessities directly from the data, allowing them to simply choose some rough parameter values in the price index formula and subsequently be good to go. One possibility for this choice would for instance be to set $\varepsilon = \gamma$ and then pick a value around 0.2 to 0.3, which several papers in the structural change literature suggest is reasonable for macroeconomic data. All in all, this should provide a quick and estimation-free route to obtain nonhomothetic deflators and inflation measures.

## References


Alder, S., Boppart, T., and Müller, A. (2022). "A theory of structural change that can fit the data". *American Economic Journal: Macroeconomics*, vol. 14(2), pp. 160–206. (Cited on pp. 1, 9.)

Argente, D., and Lee, M. (2021). "Cost of living inequality during the Great Recession". *Journal of the European Economic Association*, vol. 19(2), pp. 913–952. (Cited on pp. 3, 17, 18.)

Atkin, D., Faber, B., Fally, T., and Gonzalez-Navarro, M. (2020). *Measuring welfare and inequality with incomplete price information* (NBER Working Paper No. 26890). Cambridge, MA: National Bureau of Economic Research. (Cited on p. 3.)

Baqaee, D., Burstein, A., and Koike-Mori, Y. (2022). *A fixed point approach to measuring welfare* (NBER Working Paper No. 30549). Cambridge, MA: National Bureau of Economic Research. (Cited on p. 3.)

Boppart, T. (2014). "Structural change and the Kaldor facts in a growth model with relative price effects and non-Gorman preferences". *Econometrica*, vol. 82(6), pp. 2167–2196. (Cited on pp. 1, 3, 9, 13.)

Broda, C., and Romalis, J. (2009). *The welfare implications of rising price dispersion*, Booth School of Business, University of Chicago. Unpublished manuscript. (Cited on p. 3.)

Comin, D., Lashkari, D., and Mestieri, M. (2021). "Structural change with long-run income and price effects". *Econometrica*, vol. 89(1), pp. 311–374. (Cited on pp. 9, 13.)

Cravino, J., Levchenko, A. A., and Rojas, M. (2022). "Population aging and structural transformation". *American Economic Journal: Macroeconomics*, vol. 14(4), pp. 479–498. (Cited on p. 1.)

Diewert, W. E. (1976). "Exact and superlative index numbers". *Journal of Econometrics*, vol. 4(2), pp. 115–145. (Cited on pp. 2, 6.)

Diewert, W. E. (1978). "Superlative index numbers and consistency in aggregation". *Econometrica*, vol. 46(4), pp. 883–900. (Cited on p. 2.)

Engel, E. (1857). "Die vorherrschenden Gewerbszweige in den Gerichtsämtern mit Beziehung auf die Productions- und Consumtionsverhältnisse des Königreichs Sachsen". *Zeitschrift des Statistischen Bureaus des Königlich Sächsischen Ministeriums des Innern*, vol. 3(8–9), pp. 153–182. (Cited on p. 1.)

Fan, T., Peters, M., and Zilibotti, F. (2021). *Growing like India: The unequal effects of service-led growth* (NBER Working Paper No. 28551). Cambridge, MA: National Bureau of Economic Research. (Cited on pp. 1, 13.)

Feenstra, R. C. (1994). "New product varieties and the measurement of international prices". *American Economic Review*, vol. 84(1), pp. 157–177. (Cited on p. 2.)

Feenstra, R. C., and Reinsdorf, M. B. (2000). "An exact price index for the almost ideal demand system". *Economics Letters*, vol. 66(2), pp. 159–162. (Cited on p. 2.)

Fixler, D., Gindelsky, M., and Johnson, D. (2020). *Measuring inequality in the national accounts* (BEA Working Paper No. 2020-3). Washington, DC: U.S. Bureau of Economic Analysis. (Cited on p. 2.)

Fixler, D., Johnson, D., Craig, A., and Furlong, K. (2017). "A consistent data series to evaluate growth and inequality in the national accounts". *Review of Income and Wealth*, vol. 63(S2), pp. S437–S459. (Cited on p. 2.)

Garner, T., Martin, R. S., Matsumoto, B., and Curtin, S. (2022). *Distribution of U.S. Personal Consumption Expenditures for 2019: A prototype based on Consumer Expenditure Survey data* (BLS Working Papers No. 557). Washington, DC: U.S. Bureau of Labor Statistics. (Cited on pp. 2, 7, 8, 13–15.)

Herrendorf, B., Rogerson, R., and Valentinyi, Á. (2014). "Growth and structural transformation". In P. Aghion and S. N. Durlauf (Eds.), *Handbook of Economic Growth* (pp. 855–941, Vol. 2B). Elsevier. (Cited on p. 9.)

Hobijn, B., and Lagakos, D. (2005). "Inflation inequality in the United States". *Review of Income and Wealth*, vol. 51(4), pp. 581–606. (Cited on pp. 3, 16.)

Hochmuth, P., Pettersson, M., and Weissert, C. J. (2023). *A nonhomothetic price index and cost-of-living inequality*. Unpublished manuscript. (Cited on pp. 1–3, 5–9, 20.)

Jaravel, X. (2019). "The unequal gains from product innovations: Evidence from the U.S. retail sector". *Quarterly Journal of Economics*, vol. 134(2), pp. 715–783. (Cited on pp. 3, 17, 18.)

Jaravel, X. (2021). "Inflation inequality: Measurement, causes, and policy implications". *Annual Review of Economics*, vol. 13, pp. 599–629. (Cited on p. 3.)

Jaravel, X., and Lashkari, D. (2023). *Measuring growth in consumer welfare with income-dependent preferences*. Unpublished manuscript. (Cited on pp. 2, 3, 19.)

Kaplan, G., and Schulhofer-Wohl, S. (2017). "Inflation at the household level". *Journal of Monetary Economics*, vol. 91, pp. 19–38. (Cited on pp. 3, 16, 18.)

Klick, J., and Stockburger, A. (2021). *Experimental CPI for lower and higher income households* (BLS Working Papers No. 537). Washington, DC: U.S. Bureau of Labor Statistics. (Cited on pp. 3, 17, 18.)

Konüs, A. A. (1939). "The problem of the true index of the cost of living" (J. Bronfenbrenner, Trans.). *Econometrica*, vol. 7(1), pp. 10–29. (Original work published 1924). (Cited on pp. 2, 4, 5.)

Lauper, C., and Mangiante, G. (2021). *Monetary policy shocks and inflation heterogeneity*, University of Lausanne. Unpublished manuscript. (Cited on p. 3.)

McGranahan, L., and Paulson, A. (2005). *Constructing the Chicago Fed income based economic index – consumer price index: Inflation experiences by demographic group: 1983–2005* (Working Paper No. 2005–20). Chicago, IL: Federal Reserve Bank of Chicago. (Cited on p. 3.)

Muellbauer, J. (1975). "Aggregation, income distribution and consumer demand". *Review of Economic Studies*, vol. 42(4), pp. 525–543. (Cited on pp. 1, 3.)

Muellbauer, J. (1976). "Community preferences and the representative consumer". *Econometrica*, vol. 44(5), pp. 979–999. (Cited on pp. 1, 3.)

Orchard, J. (2022). *Cyclical demand shifts and cost of living inequality*, Federal Reserve Board. Unpublished manuscript. (Cited on pp. 3, 8, 9.)

Piketty, T., Saez, E., and Zucman, G. (2018). "Distributional national accounts: Methods and estimates for the United States". *Quarterly Journal of Economics*, vol. 133(2), pp. 553–609. (Cited on p. 2.)

Redding, S. J., and Weinstein, D. E. (2020). "Measuring aggregate price indices with taste shocks: Theory and evidence for CES preferences". *Quarterly Journal of Economics*, vol. 135(1), pp. 503–560. (Cited on p. 2.)

Samuelson, P. A., and Swamy, S. (1974). "Invariant economic index numbers and canonical duality: Survey and synthesis". *American Economic Review*, vol. 64(4), pp. 566–593. (Cited on pp. 2, 4.)

Sato, K. (1976). "The ideal log-change index number". *Review of Economics and Statistics*, vol. 58(2), pp. 223–228. (Cited on p. 4.)

Sitthiyot, T., and Holasut, K. (2021). "A simple method for estimating the Lorenz curve". *Humanities and Social Sciences Communications*, vol. 8(268), pp. 1–9. (Cited on p. 15.)

Törnqvist, L. (1936). "The Bank of Finland's consumption price index". *Bank of Finland Monthly Bulletin*, vol. 16(10), pp. 27–34. (Cited on pp. 1, 6, 7.)

Vartia, Y. O. (1976). "Ideal log-change index numbers". *Scandinavian Journal of Statistics*, vol. 3(3), pp. 121–126. (Cited on p. 4.)

Wachter, J. A., and Yogo, M. (2010). "Why do household portfolio shares rise in wealth?" *Review of Financial Studies*, vol. 23(11), pp. 3929–3965. (Cited on pp. 8, 9.)