

BLS WORKING PAPERS

U.S. Department of Labor
U.S. Bureau of Labor Statistics
Office of Prices and Living Conditions



Democratic Aggregation: Issues and Implications for Consumer Price Indexes

Robert S. Martin, U.S. Bureau of Labor Statistics

Working Paper 600
November 17, 2022

Democratic Aggregation: Issues and Implications for Consumer Price Indexes*

Robert S. Martin[†]

November 17, 2022

Abstract

Most consumer price indexes (CPI) implicitly weight households by their total expenditure. These may differ from indexes that weight households equally. I calculate so-called “democratic” counterparts to the Bureau of Labor Statistic’s CPI and Chained CPI products. The building blocks are household-level Lowe and Tornqvist indexes constructed using matched Consumer Expenditure Survey (CE) Diary and Interview microdata, along with CPI elementary item-area indexes. The impact of democratic aggregation is greater for the 2002-2008 period than it is for the 2009-2021 period, when it is often negligible. The impact also appears greater when the index uses higher frequency expenditure weights, like a monthly chained Tornqvist. However, this appears to be driven by durable goods for which micro expenditure data likely differ considerably from flow-of-service consumption.

Keywords: Cost of living index; price index; inflation; democratic aggregation; consumer heterogeneity

JEL Codes: C43, E31

*I am grateful to David Johnson, William Johnson, Josh Klick, Sara Stanley, Anya Stockburger, Rob Cage, Brian Adams, Thesia Garner, and many others for helpful comments. All errors are my own.

[†]Division of Price and Index Number Research, Bureau of Labor Statistics, 2 Massachusetts Ave, NE, Washington, DC 20212, USA. Email: Martin.Robert@bls.gov

1 Introduction

Most statistical agencies calculate consumer price indexes (CPI) using market-level expenditures and prices, corresponding to a representative consumer model. Such indexes are sometimes called “plutocratic” because they implicitly weight households by their total expenditure (Pollak, 1989; ILO, 2004). If inflation varies systematically with household expenditures, then a traditional CPI may differ from an aggregate that equally-weights households, so-called “democratic” (Prais, 1959). Plutocratic indexes may be appropriate indicators of overall macroeconomic conditions as they weight each dollar of expenditure equally. However, when constructing an indicator of consumer experiences (e.g., for adjusting payments), then equal weighting is attractive (Astin and Leyland, 2015).

This paper estimates democratic counterparts to the BLS’s CPI and Chained CPI (C-CPI) for all urban consumers. I also compare plutocratic-democratic inflation gaps between the Lowe and Tornqvist formulas and for different weight frequencies.¹ To my knowledge, this is the first paper to estimate an all-items CPI using a Tornqvist formula that is democratic in the sense of weighting households equally.² Previous studies of democratic CPIs have tended to focus on fixed-weight indexes like the Laspeyres and Lowe. The building blocks for my main analysis are Lowe and Tornqvist indexes constructed for individual consumer units (roughly equivalent to households) from the Consumer Expenditure Survey (CE).³ To construct consumer unit-level expenditure baskets, I modify a statistical matching method proposed by Hobijn, Mayer, et al. (2009).⁴ The democratic Lowe and Tornqvist indexes are then formed as equally-weighted averages of the respective individual indexes. I also consider

¹The CPI-U uses the Lowe (also called a modified Laspeyres) formula at the upper level of aggregation, which is based on fixed expenditure weights from a prior biennial reference period. The C-CPI-U uses the Tornqvist formula along with concurrent, monthly expenditure weights (Bureau of Labor Statistics, 2020).

²Diewert (2005) considers equally-weighted countries in calculating price parities with a formula which includes the Tornqvist as a special case. Rambaldi and Fletcher (2012), Rambaldi and Rao (2013), and Hill and Scholz (2018) use democratic Tornqvist formulas to estimate hedonic house price indexes.

³For exposition purposes, I refer generically to “households” as any basic unit of analysis we may pick for which we assume joint economic decision making. I reference “consumer units” specifically as the micro units sampled in the CE.

⁴I follow the BLS CPI expenditure categories, which include rental equivalence, a measure of implicit consumption of owner-occupied housing.

indexes based on groupings of similar consumer units.

I estimate that from December 2002 to June 2021, a democratic version of the CPI-U exceeds its plutocratic counterpart by about 0.08 percentage points per year, on average, with differences tending to decline over the sample period. The results for the Tornqvist index are quite striking. Using consumer unit-based weights, I find a democratic C-CPI-U is higher than the plutocratic by an average of 0.19 percentage points per year. The pattern of higher democratic indexes is consistent with a tendency for individual inflation to co-vary negatively with household expenditure level, which matches the relationship observed between inflation and income from Jaravel (2018), Argente and Lee (2021), Kaplan and Schulhofer-Wohl (2017) and Klick and Stockburger (2021), who calculate plutocratic formulas for income quantiles. I also find differences between plutocratic and democratic indexes tend to be larger over the 2002-2008 period than over the 2009-2021 period.

The gap for the Tornqvist formula is large in price index terms. It is more than double the gap for the Lowe index and is comparable to well-known price index biases like those discussed in Moulton (2018). However, the results in Section 4 suggest much of the effect is mechanical and related to durable goods categories for which spending at the household level is likely more volatile than implicit consumption. I recommend that either statistical agencies construct household-level flow-of service values for big-ticket items like automobiles, or base democratic measures on more aggregate “pseudo-households” within which average spending may closer approximate consumption values in a given period.⁵ Coarse versions of the latter, based on quintiles of annual expenditure or annual before-tax income, result in democratic indexes which average only about 0.05 percentage points per year higher than their plutocratic counterparts. However, I do occasionally find short-run differences on the order of 0.1-0.2 percentage points per year.

This paper relates to the literature on inflation heterogeneity across individuals and sub-

⁵In the U.S. CPI, flow of service values for household appliances like refrigerators and stoves are already included implicitly part of owner-equivalent rent. Flow-of-service values for automobiles are in development at BLS as part of a broader consumption measure. These include a discount rate (opportunity cost), imputed depreciation, vehicle property taxes and vehicle interest rates (Armstrong et al., 2022).

populations such as Hobijn and Lagakos (2005) and Kaplan and Schulhofer-Wohl (2017). Jaravel (2021) gives a broad review. A review of the cost-of-living index theory as applied to heterogeneous groups can be found in Pollak (1989) and ILO (2004). Ley (2005) contains a substantial international review of price index aggregation across households, as well as a mathematical description of how household budget differences may lead to gaps between plutocratic and democratic Lowe indexes. Democratic indexes using U.S. consumer expenditure data have been studied previously in Kokoski (2000), Hobijn, Mayer, et al. (2009), and Cage, Klick, and W. Johnson (2018), among others, and is an active area of research at the Bureau of Labor Statistics (BLS).

A limitation of my study (and many previous ones) is that the BLS only samples prices to represent the urban population in a given geographic area—no available information links prices with individual households or groups. As a consequence, I must assume identical prices paid and expenditure shares within the item-stratum (e.g., men’s suits). Within-stratum heterogeneity has been shown to be the major driver of nondurable goods inflation heterogeneity across households (Kaplan and Schulhofer-Wohl, 2017), and income deciles (Jaravel, 2018). There is indirect evidence that within-stratum inflation at the individual or group-level is also negatively correlated with expenditure. Indeed, using transactions data for consumer nondurables, Jaravel (2018) finds greater differences between top and bottom income-decile specific price indexes when they are computed using barcode-by-income level detail in prices and expenditures than when they are computed starting from more aggregated basic price indexes, which would more closely match BLS methodology. Given positive correlation between income and expenditures, I expect the results in this paper to underestimate (in terms of magnitude) for the true gaps between plutocratic and democratic indexes.

2 Plutocratic and Democratic Price Indexes

The BLS constructs price indexes in two stages. First, elementary indexes are computed for narrowly defined item-area strata, such as men’s suits in Pittsburgh (Bureau of Labor Statistics, 2020). These are then combined using either the Lowe (CPI-U) or Tornqvist (C-CPI-U) formulas. I consider only the second stage of aggregation as this is feasible using the CE to create different sets of weights. This section describes the relationships between household-level expenditure data and market level price indexes.⁶ Similar expressions for the Lowe indexes, can be found in National Research Council (2002), Ley (2005), and elsewhere. For ease of exposition, I make some simplifications for notation. First, I refer to individual prices instead of elementary price indexes as the subjects of either the Lowe or Tornqvist aggregation. The formulas also omit CE sampling weights for households, but these are used in all empirical analysis. Finally, I do not explicitly incorporate geographic areas into notation, but one can consider the item indexes $i = 1, \dots, N$ as representing item-area pairs.

Let the subscripts i and h denote elementary items and households, respectively, and let t denote a time period, usually a month. The price index compares prices in a period t against some reference period, weighting them by quantities or expenditures from a potentially different reference period. For the CPI, the price reference period is fixed (for two years) at v , while the quantity reference period is biennial and denoted b . In current practice, b ends twelve months prior to v , and both are updated every two years. For example, for $t = \text{January 2018} \dots \text{December 2019}$, $v = \text{December 2017}$ and $b = \text{2015-2016}$. As for the C-CPI, the index for month t has month $t - 1$ as its price reference period, and the weights are derived from expenditures in both t and $t - 1$.

As discussed in Section 1, we assume that all households within the same geographic region face the same price change for elementary items. Let q_{ith} represent the elemen-

⁶The use of “household” in this section is intended to be generic. It is whatever base unit for which we implicitly assumes makes joint economic decisions. If we consider CE consumer units as households, then the relationships described in this section are hypothetical since the CE comprises two samples to estimate total expenditure.

tary quantity aggregate purchased at price p_{it} . Household-specific expenditures are denoted $e_{ith} = p_{it}q_{ith}$ and expenditure shares $s_{ith} = e_{ith} / \sum_{j=1}^N e_{jth}$. Index formulas are traditionally written for a representative consumer, i.e., in terms of aggregate quantities $q_{it} = \sum_{h=1}^H q_{ith}$, expenditures $e_{it} = \sum_{h=1}^H e_{ith}$, and expenditure shares $s_{it} = e_{it} / \sum_{i=1}^N e_{it}$. Equivalently, the representative consumer's expenditure shares can be written in terms of average quantities $H^{-1} \sum_{h=1}^H q_{ith}$ or average expenditures $H^{-1} \sum_{h=1}^H e_{ith}$.

The Lowe index is an arithmetic average of price relatives p_{it}/p_{iv} . These are weighted using hybrid (also called price-updated) expenditure shares, which combine period v prices and period b quantities. For a household, we can write it as $P_{Lo,h} = \sum_{i=1}^N s_{i\{v,b\}h} \frac{p_{it}}{p_{iv}}$, where $s_{i\{v,b\}h} = p_{iv}q_{ibh} / \sum_{j=1}^N p_{jv}q_{jbh}$ is the hybrid expenditure share for the household. Using the CE, we only observe consumer units for a sub-period of b (i.e., a twelve-month span). I abstract from this issue for now, but it is discussed further in Section 3. The plutocratic Lowe index is denoted P_{PLo} and written in Eq. 1. With common prices, this corresponds to the usual market-level Lowe index using market expenditure shares $s_{i\{v,b\}} = \sum_{h=1}^H p_{iv}q_{ibh} / \sum_{k=1}^H \sum_{j=1}^N p_{jv}q_{jbh}$. The third line of Eq. 1 shows the well-known result that this index can be written as an expenditure-weighted (plutocratic) average of the household-specific Lowe indexes.

$$\begin{aligned}
P_{PLo} &= \sum_{i=1}^N s_{i\{v,b\}} \frac{p_{it}}{p_{iv}} \\
&= \sum_{h=1}^H S_{\{v,b\}h} \sum_{i=1}^N s_{i\{v,b\}h} \frac{p_{it}}{p_{iv}} \\
&= \sum_{h=1}^H S_{\{v,b\}h} P_{Lo,h}
\end{aligned} \tag{1}$$

The second line of the equation follows from re-writing the aggregate shares as $s_{i\{v,b\}} = \sum_{h=1}^H S_{\{v,b\}h} s_{i\{v,b\}h}$, where $S_{\{v,b\}h} = \sum_{i=1}^N p_{iv}q_{ibh} / \sum_{k=1}^H \sum_{i=1}^N p_{iv}q_{ibk}$ is household h 's share of total consumer expenditure. This shows how the formula gives more weight to households with higher expenditures, hence the plutocratic designation.

The C-CPI-U, in contrast, uses the Tornqvist formula, which is a geometric mean of monthly price relatives $p_{it}/p_{i,t-1}$, weighted by the arithmetic average of the month t and $t - 1$ expenditure shares. For the household, this is written $P_{T,h} = \prod_{i=1}^N \left(\frac{p_{it}}{p_{i,t-1}} \right)^{w_{ith}}$, where $w_{ith} = .5(s_{i,t-1,h} + s_{ith})$. The Tornqvist reflects substitutions in demand due to relative price change and is considered a better approximation to the consumer's cost-of-living index (COLI). In economic theory, the COLI is the change in expenditure required for the consumer to be indifferent between facing current prices and reference period prices (Konüs, 1924; Diewert, 1976). With common prices, the plutocratic Tornqvist can be written a few ways, as shown in Eq. 2.

$$\begin{aligned}
P_{PT} &= \prod_{i=1}^N \left(\frac{p_{it}}{p_{i,t-1}} \right)^{w_{it}} \\
&= \prod_{h=1}^H \prod_{i=1}^N \left(\frac{p_{it}}{p_{i,t-1}} \right)^{w_{ith}^*} \\
&= \left\{ \prod_{h=1}^H \left[\prod_{i=1}^N \left(\frac{p_{it}}{p_{i,t-1}} \right)^{s_{i,t-1,h}} \right]^{S_{t-1,h}} \right\}^{.5} \left\{ \prod_{h=1}^H \left[\prod_{i=1}^N \left(\frac{p_{it}}{p_{i,t-1}} \right)^{s_{ith}} \right]^{S_{th}} \right\}^{.5}.
\end{aligned} \tag{2}$$

The first line shows the market level Tornqvist index using aggregate expenditure weights $w_{it} = .5(s_{i,t-1} + s_{it})$, while the second line shows an aggregation across households and items using weights $w_{ith}^* = .5(S_{t-1,h}s_{i,t-1,h} + S_{th}s_{ith})$. This follows from the relationship $s_{it} = \sum_{h=1}^H S_{th}s_{ith}$, where $S_{th} = \sum_{i=1}^N e_{ith} / \sum_{k=1}^H \sum_{i=1}^N e_{itk}$ is household h 's share of total expenditure in month t . Unlike the Lowe, the market-level Tornqvist cannot generally be written as an average of household-level Tornqvist indexes. The index can be written in terms of household price indexes, as on line 3, but they are not Tornqvist. The terms in square brackets are household-specific geometric Laspeyres and geometric Paasche indexes, which are then averaged geometrically across h using either S_{th} or $S_{t-1,h}$ as weights.

Democratic Lowe and Tornqvist indexes are calculated simply by giving each household

equal weight. For the Lowe index, we use an arithmetic average, as in Eq. 3.

$$P_{DLo} = \frac{1}{H} \sum_{h=1}^H \sum_{i=1}^N s_{i\{v,b\}h} \frac{p_{it}}{p_{iv}} = \sum_{i=1}^N \bar{s}_{i\{v,b\}} \frac{p_{it}}{p_{iv}}. \quad (3)$$

The last term of Eq. 3 shows the democratic Lowe looks very similar to the usual Lowe formula, but uses equally-weighted expenditure shares, given by $\bar{s}_{i\{v,b\}} = H^{-1} \sum_{h=1}^H s_{i\{v,b\}h}$.

I define the democratic Tornqvist as an equally-weighted geometric average of household Tornqvist indexes.

$$P_{DT} = \left(\prod_{h=1}^H \prod_{i=1}^N \left(\frac{p_{it}}{p_{i,t-1}} \right)^{w_{ith}} \right)^{1/H} = \prod_{i=1}^N \left(\frac{p_{it}}{p_{i,t-1}} \right)^{\bar{w}_{it}}, \quad (4)$$

where $\bar{w}_{it} = H^{-1} \sum_{h=1}^H w_{ith}$. The final term shows that by using a geometric mean, the democratic Tornqvist can be written similarly to the market-level Tornqvist, but using democratic weights $\bar{w}_{it}$ instead of the usual w_{it} . While the plutocratic formula's weights can be written in terms of average (across households) expenditures, the democratic formulas are based on average expenditure shares.

Eq. 4 differs from the democratic Laspeyres and Paasche indexes defined in ILO (2004) as equally-weighted arithmetic averages of household index. I use the geometric mean because it is the same type of mean function used in the plutocratic formula to average across households. Jensen's inequality implies that an arithmetic average of household indexes exceeds the geometric average unless the $P_{T,h}$ are equal across h . Appendix A.3 shows using an arithmetic mean would increase the index by about 0.03 percentage points per year.

Index levels, or long-term changes, are formed by chaining (multiplying) shorter-term price changes. For the Lowe case, let $P(v, t)$ be either the plutocratic or democratic covering price change from v to t , as given by either Eq. 1 or Eq. 3. Then for index levels $I(t)$, we have $I(t) = I(t-1) \times [P(v, t)/P(v, t-1)]$. Similarly, we convert the monthly Tornqvist changes (Eq. 2 or Eq. 4) into index levels with the relationship $I(t) = I(t-1) \times P(t-1, t)$.

2.1 Plutocratic Gaps

Ley (2005) refers to the difference between the plutocratic and democratic index as the “plutocratic gap,” written for the Lowe formula as

$$\begin{aligned} P_{PLo} - P_{DLo} &= \sum_{i=1}^N (s_{i\{v,b\}} - \bar{s}_{i\{v,b\}}) \frac{p_{it}}{p_{iv}} \\ &= \sum_{i=1}^N (s_{i\{v,b\}} - \bar{s}_{i\{v,b\}}) (R_{it} - \bar{R}_t), \end{aligned} \quad (5)$$

where $R_{it} = p_{it}/p_{iv}$ and $\bar{R}_t = N^{-1} \sum_{i=1}^N R_{it}$. For the Tornqvist, I apply Ley (2005) to write the gap in natural logs. Appendix A.1 has more details.

$$\begin{aligned} \ln P_{PT} - \ln P_{DT} &= \sum_{i=1}^N (w_{it} - \bar{w}_{it}) \ln \left(\frac{p_{it}}{p_{i,t-1}} \right) \\ &= \sum_{i=1}^N (w_{it} - \bar{w}_{it}) (r_{it} - \bar{r}_t), \end{aligned} \quad (6)$$

where $r_{it} = \ln(p_{it}/p_{i,t-1})$ and $\bar{r}_t = N^{-1} \sum_{i=1}^N r_{it}$. The second equalities of Eq.’s 5 and 6 follow because the differences in weights $s_{i\{v,b\}} - \bar{s}_{i\{v,b\}}$ and $w_{it} - \bar{w}_{it}$ sum to zero and $\bar{R}_t$ and $\bar{r}_t$ are constant across i .

The plutocratic gaps equal $N - 1$ times the sample covariance between the weight differences and item-level price changes, as the differences between democratic and plutocratic weights are mean zero by construction. Ley (2005) derives conditions for the plutocratic gap to exist. First, there must be dispersion in expenditures across households and price change across items, which is generally true (see, e.g., Bureau of Labor Statistics (2019); Bureau of Labor Statistics (2021)). In addition, differences between plutocratic and democratic weights, $s_{i\{v,b\}} - \bar{s}_{i\{v,b\}}$, or $w_{it} - \bar{w}_{it}$, must be correlated with price changes for individual. It is not sufficient, for example, that poorer households spend a greater proportion on food than richer households. This must be accompanied by inflation rates for food items that are systematically higher or lower than average inflation.

There are a couple reasons to expect the Lowe and Tornqvist formulas might have different plutocratic gaps. First, the Tornqvist index approximates a cost-of-living index, while the Laspeyres is an upper bound (Diewert, 1976; Konüs, 1924).⁷ A difference like $P_{PLo} - P_{PT}$ is sometimes (e.g., Aizcorbe and Jackman (1993)) interpreted as a “substitution effect” in that the Tornqvist formula reflects expenditure shifts in response to relative price changes while the Laspeyres and Lowe do not. T. Garner, D. Johnson, and Kokoski (1996) and Argente and Lee (2021) suggested lower income or impoverished households have more limited ability to substitute consumption. If so, then household-level differences $P_{Lo,h} - P_{T,h}$ may increase with total expenditure, which would correspond to a higher Tornqvist plutocratic gap.

The second reason is more mechanical and potentially less desirable for a consumer price index. Infrequently-purchased items can receive a lower weight in a democratic indexes even if long-run purchases are homogeneous. Further, purchases will necessarily be less frequent at a monthly frequency (the Tornqvist) than at an annual or biennial frequency (the Lowe).⁸ To understand how this occurs, suppose there are only two items and consider their expenditure shares. With simplified notation, the expenditures at the household level are given by e_{ih} , $i = 1, 2$. The plutocratic share is

$$s_i = \frac{H^{-1} \sum_{h=1}^H e_{ih}}{H^{-1} \sum_{h=1}^H e_{1h} + H^{-1} \sum_{h=1}^H e_{2h}}, \quad i = 1, 2 \quad (7)$$

while the democratic share is

$$\bar{s}_i = H^{-1} \sum_{h=1}^H \frac{e_{ih}}{e_{1h} + e_{2h}}, \quad i = 1, 2. \quad (8)$$

Suppose the first item is always purchased, while the second item is purchased by the household with probability ρ . Suppose further that conditional on purchase, quantities are con-

⁷While the Lowe is technically a modified Laspeyres formula, the version in this paper uses lagged expenditure information like the Laspeyres index.

⁸As discussed in Section 3, the lowest frequency household-level weights we can compute for the Lowe index are annual.

stant across households. Formally, we assume

$$e_{1h} = e_1, \forall h. \quad (9)$$

$$e_{2h} = d_h e_2, d_h \sim \text{Bernoulli}(\rho). \quad (10)$$

From Slutsky's theorem and the expected value of a Bernoulli random variable, we have the following asymptotic results as H approaches infinity.⁹

$$s_1 = \frac{e_1}{e_1 + \rho e_2} + o_p(1). \quad (11)$$

$$s_2 = \frac{\rho e_2}{e_1 + \rho e_2} + o_p(1). \quad (12)$$

$$\bar{s}_1 = \frac{e_1 + (1 - \rho)e_2}{e_1 + e_2} + o_p(1). \quad (13)$$

$$\bar{s}_2 = \frac{\rho e_2}{e_1 + e_2} + o_p(1). \quad (14)$$

Note that by construction, if $\rho = 1$, then there is no difference between the democratic and plutocratic shares. However, for $\rho < 1$, then we have (with probability approaching 1) $\bar{s}_2 < s_2$ and consequently, $\bar{s}_1 > s_1$, since shares must sum to one. Therefore, the democratic shares place lower weight on the less frequently purchased item (e.g., an automobile) while placing higher weight on the more frequently purchased item (e.g., rent). The consequences are more severe for higher frequency indexes as this lowers the probability of purchase in the period. Section 4 will show this can be significant when infrequently-purchased items have large expenditures, like automobiles. As durable goods, purchases of automobiles are more volatile than the household's implicit consumption of the services they provide, and this is necessarily exacerbated at the household level. Ideally, household-level weights for durable goods would reflect consumption rather than acquisition.

⁹The term $o_p(1)$ means a term that converges in probability to zero.

2.1.1 Pseudo-Households

Section 2 thus far has presumed actual household (or consumer unit) microdata are the basis for computing democratic weights, as in Kokoski (2000), Ley (2005) or Hobijn, Mayer, et al. (2009). Household-based indexes have a high data requirement, however. For instance, they require detailed data on the entire expenditure basket, which may not be available. Additionally, the above discussion suggests reliable durable good consumption measures at the household level are superior to raw expenditures. An alternative mentioned in ILO (2004) and elsewhere, and implemented by Cage, Klick, and W. Johnson (2018) is to first group households into coarser units which are then treated as single agents. Cage, Klick, and W. Johnson (2018), for example, groups households based on similarity in expenditure shares for broad categories. This may be attractive if average expenditures for durable goods within pseudo-households approximate consumption values better than consumer-unit specific acquisition costs. The pseudo-households should be comprised of otherwise relatively homogeneous consumer units, which ensures within-group plutocratic gaps can be ignored.¹⁰ I choose a few relatively coarse groupings to explore as pseudo-household definitions for democratic indexes in Section 4: quintiles of monthly expenditure, quintiles of annual expenditure, and quintiles of annual before-tax income. These definitions are less sophisticated than the procedure used by Cage, Klick, and W. Johnson (2018) for annually-weighted indexes. They use a machine learning procedure to cluster consumer units based on annual expenditure shares for 15 expenditure categories. Extending this procedure to a monthly-weighted index may be challenging, as for example, monthly expenditure shares on transportation-related categories will still be heavily influenced by infrequent auto purchases.

To explain why more homogeneous household groupings are preferred, I decompose the Lowe plutocratic gap into between-group and within-group components, with the groups

¹⁰Cage, Klick, and W. Johnson (2018) also argue the first stage grouping by expenditure shares makes the assumption of homotheticity more plausible, though such a grouping would not rule out income effects on expenditure shifts over time. In this section, I am making a different argument related specifically to durable goods expenditures.

corresponding pseudo-households.¹¹ Suppose we partition the H households into G groups indexed by g . Let $S_{P,g}$ be group g 's share of total expenditure and $S_{D,g}$ its share of the total population. We can then write the plutocratic gap as:¹²

$$P_{PLo} - P_{DLo} = \sum_{g=1}^G S_{P,g} P_{PLo,g} - \sum_{g=1}^G S_{D,g} P_{DLo,g} \quad (15)$$

$$= \sum_{g=1}^G (S_{P,g} - S_{D,g}) \bar{P}_g + \sum_{g=1}^G \bar{S}_g (P_{PLo,g} - P_{DLo,g}), \quad (16)$$

where $\bar{P}_g = .5 (P_{PLo,g} + P_{DLo,g})$, $\bar{S}_g = .5 (S_{P,g} + S_{D,g})$, and $P_{PLo,g}$ and $P_{DLo,g}$ are the plutocratic and democratic Lowe indexes, respectively, computed over households in group g .

The first term of Eq. 16 is the between-group component. It captures differences in inflation and expenditures across groups and is the plutocratic gap that would prevail if within-group spending patterns were either homogeneous or uncorrelated with price changes. The second term of Eq. 16 is the within-group component. It captures the differences in inflation within the group and is the plutocratic gap that would prevail if average inflation was the same across groups or if group-level spending was proportional to its share of the population.

A democratic index based on pseudo-households is necessarily formed by expenditure-weighting households in the first stage. Therefore, it makes sense to choose groupings which minimize the “within” component of the plutocratic gap. One way to accomplish this is to choose groupings which minimize within-group variance in household expenditure, which might be unobserved if we think consumption expenditure of durable goods is mismeasured.

¹¹I focus on the Lowe index for simplicity. With respect to the household dimension, the consistency in aggregation property of index numbers means we can first aggregate households into groups according to their share of group expenditure (or share of the group population), and then aggregate the group indexes into a full population index according to the group's share of total expenditure (or share of total population). This is true of the democratic and plutocratic Lowe, as well as the democratic Tornqvist, but not the plutocratic Tornqvist. With respect to the item dimension, the Tornqvist formula is approximately consistent in aggregation (Diewert, 1978), implying that the intuition of the following discussion still applies, even if the formulas do not exactly.

¹²This follows Eq. 1 of Jaravel (2018), who similarly decomposes the rich-poor inflation gap into between-stratum and within-stratum components.

Why should this be the aim? From Ley (2005), the plutocratic gap consists of a dispersion term (the square of the coefficient of variation of total household expenditures), multiplied by a term that captures the covariance between item-level expenditure elasticities and price changes, reflecting heterogeneity in expenditure shares. Of these two factors, the former (within-group dispersion in total expenditures) seems to be the most straightforward to reduce. One potential method is to use quantiles of the expenditure distribution itself, as I explore in Section 4. Ideally the period for expenditures should be long enough so that group membership is not overly sensitive to a single large purchase; otherwise the intended effect in reducing dispersion in infrequently-purchased durable goods shares will not be achieved, as will be the case when the quintiles are based on monthly expenditures. Greater homogeneity in all shares will also reduce the within-group gap. To use a clustering method like in Cage, Klick, and W. Johnson (2018) for the Tornqvist index, however, one may wish to exclude the questionable durable goods categories from the shares being matched on. One would also need to use similarity measures that match on month $t-1$ and t expenditure shares separately, so that the resulting indexes still reflect expenditure shifts at the pseudo-household level.

3 Data and Processing

This section briefly describes the data sources used in calculation of the CPI and highlight relevant features for constructing household-weighted measures. More complete information is available in Bureau of Labor Statistics (2018) and Bureau of Labor Statistics (2020). I also briefly describe how I link CE Diary and Interview observations and process the data for consumer unit-level price indexes.

3.1 CPI Elementary Item-Area Indexes

Since 2018, the lowest level of aggregation in the CPI consists of 243 items and 32 geographic areas for a total of 7,776 elementary indexes, with similar dimensions in earlier

periods. Within each item-area, individual price quotes are sampled to be representative of the products purchased and outlets patronized by urban consumers. These are then aggregated using either a modified Laspeyres or geometric mean formula, depending on the product category. The same elementary indexes are the basic inputs to both the CPI and C-CPI.

The CPI undergoes periodic changes in the item and area samples. To simplify processing, I create a constant item classification structure over time by combining a few elementary aggregates. I also bound the monthly item-area index relatives to be between 0.05 and 20, which amounts to censoring a few monthly price changes.

3.2 Consumer Expenditure Survey

In the second stage of CPI aggregation, elementary indexes are combined using weights derived from the Consumer Expenditure Survey (CE). The CE is a nationwide survey designed to measure the full range of spending for the noninstitutionalized civilian population. The survey also collects information on income and demographic characteristics. The sample unit is known as a consumer unit (CU). While often corresponding to a household, a CU is a group of individuals who make joint expenditure decisions (Bureau of Labor Statistics, 2018).

The CE consists of two separate surveys—the Interview and the Diary—comprising different samples and categories of expenditure. The Interview survey has a three month recall period and is designed to cover major purchases (e.g., major appliances) and recurring items (e.g., rent, utilities). Consumer units are interviewed once every three months for four consecutive quarters on a rolling basis. The reference period for expenditures for each interview is the prior three calendar months. Roughly 7,000 consumer units complete the interview in a given quarter, with approximately 25% of the sample turning over each quarter. In contrast, the Diary consists of both a smaller sample and shorter recall period. Approximately 1,700 consumer units provide usable diaries in a given quarter, recording expenditures for

two consecutive one-week periods. The survey covers more minor and frequent expenditures, such as food and apparel. The CE treats separate Interview and Diary waves as independent. I do the same for the weekly diaries, but I link interviews for the same consumer unit across collection quarters. The latter allows for a larger sample when constructing Tornqvist indexes because I can use adjacent months a consumer unit’s expenditure even if they were collected in different quarters. It also allows me to construct annual expenditures for the subset of consumer units who completed four interviews. For my analysis, I use the CE sampling weights, FINLWT21, which treat Interviews as independent.¹³

Table 1 gives the share of CPI-covered expenditure by major category and survey source. The underlying aggregations incorporate sampling weights and adjustments made by CPI for recall period differences across the two surveys. Roughly 75% of expenditure covered by the CPI is sourced from the Interview, while the remaining 25% of CPI expenditures are sourced from the Diary. The Interview provides the vast majority of expenditures for Housing, Transportation, Medical Care, and Education and Communication, while the Diary is the predominant source for Food and Beverages and Apparel. Significant proportions of Recreation and Other Goods and Services come from each survey.

The Interview survey is the natural starting point for constructing household-based measures of inflation, as participation lasts up to one year (versus two weeks for the diary) and coverage includes most CPI-eligible expenditure categories. I use all available consumer units for which I can compute individual Tornqvist indexes. As these depend on current and prior month expenditures, I can not compute them until a consumer unit’s second month in the sample. As for the Lowe indexes, a challenge is that consumer units contribute a maximum of twelve months of expenditures during a single biennial reference period. Moreover, variation in total expenditures across households is part of what determines the difference between democratic and plutocratic expenditure shares (Ley, 2005). Since I base the Lowe

¹³Ideally, longitudinal weights would be created for each sample of consumer units contributing to the $\{t - 1, t\}$ Tornqvist relatives and the samples contributing to the reference period b Lowe indexes. This is not a trivial process (T. I. Garner et al., 2022), and I leave it to future research.

indexes on the same biennial reference periods as the published CPI, I also restrict the sample to consumer units who completed all four interviews and have four complete quarters of spending within the biennial reference period. This is to reduce the variation in total expenditures at the consumer unit level stemming from either normal sample rotation or attrition. For the 2013-2014 reference period, for example, I take consumer units who participated in four interviews and had complete reference quarters occurring between January 2013 and December 2014. I sum their expenditures by item, and then the consumer unit-item level expenditures are price-updated to the pivot month, December 2015. These shares are combined with elementary index relatives based to December 2015 and ending between January 2016 and December 2017. While all consumer unit-level shares are drawn from within the same biennial reference period, the shares themselves are necessarily at an annual frequency. In the appendix, Figure A4 shows the impact of consumer unit selection is quite small for the plutocratic Lowe and Tornqvist indexes.

Expenditures on categories usually collected in the Diary are imputed using a procedure described in the next subsection. For those categories collected in the Interview, my initial analysis assumes missing values represent expenditures of zero. Missing expenditures at the consumer unit level also (implicitly) receive a zero weight in the plutocratic formula. If differential missingness across consumer units is linked to preferences or income, then we may want this reflected in democratic indexes. For instance, not all households buy children's clothes, pet supplies, airline tickets, or cable television services. For durable goods like automobiles, however, missing expenditures do not necessarily mean the household did not consume the implicit service provided by the good. As discussed in Section 2.1, infrequently purchased items will generate weight differences between plutocratic and democratic indexes even if average expenditures are homogeneous. Section 4 explores the empirical consequences, which turn out to be substantial for monthly weighted indexes.

I also set negative expenditures at the CU-item-month level to zero, so that the weights for individual indexes are all weakly positive. Negative expenditures are occasionally recorded

for certain items, like medical care, for which the consumer unit receives reimbursement. Positive expenditure shares at the household level seems a reasonable requirement for a democratic index. Nevertheless, while Figure A4 in the appendix suggests that bounding CU-level expenditures to be positive has a small negative effect on average, in the earliest years of the sample, it lowers the Plutocratic Lowe twelve month percent changes by around a tenth of a point on average. For this reason I compute plutocratic gaps using plutocratic and democratic indexes that are both based on the same bounding rule. Finally, the CE implemented an area sample revision starting in January 2015. Even when a CU in a continuing area remained in the sample, their identifier changed. As a Tornqvist index requires expenditures in adjacent months, I link 1,315 CU's from continuing index areas with their previous identifiers by matching income, core-based statistical area, age, housing tenure, race, sex, occupation, and family size. As a result, sample of CUs for which I can compute a Tornqvist index in January 2015 is small, only about 25% of the sample for February 2015. Nevertheless, there is no structural break in the timeseries for this period.¹⁴

Finally, I also omit two weight-smoothing procedures used in the CPI and C-CPI—composite estimation for the biennial item-area Lowe weights and ratio-allocation for the monthly item-area Tornqvist weights. Both of these are designed to lower the sampling variance of weights for specific item-areas. Figure A4 in the Appendix shows my national-level CPI-U and C-CPI-U replications closely match the published indexes even without these procedures. These procedures are designed to reduce sampling error, but I leave it to future research to extend them to democratic indexes.

3.3 Matching Procedure

As mentioned, I start with the expenditures which the CPI program sources from the Interview sample. I then implement a statistical matching method based on Hobijn, Mayer,

¹⁴Furthermore, the democratic Tornqvist yields a very similar relative in this month as the democratic Fisher of ILO (2004), which can be constructed to use all observations regardless of whether they appear in adjacent months. The results are available from the author by request.

et al. (2009). Similar donors from the Diary provide the missing data for each Interview consumer unit, where similarity is determined by a model of average monthly expenditure as a function of demographic characteristics. The model provides a convenient method of weighting a relatively large number of characteristics by doing so according to which linear combination most strongly predicts expenditures.

The matching procedure is summarized as follows. For each month and subpopulation (when computing subpopulation indexes), I estimate a regression of average expenditure on a vector of characteristics for the interview sample. Characteristics include Census region, urban/rural, age, race, sex, and education of the reference person, family size, and the prior year’s annual income, and I estimate the model using the least absolute shrinkage and selection operator (LASSO). I then compute predicted values for the Interview observations and Diary observations from the same month. Defining distance as the absolute difference between these predicted values, I randomly pick one Diary from the closest 20 observation, repeating the process for each Interview observations. The donor Diary’s expenditures are then assigned to the Interview, with a scale adjustment to convert from a one-week to a one-month reference period. Since the Interview sample is much larger than the Diary on a per-month basis, draws are performed with replacement. As a result, each Diary is matched with several Interviews.

Appendix A.2 has more details on the matching process and the properties of the matched dataset. Subsequent index calculation in Section 4 uses the synthetic consumer unit observations for both plutocratic and democratic aggregates, in order to control for the effect of the matching process and isolate the effect of equally-weighting households. Appendix A.3 includes comparisons between the published CPI-U and C-CPI-U and the plutocratic replications using the matched data, finding them to be very close on average.

In some cases, both surveys collect comparable expenditure data. In cases of overlapping coverage, CPI chooses a single source based on factors such as classification detail and sampling variance. As a consequence, a significant portion of the 25% of expenditure sourced

from the Diary is actually available from the Interview in some form. In particular, the Interview collects a handful of coarse food spending totals, such as food-at-home. Appendix Table A1 has more details. One potential alternative to matching, along the lines of Cage, Klick, and W. Johnson (2018), is to source as much as possible from the Interview survey regardless of classification detail or the CPI program’s actual source selection. This results in indexes that reflect roughly 95% of CPI-eligible expenditures. However, they incorporate only a few coarse sub-aggregate indexes for food and beverages rather than the full detail of elementary indexes available. Using this method I find many qualitative aspects of the democratic-plutocratic comparison match my main results (see Figure A7 in the appendix). However, I find it produces plutocratic indexes which differ more significantly from the official CPI and C-CPI than does my replication using the synthetic matched sample. See Appendix A.3 for more details.¹⁵

4 Results

4.1 Indexes Based on Consumer Units

I first describe the substantial heterogeneity in expenditures and inflation across consumer units before presenting the democratic index estimates. As discussed in Section 2.1, a necessary condition for a plutocratic gap to exist is that the expenditures are unevenly distributed across households. Figure 1 shows this is certainly true using CE data. Looking at the expenditures underlying the biennial Lowe index, the top 20% of consumer units accounts for 39% of expenditures over the 2002-2021 sample period, while the bottom 20% accounts for only 7.8%. For the monthly Tornqvist, expenditures are less equal at the higher frequency, as the top 20% of accounts for 45.7%, while the bottom 20% accounts for only 6.1%.

¹⁵Another potential alternative, along the lines of Cage, T. Garner, and Ruiz-Castillo (2002), is to use the Diary to compute average shares of expenditure on detailed food items (e.g., Breakfast Cereal) out of total food spending for some subgroup, and then apply these shares to the total food spending of Interview consumer units within the same subgroup.

Figure 2 describes how expenditure patterns tend to differ at different points in the expenditure distribution, focusing on the samples of consumer units contributing to the 2018-19 index period. The stacked bars reflect average expenditure weights within each quintile of the expenditure distribution. Item-level weights are first summed within the eight major categories shown. For the Lowe, the underlying consumer unit expenditures (which occurred sometime during the 2015-16 weight reference period) are annual and price-updated to December 2017 (see Section 2). For the Tornqvist, the weights are the average of the consumer units current month share and prior month share. Food weights fall with total expenditure, reflecting a pattern observed at least since Engel in 1857 (Houthakker, 1957). Housing weights also tend to decrease with total expenditure shares, though not monotonically. In contrast, weights for Transportation, Medical Care, and Education and Communication tend to rise with total expenditure, while weights for Apparel are fairly flat. These patterns are broadly consistent with Figure 16 of Cage, Klick, and W. Johnson (2018), which examines annual spending shares by income level and age group for slightly finer categories during the 2013-14 reference period.

Similar to Hobijn, Mayer, et al. (2009) and Kaplan and Schulhofer-Wohl (2017), I find inflation is quite dispersed across consumer units. Figure 3 shows the distribution of Lowe indexes measuring inflation from December 2017 to December 2019 for the 2016-17 sample of consumer units with four interviews. While the plutocratic and democratic means both imply about 4.2% inflation, the standard deviation is about 2 percentage points, with observations ranging from about -5% to 25%. Previous studies such as Kaplan and Schulhofer-Wohl (2017) and Cage, Klick, and W. Johnson (2018) also found considerable dispersion in household-level indexes.¹⁶ Figure 4 summarizes the monthly Tornqvist results over the 2018-19 period, where the consumer level indexes measure one-month inflation. The two means track each other

¹⁶For comparison, Cage, Klick, and W. Johnson (2018) compute Lowe indexes covering December 2014 to December 2017 for the sample of consumer units with four interviews during 2013-14 and find average inflation of 4.9% with a standard deviation of 1.6 percentage points. Their method is based on coarser Interview aggregates for food spending, which might explain the lower dispersion. Kaplan and Schulhofer-Wohl (2017) use household scanner data for food and nonfood grocery items over 2004-13 and find an annual interquartile range of 6-9 percentage points.

closely, with a tendency of the democratic mean to be slightly higher, though the differences between the plutocratic and democratic mean (-0.015 percentage points on average) are small relative to the interquartile ranges (0.555 percentage points on average). I also find that monthly Tornqvist inflation is not very persistent over time at the consumer unit level. Using a panel first-order autoregression of indexes in 2019, the autoregressive coefficient is only 0.06 with a standard error of 0.01 . On average, consumer unit movements within the distribution of monthly inflation was on average, about 33 percentile points in magnitude compared to their previous month's ranking.

Looking over a longer period, Table 2 and Figure 5 show plutocratic gaps tend to be positive for both the Lowe and Tornqvist index formulas, but that the average Tornqvist gap is about two and a half times the magnitude of the average Lowe gap. Over the period December 2002 to June 2021, the average twelve month percent change of the democratic Lowe exceeds that of the plutocratic Lowe by 0.08 percentage points per year. The Lowe findings are roughly consistent with earlier findings by Kokoski (2000) and Hobijn, Mayer, et al. (2009). Kokoski (2000) finds from 1987-1997, a democratic modified Laspeyres index exceeds the plutocratic version by 0.5 percentage points cumulative, or 0.05 percentage points per year. Looking at 1984-2004, Hobijn, Mayer, et al. (2009) find the difference between a democratic and plutocratic Laspeyres-type indexes to be “less than 3 percentage points” cumulative, or less than 0.14 percentage points per year. For the 2002-2021 period, however, however, the democratic Tornqvist exceeds the plutocratic Tornqvist by a larger magnitude, 0.19 percentage points per year on average. Gaps in monthly and twelve-month percent change vary over time both in sign and magnitude. Table 2 implies that the plutocratic gaps from 2002 to 2008 tend to be larger in magnitude than those in later periods. In 2008, for instance, the Lowe gap averages -0.33 percentage points while the Tornqvist averages -0.61 percentage points. From 2010 to 2021, average gaps in the Lowe indexes tend to be relatively small and are both positive and negative. Similarly, the Tornqvist gaps also tend to be smaller in magnitude over this time frame, but larger average gaps still occur in 2017

and 2018, and 2021 (January to June), where the gap is also positive. While the signs of the gaps are not consistent over time, in the long run, the democratic index levels tend to be higher than the plutocratic, as shown in Figure 5. In particular, the Tornqvist gap is substantial in price index terms, as it is similar in magnitude to the difference between the Lowe CPI-U and the Tornqvist C-CPI-U.

Whether the larger Tornqvist gap makes economic sense is the focus of the rest of this section. In Appendix A.3, I show the plutocratic aggregates of consumer unit indexes track the published CPI-U and C-CPI-U well, suggesting the impact of the matching and processing differences described in Section 3 are minor. I find similar qualitative patterns if indexes are computed using only the CE Interview with aggregate food spending, but the plutocratic indexes do not track the CPI-U and C-CPI-U as well. I also find using an arithmetic mean to compute the democratic Tornqvist results in a slightly larger plutocratic gap, on average -0.22 percentage points per year instead of -0.19 . Finally, I evaluate the use of simple equivalence scales based on family size and find these make relatively little difference in calculation of the democratic Tornqvist index.

Demand theory may give reason to expect a more negative plutocratic gap for the Tornqvist than for the Lowe. Such a pattern is consistent with higher expenditure households having a greater ability to substitute away from items with higher relative price change, similar to findings in T. Garner, D. Johnson, and Kokoski (1996). The Tornqvist index approximates a theoretical cost-of-living index, which reflects such substitutions, while the fixed-weight Lowe index does not. By weighting households equally, the democratic Tornqvist index may reflect less substitution, on average, which would have a tendency to push it higher than the plutocratic Tornqvist. To assess this explanation, I compare Lowe and Tornqvist indexes for different segments of the expenditure distribution. I restrict attention to consumer units which completed four Interviews, for which we observe roughly one year of spending. For each weight reference period (for the Lowe) or month (for the Tornqvist) I divide consumer units who supply indexes into quintiles based on annual expenditure. I

use annual expenditures for both so that the quintiles are roughly comparable, as monthly expenditures are more volatile and less, the latter as shown in Figure 1. Figure 6 shows the average differences between the Lowe and Tornqvist twelve month percent changes by quintile. These average differences are interpreted as the degree to which each quintile substitutes away from items with higher relative price change. Figure 6 suggests an inverse, though nonmonotonic, relationship between expenditure quintile and substitution effect, the opposite of substitution patterns which would explain the differential plutocratic gaps. For instance, weighted equally, household Lowe indexes for the first quintile average about 0.14 percentage points per year higher than the household Tornqvist indexes. The margin is 0.18 percentage points for the second quintile, 0.15 percentage points for the third quintile, 0.09 percentage point for the fourth quintile, and 0.02 percentage points for the fifth quintile.

Instead, the tendency for the Tornqvist plutocratic gaps to be larger than the Lowe appears to stem from the higher frequency of the weights, which derive from monthly expenditure shares. Figure 7 plots the difference in twelve month percent changes between the plutocratic and democratic index for the Lowe and Tornqvist formulas respectively. The figure presents the same Lowe and Tornqvist indexes as in Table 2, as well as variants of the Lowe index methodology where the household is defined differently, either as a consumer unit-quarter combination or consumer-unit month combination, still within the same (lagged) biennial weight reference period as the baseline Lowe which uses annual expenditures at the consumer unit level. These variants have the effect of making the democratic Lowe indexes based on higher frequency expenditure shares, but have no or little effect on the plutocratic Lowe indexes.¹⁷ As Figure 7 shows, increasing the weight frequency for the Lowe index increases the magnitude and volatility of the plutocratic gaps, making them track much closer to those of the Tornqvist index.

Plutocratic gaps increase with the dispersion in expenditures across households (Ley, 2005). If monthly expenditures are more dispersed than annual expenditures, one might

¹⁷For the Lowe indexes based on monthly or quarterly shares, I also use the full sample of consumer units, though this has very little effect on the indexes or plutocratic gaps.

expect there to be larger plutocratic gaps for the monthly chained Tornqvist formula than the biennial Lowe. Nevertheless, the compositions of the plutocratic gaps suggest that rather than general dispersion, there are mechanical differences between the plutocratic and democratic weights (particularly for the Tornqvist indexes) stemming from infrequent purchases like automobiles. Section 2.1 showed that even if expected expenditure patterns are homogeneous, a good that is purchased with low probability (e.g., autos) will tend to receive a lower democratic weight than plutocratic weight, while a good that is always purchased (e.g., rent or owner-occupied rent) will receive a higher democratic weight. The associated gaps are greater for higher frequency indexes, which lowers the probability of the durable goods purchase occurring.

This issue is indeed significant for the consumer unit-based Lowe and Tornqvist indexes, where the majority of differences across formulas appear driven by transportation and housing. Figure 8 decomposes the plutocratic gaps for the 2018-19 period (see Eq.'s 5 and 6) into the contributions by major group. The Lowe plutocratic gap is expressed in level differences, while the Tornqvist plutocratic gap is expressed in natural log differences, but the magnitudes are small enough to be roughly comparable. By far, the two largest contributors to the Tornqvist plutocratic gap are transportation and housing, both which have a negative contribution. Figure 9 further examines the weight differences and relative inflation levels by major group. This figure ignores within major group correlations between share differences and price changes which also contribute to the plutocratic gap, but is suggestive of the overall effects shown in Figure 8. For instance, from Figure 8, we see that food contributes relatively little to the plutocratic gaps. Panel (a) of Figure 9 shows that while both the Lowe and Tornqvist plutocratic weights for food are substantially lower than the democratic weights, panel (b) shows food inflation (coarsely speaking) differs little from average inflation during this time period. On the other hand, inflation for apparel is lower than average inflation during this period, but weight differences are minimal between the plutocratic and democratic indexes. For both index formulas, the plutocratic index places a lower weight on

housing and a higher weight on Transportation, but the differences are much larger for the Tornqvist than for the Lowe.

Figures 10 and 11 repeat Figures 8 and 9 for the five most significant item strata in terms of contributing to the differences between the Lowe and Tornqvist plutocratic gaps. Indeed, we see that rent and owner equivalent rent (for which consumer units are more likely to have expenditures each month), as well as new vehicles, used vehicles and airline fares (which are less frequently purchased) are the largest contributors. In the case of rent and owner equivalent rent, the higher frequency of the Tornqvist index increases the democratic weights. For owner equivalent rent, this leads to a positive, but smaller weight difference relative to the Lowe, while for rent, it leads to a more negative weight difference relative to the Lowe. For vehicles and airline fares, we see larger negative weight gaps for the Tornqvist than we do the Lowe, reflecting even lower weight in the democratic Tornqvist index due to infrequency of purchase. Conceptually, the weight frequency effect may not be a problem per se as it relates to airline fares, as consumption is relatively infrequent and perhaps more likely as one moves to the right of the expenditure distribution. On the other hand, the simple example from Section 2.1 suggests that the infrequent purchase of vehicles may drive the weight differences for rents as well as vehicles. Together, the two rent and two vehicle items' contributions to the Tornqvist plutocratic gap sum to -0.0036 log points for the 2018-19 period, slightly more than the total plutocratic gap in logs, which is -0.0035 . For democratic indexes, this is undesirable because vehicle weights reflect actual purchases rather than consumption values in an economic sense, the latter of which should be much smoother at the micro level.¹⁸ It is probable that a consumption measure for vehicles would lead to smaller weight differences, and therefore a smaller plutocratic gap.¹⁹ Figure 12 shows that in terms of the overall plutocratic gaps, rent and vehicles were also the largest

¹⁸For a discussion related to the volatility of purchases versus consumption of owner occupied housing at the macro level, see Diewert and Shimizu (2021).

¹⁹I find that the Tornqvist plutocratic gaps can be significantly reduced simply by editing the two vehicle expenditures to be based on within consumer unit averages or averages by income quintile and month. The results are available by request.

contributors for the 2002-03 period as well as the 2018-19 period. The 2002-03 period is interesting because both the Lowe and Tornqvist indexes have substantial plutocratic gaps, though the Tornqvist is still much larger (-1.4 percentage points cumulative for the Tornqvist versus -0.4 percentage points cumulative for the Lowe). The impact of vehicles on weight differences is even larger in 2002-03 than it is in 2018-19. In the earlier period, gasoline is also a significant contributor to the negative plutocratic gap, as speculated by Hobijn, Mayer, et al. (2009) for the 2000-2004 period, but its contribution is overshadowed by automobiles and rent, even for the Lowe index.

A consumption measure for vehicles at the consumer unit level is beyond the scope of this paper, but the following subsection explores how this issue may otherwise be mitigated by redefining the household to be something coarser than a consumer unit.

4.2 Indexes Based on Pseudo-Households

Section 2.1.1 discussed how democratic indexes based on coarser pseudo-households may be preferred to indexes based on consumer units from the standpoint of smoothing durable goods expenditures, provided within-group plutocratic gaps are negligible. Table 3 compares the plutocratic gaps for the different household definitions, including the CU-based plutocratic Lowe and Tornqvist from Table 2 for reference. The alternative household definitions in this table are quintiles of annual expenditure and quintiles of annual before-tax income. In both the Lowe and the Tornqvist case, democratic indexes based on coarser household definitions are lower than those based on consumer units, leading to smaller magnitude plutocratic gaps. For the Lowe, the plutocratic gap using quintiles of annual expenditure is -0.07 percentage points on average, while it is -0.04 on average using quintiles of annual before-tax income, versus -0.08 on average when based on consumer units. For the Tornqvist, the reductions in the plutocratic gap are more substantial. Using quintiles of annual expenditure reduces the gaps to -0.04 percentage points on average, and using quintiles of annual income reduces them to -0.05 percentage points on average, versus -0.19 when basing the indexes on consumer

units. Notably, the Lowe and Tornqvist plutocratic gaps are much closer in magnitude using coarser pseudo-households than when using consumer unit-level indexes.

Expenditure-based definitions may be preferred to income-based definitions for reasons discussed in Section 2.1.1. As income is an imperfect correlate with expenditure, within-group plutocratic gaps may still be present when the groups are based on income. Figure 13 shows that using the coarser household definitions (even that based on income) indeed reduces the impact of housing and transportation weight differences in the Tornqvist case. The Tornqvist weight differences for housing are still negative, and the weight differences for transportation are still positive, but of a much lower magnitude. In addition, they are much closer to the Lowe weight differences which are based on consumer unit weights. Notably, if we choose quintiles of monthly expenditure as the pseudo-households, the housing and transportation weight differences are still quite substantial, as these groupings are sensitive to infrequent, big-ticket purchases. For instance, in 2019, consumer units that purchased a new or used vehicle had monthly expenditure percentiles that were 29.4 points higher, on average, than consumer units without vehicle purchases.

5 Conclusion

A democratic consumer price index may serve as an additional inflation indicator of interest to researchers and policymakers, as it tracks the average price change across households rather than across dollars of expenditure in the economy. For example, a household-weighted index for low-income populations may be useful for poverty analysis. This paper shows when using the Lowe or the Tornqvist formula, such an index tends to imply higher inflation on average than the traditional plutocratic aggregation. For several years between 2002 and 2008, the plutocratic gaps are of magnitudes on par with some of the well-known biases discussed in price index literature (Moulton, 2018), but in the long-run, average less than 0.1 percentage points in magnitude for the Lowe index. A democratic Tornqvist index based

on monthly consumer unit weights implies a much larger plutocratic gap on average, but this is largely due to infrequent durable goods purchases like automobiles, which at the micro level do not reflect consumption in an economic sense. Using coarser pseudo-household definitions, I find Tornqvist plutocratic gaps to be much smaller, more on the order of the Lowe plutocratic gaps. While the later period of my analysis (2010 to 2021) featured smaller plutocratic gaps, the larger gaps in the earlier period suggest that plutocratic indexes are not always a sufficient proxy for democratic indexes. In addition, results from Jaravel (2018) suggest using common item-area indexes may understate inflation differences between households across different points in the income distribution. If so, then plutocratic gaps may be larger when starting from less aggregated price and expenditure data.

Several issues remain to be considered. To begin with, consumer unit-level weights for large, infrequently purchased durable goods are usually zero, which tends to underweight these items in a democratic index. It would be preferable to use flow-of-service approaches at the consumer unit level for all big-ticket items, such as rental equivalence or user cost. In addition, while this paper studies democratic aggregation with indexes that replicate many features of BLS methodology, some procedures are omitted. Most notably, the CPI-U and C-CPI-U incorporate weight smoothing procedures designed to reduce variance. Versions or analogs of these procedures may be worth exploring for democratic indexes as well. Furthermore, democratic weights should be combined with the price replicate samples collected by BLS for purposes of standard error estimation, though a full accounting of sampling variance as discussed by Balk and Kersten (1986) would be difficult. Finally, pseudo-household methods may be implementable without matching or with more limited imputation. For instance, by only imputing total annual expenditure for the Diary and Interview samples, quintile-based democratic indexes would be calculable using methods similar to the income quartile indexes computed by Klick and Stockburger (2021). This may be simpler to implement in a production environment.

Finally, one may raise again the question of “whose index?” as posed previously in Prais

(1959) and Ley (2005). The plutocratic gaps described in this paper are small relative to the overall distribution of household inflation as documented in Hobijn, Mayer, et al. (2009) and Kaplan and Schulhofer-Wohl (2017), or even the differences by income as presented in Klick and Stockburger (2021). Depending on the objectives of policymakers, measurements of inflation dispersion may be of interest along with inflation averages for a diverse set of subpopulations. Furthermore, other frequencies of inflation may be desirable as building blocks for a household-weighted measurements, as household-specific series also display considerable dispersion over time. Current month compared to twelve months prior would be especially interesting, but one would need longer timeseries of consumer unit-specific weights than provided by the CE.

References

- Aizcorbe, Ana M. and Patrick C. Jackman (1993). “The commodity substitution effect in CPI data, 1982-91”. In: *Monthly Labor Review* 116, p. 25. URL: <https://www.bls.gov/opub/mlr/1993/article/commodity-substitution-effect-in-cpi-data-1982-91.htm>.
- Argente, David and Munseob Lee (2021). “Cost of living inequality during the great recession”. In: *Journal of the European Economic Association* 19.2, pp. 913–952. DOI: 10.1093/jeea/jvaa018.
- Armstrong, Grayson et al. (2022). “Building a Consumption Poverty Measure: Initial Results Following Recommendations of a Federal Interagency Working Group”. In: *AEA Papers and Proceedings*. Vol. 112. American Economic Association, pp. 335–39. DOI: 10.1257/pandp.20221041.
- Astin, John and J Leyland (2015). “Towards a household inflation index”. Report commissioned by the Royal Statistical Society. URL: <https://rss.org.uk/RSS/media/News->

and-publications/Publications/Reports%20and%20guides/Astin-Leyland-HII-paper-Apr-2015.pdf.

Balk, Bert M. and H. M. P. Kersten (1986). “On the precision of consumer price indices caused by the sampling variability of budget surveys”. In: *Journal of Economic and Social Measurement* 14.1, pp. 19–35. DOI: 10.3233/JEM-1986-14102.

Bureau of Labor Statistics (2018). “Consumer Expenditures and Income”. In: *Handbook of Methods*. Bureau of Labor Statistics. URL: <https://www.bls.gov/opub/hom/cex/home.htm>.

— (2019). *Table R-1. All consumer units: Annual detailed expenditure means, standard errors, coefficients of variation, and weekly (D) or quarterly (I) percents reporting, Consumer Expenditure Survey, 2019*. Tech. rep. URL: <https://www.bls.gov/cex/tables/calendar-year/mean/cu-all-detail-2019.pdf>.

— (2020). “The Consumer Price Index”. In: *Handbook of Methods*. Bureau of Labor Statistics. URL: <https://www.bls.gov/opub/hom/cpi/home.htm>.

— (2021). *Table 2. Consumer Price Index for All Urban Consumers (CPI-U): U. S. city average, by detailed expenditure category, August 2021*. Tech. rep. URL: <https://www.bls.gov/news.release/cpi.t02.htm>.

Cage, Robert A., Thesia Garner, and Javier Ruiz-Castillo (2002). “Constructing Household Specific Consumer Price Indexes: An Analysis of Different Techniques and Methods”. BLS Working Paper 354. URL: <https://www.bls.gov/osmr/research-papers/2002/pdf/ec020030.pdf>.

Cage, Robert A., Joshua Klick, and William Johnson (2018). “Population Subgroup Price Indexes: Evidence of Heterogeneity or Measurement Error?” Meeting of the Group of

Experts on Consumer Price Indexes, United Nations Economic Commission for Europe Geneva, Switzerland, May 7-9. URL: https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.22/2018/United_States.pdf.

Diewert, W. Erwin (1976). “Exact and superlative index numbers”. In: *Journal of Econometrics* 4.2, pp. 115–145. DOI: 10.1016/0304-4076(76)90009-9.

— (1978). “Superlative index numbers and consistency in aggregation”. In: *Econometrica* 46.4, pp. 883–900. DOI: 10.2307/1909755.

— (2005). “Weighted country product dummy variable regressions and index number formulae”. In: *Review of Income and Wealth* 51.4, pp. 561–570. DOI: 10.1111/j.1475-4991.2005.00168.x.

Diewert, W. Erwin and C. Shimizu (2021). “The Treatment of Durable Goods and Housing”. In: *Consumer Price Index Theory (Draft)*. URL: https://econ2017.sites.olt.ubc.ca/files/2021/05/pdf_paper_diewert-erwin_IMFCPIChapter10.pdf.

Garner, Thesia, David Johnson, and Mary Kokoski (1996). “An experimental consumer price index for the poor”. In: *Monthly Labor Review* 119, p. 32. URL: <https://www.bls.gov/opub/mlr/1996/09/art5full.pdf>.

Garner, Thesia I. et al. (2022). “Distribution of U.S. Personal Consumption Expenditures for 2019: A Prototype Based on Consumer Expenditure Survey Data”. BLS Working Paper forthcoming.

Hill, Robert J. and Michael Scholz (2018). “Can geospatial data improve house price indexes? A hedonic imputation approach with splines”. In: *Review of Income and Wealth* 64.4, pp. 737–756. DOI: 10.1111/roiw.12303.

- Hobijn, Bart and David Lagakos (2005). “Inflation inequality in the United States”. In: *Review of Income and Wealth* 51.4, pp. 581–606. DOI: 10.1111/j.1475-4991.2005.00170.x.
- Hobijn, Bart, Kristin Mayer, et al. (2009). “Household inflation experiences in the US: a comprehensive approach”. Federal Reserve Bank of San Francisco Working Paper 2009-19. URL: <http://www.frbsf.org/publications/economics/papers/2009/wp09-19bk.pdf>.
- Houthakker, Hendrik S. (1957). “An international comparison of household expenditure patterns, commemorating the centenary of Engel’s law”. In: *Econometrica*, pp. 532–551.
- ILO (2004). *Consumer Price Index Manual: Theory and Practice*. Ed. by Peter Hill. Jointly published by ILO, IMF, OECD, UNECE, Eurostat, and the World Bank. URL: https://www.ilo.org/global/statistics-and-databases/WCMS_331153/lang--en/index.htm.
- Jaravel, Xavier (2018). “The Unequal Gains from Product Innovations: Evidence from the U.S. Retail Sector”. In: *The Quarterly Journal of Economics* 134.2, pp. 715–783. DOI: 10.1093/qje/qjy031.
- (2021). “Inflation Inequality: Measurement, Causes, and Policy Implications”. In: *Annual Review of Economics* 13. DOI: 10.1146/annurev-economics-091520-082042.
- Kaplan, Greg and Sam Schulhofer-Wohl (2017). “Inflation at the household level”. In: *Journal of Monetary Economics* 91, pp. 19–38. DOI: 10.1016/j.jmoneco.2017.08.002.
- Klick, Joshua and Anya Stockburger (2021). “Experimental CPI for lower and higher income households”. BLS Working Paper 537. URL: <https://www.bls.gov/osmr/research-papers/2021/pdf/ec210030.pdf>.

- Kokoski, Mary (2000). “Alternative CPI aggregations: two approaches”. In: *Monthly Labor Review* 123, p. 31. URL: <https://www.bls.gov/osmr/research-papers/2003/pdf/ec030070.pdf>.
- Konüs, Alexander A. (1924). “The problem of the true index of the cost of living”. In: translated in *Econometrica* (1939) 7, pp. 10-29.
- Ley, Eduardo (2005). “Whose inflation? A characterization of the CPI plutocratic gap”. In: *Oxford Economic Papers* 57.4, pp. 634–646.
- Moulton, Brent R. (2018). “The Measurement of Output, Prices, and Productivity: What’s Changed Since the Boskin Commission?” Brookings Institution Hutchins Center Report. URL: <https://www.brookings.edu/wp-content/uploads/2018/07/Moulton-report-v2.pdf>.
- National Research Council (2002). *At What Price?: Conceptualizing and Measuring Cost-of-Living and Price Indexes*. Ed. by Charles Schultze and Christopher Mackie. National Academies Press.
- Pollak, Robert A. (1989). *The Theory of the Cost-of-Living Index*. Oxford University Press on Demand.
- Prais, Sigbert J (1959). “Whose cost of living?” In: *the Review of Economic Studies* 26.2, pp. 126–134. DOI: 10.2307/2296170.
- Rambaldi, Alicia N. and Fletcher (2012). “The Effect of Alternative Estimators of the Hedonic Model on the Revisions of Hedonic Imputed Residential Property Price Indexes”. In: *International Association of Research on Income and Wealth 2012 General Conference, Boston, MA, August*, pp. 5–11. URL: <http://old.iariw.org/papers/2012/RambaldiPaper.pdf>.

Rambaldi, Alicia N. and D.S. Prasada Rao (2013). “Econometric Modeling and Estimation of Theoretically Consistent Housing Price Indexes”. Working Paper WP04/2013. URL: <https://economics.uq.edu.au/files/5157/WP042013.pdf>.

Webber, Dominic and Richard Tonkin (2013). *Statistical matching of EU-SILC and the Household Budget Survey to compare poverty estimates using income, expenditures and material deprivation*. Eurostat-Methodologies and Working Papers, Luxembourg: Publications Office. URL: <https://ec.europa.eu/eurostat/documents/3888793/5857145/KS-RA-13-007-EN.PDF/37d4ffcc-e9fc-42bc-8d4f-fc89c65ff6b1>.

Tables

Table 1: Shares of CPI-Eligible Expenditure by Source, 2019

Major Group	Diary	Interview	Total
Food and Bev.	0.1441	0.0026	0.1468
Housing	0.0233	0.3920	0.4153
Apparel	0.0251	0.0044	0.0295
Transportation	0.0010	0.1666	0.1676
Medical Care	0.0045	0.0831	0.0877
Recreation	0.0220	0.0337	0.0558
Educ. and Comm.	0.0027	0.0659	0.0685
Other	0.0099	0.0190	0.0289
Total	0.2327	0.7673	1.0000

Note: the CPI definition of “Housing” expenditure includes owner equivalent rent.

Table 2: Consumer Unit Indexes: Average 12 Month % Change, December 2002 - June 2021

	P. Lowe	D. Lowe	P. Torn.	D. Torn.	Lowe Gap	Torn. Gap
2003	2.07	2.26	1.81	2.48	−0.19	−0.67
2004	2.48	2.65	2.28	2.73	−0.17	−0.45
2005	3.29	3.36	2.94	3.26	−0.08	−0.33
2006	3.18	3.27	2.89	3.23	−0.09	−0.34
2007	2.77	2.95	2.44	2.97	−0.18	−0.53
2008	3.80	4.13	3.80	4.41	−0.33	−0.61
2009	−0.39	−0.26	−0.36	−0.44	−0.13	0.08
2010	1.68	1.75	1.59	1.49	−0.07	0.10
2011	3.16	3.25	2.99	3.14	−0.09	−0.15
2012	2.05	2.10	1.94	1.95	−0.05	−0.01
2013	1.45	1.47	1.25	1.30	−0.02	−0.05
2014	1.62	1.68	1.56	1.66	−0.06	−0.09
2015	0.16	0.08	0.01	−0.22	0.08	0.23
2016	1.26	1.20	0.95	0.97	0.06	−0.02
2017	2.16	2.22	1.67	1.98	−0.06	−0.31
2018	2.43	2.45	1.85	2.13	−0.02	−0.28
2019	1.80	1.74	1.48	1.54	0.06	−0.06
2020	1.23	1.28	1.01	1.10	−0.04	−0.09
2021	3.22	3.27	3.20	2.96	−0.05	0.24
Average	2.04	2.12	1.82	2.01	−0.08	−0.19

Note: 2021 averages cover January-June.

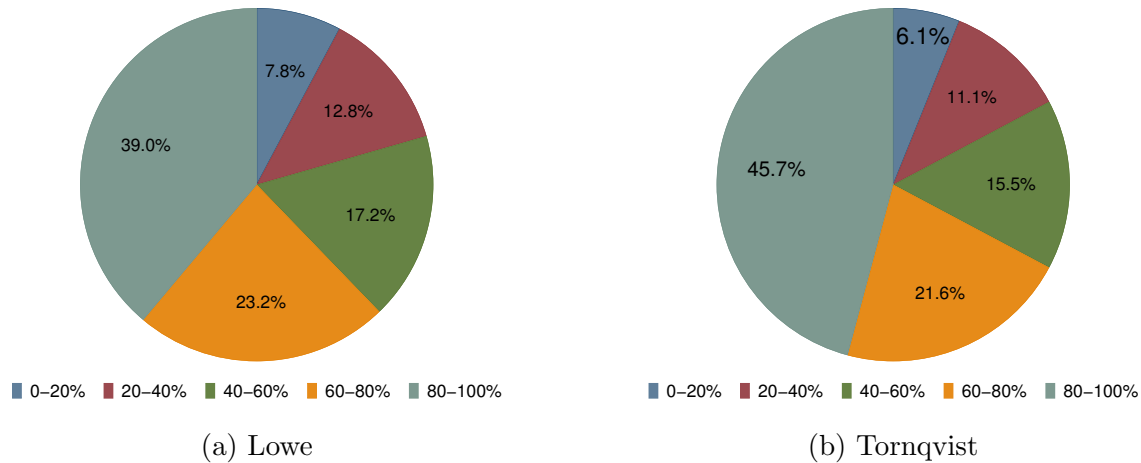
Table 3: Gaps by Household Definition, December 2002 - June 2021 (Average 12 m. % ch.)

	Lowe Plutocratic Gaps				Tornqvist Plutocratic Gaps			
	P. Lowe	CU	QAE	QAI	P. Torn	CU	QAE	QAI
2003	2.07	-0.19	-0.19	-0.03	1.81	-0.67	-0.20	-0.03
2004	2.48	-0.17	-0.17	-0.03	2.28	-0.45	-0.06	-0.05
2005	3.29	-0.08	-0.09	-0.03	2.94	-0.33	-0.07	-0.08
2006	3.18	-0.09	-0.09	-0.01	2.89	-0.34	-0.05	-0.04
2007	2.77	-0.18	-0.17	-0.06	2.44	-0.53	-0.18	-0.12
2008	3.80	-0.33	-0.31	-0.11	3.80	-0.61	-0.24	-0.12
2009	-0.39	-0.13	-0.10	-0.06	-0.36	0.08	0.00	-0.05
2010	1.68	-0.07	-0.06	-0.04	1.59	0.10	0.02	-0.03
2011	3.16	-0.09	-0.09	-0.03	2.99	-0.15	-0.04	-0.07
2012	2.05	-0.05	-0.04	-0.03	1.94	-0.01	0.01	-0.04
2013	1.45	-0.02	-0.02	-0.02	1.25	-0.05	-0.02	-0.02
2014	1.62	-0.06	-0.07	-0.05	1.56	-0.09	-0.06	-0.05
2015	0.16	0.08	0.07	-0.02	0.01	0.23	0.01	-0.02
2016	1.26	0.06	0.05	0.00	0.95	-0.02	0.03	0.02
2017	2.16	-0.06	-0.07	-0.04	1.67	-0.31	-0.09	-0.05
2018	2.43	-0.02	-0.02	-0.02	1.85	-0.28	0.07	-0.05
2019	1.80	0.06	0.05	0.00	1.48	-0.06	0.08	-0.01
2020	1.23	-0.04	-0.04	-0.07	1.01	-0.09	-0.08	-0.01
2021	3.22	-0.05	-0.06	-0.03	3.20	0.24	0.18	0.01
Average	2.04	-0.08	-0.07	-0.04	1.82	-0.19	-0.04	-0.05

Note: 2021 averages cover January-June. CU means the base unit for the democratic index is the consumer unit. QAE means the base unit is the quintile of annual expenditure. QAI means the base unit is the quintile of annual before-tax income.

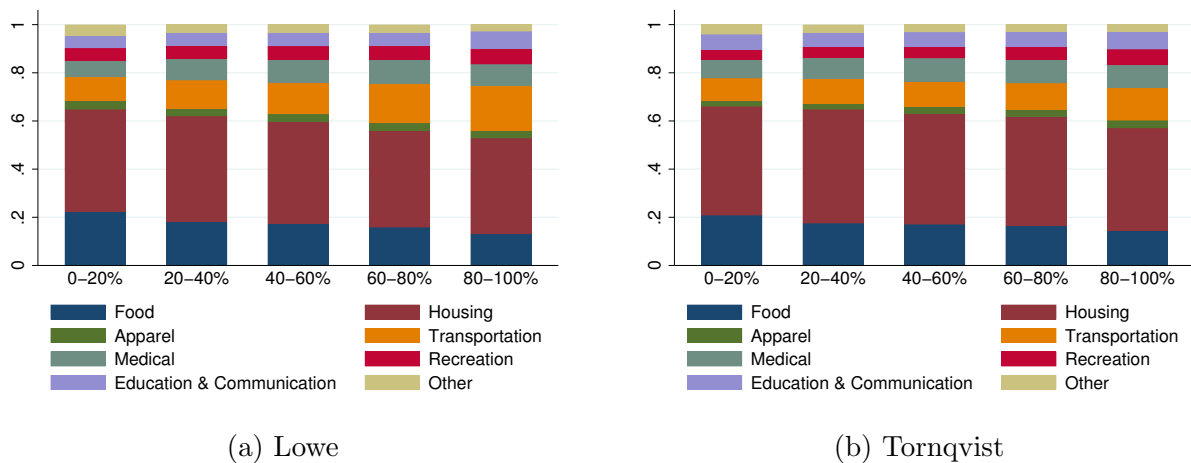
Figures

Figure 1: Consumer Unit Indexes: Reference Period Expenditure Share by Quintile



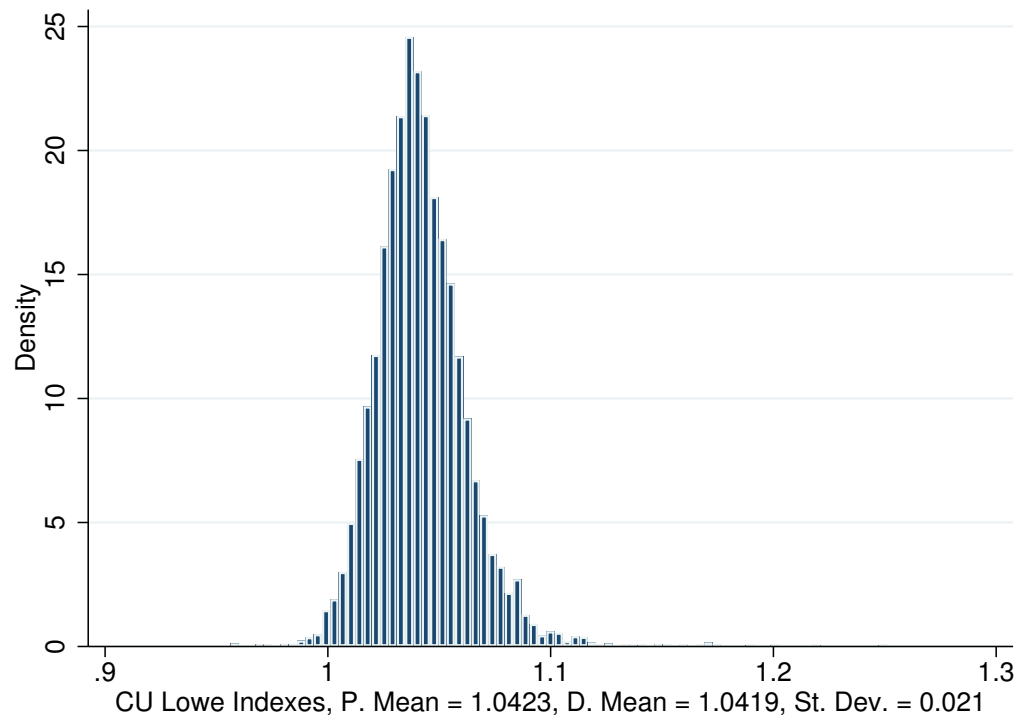
Note: Figures depict shares of total reference period expenditures from December 2001 to June 2021 by quintiles of the indexes respective reference period expenditure distributions.

Figure 2: Consumer Unit Indexes: Average Expenditure Weights by Quintile (2018-19)



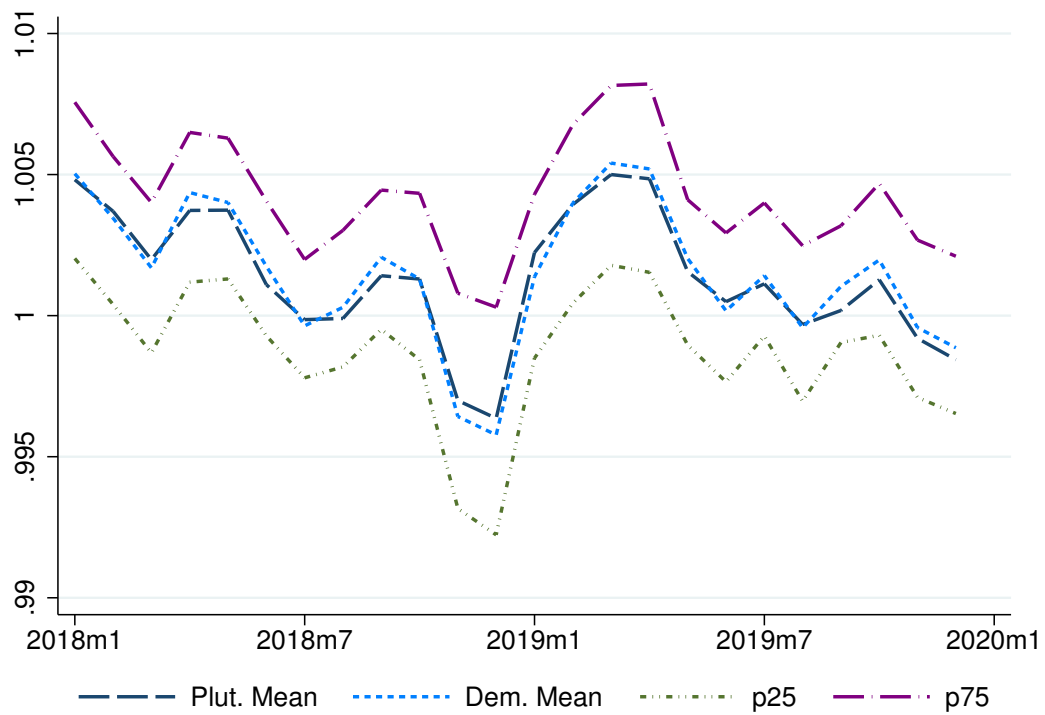
Notes: Bars represent equally-weighted (democratic) average weights of consumer units within the quintiles are of reference period expenditure. For Lowe, the weight is the annual expenditure share (price updated to December 2017) for CU's in the 2015-2016 reference period, corresponding to the 2018-19 index period. For Tornqvist, the monthly weight is itself the average of the consumer unit's current month share and prior month share. Average Tornqvist weights are pooled over January 2018 to December 2019.

Figure 3: Consumer Unit Lowe Distribution (Inflation for Dec. 2017 to Dec. 2019)



Note: The sample is consumer units with four quarters of expenditure during 2016-17 reference period.

Figure 4: Monthly Consumer Unit Tornqvist Distributions (previous month = 1.0)



Note: Plots are of means and percentiles of the distributions of household log-Tornqvist indexes by month.

Figure 5: Consumer Unit Based Plutocratic and Democratic Price Indexes (Dec. 2001 = 1.0)

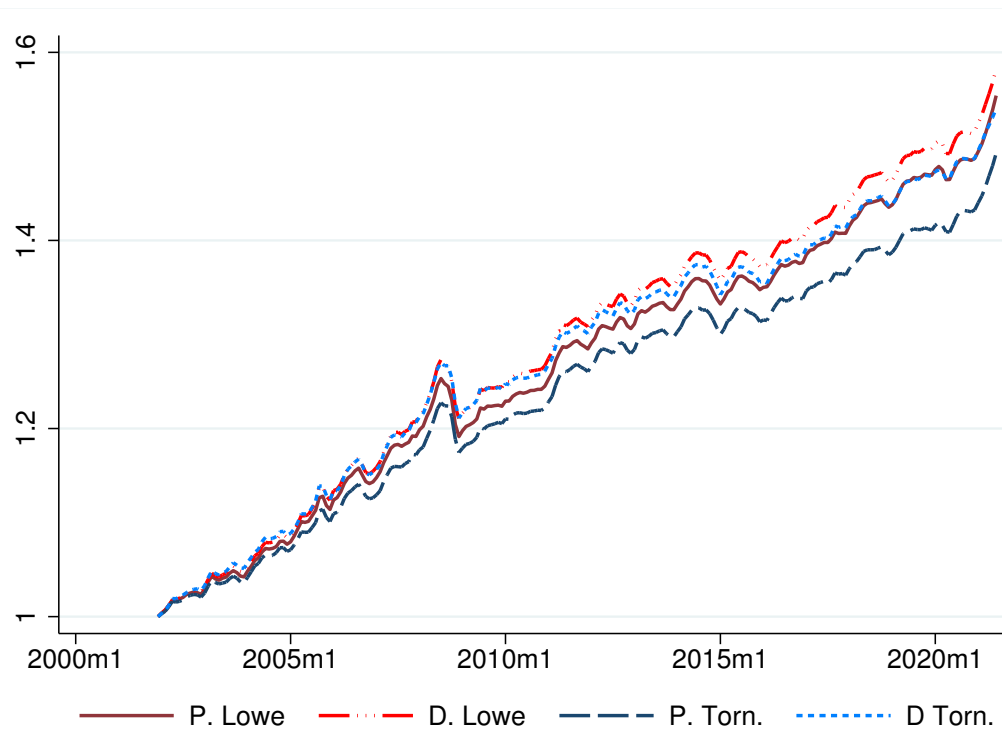
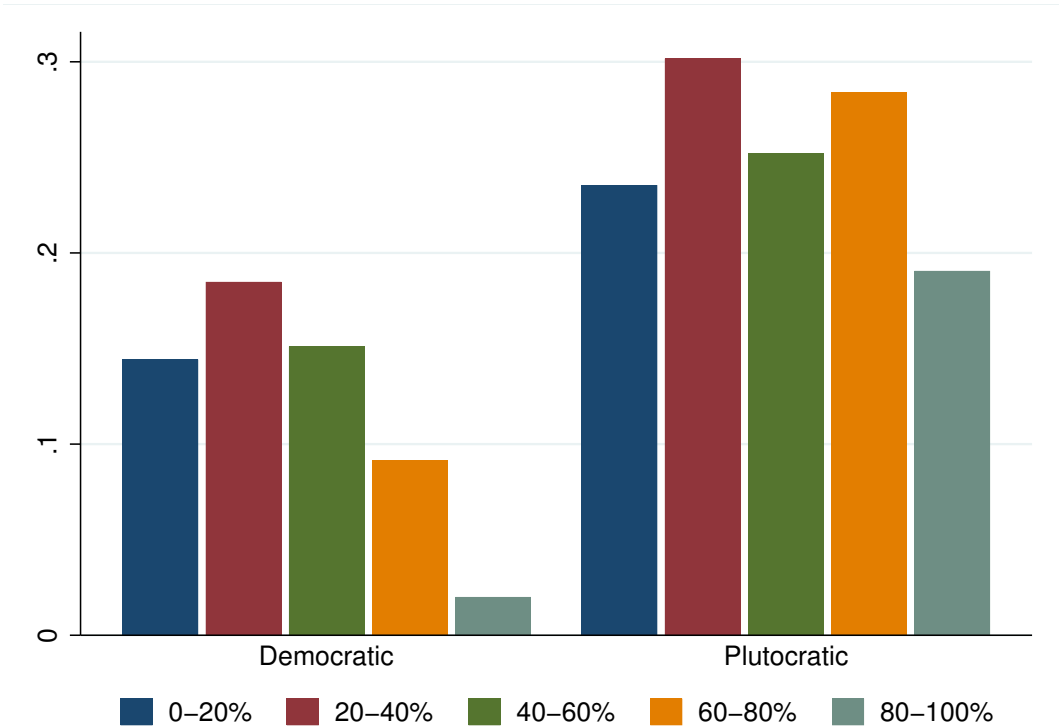


Figure 6: Consumer Unit Indexes: Average Substitution Effects (% point differences)



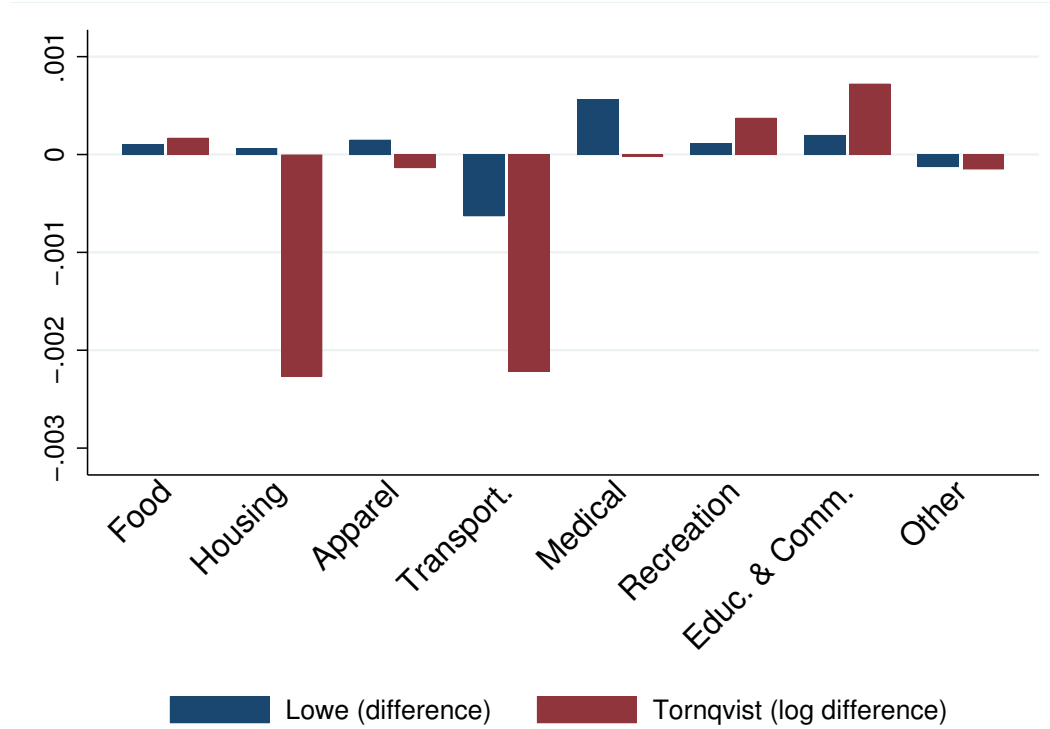
Notes: Bars represent differences between average 12 month percent changes of plutocratic and democratic Lowe and Tornqvist indexes from December 2002 to June 2021. For each index and separately by weight reference period, consumer units are first sorted into quintiles of annual expenditure, using only observations that completed four interviews.

Figure 7: Consumer Unit Indexes: Plutocratic Gaps in 12 month % Changes



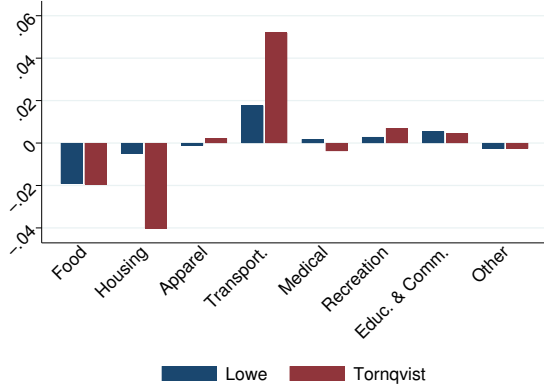
Notes: Plots depict differences between 12 month % changes for comparable plutocratic and democratic indexes based on consumer units as households. Lowe (annual CU weights) and Tornqvist (monthly CU weights) correspond to the indexes presented in Figure 5 and Table 2. Lowe (quarterly CU weights) treats CU-quarter combinations as unique households, while Lowe (monthly CU weights) treats CU-month combinations as unique households.

Figure 8: Consumer Unit Indexes: Contributions to Plutocratic Gaps (2018-19)

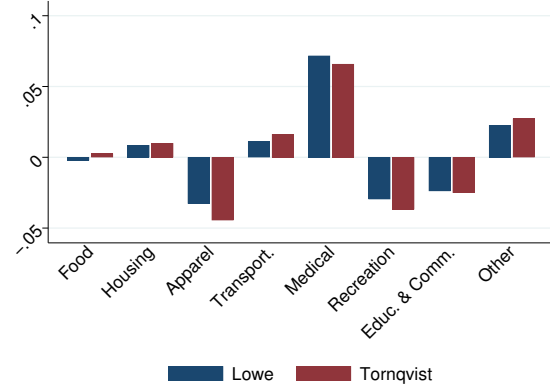


Note: For each expenditure category $\mathcal{G}$, the Lowe bars represent $\sum_{i \in \mathcal{G}} (s_{i\{v,b\}} - \bar{s}_{i\{v,b\}})(r_i - \bar{r})$ where v is 2015-2016, $r_i = p_{i,2019m12}/p_{i,2017m12}$ and $\bar{r} = N^{-1} \sum_{i=1}^N r_i$. The Tornqvist bars represent $\sum_{t=2018m1}^{2019m12} \sum_{i \in \mathcal{G}} (w_{it} - \bar{w}_{it}) (\ln r_{it} - \overline{\ln r_{it}})$, where $r_{it} = p_{it}/p_{i,t-1}$ and $\overline{\ln r_{it}} = N^{-1} \sum_{i=1}^N \ln r_{it}$.

Figure 9: Consumer Unit Indexes: Factors of Plutocratic Gaps (2018-19)



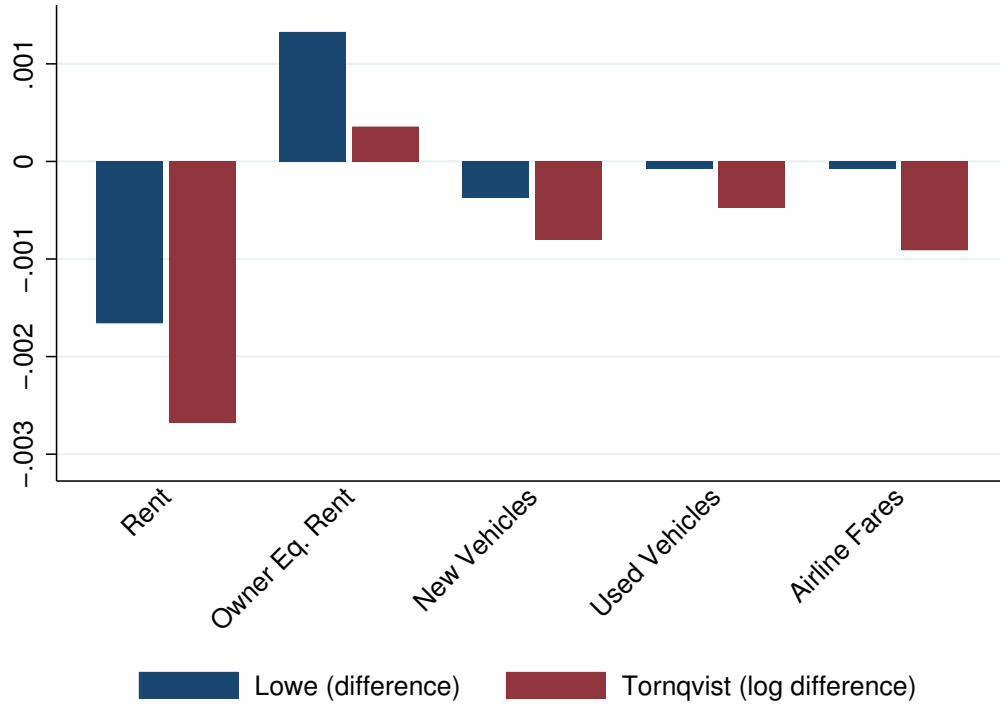
(a) Weight Differences



(b) Inflation Relative to Mean

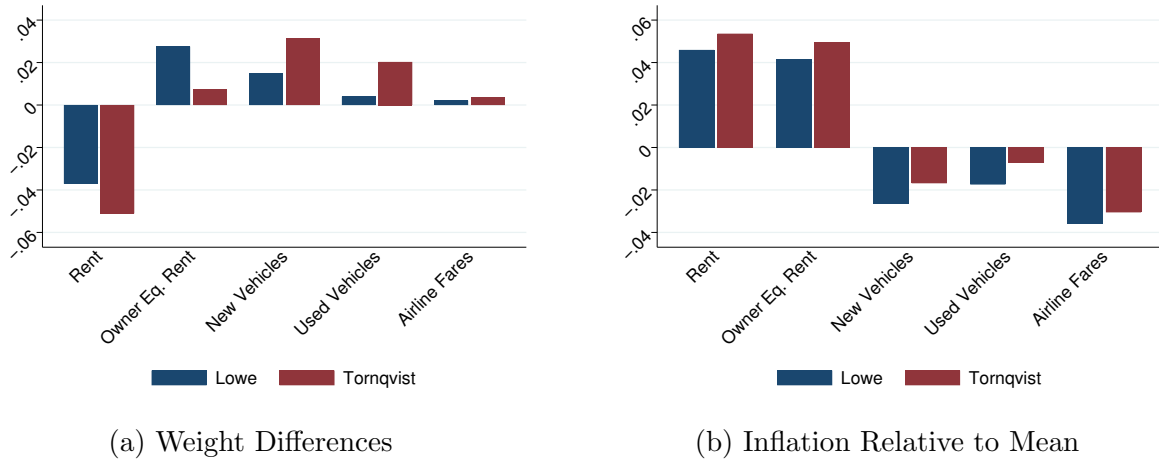
Notes: For each expenditure category $\mathcal{G}$, the Lowe weight differences are given by $\sum_{i \in \mathcal{G}} (s_{i\{v,b\}} - \bar{s}_{i\{v,b\}})$, while the inflation relative to the mean are given by $N_g^{-1} \sum_{i \in \mathcal{G}} (r_i - \bar{r})$ where N_g is the number of elementary cells in category g , b is 2015-2016, v is December 2017, $r_i = p_{i,2019m12}/p_{i,2017m12}$ and $\bar{r} = N^{-1} \sum_{i=1}^N r_i$. The Tornqvist weight differences are given by $24^{-1} \sum_{t=2018m1}^{2019m12} \sum_{i \in \mathcal{G}} (w_{it} - \bar{w}_{it})$, while the inflation relative to the mean are $\sum_{t=2018m1}^{2019m12} N_g^{-1} \sum_{i \in \mathcal{G}} (\ln r_{it} - \overline{\ln r_{it}})$ where $r_{it} = p_{it}/p_{i,t-1}$ and $\overline{\ln r_{it}} = N^{-1} \sum_{i=1}^N \ln r_{it}$.

Figure 10: Consumer Unit Indexes: Contributions to Plutocratic Gaps by Select Items (2018-19)



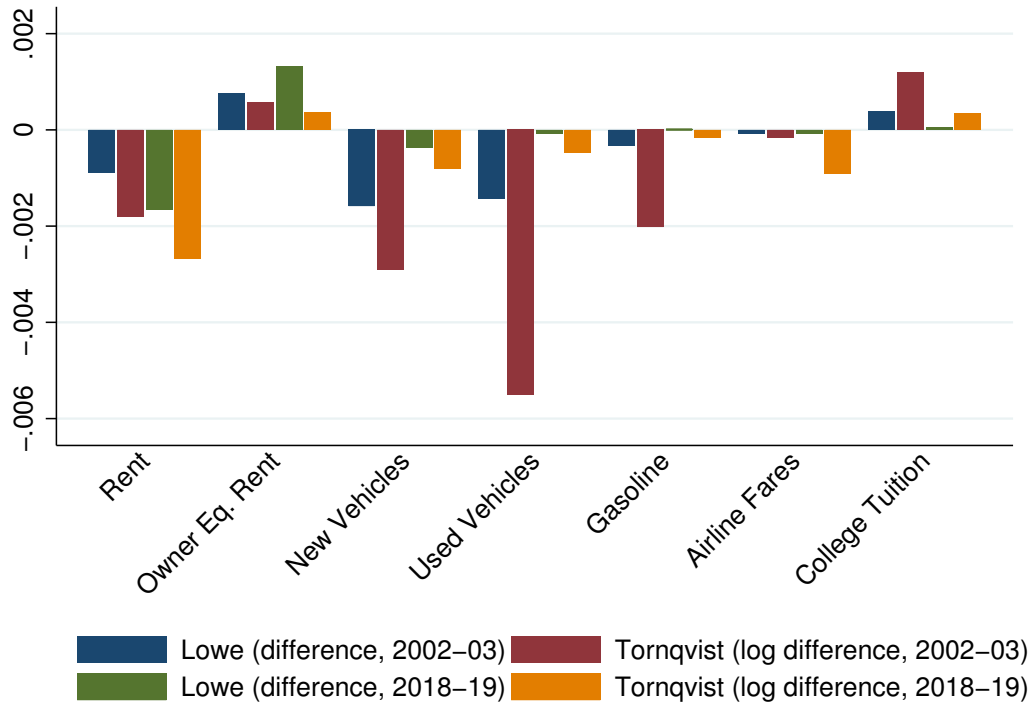
Note: For each expenditure category $\mathcal{G}$, the Lowe bars represent $\sum_{i \in \mathcal{G}} (s_{i\{v,b\}} - \bar{s}_{i\{v,b\}})(r_i - \bar{r})$ where v is 2015-2016, $r_i = p_{i,2019m12}/p_{i,2017m12}$ and $\bar{r} = N^{-1} \sum_{i=1}^N r_i$. The Tornqvist bars represent $\sum_{t=2018m1}^{2019m12} \sum_{i \in \mathcal{G}} (w_{it} - \bar{w}_{it}) (\ln r_{it} - \bar{\ln r}_{it})$, where $r_{it} = p_{it}/p_{i,t-1}$ and $\bar{\ln r}_{it} = N^{-1} \sum_{i=1}^N \ln r_{it}$.

Figure 11: Consumer Unit Indexes: Factors of Plutocratic Gaps by Select Items (2018-19)



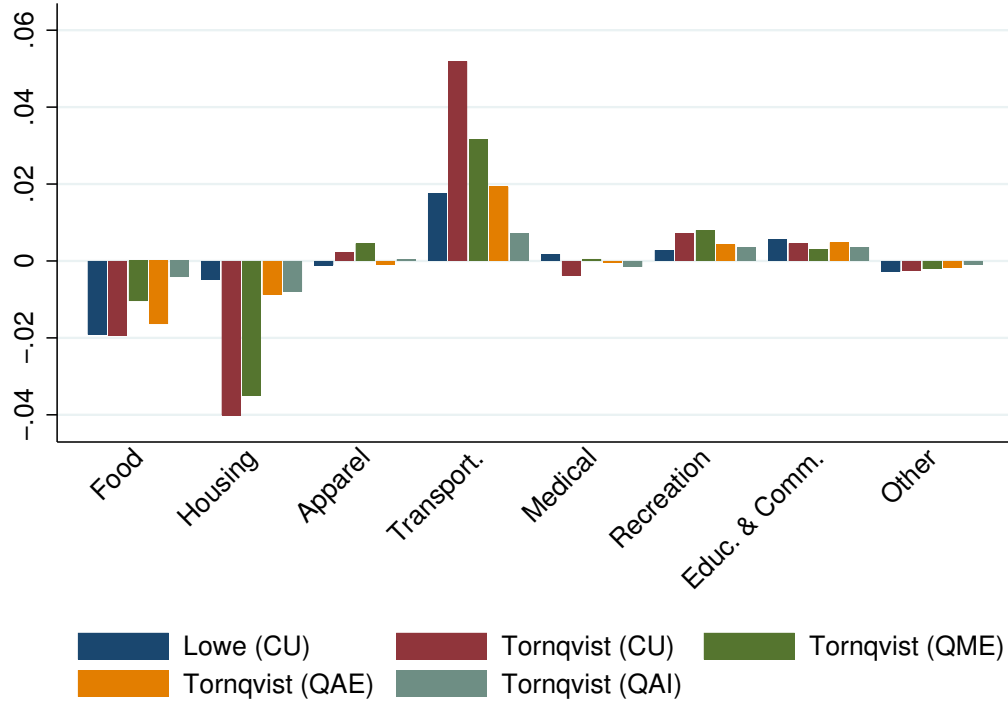
Notes: For each item-stratum $\mathcal{G}$, the Lowe weight differences are given by $\sum_{i \in \mathcal{G}} (s_{i\{v,b\}} - \bar{s}_{i\{v,b\}})$, while the inflation relative to the mean are given by $32^{-1} \sum_{i \in \mathcal{G}} (r_i - \bar{r})$ where 32 is the number of index areas, b is 2015-2016, v is December 2017, $r_i = p_{i,2019m12}/p_{i,2017m12}$ and $\bar{r} = N^{-1} \sum_{i=1}^N r_i$. The Tornqvist weight differences are given by $24^{-1} \sum_{t=2018m1}^{2019m12} \sum_{i \in \mathcal{G}} (w_{it} - \bar{w}_{it})$, while the inflation relative to the mean are $\sum_{t=2018m1}^{2019m12} 24^{-1} \sum_{i \in \mathcal{G}} (\ln r_{it} - \overline{\ln r_{it}})$ where $r_{it} = p_{it}/p_{i,t-1}$ and $\overline{\ln r_{it}} = N^{-1} \sum_{i=1}^N \ln r_{it}$.

Figure 12: Consumer Unit Indexes: Contributions to Plutocratic Gaps by Select Items



Note: For each expenditure category $\mathcal{G}$, the Lowe bars represent $\sum_{i \in \mathcal{G}} (s_{i\{v,b\}} - \bar{s}_{i\{v,b\}})(r_i - \bar{r})$ where v is 2015-2016, $r_i = p_{i,2019m12}/p_{i,2017m12}$ and $\bar{r} = N^{-1} \sum_{i=1}^N r_i$. The Tornqvist bars represent $\sum_{t=2018m1}^{2019m12} \sum_{i \in \mathcal{G}} (w_{it} - \bar{w}_{it}) (\ln r_{it} - \overline{\ln r_{it}})$, where $r_{it} = p_{it}/p_{i,t-1}$ and $\overline{\ln r_{it}} = N^{-1} \sum_{i=1}^N \ln r_{it}$.

Figure 13: Plutocratic Gap: Weight Difference Comparisons (2018-19)



Notes: Notes: For each expenditure category $\mathcal{G}$, the Lowe weight differences are given by $\sum_{i \in \mathcal{G}} (s_{i\{v,b\}} - \bar{s}_{i\{v,b\}})$, b is the 2015-16 biennial reference period, and v is December 2017. The Tornqvist weight differences are given by $24^{-1} \sum_{t=2018m1}^{2019m12} \sum_{i \in \mathcal{G}} (w_{it} - \bar{w}_{it})$. CU means the consumer unit was the base unit for the democratic index. QME means the quintile of current month expenditure is the base unit, while QAE means the quintile of annual expenditure is the base unit. QAI means the quintile of annual before-tax income is the base unit.

A Appendix

A.1 Tornqvist Plutocratic Gap Components

This section applies Ley's (2005) decomposition of the plutocratic gap to Tornqvist indexes in natural logs. The result follows directly from Ley's decomposition of the difference between plutocratic and democratic shares.

In terms of prices and expenditure shares, the gap between the plutocratic and the (geometric) democratic Tornqvist is written:

$$\begin{aligned}\ln P_{PT} - \ln P_{DT} &= \sum_{i=1}^N (w_{it} - \bar{w}_{it}) \ln \left(\frac{p_{it}}{p_{i,t-1}} \right) \\ &= \sum_{i=1}^N \frac{1}{2} [(s_{i,t-1} - \bar{s}_{i,t-1}) + (s_{it} - \bar{s}_{it})] \ln \left(\frac{p_{it}}{p_{i,t-1}} \right)\end{aligned}\quad (17)$$

where $s_{it} = \sum_{h=1}^H S_{th} s_{ith}$ and $\bar{s}_{it} = H^{-1} \sum_{h=1}^H s_{ith}$ are the plutocratic and democratic expenditure shares, respectively.

From Ley (2005), we have:

$$s_{it} - \bar{s}_{it} = \hat{\zeta}_t \hat{\beta}_{it}, \quad i = 1, \dots, N. \quad (18)$$

where $\hat{\zeta}_t$ is the square of the coefficient of variation of household expenditure on all items, and $\hat{\beta}_{it}$ is the OLS estimator of the parameter in the following regression.

$$(s_{ith} - \bar{s}_{it}) = \beta_{it} \left(\frac{e_{th} - \bar{e}_t}{\bar{e}_t} \right), \quad (19)$$

where $e_{th} = \sum_{i=1}^N p_{it} q_{ith}$ and $\bar{e}_t = H^{-1} \sum_{h=1}^H e_{th}$. Ley discusses how $\hat{\beta}_{it}$ relates to the expenditure elasticity of good i with respect to total expenditure.

Analogous to Ley (2005) Eq. 7, we then have

$$\ln P_{PT} - \ln P_{DT} = \frac{1}{2} \left[\hat{\zeta}_{t-1} N \widehat{Cov} \left(\hat{\beta}_{i,t-1}, \ln(p_{it}/p_{i,t-1}) \right) + \hat{\zeta}_t N \widehat{Cov} \left(\hat{\beta}_{it}, \ln(p_{it}/p_{i,t-1}) \right) \right], \quad (20)$$

where $\widehat{Cov} \left(\hat{\beta}_{it}, \ln(p_{it}/p_{i,t-1}) \right)$ is the sample covariance across i between $\hat{\beta}_{it}$ and $\ln(p_{it}/p_{i,t-1})$. From this, we can see Ley's requirements for the plutocratic gap: variation in household expenditures (nonzero $\hat{\zeta}_t$ and $\hat{\zeta}_{t-1}$), expenditure pattern variation across expenditure levels (nonzero $\hat{\beta}_{it}$'s and $\hat{\beta}_{i,t-1}$'s), and expenditure pattern variation that is systematically related to item-level price changes, that is nonzero $\widehat{Cov} \left(\hat{\beta}_{i,t-1}, \ln(p_{it}/p_{i,t-1}) \right)$ and $\widehat{Cov} \left(\hat{\beta}_{it}, \ln(p_{it}/p_{i,t-1}) \right)$.

A.2 Details on Matching Procedure

As mentioned, I use a statistical matching method, based on Hobijn, Mayer, et al. (2009), through which similar donors from the Diary provide the missing data for each Interview consumer unit. Similarity is determined by a model of average monthly expenditure as a function of demographic characteristics. The method implicitly places higher weight on characteristics that more strongly predict expenditures. I used a measure based on average monthly expenditure because I found it to be more highly correlated with consumer unit characteristics than raw monthly expenditures. While borrowing much from Hobijn, Mayer, et al. (2009), my approach differs in one key respect. I use the natural log of average monthly expenditure as the dependent variable for the regression model, while Hobijn, et. al. compute a price-related expenditure change variable using an annual Laspeyres index. I also tried a modified version of this procedure using an average expenditure change variable based on a 12 month modified Paasche index. I found both approaches produced very similar democratic indexes. Some features of the joint distribution between Diary expenditures and demographic characteristics were slightly better replicated in the imputed dataset when using

predicted average monthly expenditures in the distance metric. Further details are later in this section.

Denote the set of months that consumer unit h is in the sample as $S(h)$, and the number of months as T_h . Average monthly expenditure for h is defined $\bar{e}_h = T_h^{-1} \sum_{t \in S(h)} \sum_{i=1}^N e_{ith}$. I construct $\bar{e}_h$ for Interview observations using all available CPI-eligible expenditures available in the Interview. These include categories that CPI sources from the Interview and categories available in both surveys, but which the CPI sources from the Diary, such as the more aggregated food and beverage categories. Table A1 describes potential alternative source selection for Diary expenditures in greater detail. Detailed CE categories are known as Universal Classification Codes (UCC). Of the 246 UCC's sourced from the Diary in 2019, 63 were also available in the Interview. Coverage for an additional 137 food and beverage UCC's was also available, but only as a handful of coarse aggregates. In total, 83.5% of Diary-sourced expenditures belong to categories for which some Interview information is available. The actual expenditure estimates from the Interview for these categories are slightly less—totaling about 80% of Diary expenditures in 2019 (See Appendix Table A2). Referring to Table A1, $\bar{e}_h$ covers everything but the Diary-only UCC's, which account for only 4% of expenditures in 2019.

I estimate a linear regression for each month t and within a subpopulation s (e.g., all urban consumers), where s is the subpopulation of interest for the final index (e.g., all urban consumers, low-income consumers, etc.), I estimate the following linear regression over the Interview observations.²⁰

$$\ln(\bar{e}_h) \equiv y_h = \mathbf{x}_{th}\boldsymbol{\beta}_{st} + u_{ht}. \quad (21)$$

The vector $\mathbf{x}_{th}$ includes variables like Census region, urban/rural, age, race, sex, and education of the reference person, family size, and the prior year's annual income. These often do

²⁰Alternatively, it might be attractive to estimate average expenditures as a function of characteristics use the Diary sample, as these are what need to be imputed for the Interviews. I find that characteristics explain little variation in weekly Diary expenditures for the sample of Diary consumer units, perhaps due to the short (week-long) recall periods.

not change between collection quarters, and I implicitly assume that attributes true for the collection quarter are also true for the associated reference months. I used the least absolute shrinkage and selection operator (LASSO) estimate each β . LASSO is a model selection procedure that introduces a penalty term into the least squares objective function. The basic idea is to reduce over-fitting by only including those variables that have the most predictive power. Instead of estimating 25 parameters for every subpopulation in every month, LASSO essentially picks those that are most predictive of average expenditure. For matching within the all urban population, for example, the average number of parameters chosen was about 8. I found that when compared to weighted least squares, observations paired based on LASSO distance measures tended to be more likely to have matching attributes, and the covariance between Diary expenditures and characteristics like income were better preserved in the imputed data.

Let $\hat{\beta}_{st}$ be the slope estimate for subpopulation s in period t . As consumer unit characteristics are available and comparably defined in both surveys,²¹ I calculate predicted values, $\hat{y}_{ht} = \mathbf{x}_h \hat{\beta}_{st}$, for each Diary and Interview. For a given Interview observation h and Diary observation k , the distance metric is defined as

$$\delta_t(h, k) = |\hat{y}_{ht} - \hat{y}_{kt}|. \quad (22)$$

Within each month and subpopulation, I calculate $\delta_t(h, k)$ for all h, k pairs. Then for each h , I randomly select one k from the twenty smallest $\delta_t(h, k)$. The random component is intended to ensure a more even distribution of matches across Diary observations. This process is repeated for each month the Interview observation is in the sample. Prior to index calculation, the Diary expenditures are scaled by 13/3 to represent one month. When matching within the urban population, variation in $\mathbf{x}_h$ explains, on average, about 58% of the variation in y_h . The average distance between matches (in natural logs) is 0.018, which

²¹As a check, I tested for equality of means across samples for each characteristic an collection quarter. Using a 5% test, I reject the null hypothesis of equality in about 10% of cases.

implies a difference in predicted expenditure between matched observations of about 1.8%. For reference, the average difference in predicted expenditure between two randomly matched observations was about 65%. Note, only the subpopulation s is explicitly conditioned on. Variables in $\mathbf{x}_h$ will contribute to the match only to the extent to which they predict average expenditures.

I also tried a closer replication of Hobijn, et. al.'s method using price-related expenditure change as the dependent variable of the model. My version is given by

$$\Delta \bar{e}_{ht} = \bar{e}_h [1 - (P_{P,h,t,t-12})^{-1}] \quad (23)$$

where

$$P_{P,h,t,t-12} = \left[\sum_{i=1}^N (\bar{e}_{ih}/\bar{e}_h) (p_{it}/p_{i,t-12})^{-1} \right]^{-1} \quad (24)$$

Using $\Delta \bar{e}_{ht}$ as the dependent variable, I found model R-squareds averaged only about 0.14 (matching within the urban population), which is lower than what Hobijn, Mayer, et al. (2009) report. With different dependent variables, the R-squares of the regressions of $\Delta \bar{e}_{ht}$ and $\bar{e}_h$ are not comparable, but lower R-squared does not necessarily imply a lower imputation quality, as predicted expenditures are being used for matching purposes only. If the model has no predictive power, then the procedure resembles more a random match between Diary and Interview observations. Figure A3 compare features of the joint distributions of the matched datasets which make the method based on $\bar{e}_h$ look slightly more attractive. Ultimately, however, Plutocratic Tornqvist indexes using predicted $\Delta \bar{e}_{ht}$ and $\bar{e}_h$ differ from each other and the published C-CPI-U cumulatively by only 0.3% over 20 years, and differences between corresponding democratic indexes are only 0.03%, so the choice does not appear to have much of an impact on the eventual index calculation.

Evaluating match quality is difficult because the monthly expenditures on Diary categories are simply missing for the Interview sample. Nevertheless, I include some subjective diagnostic exercises, following Hobijn, Mayer, et al. (2009) and Webber and Tonkin (2013).

The following tables and figures concern matches made within the urban population. To begin with, Table A3 compares the rate at which Interviews are matched with Diaries of the same attribute. If matching were completely random, then the proportion of correct matches would equal (in large samples) the proportion of Diary observations with that attribute. Distance-based matching does not guarantee certain attributes will be correctly matched, but we expect to do at least as well as random. For virtually all cases shown here, distance-based matching delivers a Diary of the same attribute at higher rates than if the match were completely random. For ten of the attributes shown here, including income quintiles, using $\ln(\bar{e}_h)$ delivers higher match rates. In contrast, using $\Delta\bar{e}_{ht}$ delivers higher match rates for nine of these attributes, including for Census Region.

Figures A1 and A2 check that the imputed Diary expenditures (or at least the totals) have distributions that line up with the donor data. Figure A1 plots average weekly expenditures of Diary observations by month against imputed values for each dependent variable used in the distance measure. Figure A2 compares deciles of the actual distribution of average weekly expenditures against deciles of the imputed distributions. Not surprisingly, because the matching methods use actual Diary expenditures, they are able to largely replicate the volatility and distribution of this series as well as its average levels. Figure A3 checks that features of the joint distributions between average weekly Diary expenditures and certain attributes are replicated in the imputed data. The bars represent average Diary spending for different attributes. Again, none of these attributes are explicitly conditioned on in the matching process, but it is reassuring that certain patterns, e.g., rising Diary expenditures by income quintile, are preserved in the imputed data. While matching based on predicted $\ln(\bar{e}_h)$ does not uniformly outperform matching based on predicted $\Delta\bar{e}_{ht}$ with respect to these joint distributions, it does so for Income. I consider particularly important given the factors contributing to Plutocratic gaps as described in Section 2.

A.3 Additional Analysis for Consumer Unit-Based Indexes

Comparisons with Published Indexes

The democratic and plutocratic indexes calculated for this paper are based on synthetic CE consumer unit observations. As discussed in Section 3, consumer units must be in at least their second month in the sample to calculate a Tornqvist index, and I restrict the sample for the Lowe indexes to those consumer units who completed four interviews within the biennial weight reference period. Section 3 also describes other processing differences between these indexes and the published CPIs, such as the censoring of negative expenditures. Table A4 compares the average twelve-month percent changes of the plutocratic Lowe and Tornqvist indexes against the published CPI-U and C-CPI-U. On average, the matching and processing differences make a small difference in the plutocratic aggregates. The plutocratic Lowe escalates, on average, 0.05 percentage points less per year than the CPI-U, and the plutocratic Tornqvist escalates, on average 0.02 percentage points per year less than the C-CPI-U. Differences in average twelve month percent changes between the published index and its research counterpart for a specific year are occasionally larger, but these differences are usually less than the estimated plutocratic gaps.

Choice of Mean for the Democratic Tornqvist

The democratic Tornqvist is defined in this paper (Eq. 4) as an equally weighted geometric mean of consumer unit-level Tornqvist indexes. Alternatively, one could define the democratic Tornqvist using an arithmetic mean. Using a Taylor expansion, one can show that the proportional difference between the two types of means depends on the dispersion in consumer unit-level inflation. Let P_{ADT} be an equally-weighted arithmetic average of consumer unit Tornqvist indexes. From a second-order Taylor expansion of $\ln P_{T,h}$ around the point

$P_{T,h} = \bar{P}_T$ where $\bar{P}_T \equiv H^{-1} \sum_{h=1}^H P_{T,h} = P_{ADT}$. We have

$$\ln P_{T,h} \approx \ln \bar{P}_T + \frac{1}{\bar{P}_T} (P_{T,h} - \bar{P}_T) - \frac{1}{(\bar{P}_T)^2} (P_{T,h} - \bar{P}_T)^2 \quad (25)$$

Taking the sample arithmetic mean across h , the first order term drops out. Re-arranging, we have

$$\ln P_{ADT} - \ln P_{DT} \approx \frac{1}{(\bar{P}_T)^2} H^{-1} \sum_{h=1}^H (P_{T,h} - \bar{P}_T)^2 \quad (26)$$

This shows the difference in logs is approximately equal to the square of the coefficient of variation of the consumer unit Tornqvist indexes, which is weakly positive. As the arithmetic mean is generally greater than the geometric mean, one might be concerned that the choice of mean has a significant impact on the plutocratic gap. Figure A5 implies this is not the case. With a geometric mean for the democratic Tornqvist, the average gap in twelve month percent changes is -0.22 percentage points per year from December 2000 to June 2021, versus -0.25 percentage points per year when using an arithmetic mean for the democratic Tornqvist.

Indexes over CE Interview Expenditures Only

I also calculate the plutocratic and democratic indexes using only Interview-sourced expenditures. This eliminates the need to match Interview and Diary consumer units, but it means that food spending is tracked using only a few broad aggregates. Figure A6 plots the index levels for the plutocratic and democratic indexes as in Figure 5. The average plutocratic gap is slightly larger for the Lowe index, -0.1 percentage points, versus -0.08 in the main results. The average gap is about the same for the Tornqvist index, -0.19 percentage points. While this yields similar gaps as the main results, this is because the gaps are computed from plutocratic and democratic indexes based on the same data source. The underlying plutocratic indexes turn out to be further from the published CPIs when using only Inter-

view data. Figure A7 shows, for example, that a plutocratic Tornqvist index estimated only from Interview data implies cumulative inflation of about 3.1 percentage points more than the C-CPI-U over the December 1999 to June 2021 period, a compound average difference of about 0.09 percentage points per year.

Equivalence Scales

Ley (2005) also considers democratic measures which account for household size in an equivalence scale approach. This is intended to capture potential economies of scale in consumption. The simple model adopted in that paper and here assumes economies of scale depend only on household size. Denote the number of consumer unit members as n_h . We then apply consumer unit-level weights $\omega_h(\theta) \propto n_h^\theta$, where θ is a parameter which governs the impact of family size on the weight. The democratic index with consumer unit weights $1/H$ corresponds to $\theta = 0$, while $\theta = 1$ uses weights proportional to the full n_h . The latter corresponds to weighting each individual equally rather than each consumer unit. I also consider $\theta = 0.5$, which is a common equivalence scale using the square root of family size. I compute democratic Tornqvist indexes with these levels of θ . Generally, we expect higher θ to push the democratic index closer to the plutocratic in terms of the monthly Tornqvist links, but the twelve month percent changes may not follow the same pattern. Figure A8 shows the twelve month gaps in the Tornqvist indexes for the different θ . Some short-term differences do appear across the values of θ , but the effect of θ is not consistent and tends not to change the plutocratic gap qualitatively. The cumulative inflation implied by the three democratic indexes over December 1999 to June 2021 differ by less than half of one percentage point.

Appendix Tables

Table A1: CPI Expenditure Sourcing and Availability, 2019

Description	# UCC	Share of UCC	Share of Expend.
Sourced from Interview	347	0.5852	0.7673
Sourced from Diary	246	0.4148	0.2327
Only in Diary	46	0.0776	0.0384
Also available in Interview	63	0.1062	0.0502
Aggregates available in Interview	137	0.2310	0.1441
Total	593	1.0000	1.0000

UCC = Universal Classification Code (identifier used by the CE). Diary expenditures used to compute expenditure shares in fourth column.

Table A2: CE Survey Expenditures on Items CPI Sources from Diary, 2019

Month	Diary (b. dollars)	Interview (b. dollars)	Ratio
January	127.5	105.1	0.8243
February	129.1	105.2	0.8149
March	138.9	109.5	0.7883
April	156.8	111.4	0.7105
May	156.5	113.7	0.7265
June	133.4	115.8	0.8681
July	139.3	114.8	0.8241
August	136.2	115.1	0.8451
September	146.1	112.8	0.7721
October	148.2	112.2	0.7571
November	132.5	115.7	0.8732
December	163.2	129.2	0.7917
Total	1707.7	1360.5	0.7967

Table A3: Match Rates for Selected Attributes by Match Method

Attribute	Diary Proportion	Match Rates		
		Using $\Delta\bar{e}_{ht}$	Using $\ln(\bar{e}_h)$	
<i>Age</i>				
< 61	0.7321	0.7528	0.7424	
≥ 61	0.2679	0.3246	0.2983	
<i>Presence of Children</i>				
Yes	0.4061	0.4638	0.4759	
No	0.5939	0.6381	0.6454	
<i>Education Level</i>				
Less than H.S.	0.1189	0.1656	0.1960	
H.S. & Some Col.	0.5511	0.6150	0.5950	
Bachelors+	0.3300	0.5049	0.4938	
<i>Census Region</i>				
Northeast	0.1971	0.2200	0.1950	
Midwest	0.2356	0.3225	0.2442	
South	0.3227	0.3721	0.3273	
West	0.2446	0.3944	0.2503	
<i>Income Quintile</i>				
0-20%	0.2308	0.5014	0.6200	
20-40%	0.1761	0.2959	0.3323	
40-60%	0.1830	0.2731	0.2877	
60-80%	0.1945	0.2981	0.3257	
80-100%	0.2156	0.4575	0.6150	
<i>Housing Tenure</i>				
Owner w/ mort.	0.4333	0.5572	0.6340	
Owner w/o mort.	0.2133	0.2473	0.2408	
Renter	0.3213	0.5302	0.5944	

Notes: Statistics calculated following matching within the urban population. The first column of data gives the proportion of Diary observations with a given attribute. The final two columns give the proportion of Interview observations matched with a Diary with the same attribute.

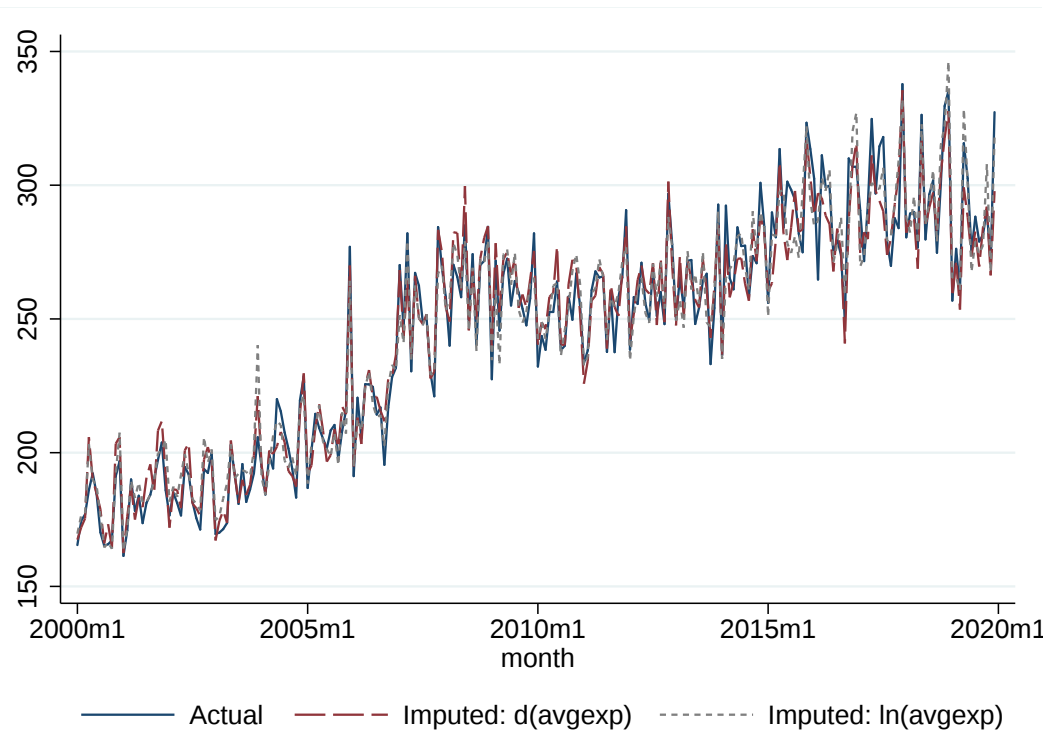
Table A4: Comparison of Published and Research Series Average 12 month Percent Changes

	CPI-U	P. Lowe	C-CPI-U	P. Torn.
2003	2.27	2.07	2.04	1.81
2004	2.68	2.48	2.50	2.28
2005	3.39	3.29	2.93	2.94
2006	3.24	3.18	2.91	2.89
2007	2.85	2.77	2.53	2.44
2008	3.85	3.80	3.74	3.80
2009	−0.34	−0.39	−0.45	−0.36
2010	1.64	1.68	1.43	1.59
2011	3.16	3.16	3.05	2.99
2012	2.07	2.05	1.95	1.94
2013	1.47	1.45	1.23	1.25
2014	1.62	1.62	1.45	1.56
2015	0.12	0.16	−0.12	0.01
2016	1.26	1.26	0.93	0.95
2017	2.13	2.16	1.77	1.67
2018	2.44	2.43	2.02	1.85
2019	1.81	1.80	1.45	1.48
2020	1.24	1.23	1.08	1.01
2021	4.69	4.54	3.35	3.20
Average	2.19	2.14	1.84	1.82

Note: Tornqvist indexes end in June 2021.

Appendix Figures

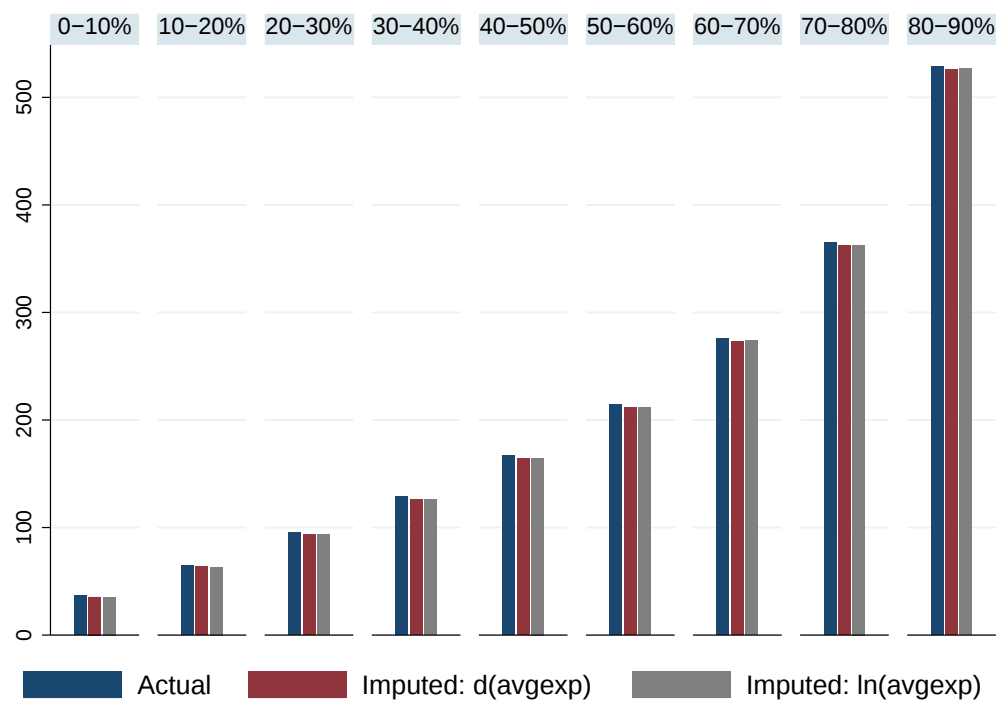
Figure A1: Actual and Imputed Average Weekly Diary Expenditures by Month



Notes:

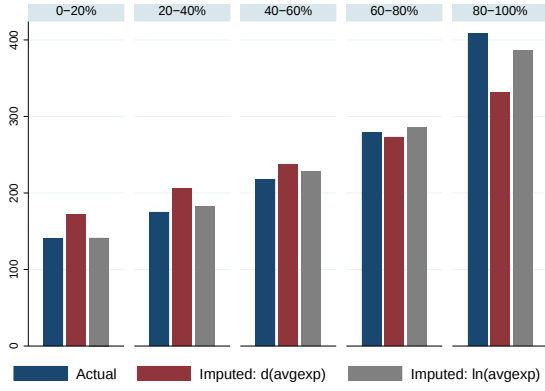
Computed following matching within the urban population. $d(avgexp)$ means $\Delta \bar{e}_{ht}$ is the dependent variable used in the distance measure. $\ln(avgexp)$ means $\ln(\bar{e}_h)$ is the dependent variable.

Figure A2: Deciles of Actual and Imputed Average Weekly Diary Expenditures

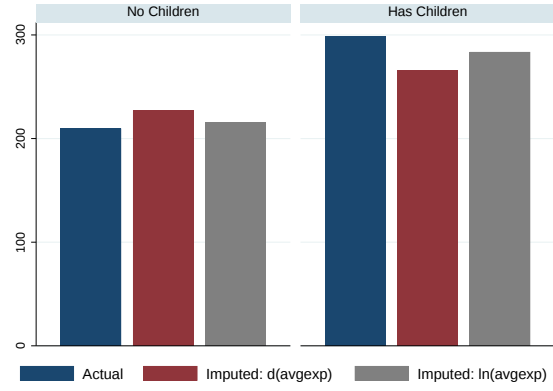


Notes: Computed following matching within the urban population. d(avgexp) means $\Delta \bar{e}_{ht}$ is the dependent variable used in the distance measure. ln(avgexp) means $\ln(\bar{e}_h)$ is the dependent variable.

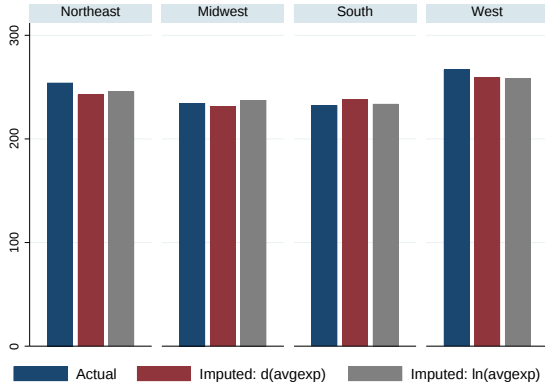
Figure A3: Average Weekly Diary-Category Expenditure By Attribute



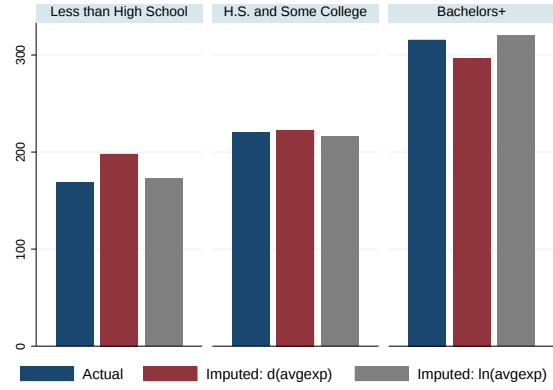
(a) Income Quintiles



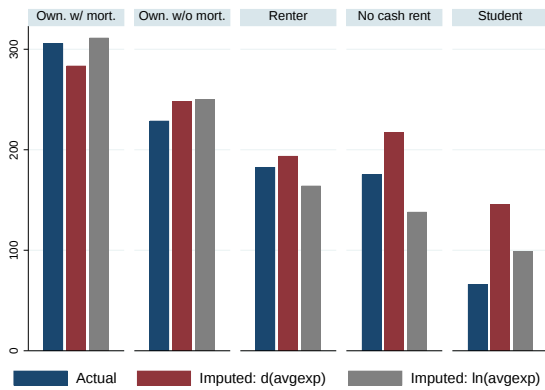
(b) Presence of Children



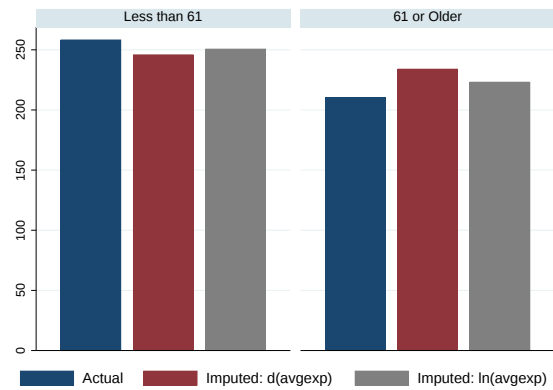
(c) Census Region



(d) Education Level



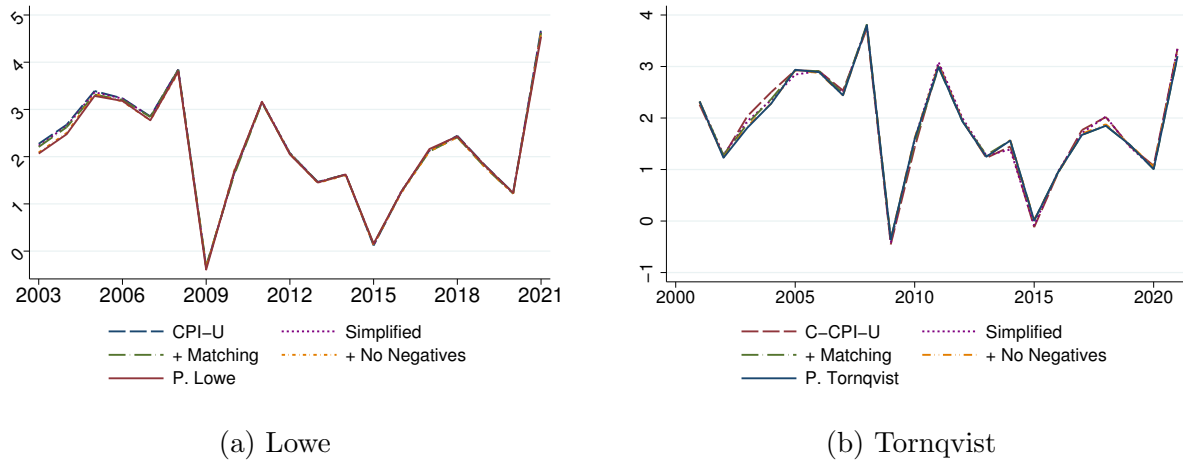
(e) Housing Tenure



(f) Age

Notes: Computed following matching within the urban population. $d(\text{avgexp})$ means $\Delta \bar{e}_{ht}$ is the dependent variable used in the distance measure. $\ln(\text{avgexp})$ means $\ln(\bar{e}_h)$ is the dependent variable.

Figure A4: Consumer Unit Indexes: Comparisons with CPI (Average 12m % Change)



Notes: Simplified means a replication procedure that omits weight-smoothing and combines a few items for the purposes of maintaining a constant item structure over time. Matching means using the simple replication procedure with the matched sample. No Negatives means taking the previous step and restricting expenditures at the CU-item-reference period level to be weakly positive. Finally, the P. Lowe index is estimated using only CUs with four interviews, while the P. Tornqvist uses only CUs with expenditures in adjacent months.

Figure A5: Democratic Tornqvist: Choice of Mean Across Consumer Units (Dec. 1999 = 1.0)

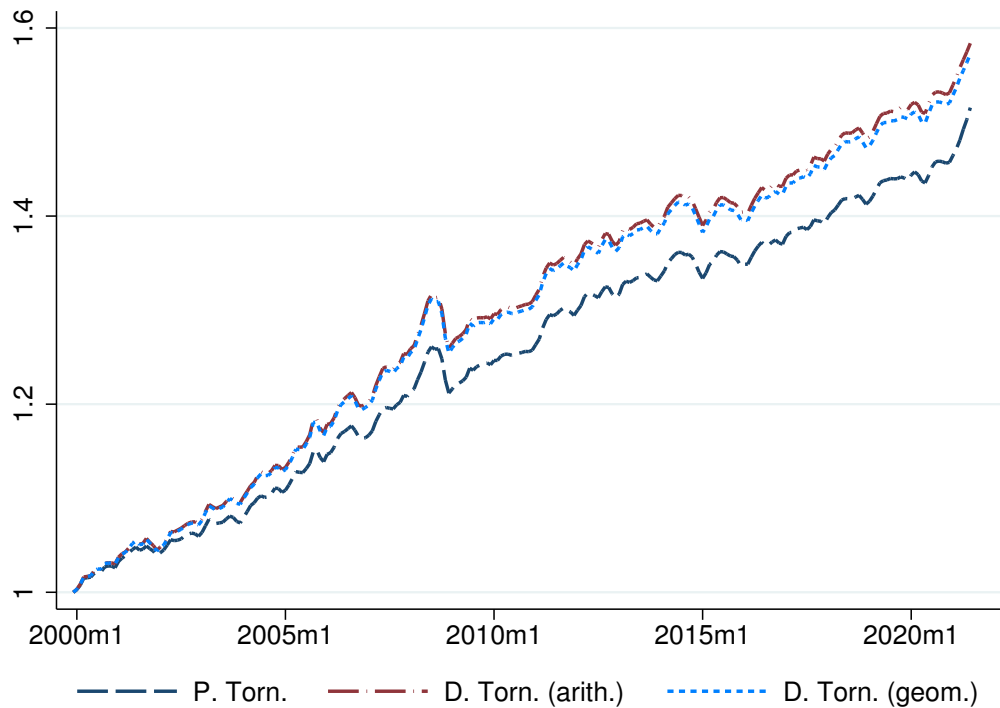
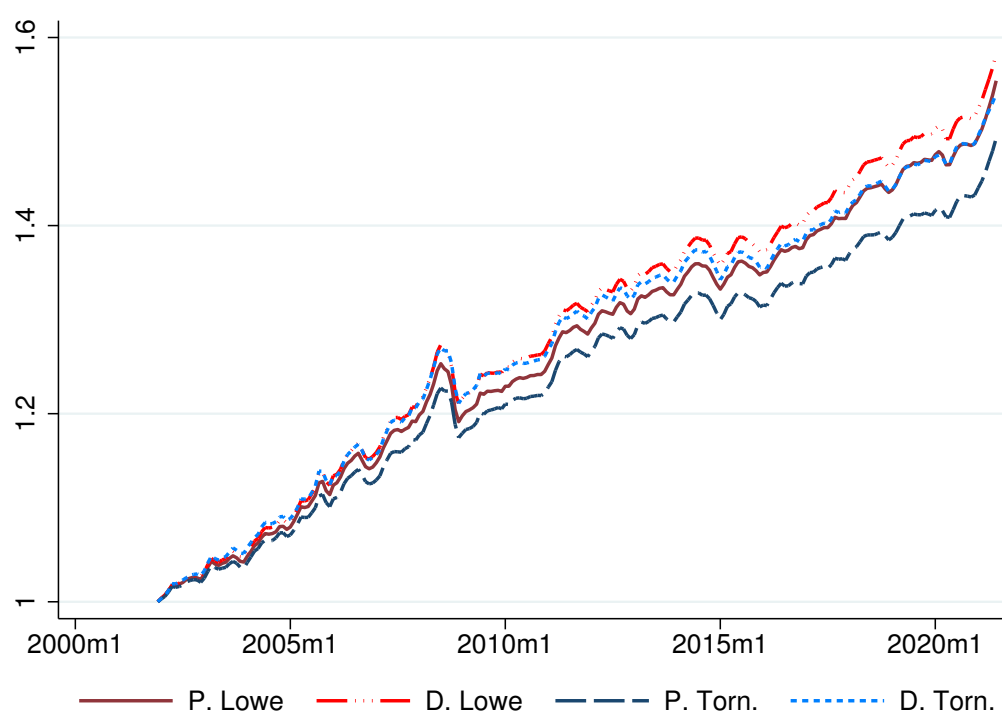
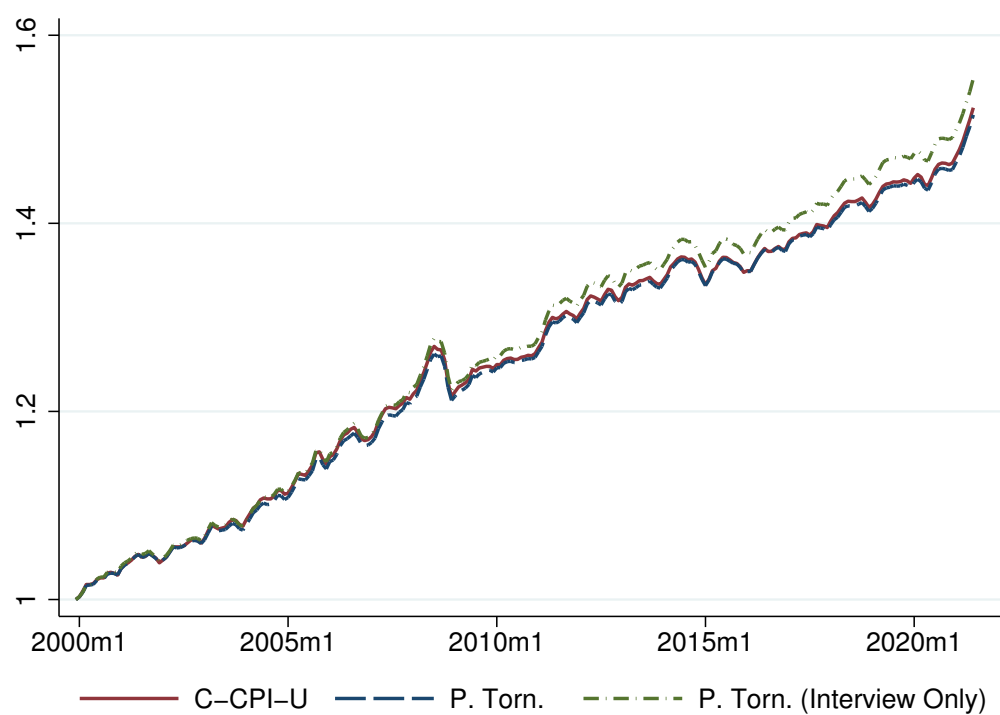


Figure A6: Consumer Unit Indexes Using Interview Only (Dec. 2001 = 1.0)



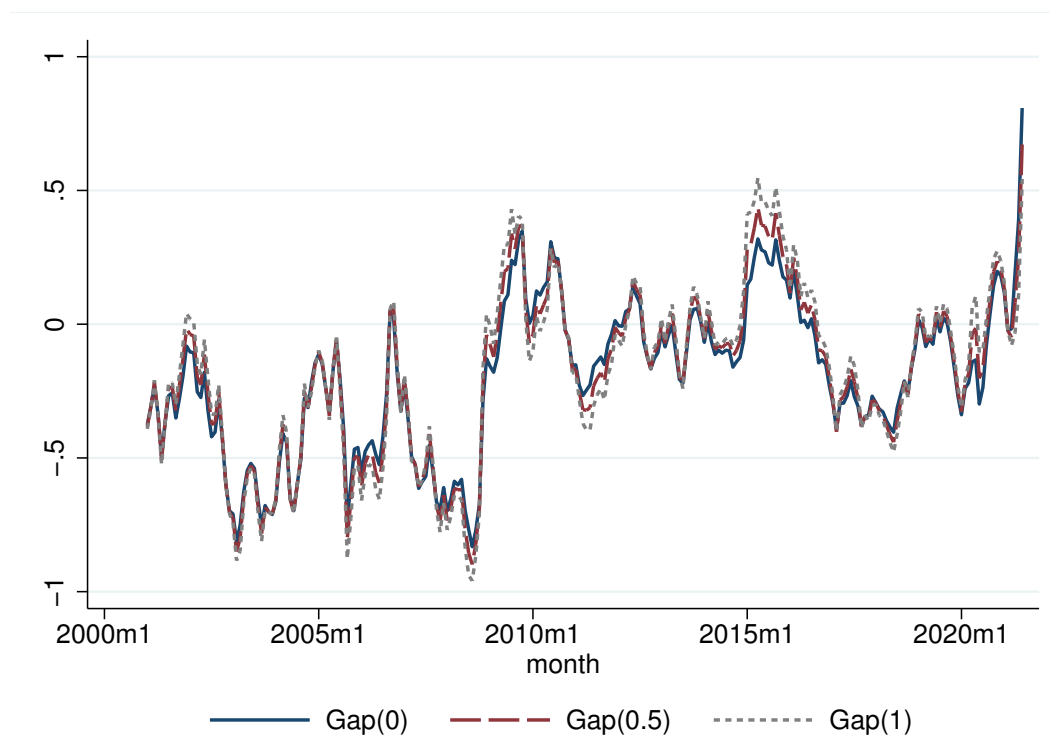
Note: Interview-only indexes use global food questions.

Figure A7: Consumer Unit Indexes: Comparing Survey Source (Dec. 1999 = 1.0)



Note: P. Torn (Interview only) index uses global food questions. P. Torn. uses matched Diary and Interview and same as Fig. 5.

Figure A8: Gaps in Tornqvist 12 month percent changes using equivalence scales



Note: Gaps are differences in 12 month percent changes between plutocratic and democratic Tornqvist. Gap(0) corresponds to the democratic Tornqvist weighting consumer units by $1/H$ as in the main results. Gap(0.5) means consumer unit weights are proportional to the square root of the family size, while Gap(1) means consumer unit weights are proportional to the family size.