Nonparametric Measurement of Long-Run Growth in Consumer Welfare

Xavier Jaravel[†] Danial Lashkari[‡]

July 4, 2022

Please click here for the most recent draft.

Abstract

How should we measure long-run changes in consumer welfare? This paper proposes a nonparametric approach that is valid under arbitrary preferences that depend on observable consumer characteristics, e.g. when expenditure shares vary with income. Our approach only requires data on the consumption baskets of a cross section of consumers facing a common set of prices. Using nominal expenditures under a constant set of prices as our money-metric for real consumption (welfare), we derive a consistent measure of its growth in terms of a correction to the conventional measures based on price index formulas. Our correction accounts for the cross-sectional dependence of the measured price indices on consumer income and other characteristics. We use nonparametric methods to approximate these corrections and provide bounds on the resulting approximation errors. Applying the approach to the measurement of growth in US real consumption per capita, we find a sizable correction to the standard measures of growth in the postwar era, a period of fast growth combined with substantial inflation gaps across income groups.

We thank Ben Faber, Thibault Fally, David Baqaee, Ariel Burstein, Ryan Chahrour, Cecile Gaubert, Shakeeb Khan, Arthur Lewbell, Jaromir Nosal, Robert Ulbricht, and David Weinstein for stimulating conversations and helpful comments and feedback. We are grateful to Sylvia Tian for excellent research assistance.

[†]London School of Economics. Email: x.jaravel@lse.ac.uk . Website: https://www.xavierjaravel.com.

[‡]Boston College. Email: danial.lashkari@bc.edu. Website: https://www.daniallashkari.com.

1 Introduction

How should we measure long-run changes in living standards? Classical demand theory shows that intuitive index number formulas, which simply aggregate observed changes in consumed quantities and prices, may provide precise measures of the change in consumer welfare. However, this powerful insight requires the crucial assumption that the composition of demand remains independent of consumer income. This so-called homotheticity assumption runs counter to the empirical regularity that demand for many goods such as food or housing systematically depends on income, a fact known since at least Engel (1857). It also belies the growing empirical evidence on sizable and systematic differences in inflation rates across income groups, typically featuring lower inflation rates for higher-income groups in the United States (e.g., Kaplan and Schulhofer-Wohl, 2017; Jaravel, 2019; Argente and Lee, 2021; Klick and Stockburger, 2021).

Despite this important and well-known theoretical limitation, classical price index formulas remain widely used in practice due to their simplicity, flexibility, and generality. Little is known about potential biases arising from the restrictive homotheticity assumption in the resulting measures of long-run growth in average standards of living. Current alternatives require us to specify and estimate the structure of the demand system–a task that is not feasible with many available datasets. For instance, Baqaee and Burstein (2021) have recently offered an approach that relies on the knowledge of the elasticities of substitution across goods to construct measures of welfare growth (see also Samuelson and Swamy, 1974).

In this paper, we develop a new approach for measuring welfare change that allows for flexible dependence of the patterns of demand on income and other sources of observed heterogeneity. Compared to classical index number theory, the only additional requirement is access to a cross-section of consumed quantities under common prices in at least one period; such data is widely available through standard surveys of consumption expenditure. Our approach nonparametrically estimates the cross-sectional dependence of measured price index formulas on consumer income. We show that knowing this dependence is sufficient to provide precise approximations for a theoretically-consistent measure of real consumption, while imposing minimal restrictions on the underlying preferences. This approach remains valid under any observable heterogeneity. We then apply our method to account for nonhomotheticity of demand in measuring long-run growth in the welfare of average US consumer from 1954 to 2019. In addition to improving the measurement of long-run growth and inflation inequality, our new approach can have important policy implications, such as the indexation of the poverty line and a more efficient targeting of welfare benefits.

We begin with a basic theory for the exact measurement of welfare change if we observe the full path of consumer choices and if we know the underlying preferences. We define real consumption as the expenditure required to achieve a certain level of welfare under constant prices fixed at a base period, and the true cost-of-living price index as the expenditure change required to maintain a given level of real consumption as prices change. When preferences are homothetic, the (instantaneous) growth in real consumption is given by the (instantaneous) growth in nominal total expenditure deflated by the (instantaneous) true price index evaluated at the current level of real consumption. We first show that, under more general preferences, we need to multiply the deflated nominal expenditure by a correction factor which is characterized by a *nonhomotheticity correction function*. This function accounts for the elasticity of the the true price index between the base and current periods with respect to real consumption. It accounts for the covariance across goods between income elasticities and the cumulative price inflation between the base and current periods. As we move away from the base period and the cumulative inflation becomes large, the correction may grow to become first-order in welfare change.

To see the intuition behind this correction, consider a setting where consumer welfare is rising over a time horizon in which price inflation rates are relatively lower for goods with high income elasticity (luxuries). Let us express real consumption in terms of constant initial period prices as our base, such that total expenditure is linear in (and identical to) real consumption in the initial period. As time passes, the *relative* cost of achieving higher levels of real consumption falls, since the goods consumed by richer consumers are getting relatively less expensive. In other words, the expenditure function — consumers' nominal expenditures as a function of real consumption — becomes more and more concave over time. Thus, a given rise in total nominal expenditure translates into increasingly larger gains in real consumption as consumers become richer. The conventional approach assumes homotheticity (and thus linearity of the expenditure function), and uses uncorrected deflated expenditure growth to measure real consumption growth. It thus ignores the gradual fall in the curvature of the expenditure function and underestimates the growth of real consumption for the initial base period. If we instead express real consumption in terms of constant final period prices as our base, the same logic implies that conventional approach overestimate the growth in all preceding periods.[1] Our nonhomotheticity correction measures the contribution of changes in the curvature of the expenditure function to account for its effect on measured growth in terms of any base period of interest.

We next provide a theory for approximating welfare change if we observe discrete observations of consumer choice and do *not* know the underlying preferences. The classical results show us that price index formulas between consecutive periods approximate the true price index that maintain the current level of real consumption. To apply the insights of our theory, we thus need

[1] In this case, since total consumer expenditure is identical to real consumption in the final period, it must be a convex function of real consumption in all prior periods. This leads to overestimating the growth of real consumption when using the final period as base.

to approximate the nonhomotheticity correction function. Our main contribution is to provide simple procedures to approximate this function based on cross-sectional data on the choices of consumers with arbitrary but identical non-homothetic preferences, and with heterogeneous income. In the base period, total nominal expenditure by definition coincides with real consumption. This allows us to nonparametrically approximate the elasticity of the expenditure function with respect to real consumption using the cross-sectional variations across consumers in price index formulas. Using this elasticity, we can infer the nonomotheticity correction and approximate real consumption in periods immediately away from the base period. We can then recursively apply the same strategy in the subsequent periods as we progress in time away from the base period and approximate real consumption over our entire period of interest. We provide two such approximations, first and second order in annual inflation and total nominal expenditure growth, and provide bounds on the error in each case using standard results in nonparametric function approximation literature.[2]

We further show that our strategy generalizes to settings where preferences systematically vary in consumer characteristics, e.g., age, family size, education, etc. When such characteristics also vary within consumer, we can account for their contributions to real consumption growth by deflating the nominal expenditure additionally by the product of the growth in the characteristic and a corresponding correction. We offer an extended methodology for approximating this correction, which is given by the elasticity of the true price index with respect to the changing characteristic, using the cross-sectional variations in the price index formulas and consumer characteristics.

We study the accuracy of our approximation strategy in measuring long-run growth and inequality using a simulation with known preference parameters. We consider the nonhomothetic CES (nhCES) preferences estimated for three main categories of sectoral goods, agriculture, manufacturing, and services, by Comin et al. (2021). We generate data on consumption choices in a synthetic sample of households, calibrating the average inflation and growth in nominal expenditure to those in the US in the 1953-2019 period. We simulate the data under different values for the covariance between price inflation rates and income elasticities across goods. Knowing the underlying expenditure function, we can compute the true price index and real consumption growth and compare them against conventional uncorrected measures and those corrected based on our methodology. High levels of covariance, when combined with high growth in real consumption, lead to sizable bias in the uncorrected measures over long time horizons. We confirm that our procedure accurately recovers the evolution of the exact index using the observed

[2]Despite its larger approximation error, the first-order approximation has the advantage that it requires access to a cross-section of consumers only in the base year. In contrast, the second-order approximation requires access to a cross-section of consumers overlapping every consecutive periods of interest.

cross-sectional data and without *any* knowledge of the underlying preference parameters.

We further apply our approach to real data from the US and quantify the magnitude of the bias in conventional measures of average real consumption growth. We build a dataset linking income-specific expenditure shares (five quintiles of income) from the Consumer Expenditure Survey (CEX) to inflation rates from the Consumer Price Index (CPI) across product categories covering the full consumption baskets of US households from 1984 to 2019. We complement this data with information on CPI inflation rates and on aggregate expenditure shares from the Bureau of Labor Statistics (BLS) over the longer horizon of 1954-2019 period.

We first focus on the period 1984-2019 where cross-sectional data is available. In line with the work cited above, using conventional uncorrected measures we find lower inflation and higher real consumption growth for high-income US households. As with the example discussed above, uncorrected measures underestimate growth in real consumption when expressed with the initial base (1984), and overestimate it when expressed with the final base (2019). We find that the magnitude of the bias in the uncorrected measure of annual growth of real consumption rises as we move away from the base period, reaching around 8% of the measured annual growth for the highest quintile. Nevertheless, since over this period growth is modest, the overall bias in uncorrected cumulative growth, both for the mean and for the highest quintile, reaches only around 1 percentage point.

We find slightly higher magnitudes in robustness checks with more granular data. We first use household-level data CEX to estimate income elasticities for a more granular set of product categories from 1990 to 2019. We also study consumer packaged goods using Nielsen scanner data (2004-2014). Overall, the robustness analyses confirm the findings obtained with our baseline dataset.

Examining the earlier postwar US experience may be of particular interest due to is substantially higher rates of real consumption growth and the fact that our nonhomotheticity correction has a compounding effect over long periods of time. Using the earliest available cross-sectional data in 1984 as the basis of our correction, we extend the analysis to the entire 1954-2019 period and express real consumption using final period (2019) prices as base. In terms of the cumulative growth over the entire period, we find the upward bias in the conventional growth measure is around 14 percentage points, from a benchmark measured cumulative growth of 141 percentage point. The bulk of this bias is coming from the 1954-1984 period featuring faster growth: the cumulative overestimation over this 30-year period is around 10 percentage points (or 32 basis points annually). When real consumption is expressed in terms of 1984 prices, the bias in the uncorrected measurement of growth is smaller over the entire period (1954 to 2019). This is because, as in the example discussed above, the direction of the bias changes sign over the periods before (1954-1984) and after (1984-2019) the base year. Nevertheless, even using 1984 prices as base, welfare comparisons using the conventional measures remain subject to error. For instance, the cumulative growth in real consumption from 1954 to 1984 is overestimated by around 4 percentage point. We conclude that accounting for nonhomotheticity does matter for consistent quantitative comparisons of welfare over long time horizons.

Prior Work Our paper builds on and contributes to three strands of literature. First, we extend the literature on index number theory (e.g., Pollak, 1990; Diewert, 1993), which has enabled transparent and consistent comparisons of the aggregate measures of consumption and production over time and space only relying on observables under the assumption of homotheticity. As emphasized by Samuelson and Swamy (e.g., 1974), many classical results do not generalize beyond settings involving homotheticity in preferences. Under nonhomotheticity, Diewert (1976) has showed that one can still rely on the conventional price index formulas to measure changes in welfare locally. However, as we show here, these results do not generalize to welfare comparisons over long time horizons.

Second, we advance a growing literature raising the point that standard price index formulas suffer from a bias due to nonhomotheticities, whose magnitude relates to the covariance between income elasticities and price changes (e.g., Fajgelbaum and Khandelwal, 2016; Atkin et al., 2020; Baqaee and Burstein, 2021). In particular, Baqaee and Burstein (2021) have recently highlighted the failure of standard divisia indices to measure welfare-relevant measures of growth in real consumption. They suggest relying on the estimates of the elasticities of substitution to account for the role of nonhomotheticity.[3] In contrast, we provide a nonparametric approach that does not require specifying the underlying demand functions. The importance of the covariance between income elasticities and inflation for measuring welfare change is also noted by Fajgelbaum and Khandelwal (2016) and Atkin et al. (2020). Fajgelbaum and Khandelwal (2016) measure changes in welfare gains from trade liberalization across different income groups in a parametric setting and under the assumption of an AIDS demand system (Deaton and Muellbauer, 1980). Atkin et al. (2020) consider the problem of welfare measurement in the absence of reliable price data, and rely on separability assumptions on the structure of preferences to infer welfare from shifts in the Engel curves. For this procedure to hold without the need for estimation of structural elasticities of substitution, they rule out exactly the covariance patterns that lead to large nonhomotheticity corrections in our framework.[4] In summary, while this literature provides parametric corrections for the bias, our contribution is to provide a nonparametric correction, which is valid

[3] Baqaee and Burstein (2021) additionally study the consequences of the endogeneity of prices in general equilibrium, as well as unobserved heterogeneity, e.g. taste shocks. The latter effects have also been recently considered by Redding and Weinstein (2020).

[4] Atkin et al. (2020) also analyze the case when relative prices change, in which case their procedure requires computing (unobserved) compensated shifts in expenditure shares due to the change in relative prices (see their equation (2)).

under arbitrary preferences where all heterogeneity is in terms of observables.

Third, we contribute to the literature on the measurement of inflation inequality (e.g., Hobijn and Lagakos, 2005; McGranahan and Paulson, 2006; Kaplan and Schulhofer-Wohl, 2017; Jaravel, 2019; Argente and Lee, 2021). Prior work on inflation inequality has posited the existence of separate homothetic indices for different income groups. We apply our methodology to provide estimates of inflation inequality that are robust to potential biases arising from nonhomotheticities. Using a new linked dataset covering the period 1953-2019 in the United States, we apply our methodology to the measurement of short, medium, and long run growth in real consumption, on average and across the income distribution, and we quantify the magnitude of the bias in the conventional measures in each case.

2 Exact Measurement of Welfare Change

In this section, we assume that the underlying preference parameters are known and we derive the correction for nonhomotheticities. Section 2.1 introduces the notation used throughout the paper and defines the standard concepts used for the measurement of welfare, cost of living, and real consumption. Section 2.2 derives our nonhomotheticity correction and Section 2.3 extends the analysis to accounting for consumer characteristics that matter for preferences.

2.1 Definitions

2.1.1 Real Consumption

Consider a setting where we observe the composition of all consumption expenditures of a consumer whose total expenditure evolves over time. We assume that the consumer faces the path of prices $\boldsymbol{p}_t$ and quantities $\boldsymbol{p}_t$ for I products, where $\mathbf{0} < \boldsymbol{p}_t \in \mathbb{R}^I$ and $\mathbf{0} \leq \boldsymbol{q}_t \in \mathbb{R}^I$, over the interval $t \in [0, T]$. Correspondingly, we let $y_t \equiv \sum_i p_{i,t} q_{i,t}$ and $s_{i,t} \equiv p_{i,t} q_{i,t} / y_t$ denote the total expenditure and the share of good i in total expenditure at time t, respectively. We assume that consumption choices of the consumer are rationalized by a utility function $U(\boldsymbol{q})$ and a corresponding second-order continuously differentiable expenditure function $E(u; \boldsymbol{p})$ that gives the expenditure required to achieve utility u under prices $\boldsymbol{p}$.

We begin by our concept of real consumption as a money metric for consistent measurement of welfare over time.

Definition 1 (Real Consumption)**.** For given base vector of prices $\boldsymbol{p}_b$ (with $0 \leq b \leq T$), define *real consumption under period-b constant prices* as a monotonic transformation $M_b(\cdot)$ of utility u given by

$$c^b = M_b(u) \equiv E(u; \boldsymbol{p}_b). \tag{1}$$

Equation (1) constitutes our core measure of welfare for a household with utility u, which gives the minimum expenditure needed to achieve that level of utility under the constant vector of prices prevailing at time b. Since real consumption is defined with reference to base b, we should include it in our notation for real consumption as c^b. However, for brevity we may at times drop the superscript to simplify the expression whenever it is clear from context that the base b is fixed.

Consider now comparing welfare between an initial period t_0 and a final period t, such that $0 \leq t_0, t \leq T$. As examples for the choices of the base year, we may choose the initial year $\boldsymbol{p}_{t_0}$. In this case, the comparison between the final real consumption at time t and the initial one leads to the notion of Equivalent Variation (EV). On the other hand, if we instead choose the base year to the be the final period t, the comparison between the final and the initial real consumption gives us the notion of Compensating Variation (CV):

$$EV \equiv \frac{c_t^{t_0}}{c_{t_0}^{t_0}} = \frac{E\left(u_t; \boldsymbol{p}_{t_0}\right)}{E\left(u_{t_0}; \boldsymbol{p}_{t_0}\right)}, \qquad CV \equiv \frac{c_t^t}{c_{t_0}^t} = \frac{E\left(u_t; \boldsymbol{p}_t\right)}{E\left(u_{t_0}; \boldsymbol{p}_t\right)}. \tag{2}$$

Equivalent variation measures the growth in expenditure that would be equivalent to the change in welfare for the consumer over the full period, keeping prices fixed at their initial-period level. In contrast, compensating variation measures the growth in expenditure that would compensate the consumer for the welfare change over the full period, keeping prices fixed at their final-period level.

While equivalent and compensating variation by construction only allow two-period welfare comparisons, Definition 1 allow us to consistently compare utility across multiple periods in terms of real consumption with a constant base vector of prices. We can also express the expenditure function and the Hicksian demand in terms of this utility metric.

Definition 2. For measures of real consumption under period-b constant prices, define the expenditure function and expenditure shares as

$$\widetilde{E}^b(c; \boldsymbol{p}) \equiv E\left(M_b^{-1}(c); \boldsymbol{p}\right), \tag{3}$$

$$\widetilde{\Omega}_i^b(c; \boldsymbol{p}) \equiv \frac{\partial \log \widetilde{E}^b(c; \boldsymbol{p})}{\partial \log p_i}. \tag{4}$$

The expenditure function in Equation (3) gives the expenditure required to achieve real consumption c (defined under base period b) under the vector of prices $\boldsymbol{p}$. By definition, we have $c = \widetilde{E}^b(c; \boldsymbol{p}_b)$ for all c. Moreover, the function evaluated at the consumer's value of real consumption c_t^b at time t gives the time-t expenditure, that is, $y_t = \widetilde{E}^b\left(c_t^b; \boldsymbol{p}_t\right)$.

Equation (4) uses Shephard's lemma to define a representation of the Hicksian demand as the expenditure share of a consumer with real consumption c (defined under base period b) for good i under the vector of prices $\boldsymbol{p}$. Again, this function evaluated at the consumer's value of real consumption c_t^b at time t gives the time-t expenditure share for good i, that is, $s_{it} = \widetilde{\Omega}_i^b\left(c_t^b; \boldsymbol{p}_t\right)$.

2.1.2 True Price Index

We next use build on Definition 1 to define the inflation in cost-of-living as a function of the consumer's real consumption.

Definition 3. Define the instantaneous inflation in the cost-of-living for a consumer with real consumption c (defined under base period b) at time t as

$$P_t^b(c) \equiv \exp\left[\sum_i \widetilde{\Omega}_i^b(c; \boldsymbol{p}_t) \frac{d \log p_{it}}{dt}\right]. \tag{5}$$

Equation (5) defines the instantaneous inflation in the costs of maintaining real consumption c, under base period b, as the weighted mean of price growth across goods, where we use the expenditure shares at that level of real consumption as weights. When evaluating the inflation experienced for the consumer at time t for their concurrent level of real consumption c_t^b, the weights correspond to the observed expenditure shares at time t:

$$P_t^b\left(c_t^b\right) = D_t \equiv \exp\left(\sum_i s_{it} \frac{d \log p_{it}}{dt}\right). \tag{6}$$

Along the path of consumer choice, Equation (6) shows that the instantaneous inflation in the cost-of-living defined in Defintion 3 leads to the standard definition of the Divisia price index D_t (Hulten, 1973). Note that the instantaneous inflation evaluated at concurrent level of real consumption does *not* depend on the choice of the constant base vector of prices.

While Defintion 3 provided an instantaneous measure of inflation, we now define an index for the inflation in the cost-of-living over a finite interval of time.

Definition 4 (True Price Index)**.** Define the "True" cost-of-living price index $\mathscr{P}_{t_0,t}^b(c)$ for a consumer with real consumption c (defined under base period b) between periods t_0 and t ($0 \leq t_0, t \leq T$) as:

$$\mathscr{P}_{t_0,t}^b(c) \equiv \frac{\widetilde{E}^b(c; \boldsymbol{p}_t)}{\widetilde{E}^b\left(c; \boldsymbol{p}_{t_0}\right)} = \exp\left[\int_{t_0}^t \log P_\tau^b(c)\, d\tau\right]. \tag{7}$$

The true price index defined in Equation (7) is a "structural" index, in the sense that its computation requires the knowledge of the underlying structure of preferences. Using Definitions

1 and 4, we can write the following relationship between real consumption growth and the true price index over the horizon $[t_0, t]$:[5]

$$\frac{c_t^b}{c_{t_0}^b} = \frac{y_t / y_{t_0}}{\mathscr{P}_{t_0,b}^{b}\left(c_{t_0}\right) \times \mathscr{P}_{b,t}^{b}\left(c_t\right)}. \tag{8}$$

Equation (8) shows that we can compute the growth in real consumption for any base period b by deflating the growth in the nominal consumer expenditure by a composite true price index. This composite price index is the product of the true price index between the initial period t_0 and the base period b, and the true price index between the base period b and the final period t. Crucially, the former index has to account for the inflation in the cost of living at the initial level of real consumption c_{t_0}, while the latter has to account for that at the final level of real consumption c_t. For the specific choices of initial and final periods as based, Equation (8) reduces to

$$EV = \frac{c_t^{t_0}}{c_{t_0}^{t_0}} = \frac{y_t / y_{t_0}}{\mathscr{P}_{t_0,t}\left(c_t\right)}, \qquad CV = \frac{c_t^{t}}{c_{t_0}^{t}} = \frac{y_t / y_{t_0}}{\mathscr{P}_{t_0,t}\left(c_{t_0}\right)},$$

suggesting that the equivalent and compensating variations correspond to the nominal expenditure growth deflated by the true price index for the final and the intial levels of real consumption, respectively.

Homothetic Preferences Let us now consider the restriction that the underlying preferences are homothetic, that is, the composition of demand does not depend on the level of utility, $\widetilde{\Omega}_i^b(c;\boldsymbol{p}) \equiv \Omega_i(\boldsymbol{p})$ for all i and for all base periods b.[6] In this case, Equation (5) suggests that the instantaneous inflation is identical for all levels of real consumption: $P_t^b(c) = \sum_i \Omega_i(\boldsymbol{p}) \frac{d \log p_{it}}{dt}$ for all c. Correspondingly, from Definition 4, the true price index $\mathscr{P}_{t_0,t}^{b}(c)$ between any two time periods t_0 and t takes the same value independent of the level of real consumption c and the choice of the base period b. Equation (8) then implies:[7]

$$\frac{c_t^b}{c_{t_0}^b} = \frac{y_t / y_{t_0}}{\mathscr{P}_{t_0,t}^{b}(c)}, \qquad \text{for any } c, b.$$

[5]See proof in Appendix A.3.

[6]The utility function $U(\cdot)$ is homothetic If (and only if) we can write the expenditure function as $E(u;\boldsymbol{p}) = P(\boldsymbol{p}) \cdot F(u)$,for some unit cost function $P(\cdot)$ and some canonical homothetic cardinalization $F(\cdot)$ of utility (Diewert, 1993).

[7]Homotheticity is a necessary and sufficient condition for the true price index $\mathscr{P}_{t_0,t}^{b}(c)$ to tbe independent of c and for the growth in real consumption $c_t^b / c_{t_0}^b$ to be independent of the base b. Samuelson and Swamy (1974) refer to this result as the *homogeneity theorem.*

2.2 Nonhomotheticity Correction

Using the definitions in Section 2.1.1, our first results characterizes the growth in consumer real consumption along the path as the product of the deflated growth in the (nominal) consumer expenditure and a multiplicative factor that accounts for the effect of nonhomotheticity.

Lemma 1. *The instantaneous growth in real consumption of the consumer, as defined under period-b constant prices, satisfies*

$$\frac{d\log c_t^b}{dt} = \frac{1}{1+\Lambda_t^b\left(c_t^b\right)}\left(\frac{d\log y_t}{dt} - \log D_t\right), \tag{9}$$

where D_t stands for the Divisa index given by Equation (6) and where the nonhomotheticity correction function $\Lambda_t^b(c)$ is characterized by the initial condition $\Lambda_b^b(c) = 0$ and evolves over time according to:

$$\frac{\partial \Lambda_t^b(c)}{\partial t} = \frac{\partial \log P^b(c)}{\partial \log c}. \tag{10}$$

Proof. We can write the growth in the consumer expenditure as

$$\frac{d\log y_t}{dt} = \frac{d\log \widetilde{E}(c_t;\boldsymbol{p}_t)}{dt} = \sum_i \frac{\partial \log \widetilde{E}(c_t;\boldsymbol{p}_t)}{\partial \log p_{it}} \frac{d\log p_{it}}{dt} + \frac{\partial \log \widetilde{E}(c_t;\boldsymbol{p}_t)}{\partial \log c_t} \frac{d\log c_t}{dt},$$

where we have suppressed the base period superscripts b to simplify the expression. Substituting for the partial derivative of the expenditure function with respect to price from Equation (4) and defining $1+\Lambda_t(c)$ as the partial derivative of the expenditure function with respect to real consumption leads to Equation (9).

The initial condition $\Lambda_b^b(c) = 0$ follows from $c = \widetilde{E}^b(c;\boldsymbol{p}_b)$, and Equation (10) from the observation that

$$\frac{\partial \Lambda_t(c)}{\partial t} = \frac{\partial^2 \log \widetilde{E}(c;\boldsymbol{p}_t)}{\partial t\,\partial \log c} = \frac{\partial}{\partial \log c}\left(\sum_i \frac{\partial \log \widetilde{E}(c;\boldsymbol{p}_t)}{\partial \log p_{it}} \frac{d\log p_{it}}{dt}\right),$$

where in the second equality we have switched the order of derivatives. Using again the definition in Equation (4) leads to the desired result. □

To understand the implications of Lemma 1, let us first consider imposing the restriction that the underlying preferences are homothetic. In this case, as we saw the inflation in the cost-of-living defined in Equation (5) is independent of real consumption and thus the nonhomotheticity correction function is zero. Hence, we find the standard result that the growth in real consumption is simply given by the growth in nominal expenditure deflated by the integral of the Divisia price index:

$$\log\left(\frac{c_t^b}{c_{t_0}^b}\right)=\log\left(\frac{y_t}{y_{t_0}}\right)-\int_{t_0}^{t}\log D_\tau d\tau. \tag{11}$$

Importantly, Equation (11) shows that the growth in real consumption is identical regardless of the choice of the base period, as the right hand side of the equation does not depend on b.

Once we deviate from homotheticity, we do indeed have to account for the nonhomotheticity correction function Λ_t in Equation (9). Thus, Equation (11) provides an exact characterization of real consumption growth only at the base period b where we have $\Lambda_b \equiv 0$ by definition. Moving forward in time from the base period $t > b$, Equation (10) shows that the nonhomotheticity correction rises if inflation is higher at higher levels of real consumption. In such cases, raising one's real consumption is becoming more expensive over time, and and thus the exact measure of real consumption growth is smaller than what is suggested by Equation (11). In contrast, if inflation is higher at lower levels of real consumption, raising one's real consumption is becoming less expensive over time, and thus the exact measure of real consumption growth exceeds what is suggested by Equation (11).

Under what conditions does the nonhomotheticity correction requires a sizable adjustment of the simple deflation formula? We can alternatively write the nonhomotheticity correction function at time t as the elasticity of the true price index between the base period b and t with respect to real consumption:

$$\Lambda_t^b(c)=\frac{\partial\log\mathscr{P}_{b,t}^b(c)}{\partial\log c}. \tag{12}$$

Thus, the nonhomotheticity correction will be small when the current period t is close enough to the base period b such that the cumulative inflation, proxied by the true cost-of-living price index, is small. Alternatively, we can express Equation (10) as the weighted covariance between the elasticity of demand with respect to real consumption $\widetilde{\eta}_i(c;\boldsymbol{p})\equiv\frac{\partial\log\widetilde{\Omega}_i(c;\boldsymbol{p}_t)}{\partial\log c}$ and the inflation in prices across goods

$$\frac{\partial\Lambda_t(c)}{\partial t}=\sum_i\widetilde{\Omega}_i(c;\boldsymbol{p}_t)\,\widetilde{\eta}_i(c;\boldsymbol{p})\,\frac{d\log p_{it}}{dt},$$

where we have again dropped the base period superscripts b to simplify the expression. Thus, the nonhomotheticity correction is small when price inflation is uncorrelated with income elasticities across goods, even if the average size of price inflation is large. Thus, we conclude that the nonhomotheticity correction is likely to be sizable when preferences are nonhomothetic, price inflation is large and correlated with income elasticities across goods, and we are expressing real consumption in terms of a base period that is distant from the current period.

We note that the role of the covariance between income elasticities and price changes has been highlighted in prior work (e.g., Fajgelbaum and Khandelwal, 2016; Atkin et al., 2020; Baqaee and Burstein, 2021). As we will see in Section 3 below, our main contribution is to provide a nonparametric approximation for the nonhomotheticity correction that is valid under arbitrary preferences (as well as heterogeneity in terms of observables). Before doing so, we highlight another important property of the nonhomotheticity correction, and we generalize Lemma 1 to account for changes in characteristics in Section 2.3.

Real Consumption Growth and the Choice of Constant Prices How does the choice of the base period that fixes constant prices affect growth in the corresponding measure of real consumption? The following lemma characterizes the systematic relationship between the choice of the base period and the measurement of real consumption.

Lemma 2. *Consider two base periods $b_1 < b_2$. The instantaneous growth in real consumption as measured by constant prices in period b_2 relative to that by constant prices in period b_1 satisfy*

$$\frac{d\log c_t^{b_2}}{d\log c_t^{b_1}} = 1 + \Lambda_{b_2}^{b_1}\left(c_t^{b_1}\right) = 1 + \left.\frac{\partial \log \mathscr{P}_{b_1,b_2}^{b_1}(c)}{\partial \log c}\right|_{c=c_t^{b_1}}. \tag{13}$$

Proof. See Appendix A.3. □

Lemma 2 shows that the gap in two measures of growth at time t based on two base periods b_1 and b_2 depend on the nonhomotheticity correction between the two periods b_1 and b_2. Let us consider a scenario where $b_1 < b_2$, and where prices are on the rise, and price inflation negatively covaries with income elasticities across goods between b_1 and b_2. In this case $\Lambda_{b_2}^{b_1} < 0$, the expression on the right hand side of Equation (13) suggests that real consumption growth is lower when measured from the perspective of the later period b_2.

To gain some intuition about the economics behind this result, let us consider a simple economy with two goods: burgers and mobile phones. Assume that mobile phones are more income elastic than burgers and that over a period of time, say, from 1970 to 2020, the relative price of mobile phones falls substantially relative to burgers. If we measure the growth in real consumption in this economy from the perspective of prices held constant in some early base period, for example 1970, then real consumption growth over this fifty year period is larger when preference nonhomotheticity is taken into account. The reason is that consumers become richer over time, which leads to an increase in the propensity to spend on mobile phones, precisely when the relative price of mobile phones is falling. Thus, in this example conventional measures of real consumption growth are biased downward because they don not account for the fact that the income-elastic goods become relatively cheaper at the same time as they become relatively more important from the point of view of consumer preferences.

In contrast, looking backward in time from the perspective of prices held fixed at a later period, for example 2020, then real consumption growth during the period is smaller when accounting for the nonhomotheticitiy correction. The reason is that lower incomes in the past require the consumer to consume relatively more burgers, due to its lower income elasticity, compared to the later periods. Going backward in time, consumers become poorer and spend relatively more on the income-inelastic good, which is assumed to become relatively cheaper in our example. Thus, the fall in income is dampened by the fact that burgers are relatively cheaper while consumer demand for burgers have increased. In this example, conventional measures of real consumption growth are therefore biased upward.

This example shows that the sign of the bias induced by the nonhomotheticity correction inherently depends on the choice of the base period.[8]

2.3 Correction for Change in Consumer Characteristics

In this section, we extend the results of Section 2.2 to a setting that includes additional sources of observed consumer characteristics, beyond income, are changing over time. Examples of such characteristics include the age and education of the consumer, the number of household members, and the location of the customer in space. Assume that we observe a vector of consumer characteristics (covariate) $\boldsymbol{x}_t \in \mathbb{R}_+^D$ at time t.[9] We assume that consumer preferences are characterized by a well-behaved utility function $u = U(\boldsymbol{q};\boldsymbol{x})$ depends on the consumer characteristics and let $y = E(u;\boldsymbol{p};\boldsymbol{x})$ denote the corresponding expenditure function, and that we again observe the path of prices and quantities for a consumer with (potentially) changing characteristics. We first define our generalized concept of real consumption in this environment.

Definition 5 (Generalized Real Consumption)**.** For given base vector of prices $\boldsymbol{p}_b$ (with $0 \leq b \leq T$), define *real consumption under period-b constant prices* for a consumer with utility u and characteristics $\boldsymbol{x}$ as a monotonic transformation $M_b(u,\boldsymbol{x})$ of utility given by

$$c^b = M_b(u,\boldsymbol{x}) \equiv E(u;\boldsymbol{p}_b;\boldsymbol{x}). \tag{14}$$

Recall that Definition 1 gave us a money metric that allows us to compare the expenditure required to achieve different levels of welfare for consumers (with identical preferences) under a constant vector of prices. Definition 5 generalizes this concept to a setting in which preferences potentially depend on consumer characteristics. While we cannot compare the welfare of

[8]To the best of our knowledge, this point has not been made in prior work on measuring welfare change in the presence of preference nonhomotheticity.

[9]The assumption that the elements of the vector are positive valued is without loss of generality, as we can always transform the characteristic space in such a way that this condition holds.

consumers with different preferences, we can still compare the expenditure required to achieve different levels of welfare for such consumers when they face the same vector of prices. Therefore, we can say the real income of a consumer with preferences $\boldsymbol{x}_t$ with utility u_t is higher than that of household with preferences $\boldsymbol{x}_{t_0}$ and utility u_{t_0} by $c_t^b - c_{t_0}^b \equiv M_b(u_t; \boldsymbol{x}_t) - M_b\left(u_{t_0}; \boldsymbol{x}_{t_0}\right)$ under the base vector of prices $\boldsymbol{p}_b$.

Let us investigate our the definitions above under two trivially special cases. First, if household preferences do not change, i.e., $\boldsymbol{x}_t \equiv \boldsymbol{x}_{t_0}$, then the definition above reduces to that given under homogeneous preferences. Second, if the prices do not change, i.e., $\boldsymbol{p}_t \equiv \boldsymbol{p}_{t_0}$, the growth in real consumption simply accounts for the growth in nominal expenditure even if consumer characteristics change, $c_t^b / c_{t_0}^b \equiv y_t / y_{t_0}$. Intuitively, what matters is the covariance between the change in prices and the chance in consumers' characteristics, as we show formally below.

In parallel to the definitions introduced in Section 2.1.1, we define the expenditure function and the Hisksian expenditure share function in terms of real consumption defined by Equation (14) as $\widetilde{E}^b(c; \boldsymbol{p}, \boldsymbol{x}) \equiv E\left(M_b^{-1}(c; \boldsymbol{x}); \boldsymbol{p}, \boldsymbol{x}\right)$ and $\widetilde{\Omega}_i^b(c; \boldsymbol{p}, \boldsymbol{x}) \equiv \Omega_i\left(M_b^{-1}(c; \boldsymbol{x}); \boldsymbol{p}, \boldsymbol{x}\right)$ where $\Omega_i(u; \boldsymbol{p}, \boldsymbol{x}) \equiv \partial \log E / \partial \log p_i$. We are now prepared to characterize

Lemma 3. *The instantaneous growth in real consumption of the consumer, as defined under period-b constant prices, satisfies*

$$\frac{d \log c_t^b}{dt} = \frac{1}{1 + \Lambda_t^b(c_t; \boldsymbol{x}_t)} \left[\frac{d \log y_t}{dt} - \log D_t - \sum_d \Gamma_{d,t}^b(c_t; \boldsymbol{x}_t) \frac{d \log x_{dt}}{dt} \right], \tag{15}$$

where we have again defined the Divisia index $D_t \equiv \sum_i s_{it} \frac{d \log p_{it}}{dt}$, and where the nonhomotheticity correction function $\Lambda_t(c; \boldsymbol{x})$ and the characteristic-d correction function $\Gamma_{dt}(c; \boldsymbol{x})$ are given by

$$\Lambda_t^b(c; \boldsymbol{x}) \equiv \frac{\partial \log \mathscr{P}_{b,t}^b(c; \boldsymbol{x})}{\partial \log c}, \qquad \Gamma_{d,t}^b(c; \boldsymbol{x}) \equiv \frac{\partial \log \mathscr{P}_{b,t}^b(c; \boldsymbol{x})}{\partial \log x_d}, \tag{16}$$

with the true price index $\mathscr{P}_{t_0,t}^b(c; \boldsymbol{x})$ given by

$$\mathscr{P}_{t_0,t}^b(c; \boldsymbol{x}) \equiv \frac{\widetilde{E}^b(c; \boldsymbol{p}_t, \boldsymbol{x})}{\widetilde{E}^b\left(c; \boldsymbol{p}_{t_0}, \boldsymbol{x}\right)}. \tag{17}$$

Proof. See Appendix A.3. □

Lemma 3 defines the true price index in Equation (17), which is the generalization of the definition in Equation (7) under the case with no heterogeneity in characteristics. This index measures the growth from period t_0 to t in the cost-of-living corresponding to a constant level of real consumption c for a consumer with a constant vector of characteristics $\boldsymbol{x}$. As before, we can define an instantaneous inflation function $P_t^b(c;\boldsymbol{x})$ as

$$P_t^b(c;\boldsymbol{x}) \equiv \sum_i \widetilde{\Omega}_i^b(c;\boldsymbol{p}_t,\boldsymbol{x}) \frac{d\log p_{it}}{dt},$$

and express the true price index as $\log \mathscr{P}_{t_0,t}^b(c;\boldsymbol{x}) = \int_{t_0}^t \log P_\tau^b(c;\boldsymbol{x})\, d\tau$.

Lemma 3 introduces a characteristic correction function index for the degree to which changes in the preferences of the consumer affect their real consumption. This index aggregates the relative changes in across characteristics using the values of the corresponding correction functions as weights. Just like the case of the nonhomotheticity correction function in Lemma 1, these characteristic correction functions in turn account for the cumulative cross-product covariance between price inflations and the elasticities of demand with respect to each characteristic:

$$\Gamma_{d,t}^b(c;\boldsymbol{x}) = \int_b^t \left[\sum_{i=1}^I \widetilde{\Omega}_i^b(c;\boldsymbol{p}_\tau,\boldsymbol{x})\, \zeta_{i,d}^b(c;\boldsymbol{p}_\tau,\boldsymbol{x})\, \frac{d\log p_{i,\tau}}{d\tau}\right] d\tau,$$

where $\zeta_{i,d}(q;\boldsymbol{p}_t,\boldsymbol{\theta}) \equiv \frac{\partial \log \widetilde{\Omega}_i^b(c;\boldsymbol{p}_\tau,\boldsymbol{x})}{\partial \log x_d}$ accounts for the elasticity of the expenditure share of good-i with respect to characteristic d.

To see the intuition behind these results, consider an aging consumer and assume that price inflation is on average higher for goods that are elastic with respect to age. As time progresses the expenditure required to maintain the same level of real consumption for the consumer rises due to the reallocation of expenditure toward goods with prices that are rising faster. In this environment, holding prices fixed at their base period, Equation (15) shows that we need to deflate the growth in nominal expenditure by an additional term $\frac{\partial \log \mathscr{P}_{b,t}^b(c_t;\boldsymbol{x}_t)}{\partial age_t}\frac{dage_t}{dt}$ to account for the effect of aging on real consumption growth. In this example conventional measures of real consumption growth are biased upward because they do not account for the fact that, as people age, the relative price of the products they favor increase. As in the case of nonhomotheticity, the sign of the bias inherently depends on the choice of the base period for prices. Holding prices fixed in the final period to express real consumption, conventional measures of real consumption growth are now biased downward since, going backward in time, consumers are getting younger and the relative prices of the products the favor is falling.

3 Approximating Welfare Change with Price Indices

In this section, we use standard price index formulas to construct approximations for the true price indices and measures of real consumption growth inroduced in the previous section without the knowledge of the underlying preferences. We thus generalize the results of the classical index number theory and provide a methodology to approximate the nonhomotheticity correction using cross-sectional data.

3.1 Price Index Formulas

We first define the common price index formulas, which can be computed without the knowledge of the underlying structure on consumer preferences in terms of observed expenditures and prices.

Definition 6 (Price Index Formulas)**.** A price index "formula" is a positive-valued function $\mathbb{P}\left(\boldsymbol{p}_{t_0}, \boldsymbol{q}_{t_0}; \boldsymbol{p}_t, \boldsymbol{q}_t\right)$ of a pair of initial and final vectors of prices and quantities such that

$$\mathbb{P}\left(\boldsymbol{p}_{t_0}, \boldsymbol{q}_{t_0}; \alpha \boldsymbol{p}_{t_0}, \boldsymbol{q}_{t_0}\right) \equiv \alpha,$$

for any $\alpha > 0$.

Price index formulas allow us to aggregate the changes in a vector of prices and quantities as a single price index. In constrast to the true price index defined in Equation (7), index formulas are only functions of observables and we do not require the knowledge of the underlying to be able to compute them. The most common examples include Laspeyres $\mathbb{P}_L$ and Paasche $\mathbb{P}_P$ indices that only use the vector of quantities in the initial and final periods, respectively:

$$\mathbb{P}_L\left(\boldsymbol{p}_{t_0}, \boldsymbol{q}_{t_0}; \boldsymbol{p}_t, \boldsymbol{q}_t\right) \equiv \sum_i s_{i t_0} \left(\frac{p_{it}}{p_{t_0}}\right), \qquad \mathbb{P}_P\left(\boldsymbol{p}_{t_0}, \boldsymbol{q}_{t_0}; \boldsymbol{p}_t, \boldsymbol{q}_t\right) \equiv \left(\sum_i s_{it} \left(\frac{p_{t_0}}{p_t}\right)\right)^{-1}, \tag{18}$$

and the Fisher index formula that is defined as the geometric mean of the two indices $\mathbb{P}_F \equiv (\mathbb{P}_P \cdot \mathbb{P}_L)^{\frac{1}{2}}$. In this paper, we will focus our attention on the geometric $\mathbb{P}_G$ and Törqvist $\mathbb{P}_T$ index formulas defined as

$$\mathbb{P}_G\left(\boldsymbol{p}_{t_0}, \boldsymbol{q}_{t_0}; \boldsymbol{p}_t, \boldsymbol{q}_t\right) \equiv \prod_i \left(\frac{p_{i,t}}{p_{i,t_0}}\right)^{s_{i,t_0}}, \qquad \mathbb{P}_T\left(\boldsymbol{p}_{t_0}, \boldsymbol{q}_{t_0}; \boldsymbol{p}_t, \boldsymbol{q}_t\right) \equiv \prod_{i=1}^{I} \left(\frac{p_{i,t}}{p_{i,t_0}}\right)^{\frac{1}{2}\left(s_{i,t_0}+s_{i,t}\right)}. \tag{19}$$

It is well-known that, under arbitrary preferences, the above price index formulas provide approximations of the true price index for specific levels of real consumption. We combine and consolidate these results in the following lemma.

Lemma 4. *Assume that the expenditure function $E\left(\cdot;\cdot\right)$ is third-order continuously differentiable in* *all its arguments and define*

$$\Delta_p \equiv \max_i \left\{ \left| \log\left(\frac{p_{i,t}}{p_{i,t_0}} \right) \right| \right\}, \qquad \Delta_y \equiv \left| \log\left(\frac{y_{i,t}}{y_{i,t_0}} \right) \right|. \tag{20}$$

For any underlying preferences, any first-order price index formula approximates the true price index $\mathscr{P}^b_{t_0,t}(c)$ *for* $c = c_\tau$ *with* $\tau \in [t_0, t]$ *as*

$$\log \mathscr{P}^b_{t_0,t}(c) = \log \mathbb{P}_I\left(\boldsymbol{p}_{t_0}, \boldsymbol{q}_{t_0}; \boldsymbol{p}_t, \boldsymbol{q}_t\right) + O\left(\Delta^2\right), \qquad I \in \{G, L, P\}, \tag{21}$$

where $\Delta \equiv \Delta_p$ *if the underlying preferences are homothetic and* $\Delta \equiv \Delta_p + \Delta_y$*, otherwise. Moreover, any second-order price index formula approximates the true price index* $\mathscr{P}^b_{t_0,t}(c)$ *with* $c = \sqrt{c^b_{t_0} \cdot c^b_t}$ *as*

$$\log \mathscr{P}^b_{t_0,t}(c) = \log \mathbb{P}_T\left(\boldsymbol{p}_{t_0}, \boldsymbol{q}_{t_0}; \boldsymbol{p}_t, \boldsymbol{q}_t\right) + O\left(\Delta^3\right), \qquad I \in \{T, F, S\}, \tag{22}$$

where the Sato-Vartia index formula $\mathbb{P}_S = \prod_i \left(\frac{p_{i,t}}{p_{i,t_0}}\right)^{\omega_i}$ *where Sato-Vartia weights are proportional to* $\omega_i \propto s_{it}/s_{it_0}/\log\left(s_{it}/s_{it_0}\right)$ *and sum to 1.*

Proof. See Appendix A.3. □

Lemma 4 shows that price index formulas provide local approximations to the true price index, in the sense that the corresponding real consumption level is fixed in between the initial and final periods. Recall, however, that if we assume homothetic preferences the true price index does not depend on the level of real consumption. Thus, under homotheticity, a corollary of the lemma is that Equations (21) and (22) hold *for all levels of real consumption* c and under the stronger error bound Δ_p.

Another implication of Lemma 4 is the classification of price index formulas into two groups: the first group (composed of geometric, Laspeyres, and Paasche index formulas) provide a first-order approximation to the true price index while the second group (composed of Törqvist, Fisher, and Sato-Vartia) provide a second-order approximation.[10] To reflect the accuracy of the approximations for each group, we refer to the first group of index formulas as *first order* index formulas and to the second group as *second order* index formulas.

Thus, if we are willing to restrict our attention to homothetic preferences, the above result provides us with a powerful tool to approximate real consumption growth over long time horizons. In practice, we typically observe a sequence of prices $\{\boldsymbol{p}_t\}_{t=0}^T$ and quantities $\{\boldsymbol{q}_t\}_{t=0}^T$ instead

[10] To prove Lemma 4, we first establish that it holds for geometric and Törqvist indices, and then use Lemma 7 (presented in Appendix A.3.2) that shows that the Laspeyres and Paasche (Fisher and Sato-Vartia) indices are equivalent to the geometric (Törqvist) index to the first (second) order of approximation.

of the full paths of prices and quantities. We can now compare welfare over long horizons for any b by deflating the change in nominal expenditure by a chained index formula according to:

$$\log\left(\frac{c_T^b}{c_0^b}\right)=\log\left(\frac{y_T}{y_0}\right)-\sum_{t=0}^{T-1}\log\mathbb{P}_{I_1}\left(\boldsymbol{p}_t,\boldsymbol{q}_t;\boldsymbol{p}_{t+1},\boldsymbol{q}_{t+1}\right)+O\left(\frac{\Delta}{T}\right),\qquad I_1\in\{G,L,P\},$$
$$=\log\left(\frac{y_T}{y_0}\right)-\sum_{t=0}^{T-1}\log\mathbb{P}_{I_2}\left(\boldsymbol{p}_t,\boldsymbol{q}_t;\boldsymbol{p}_{t+1},\boldsymbol{q}_{t+1}\right)+O\left(\frac{\Delta}{T^2}\right),\qquad I_2\in\{T,F,S\},$$

where $\Delta\equiv\max_i\{\log(p_{iT}/p_{i0})\}$. The error in the approximation goes to zero as the frequency of sampling T increases.

When we allow for nonhomotheticity, we cannot rely on Lemma 4 to approximate change in real consumption in terms of constant base periods. As we saw in Equation (8), we need to deflate the nominal expenditure growth by a composite true price index that is evaluated at fixed levels of real consumption corresponding to the terminal periods (initial and/or final). However, the lemma shows that the chained price index formulas approximate the true indices at levels of real consumption that change over time, and may thus lead to inconsistent measures of real consumption growth. In the next section, we tackle this problem by approximating the nonhomothetic correction function.

Discussion Lemma 4 classifies common price index formulas based on the accuracy of the approximations they provide for true price indices under arbitrary underlying preferences. This approach is in contrast to the standard treatment of index formulas that classifies them based on the underlying family of preferences for which they provide *exact* measures of true price indices (Diewert, 1993). For instance, the Törnqvist price index is exact for the family of preferences that lead to a translog unit cost function.[11] Thus, unlike our approach, the concept of exact price indices still requires us to take a stance on the underlying form of the preference functions.

One crucial step is to define, as in Diewert (1976), the Fisher and Törqvist price indices as *superlative* price indices, on the grounds that they are exact for families of preferences that can provide a second-order approximation to other homothetic preferences, e.g., the translog family. In line with this insight, Diewert (1978) has shown that alternative choices of superlative indices, when chained, lead to very similar estimates for the changes in cost-of-living and real consumption in practice. Lemma 4 formalizes these classical insights from a different angle. Instead of establishing the exactness of different index formulas for distinct families of preferences that may approximate general preferences, the lemma provides bounds on the approximation error of the

[11] As for other examples, the Laspeyres and Paasche indices are exact for Leontief utility functions, and the geometric and Sato-Vartia index formulas are exact for Cobb-Douglas and CES utility functions. The Fisher price index is exact for the family of preferences that lead to quadratic unit cost functions.

reduced-form indices for arbitrary preferences.[12]

As mentioned, these classical results do not allow us to provide precise approximations of real consumption growth over long time horizons beyond the case of homothetic preferences.[13] We next provide a methodology to tackle this problem.

3.2 Approximating the Nonhomotheticity Correction using Price Index Formulas

Consider a setting where we observe the composition of consumption expenditures for a collection of N consumers with heterogeneous levels of income over time. We assume that the consumers face the same sequence of prices $\{\boldsymbol{p}_t\}_{t=0}^T$ and the consumption patterns of consumer n is characterized by the sequence of vector of quantities $\{\boldsymbol{q}_t^n\}_{t=0}^T$. Correspondingly, we let y_t^n and $s_{i,t}^n \equiv p_{i,t} q_{i,t}^n / y_t^n$ denote the total expenditure and the share of good i in total expenditure for consumer n at time t, respectively. We assume that consumers have identical preferences characterized by utility function $U(\boldsymbol{q})$ and twice continuously differentiable expenditure function $E(u; \boldsymbol{p})$.

We first impose the following distributional assumption on the underlying distribution of welfare in the sample of consumers.[14]

Assumption 1. *For all $t \geq 0$, the real consumption across consumers has a probability distribution function that is bounded away from zero over an interval $[\underline{c}, \overline{c}]$.*

First-Order Algorithm We can now introduce an algorithm that starts from the base period b and computes the real consumption of consumers period-by-period forward (and backward) over time. The core step involves a nonparametric series-function approximation, in which we fit the cross-sectional variations in the observed price index formulas across consumers in each period to a sequence of log-power functions $f_k(z) \equiv \{z^k\}_{k=0}^{K_N}$.[15] This nonparametric functional approximation allows us to construct an approximation $\widehat{\mathscr{P}}_{b,t}(c)$ for the Equation (7) and a corresponding approximation $\widehat{\Lambda}_t(c) \approx \Lambda_t(c)$ for the nonhomotheticity correction function in Equation (12).

[12] In line with Equation (22), Diewert (1976) shows that Törqvist index is exact for the family of nonhomothetic preferences characterized by a translog expenditure function, for the true index under the level of real consumption specified in Lemma 4.

[13] Samuelson and Swamy (1974) discuss several examples of such results and provide examples that show how they fail under nonhomotheticity.

[14] In Appendix A.3, we offer an alternative set of assumptions that do not impose probabilistic restrictions on the sample of consumers.

[15] One can apply alternative series-function approximations, using alternative basis functions such as Fourier, Spline, or Wavelets. The results here generalize to such alternative nonparametric methods subject to modified regularity assumptions on the expenditure function and the distribution of real consumption across consumers (Newey, 1997).

Algorithm 1. *Consider a sequence of power functions* $\{f_k(z) \equiv z^k\}_{k=0}^{K_N}$ *for some* K_N *that depends on* N*, where* N *is the number of consumers in the cross-section. Let* $\widehat{c}_b^n \equiv y_b^n$*, define a function* $\widehat{\mathscr{P}}_{b,b}(c) \equiv 1$*. For each* $t \geq b$*, take* $\{\widehat{c}_t^n\}_n$ *and function* $\widehat{\mathscr{P}}_{b,t}(\cdot)$ *as known and apply the following two steps to update to the next period.*

1. *Compute the next period function* $\widehat{\mathscr{P}}_{b,t+1}^{(1)}(\cdot)$*:*

$$\log \widehat{\mathscr{P}}_{b,t+1}(c) \equiv \log \widehat{\mathscr{P}}_{b,t}(c) + \sum_{k=0}^{K_N} \widehat{\alpha}_{k,t} f_k(\log c), \tag{23}$$

where the coefficients $(\widehat{\alpha}_{k,t})_{k=0}^{K_N}$ *solve the following problem:*

$$\min_{(\alpha_{k,t})_{k=0}^{K}} \sum_{n=1}^{N} \left(\pi_t^n - \sum_{k=0}^{K_N} \alpha_{k,t} f_k(\log \widehat{c}_t^n) \right)^2. \tag{24}$$

where $\{\pi_t^n\}_n$ *are the set of first-order price index formulas:*

$$\pi_t^n \equiv \log \mathbb{P}_I(\boldsymbol{p}_t, \boldsymbol{q}_t^n; \boldsymbol{p}_{t+1}, \boldsymbol{q}_{t+1}^n), \qquad I \in \{L, P, G\}. \tag{25}$$

2. *Compute the real consumption in the next period for each consumer:*

$$\log \widehat{c}_{t+1}^n = \log \widehat{c}_t^n + \frac{1}{1 + \widehat{\Lambda}_{t+1}(\widehat{c}_t^n)} \left(\log(y_{t+1}^n / y_t^n) - \pi_t^n \right), \tag{26}$$

where we have defined the approximate nonhomotheticity correction function as:

$$\widehat{\Lambda}_{t+1}(c) \equiv \sum_{k=0}^{K_N} \left(\sum_{\tau=b+1}^{t+1} \widehat{\alpha}_{k,\tau} \right) f_k'(\log c). \tag{27}$$

The following proposition establishes that Algorithm 1 indeed provides an approximation to the growth in real consumption for all $t \geq b$ as K_N and N go to infinity, under appropriate regularity assumptions.

Proposition 1. *If Assumption 1 holds, if* $\Delta \equiv \Delta_p + \Delta_y < 1$ *with* Δ_p *defined as in Equation (20) and* $\Delta_y \equiv \max_{n,i} \left\{ \log(y_{t+1}^n / y_t^n) \right\}$*, and if the expenditure function* $\log E(\cdot;\cdot)$ *is continuously differentiable of order* $m \geq 5$*, then as* N *and* K_N *grow toward infinity, the sequences of real consumptions constructed by Algorithm 1 satisfy:*

$$\log\left(\frac{c_{t+1}^n}{c_t^n}\right) = \log\left(\frac{\widehat{c}_{t+1}^n}{\widehat{c}_t^n}\right) + O(\Delta^2) + O_p\left(K_N^3 \left(\sqrt{\frac{K_N}{N}} \cdot \Delta^4 + K_N^{1-m} \right) \Delta \right). \tag{28}$$

Proof. See Appendix A.3. □

Proposition 1 shows three sources of approximation error in the results produced by Algorithm 1: 1) the original Taylor-series approximation error in the reduced-form price index, which is second-order in Δ, 2) the error due to the approximation of the true price index function $\mathscr{P}_{b,t}(c)$ based on the cross-section of consumers, which falls as we observe more consumers N, and we choose K_N such that $K_N^7/N \to 0$, and 3) the error due to functional approximation using a finite set of basis functions, which falls as we choose a more flexible set of basis functions by increasing K_N and thus reduce the term K_N^{4-m}.

Bias in Conventional Growth Measurements What is the implication for the bias in conventional measurements of welfare change? Let $g_t^n \equiv \log\left(y_{t+1}^n/y_t^n\right) - \pi_t^n$ stand for the uncorrected measure of real cnosumption growth for consumer n. We can express the bias in this measure of real consumption growth relative to the corrected measure found based on Algorithm 1 as

$$g_t^n - \log\left(\frac{c_{t+1}^n}{c_t^n}\right) = \lambda_t^n \times g_t^n, \tag{29}$$

where the value of the *annual* growth bias $\lambda_{b,t}^n$ is a monotonic transformation of the nonhomotheticity correction function:

$$\lambda_t^n \equiv \frac{\Lambda_{t+1}(c_t^n)}{1+\Lambda_{t+1}(c_t^n)}. \tag{30}$$

Equation (29) shows the overall size of the bias depends on two distinct forces: the size of the nonhomotheticity correction function and the size of the measured growth in real consumption. Thus, the bias is likely to be large in environments with a large covariance between price inflation and income elasticities and with fast growth in real income.[16]

How does this bias accumulate when measuring growth over longer horizons from some period t_0 to t using the chain rule? Let $G_t^n \equiv \sum_{\tau=0}^{t-1} g_\tau^n$ denote the cumulative real consumption growth inferred based on the uncorrected chained measures. We can define a cumulative weighted-sum $\lambda_{C,t}^n$ of the annual corrected bias

$$\lambda_{C,t}^n \equiv \sum_{\tau=t_0}^{t-1} \left(\frac{g_\tau^n}{G_\tau^n}\right) \cdot \lambda_\tau^n, \tag{31}$$

such that the bias as a share of measured growth can be expressed as

$$G_\tau^n - \log\left(\frac{c_t^n}{c_0^n}\right) = \lambda_{C,t}^n \cdot G_\tau^n. \tag{32}$$

[16] Expressed as a fraction of the standard measure of real consumption growth, the relative bias only depends on the covariance between price inflation and income elasticity.

As this definition makes clear, the overall bias is higher to the extent that over time periods of fast growth coincide more closely with periods of high covariance between inflation and income elasticities.

Algorithm 1 provides a simple and intuitive approach to correcting for the effects of nonhomotheticity in inferring the true price index based on reduced-form price indices. However, our approximation error bounds here are weaker compared to the baseline case under homotheticity in Equation (22) when we use second-order price index formulas. We next offer an extended iterative method that achieves a second order approximation of the real consumption growth.

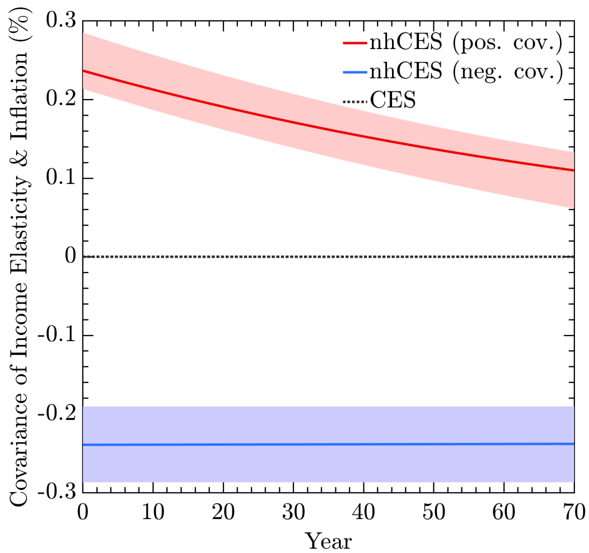
Second-Order Algorithm Algorithm 1 relies on the value of the nonhomotheticity correction evaluated at the previous period's level of real consumption to approximate the current level of real consumption. To construct a second-order approximation for the current real consumption, we need to use the values of the nonhomotheticity correction evaluated both at the previous and current levels of real consumption. Thus, unlike the case of Algorithm 1, we need to solve a fixed-point problem to update the value of real consumption in each period. Algorithm 2 in Appendix A.1 introduced below uses a recursive structure to achieve this second order approximation. Proposition 2 in Appendix A.1 establishes the tighter error bounds achieved by this second-order algorith. This result offers a substantial generalization of the index number theory to the cases involving nonhomotheticity, by showing how the algorithm approximates true price indices for arbitrary preferences solely based on the cross-sectional variations in the observed price index formulas across consumers.

Observable Heterogeneity in Consumer Characteristics We can further generalize Algorithm 1 to additionally account for variations in observable consumer characteristics and to approximate the characteristic correction function introduced in Section 2.3. Algorithms 3 and 4 in Appendix A.1 achieve these generalizations based on first order and second order price index formulas, respectively. The idea is fairly similar to that of Algorithm 1: starting in the base period, we nonparametrically estimate the relationship between the measured price index formulas across consumers and their total expenditures and other characteristics. We then use the estimated relationship with total expenditure and with other characteristics to approximate the corresponding correction functions. Propositions 3 and 4 in Appendix A.1 establish the error bounds on the approximation error for these extensions of our method.

3.3 Simulation

In this section, we perform a simple simulation to illustrate and validate the performance of our algorithms in accounting for the effect of nonhomotheticity on the relationship between true

Figure 1: Example: The Evolution of the Income Elasticity-Inflation Covariance



Note: The figure compares the evolution of the covariance between the expenditure elasticities of sectoral consumption and price inflation based on the example in Section 3.3 for the two cases with positive and negative covariances against the case with homothetic preferences. The shaded area shows the standard deviation of the covariance around the mean in the population of 1,000 households.

price indices and common price index formulas over long time horizons. Comin et al. (2021) have shown that the nonhomothetic CES (nhCES) preferences provide a suitable account of the cross-sectional relationship between household income and the composition of expenditure among three main sectors of the economy: agriculture, manufacturing, and services. Following their specification, we assume that the expenditure function satisfies:

$$E(u;\boldsymbol{p}_t) \equiv \left(\sum_{i\in\{a,m,s\}} \omega_i \left(u^{\varepsilon_i} p_{i,t} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}. \tag{33}$$

We use the same parameters as in Comin et al. (2021): $(\sigma, \varepsilon_a, \varepsilon_m, \varepsilon_s) = (0.26, 0.2, 1, 1.65)$, suggesting that services are luxuries (income elasticities exceeding unity) and agricultural goods are necessities (income elasticities lower than unity). We consider a population of a thousand households with an initial distribution of expenditure with a log-normal distribution with a mean corresponding to the average US per-capita nominal consumption expenditure of 3,138 in 1953 and a standard deviation of log expenditure of 0.5 (Battistin et al., 2009). We consider a horizon of 70 years and assume that over this horizon nominal expenditure grows at the constant rate of 4.48% per year (in line with the US data from the period 1953-2019). In each of the cases discussed below, we choose the fixed sectoral demand shifters ω_i in Equation (33) in such a way that in the first period the composition of aggregate expenditure fits the US average shares of sectoral consumption in the three sectors in 1953.[17] We compare the nonhomothetic specification above against a homothetic CES specification with $(\sigma, \varepsilon_a, \varepsilon_m, \varepsilon_s) = (0.26, 1, 1, 1)$.

To examine the role of the covariance between price inflation and income elasticities, we apply the following strategy. We set the inflation rate in manufacturing to be the average inflation rate in the US over the period 1953-2019 of 3.19%. We then consider two cases with positive and negative covariances: the inflation rates in service and agriculture are 1% higher or lower than manufacturing, respectively, in the positive case and the reverse is true in the negative case. Later in the section, we will further consider a wider range of values for these covariances.

The nonhomothetic specification implies that the income elasticity is highest for services and lowest for agriculture. Given the assumption that price inflation is highest in services and lowest in agriculture, the implication is that income elasticities positively covary with price inflation across goods. Figure 1 shows the evolution of the covariance between price inflation and the elasticity of demand for each of three goods with respect to the total expenditure of the household, where each good is weighted by its corresponding expenditure share.[18] This covariance is trivially zero in the case of homothetic preferences since the expenditure elasticities are identical and unity for all three goods.[19]

For the case with the positive covariance, Figures 2a and 2b compare the expenditure function in terms of real consumption $\widetilde{E}^b(c; \boldsymbol{p}_t)$ between the nonhomothetic and homothetic specifications, with the initial and the last periods as the base, respectively. The figures further illustrate how the corresponding expenditure functions change over time. In the homothetic case, the expenditure function always has a log-linear form. Due to the overall inflation in prices, the expenditure function uniformly shifts upward over time for the homothetic CES preferences.

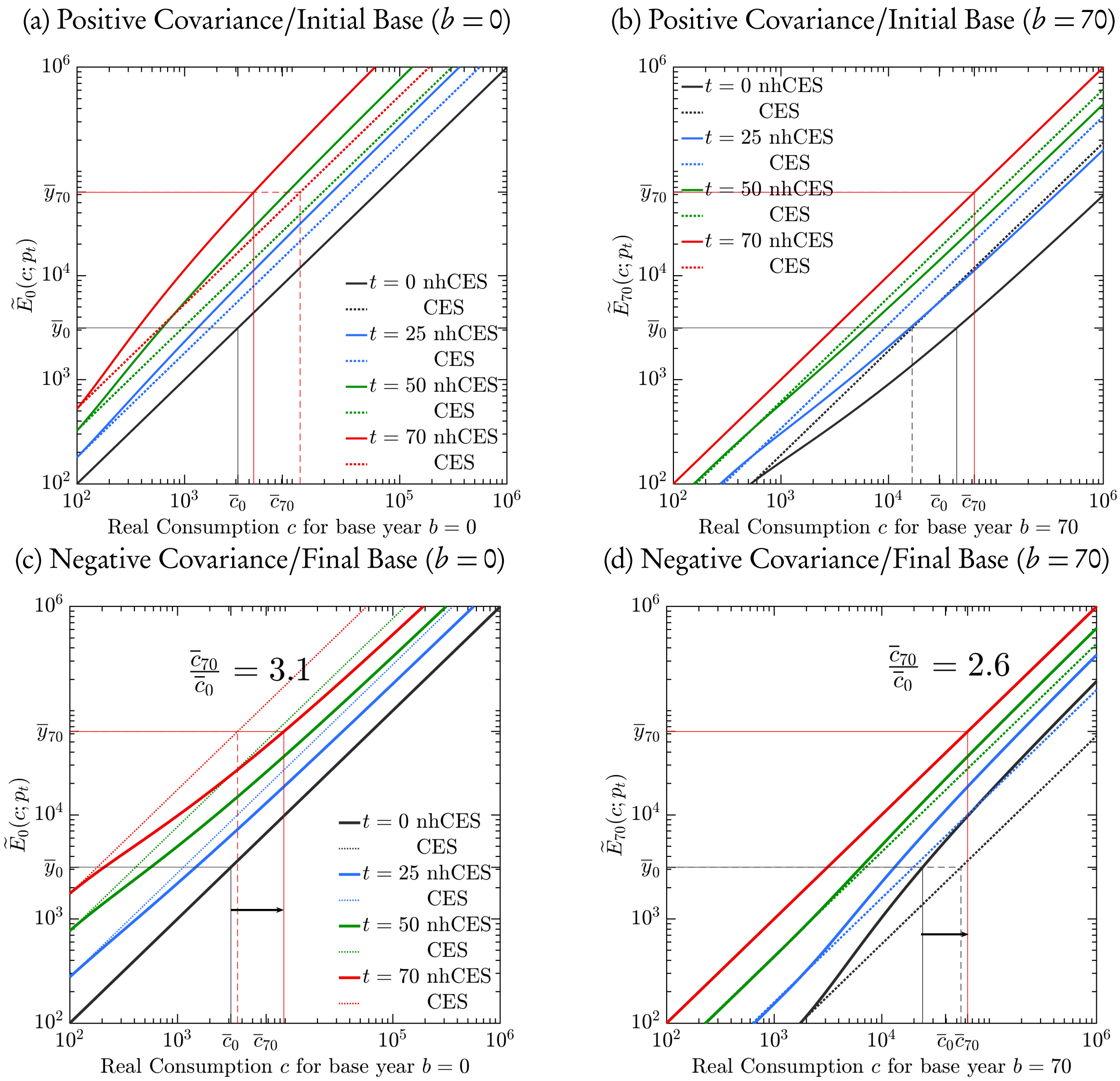
Moving to the nonhomothetic specification, first let us consider the initial period as the base as in Figure 2a. By definition, the expenditure function begins with the same log-linear form in the initial base period. As time passes, the costs of achieving higher levels of real consumption increasing rises relatively faster since to achieve these levels households want to shift their consumption toward goods with faster growing prices (higher inflation). Thus, the expenditure

[17]The corresponding shares in the US based on the BLS data are 0.14, 0.27, and 0.59 for agriculture, manufacturing, and services, respectively.

[18]The expenditure elasticity of demand for good i for preferences in Equation (33) is given by $\eta_{it} = \sigma + (1 - \sigma)\epsilon_i / \overline{\epsilon}_t$, where $\overline{\epsilon}_t$ is the expenditure share weighted average of the ϵ_i parameters across the three goods, based on the expenditure shares in time t. The reason for the downward trend in the case with positive covariance is that, due to gross complementarily among the goods, the expenditure shares of households strongly shifts toward services due to both higher income elasticities and rising prices over the period. The rising concentration of consumption in services mechanically reduces the weighted covariance between price inflation and income elasticities.

[19]Figures A1a and A1b in Appendix C.1 show how the chained geometric and Törqvist price index formulas for households at different quintiles of initial income deviate from the population average over time. As expected, these index formulas are increasing in income in the case with positive covariance and decreasing in the case with negative income elasticity-covariance covariance.

Figure 2: Example: The Expenditure Function $\widetilde{E}^b(\cdot;\boldsymbol{p}_t)$

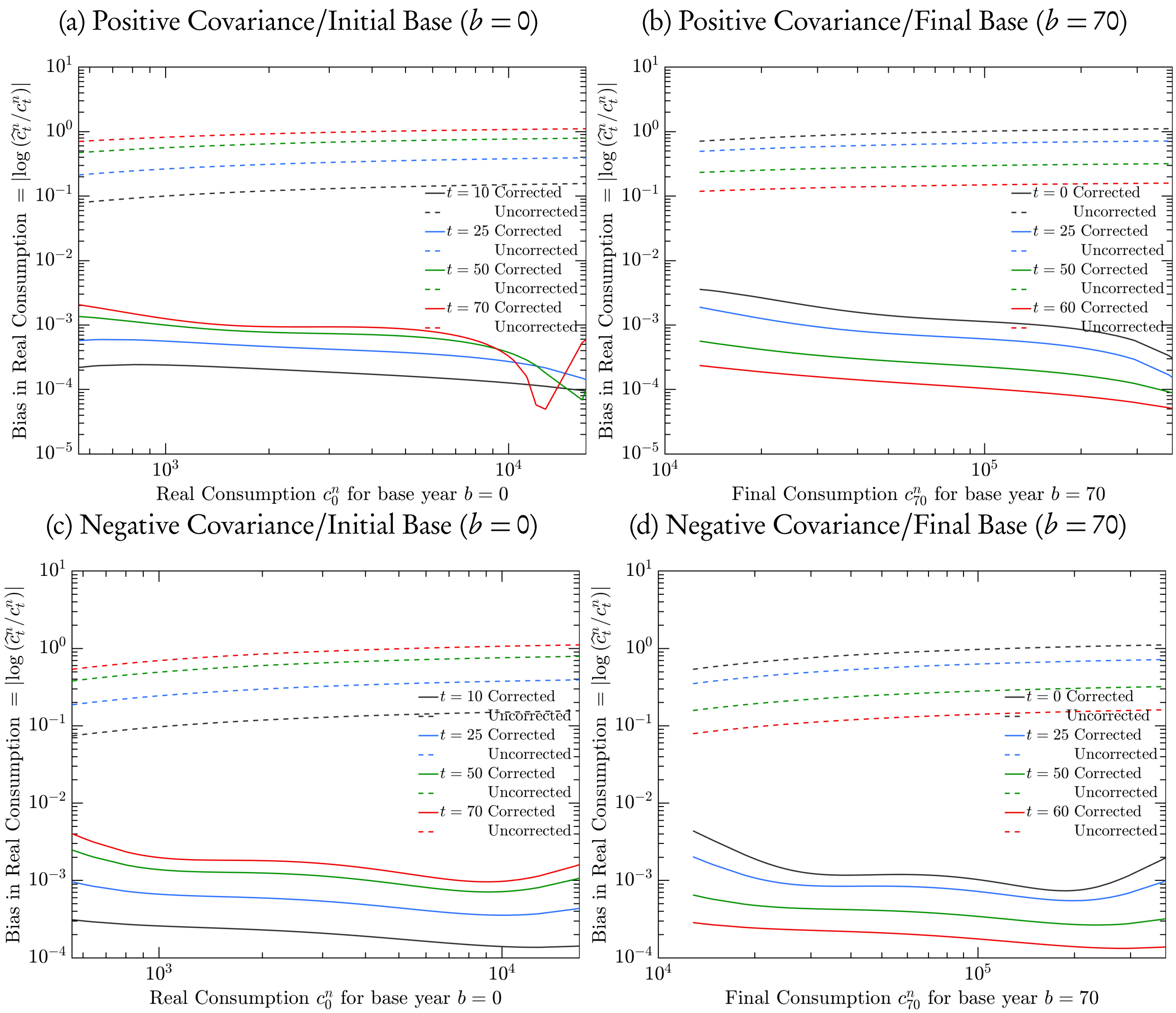


Note: The figure shows the expenditure function defined in terms of real consumption with the initial year as the base, for the preferences defined in Equation (33) with parameters corresponding to a nonhomothetic CES $(\sigma, \varepsilon_a, \varepsilon_m, \varepsilon_s) = (0.26, 0.2, 1, 1.65)$ (nhCES) and homothetic CES $(\sigma, \varepsilon_a, \varepsilon_m, \varepsilon_s) = (0.26, 1, 1, 1)$ functions. Panels (a) and (b) show the results for initial and final periods as the base for the case with positive income elasticity-inflation covariance, respectively. Panels (c) and (d) show the same results for the case with negative income elasticity-inflation covariance.

function defined in terms of real consumption $\widetilde{E}^b(c;\boldsymbol{p}_t)$ increasingly deviates from linearity and becomes more convex as time passes. As the figure shows, the upward shift in the expenditure function is larger compared to the homothetic case for higher levels of real consumption.

Next, consider the final period as base as in Figure 2b. By definition, in this case the expenditure function has a log-linear form in the final period. As we move backward in time, the costs of achieving higher levels of real consumption falls relatively faster since to achieve the corresponding levels of welfare households want to shift their consumption toward goods that are falling

Figure 3: nhCES Example: Nonparametric Approximation of Real Consumption



Note: The figures compare the error in the approximate value of real consumption between the geometric price index formula and the one corrected based on the first-order Algorithm 1. The correct value of real consumption is calculated based on the underlying parameters of the nhCES preferences. The panels show the error for the choices of base period (a) $b = 0$ and (b) $b = 70$ with the positive income elasticity-inflation covariance and (c) $b = 0$ and (d) $b = 70$ with the negative covariance.

faster (higher inflation). Thus, the expenditure function increasingly deviates from linearity and becomes more concave as we move toward the initial period. Crucially, regardless of the choice of the base period, with nonhomothetic preferences and with a positive income elasticity-inflation covariance, the expenditure function is more convex in the later periods.

The figures further illustrate the determination of the growth in real consumption, in the case of the observed average nominal expenditures in the US data, which rises from $\overline{y}_0 = 3,138$ to $\overline{y}_{70} = 63,036$ dollars. In Figure 2a, the real consumption is identical to the nominal expenditure in the base period ($t = 0$) for both preferences, but maps to two distinct values in the final period depending on whether the preferences are homothetic or nonhomothetic. The situation is reversed in Figure 2b, whereby the real consumption is identical to the nominal expenditure in the final period ($t = 70$) for both preferences, but maps to two distinct values in the initial period. Due to the positive covariance between inflation and income elasticities, regardless of the choice of the base period, the growth in the average real consumption is *lower* for households whose preferences are characterized by nonhomotheticity. This is despite the fact that households share the same evolution of nominal expenditures and sectoral prices under both sets of preferences.

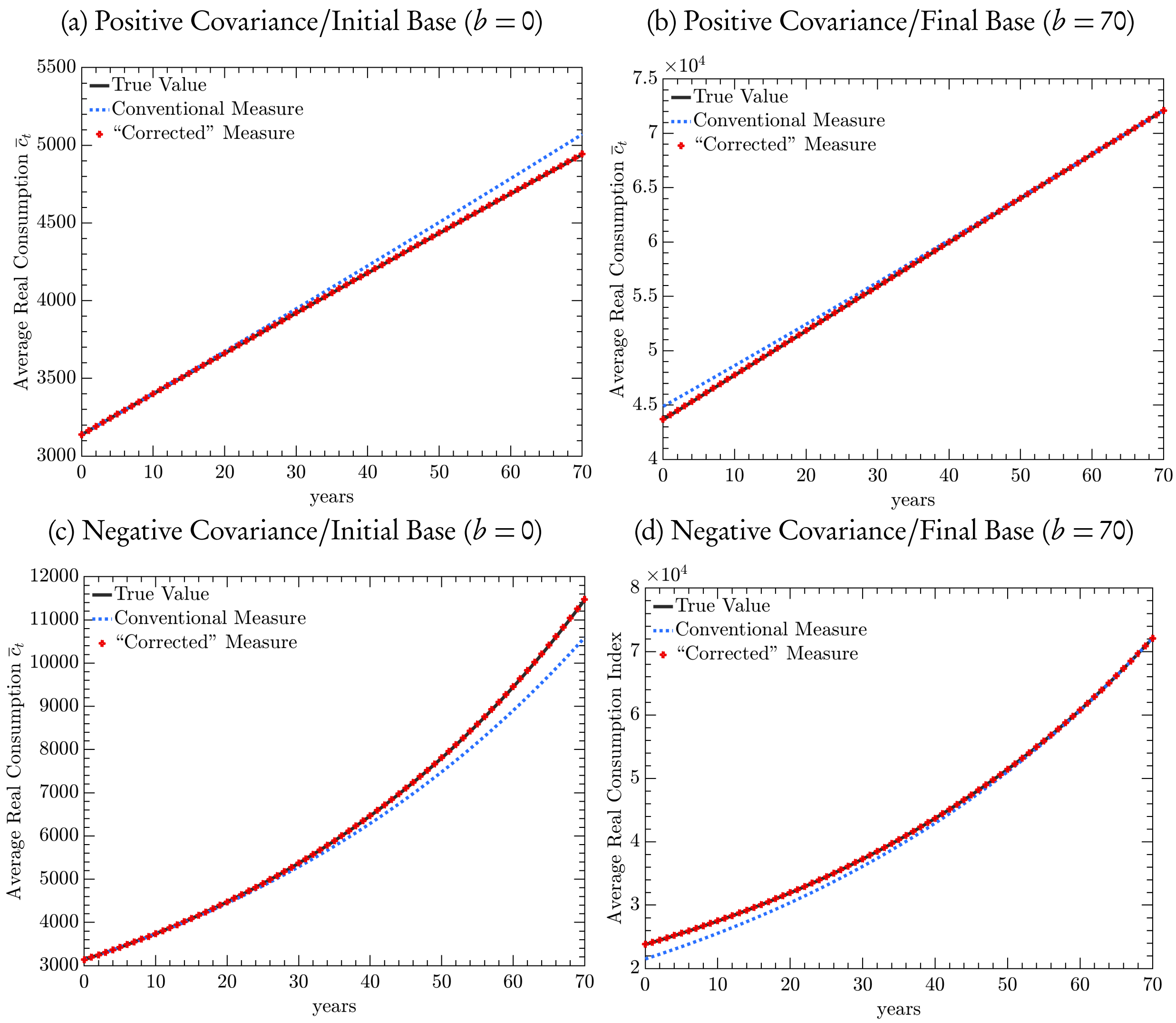
Figures 2c and 2d examine the same patterns in the case with a negative covariance between price inflations and income elasticities. In this case, the expenditure function becomes more concave as time progresses, since now consumers shift the composition of their expenditures toward goods that have lower price inflations. With the initial period as base, the expenditure function begins with a log linear form and becomes more concave as we move forward in time. With the final period as base, the expenditure function ends with a log linear form in the last period and becomes more convex as we move backward in time. Regardless of the choice the base, the growth in average real consumption is *higher* for households whose preferences are characterized by nonhomotheticity.

Figures 3a-3d compare the error in the approximations of real consumption growth with and without the correction using the first-order nonhomotheticity correction following Algorithm 1 for different base periods and income elasticity-inflation covariances.[20] Here, we use the underlying preference parameters to compute the correct value of the real consumption $c_t^{b,n}$ for each household n at each point in time t, and compare that value with the approximate value $\widehat{c}_t^{b,n}$ found by either of the two strategies in each case. As we can see, the standard approach leads to substantially larger errors in the inferred measures of real consumption. After 70 years, this error grows for some households to be of the same order of magnitude as the correct real consumption. Applying the simple first-order correction of Algorithm 1 reduces the error by several orders of magnitude. Figures A2a-A2d in Appendix C compare the sizes of the error when using the first-order approximation approach of Algorithm 1 and that found by the recursive approach of Algorithm 2 in each case, showing that the second-order approximation generally leads to lower approximation errors.

Figures 4a-4d examine the implications for aggregation across households. We compare the

[20]Figures A3a-A3d in Appendix (C.1) show the sizes of the nonhomotheticity correction $\left(1+\Lambda_t^b\left(c_t^b\right)\right)^{-1}$ in Equation (9) for different signs of covariance and choices of base periods. While ratio is unity under homotheticity, it becomes less or greater than unity depending on the choice of the base year and the sign of the covariance in the nonhomothetic case in line with the results of Lemma 2. For instance, with positive covariance, the uncorrected measure of growth based on index formulas overestimates real consumption growth with the initial period as the base, and underestimates it with the final period as the base. The size of this bias is larger for poorer households and grows substantially for all households over time. The former result is in line with the fact that poorer household face a higher dispersion in the sectoral composition of their expenditures, leading to a higher expenditure-weighted covariance of price inflation and income elasticities.

Figure 4: nhCES Example: The Evolution of Average Real Consumption



(a) Positive Covariance/Initial Base ($b = 0$)

(b) Positive Covariance/Final Base ($b = 70$)

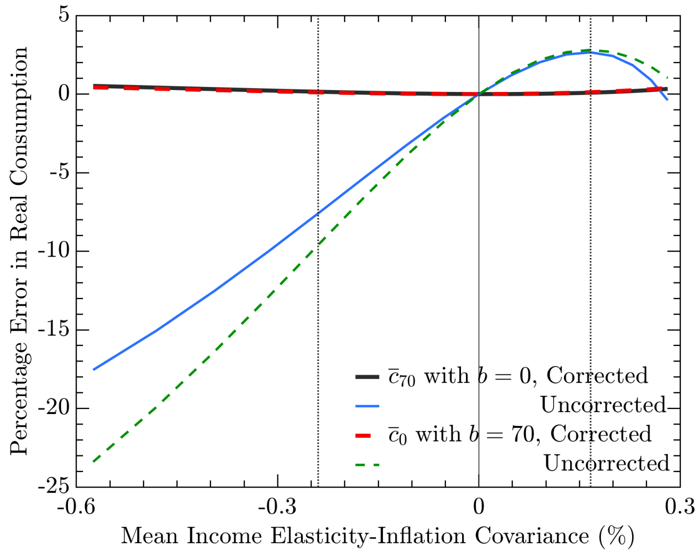
(c) Negative Covariance/Initial Base ($b = 0$)

(d) Negative Covariance/Final Base ($b = 70$)

Note: The figures compare the evolution of the correct value of average real consumption with three different approaches to approximating this value. The standard approach relies on the Törnqvist index defined in terms of aggregate consumption expenditure shares. The average household-level approach uses household-specific expenditure shares, and the "corrected" approach additionally uses the second-order nonhomotheticity correction. The panels show the resulting series for the choices of base period (a) $b = 0$ and (b) $b = 70$ with the positive income elasticity-inflation covariance and (c) $b = 0$ and (d) $b = 70$ with the negative covariance.

evolution of measures of real consumption over time with a number of different approximations. First, we see that the conventional approach based on chaining uncorrected measures of nominal expenditure growth deflated by the Törnqvist index leads to sizable bias depending on the choice of the base period or the covariance between price inflation and income elasticities. Applying our second-order nonhomotheticity correction yields results that are virtually indistinguishable from the correct evolution of real consumption found based on the underlying preferences. Thus, our approach accurately recovers the evolution of the exact index *without the knowledge of the parameters of the demand system.*

Figure 5: Example: Real Consumption Error and Income Elasticity-Inflation Covariance



Note: The figure compares the error in the corrected and uncorrected approximations of the average final and initial real consumption for the initial and final periods as base, respectively, as a function of the mean covariance between price inflations and expenditure elasticities over the period.

To show how the results extend to other ranges of the values of covariance between price inflations and expenditure elasticities, we perform one last exercise with our illustrative example. We consider alternative trends in prices varying the deviations between inflation in services and agriculture from that in manufacturing (fixed to the average level of 3.19%) symmetrically from -2% to +2%. In each case, we apply the same analysis as before, comparing the chained measures of deflated nominal consumption growth with and without our correction. Figure 5 shows how the error in the approximated values of average real consumption in the final period (when the initial period is taken as base) and the initial period (when the final period is taken as base) vary with the corresponding mean of the covariance between income elasticities and price inflations over the period. The figure also indicates the two cases corresponding to the the positive and negative covariance settings studied so far with dotted black lines.

As expected, when income elasticities are uncorrelated with price inflations, the uncorrected measures approximate the correct values with negligible errors. However, as the covariance deviates from zero, the bias in the uncorrected measures grows. As the covariance falls to around -0.6% per year, the error in the uncorrected measure grows to around 20% of the average real consumption across households. As the covariance grows above zero, the error initially rises but ultimately begins to fall for large and positive values of covariance. This is because those scenarios lead to negligible growth in average household real consumption, which mechanically reduces the size of the bias in the reduced-form indices as we saw in Equation (29).[21] In contrast, the error

[21] Figure A4 in Appendix C shows the overall growth in average real consumption over the period as a function of

in the approximation achieved with our nonhomotheticity correction remains close to zero over the entire range of values of the covariance.

4 Empirics

In this section, we apply our approach to real data from the US and quantify the magnitude of the bias in conventional measures of real consumption growth.

4.1 Data

To assess the importance of correcting for nonhomotheticity on the measurement of inequality and long-run growth in welfare, we build a dataset providing total expenditures and expenditure shares for a consistent set of products over time, covering the entire consumption basket of households in the United States. Such a dataset is not readily available from public statistics, due to two challenges. First, the product classification in surveys of consumer expenditures, such as the Consumer Expenditure Survey (CEX), is not the same as the classification used in the price surveys of the Bureau of Labor Statistics. Second, the definition of product categories changes over time in both the expenditure surveys and price surveys.

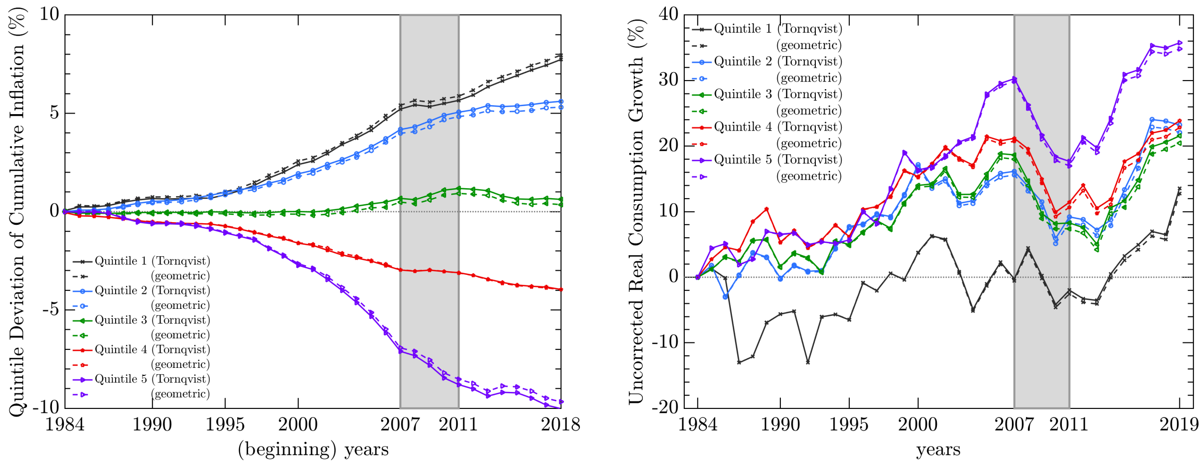
To address these challenges, we build consistent categories over time and a crosswalk delivering a linked dataset of the Consumer Expenditure Survey (CEX) to the Consumer Price Index (CPI) categories. Our preferred dataset tracks 19 categories from 1953 to 2019. Expenditure shares by income quintiles are available each year from 1984 onward. Prior to this date, we use our first-order approximation to the correction for nonhomotheticities, using expenditure shares observed in the 1980s. In addition to implementing the correction for nonhomotheticities, this dataset allows us to compute the inequality in price index formulas over a long time horizon, thus extending prior estimates that have focused on much shorter time series. Finally, to measure the growth rates of consumption by income quintiles, we use the CEX from 1984 onward. Prior to 1984, due to data limitations, for all income groups we use the aggregate growth rate of consumption expenditure per capita, as measured by the Bureau of Economic Analysis (BEA). Appendix B provides a detailed description of the data sources and crosswalks used for our main linked datasets as well as for sensitivity analysis.

To assess the sensitivity of our findings to other data construction choices, we build and study three alternative datasets. First, instead of relying only on the expenditure weights from the Consumer Expenditure Survey, we use the aggregate category-level consumption weights used by the Bureau of Labor Statistics in the official CPI index. To obtain corrected expenditure

the income elasticity-inflation covariance. For positive values of the covariance, the growth diminishes toward zero.

Figure 6: CEX and BLS Data: Measures of Price and Real Consumption

(a) Quintile Deviation of Price Indices (b) Quintile Uncorrected Real Consumption



Note: Panels (a) and (b) show the evolution of cumulative price index formulas and uncorrected measures of real consumption for each quintile of income over the period, respectively. The great recession has been indicated in grey background.

shares for each income quintile, we rescale these aggregate weights across income quintiles using CEX expenditure weight. This approach allows us to perfectly match the aggregate expenditure weights used in the official CPI index, but we can implement it only at the level of 10 broad expenditure categories from 1953 to 2020.

In a second robustness check, we use the CEX household-level data from 1990 to 2019 to estimate income elasticities for a more granular set of product categories. We obtain 61 consistent product categories, linked to CPI price data, for the period 1990-2020. We compute prices indices at the household level and implement the non-parametric non-homotheticity correction from our algorithm. We also allow for covariates, controlling for household age, education, and family size. This approach has the advantage of providing more granular product categories and household characteristics, at the cost of a shorter time period.

In a third robustness check, we implement our nonhomothetic correction for a subset of expenditures for which product-level data is available, using Nielsen data covering consumer packaged goods, or about 15% of aggregate expenditure. This robustness check is motivated by prior work showing that most of the heterogeneity in inflation rates arises at the product level, within detailed product categories (Jaravel, 2019). We assess whether using product-level data meaningfully affects the size of the bias we estimate, at the cost of restricting attention to a subset of total expenditure. To implement this robustness check, we work with the Nielsen data from 2004 to 2015.

4.2 Main Results

Measuring Growth in the Short and Medium Run Figure 6a compares the deviation of the chained price index formulas for different quintiles of income from 1984 to 2019 from the population mean over time. First, we find that price index formulas are decreasing in income over the period, i.e. there is *a negative* covariance between income elasticities and price inflation across goods. Over the course of these 35 years, a gap of around 20 of percentage points has opened up in the chained reduced-form indices between the lowest and the highest quintiles of income. This finding is consistent with the growing literature on "inflation inequality," the fact that inflation rates are higher for lower-income households (e.g., Kaplan and Schulhofer-Wohl, 2017; Jaravel, 2019; Argente and Lee, 2021); our data shows that this trend persists over several decades. Based on our discussions in Section 2, we should expect that uncorrected measures of real consumption underestimate the values of final and initial real consumption when we consider the initial and final period prices as the basis for the welfare comparisons.[22]

Figure 6b compares the evolution of the corresponding uncorrected measures of real consumption, found by deflating nominal expenditure by the index formulas shown in Figure 6a across quintiles over the period. The figure clearly shows that the consumption inequality has risen for nearly three decades (1984-2014), over which real consumption stagnated for households in the lowest quintile while growing for richer households. Whereas the real consumption of households in the highest quintile has grown by around 35% over the entire period, the corresponding growth for the lowest quintile is only around 13%.

Figures 7a and 7b show the evolution of the annual bias in reduced-form indices of real consumption growth for initial (1984) and final (2019) periods as base, respectively. The sizes of these biases are expressed as a share of measured growth, as given by λ_t^n defined in Equation (29) for each quintile n at each time t throughout the period.[23] Here, we have used a quadratic function for fitting the cross-sectional variations in price indices.[24]

More importantly, the values of bias are negative and positive depending on whether the base in the initial or the final period, respectively. The signs of the measured biases are in line with the discussion of our illustrative example in Section 3.3, for the case with negative covariance of

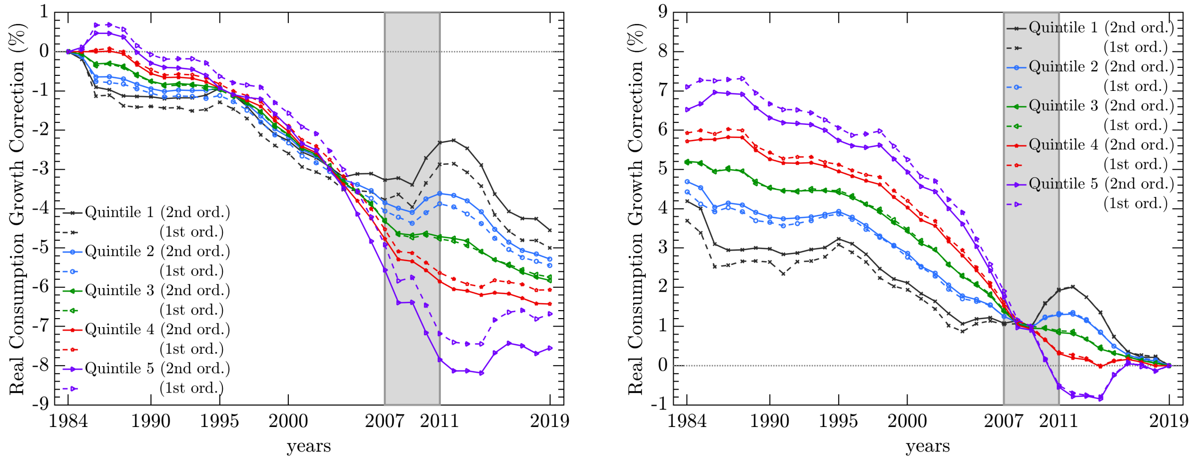
[22]Note that, although the cumulative level of inflation inequality shown in Figure 6a is economically meaningful, it is smaller than the deviations we considered in the illustrative example in Figures A1a and A1b.

[23]In the case of the second order approximation, the value of the bias is given by $\lambda_t^n \equiv \frac{\frac{1}{2}\Lambda_t(\widehat{c}_t^n)+\frac{1}{2}\Lambda_{t+1}(\widehat{c}_{t+1}^n)}{1+\frac{1}{2}\Lambda_t(\widehat{c}_t^n)+\frac{1}{2}\Lambda_{t+1}(\widehat{c}_{t+1}^n)}$.

[24]Figures A5a and A5b show in Appendix C the values of the first and second order nonhomotheticity corrections $\left[1+\Lambda_{t+1}(\widehat{c}_t^n)\right]^{-1}$ and $\left[1+\frac{1}{2}\Lambda_t(\widehat{c}_t^n)+\frac{1}{2}\Lambda_{t+1}(\widehat{c}_{t+1}^n)\right]^{-1}$ for initial (1984) and final (2019) periods as base, respectively. Figures A6a and A6b in Appendix C show the first and second order nonhomotheticity corrections where we instead use a linear approximation for fitting the cross-sectional variations in price indices, in which case the correction is identical for all quintiles. The figure shows that the values of the bias are fairly similar for both approximations.

Figure 7: CEX and BLS Data: Bias in Uncorrected Real Consumption Growth

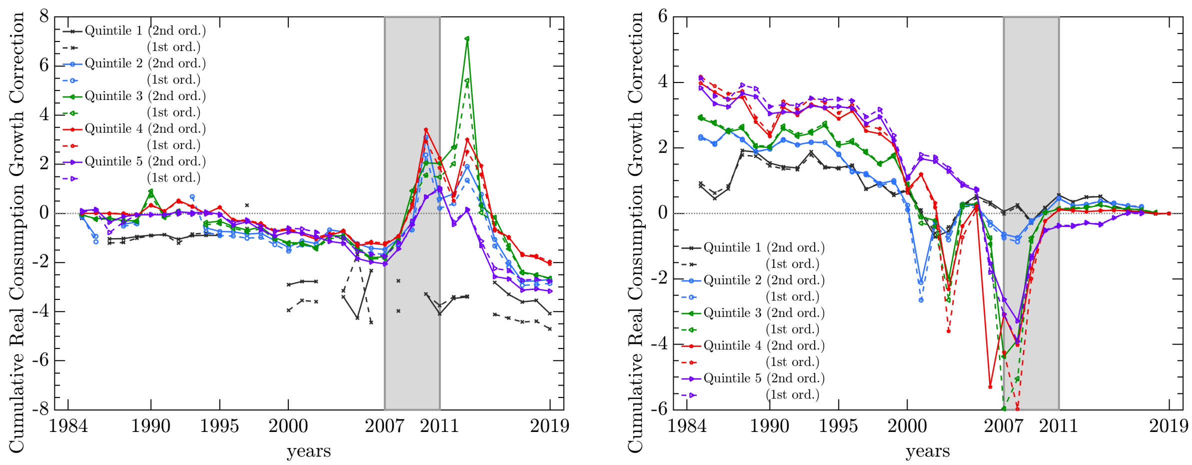
(a) Initial Period as Base ($b = 1984$) (b) Final Period as Base ($b = 2019$)



Note: Panels (a) and (b) show how the evolution of the annual bias in the uncorrected measures of real consumption growth λ_t^n, defined in Equation (29), for different quintiles of income for the initial and final years as the base period.

Figure 8: CEX and BLS Data: Cumulative Bias in Uncorrected Real Consumption Growth

(a) Initial Period as Base ($b = 1984$) (b) Final Period as Base ($b = 2019$)



Note: Panels (a) and (b) show how the evolution of the cumulative bias $\lambda^{c,n}_{1984,(1984,t)}{}^{n}$ and $\lambda^{c,n}_{2019,(t,2019)}{}^{n}$ in the reduced-form measures of real consumption growth $\lambda^n_{b,t}$, defined in Equation (29), for different quintiles of income for the initial and final years as the base period. To avoid showing outlier cases, we have dropped quintile-year observations for which the overall real consumption growth, the denominator in Equation (32), is smaller than 0.01.

income elasticities and price inflations (see Figures A3c and A3d). The magnitudes of the bias (and the required correction) grow over time, particularly for households in higher income quintiles. For the highest income quintile, the uncorrected measures underestimate the annual growth in the final year and overestimate it in the initial year by around 8%, when the base period is taken to be the initial and the final periods, respectively. The sizes of the biases for the lowest quintile falls to roughly 4% of the measured growth.

Figures 8a and 8b show the corresponding values of cumulative bias (and correction) $\lambda^{1984,}_{C,(1984,t)}$ and $\lambda^{2019,n}_{C,(t,2019)}$ for the initial and final base vectors of prices, defined in Equations (31) and (??). As we may expect, the cumulative bias grows in tandem with the annual bias over the period, albeit at a slower pace due to the fact that the former is a weighted average of the latter. The patterns become disorderly during the great recession since the weights in Equation (31) become negative due to the fall in real consumption over this period. But over the entire horizon, the uncorrected measures underestimate growth by up to 4% with respect to the initial base prices, while overestimating it by up to 4% with respect to the final base prices.

Figures 9a and 9b show the evolution of the relative correction in each quintile's *level* of real consumption. These values correspond to the relative size of the corrected measures of real consumption compared to the uncorrected measures, when expressed in terms of the initial and final periods, respectively.[25] The figure shows the relative values of real consumption based on both the first and second order approximations. With few exceptions, the uncorrected measures *underestimate* the value of real consumption. For instance, when we express real consumption in terms of constant 1984 base prices, the magnitude of this underestimation grows exponentially over time all the way until the great recession, where it falls substantially as the real consumption of households also collapses. It then rebounds quickly over the last decade, growing to about 1 percentage point for the highest quintile.[26]

Figures 10a and 10b compare the evolution of *average* real consumption across quintiles as given by corrected and uncorrected measures, when expressed relative to the uncorrected measure of *aggregate* real consumption.[27] The former measures average real consumption across quintiles of income based on each the composition of expenditure in each quintile. The latter measure is real consumption based on the aggregate composition of consumption expenditure. Relative to the aggregate measure that ignores heterogeneity, the average real consumption found using quintile-level, uncorrected indices imply higher growth over the period, i.e., higher final real consumption with initial base prices and lower initial real consumption with final base prices. This pattern is driven by the fact that households at higher quintiles have witnessed disproportionately higher growth in their real consumption. More importantly, our nonhomotheticity correction suggests that the uncorrected measures, even when accounting for heterogeneity, underestimate average real consumption. The gap between the corrected and uncorrected approximations of

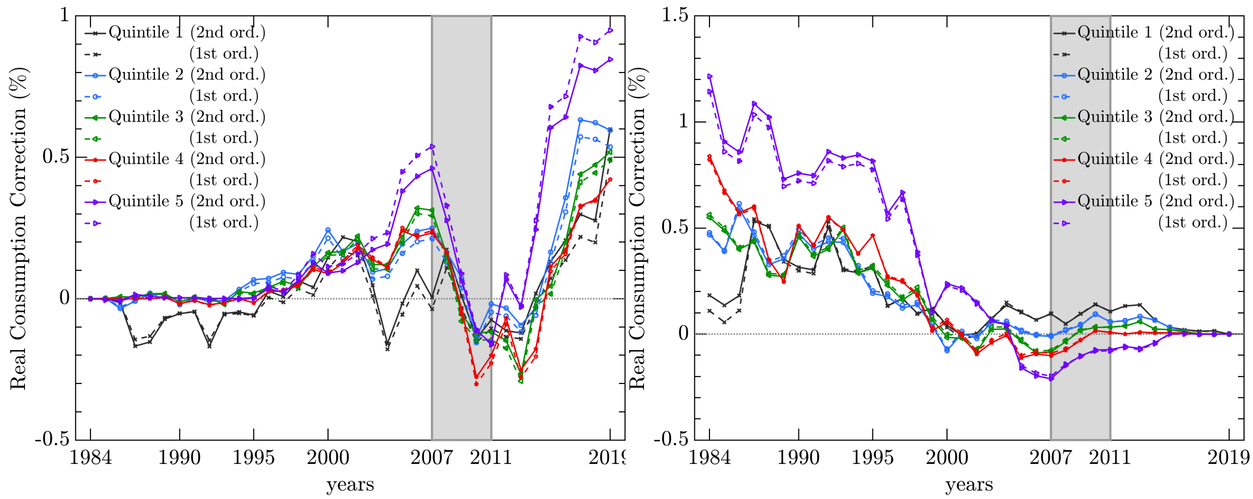
[25]Figures A7a and A7a in Appendix C show the same results when we use a linear approximation for fitting the cross-sectional variations in price indices.

[26]The contribution of the correction to the measurement of inequality in real consumption is negligible in the case with 1984 prices as base. With 2019 prices as base, the correction lowers the measured rise in inequality of real consumption by around 5% (1 percentage points lower from a baseline of 19.5 percentage points).

[27]Figures A8a and A8b in Appendix C presents the same graph using a linear, as opposed to a quadratic, approximation for fitting the cross-sectional variations in price indices

Figure 9: CEX and BLS Data: Corrected relative to Uncorrected Real Consumption

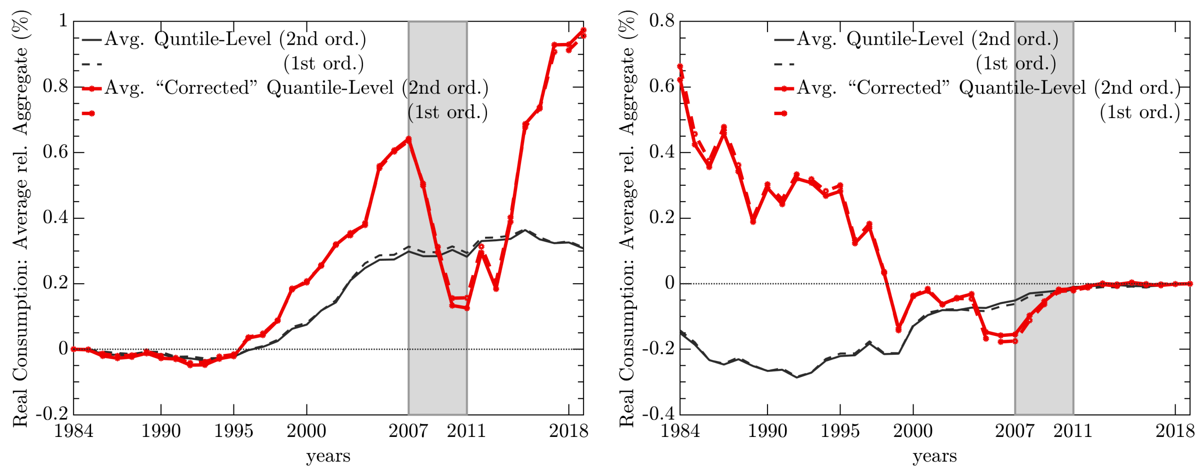
(a) Initial Period as Base ($b = 1984$) (b) Final Period as Base ($b = 2019$)



Note: Panels (a) and (b) show how the ratio of the corrected to the uncorrected measures of real consumption vary over time for each quintile of initial real consumption for the initial and final years as the base period, respectively. The great recession has been indicated in grey background.

Figure 10: CEX and BLS Data: The Evolution of Average Real Consumption

(a) Initial Period as Base ($b = 1984$) (b) Final Period as Base ($b = 2019$)

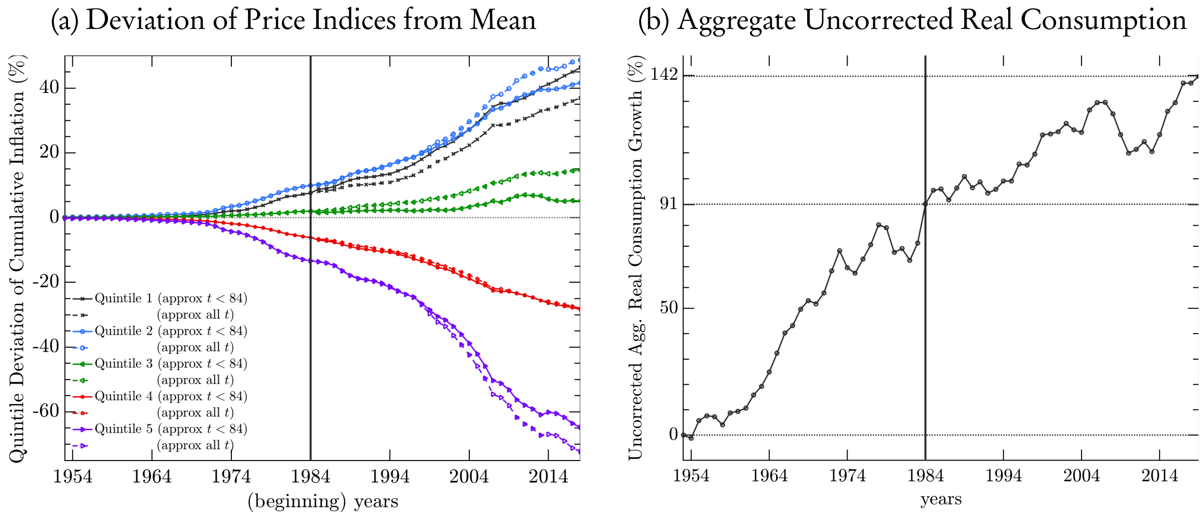


Note: The figure shows the evolution the average corrected and uncorrected measures of real consumption across quintiles relative to the measure of aggregate real consumption that ignores income heterogeneity. The latter defines the price index formulas using aggregate consumption expenditure shares to deflated nominal expenditure. Panels (a) and (b) show the correction for the initial and final years as the base period, respectively. The price index formulas used for the 2nd and 1st order approximations are geometric and Tornqvist indices, respectively.

average real consumption grows to around 0.8% over the period.

Measuring Growth in the Long Run Our next exercise is to extend the horizon of our comparisons of real consumption all the way back to the early 1950s. As previously indicated, we do not have access to systematic disaggregated data on the household-level composition of consump-

Figure 11: CEX and BLS Data: Price Indices and Deflated Expenditure Over the Extended Horizon (1954-2019)



Note: Panel (a) shows the deviation of cumulative geometric price index formulas for different quintiles from the mean, where we have used the 1984 cross-section to define the expenditure shares for each quintile (approx all t). For comparison, the figure also shows the chained indices using all the cross-sections after 1984 (approx $t < 84$). Panel (b) shows the growth in uncorrected real consumption using the aggregate composition of consumption expenditure. The growth in years 1984 and 2019 are indicated.

tion expenditure before 1984. To utilize our approximation of nonhomotheticity correction, we apply the following approximation to express real consumption in terms of constant 1984 prices:

$$\log\left(\frac{\overline{c}_t^{1984}}{\overline{c}_{t+1}^{1984}}\right) \approx \frac{1}{1+\widehat{\Lambda}_t^{1984}\left(\overline{c}_{t+1}\right)} \log\left(\frac{\overline{y}_t/\overline{y}_{t+1}}{\mathbb{P}_G\left(\boldsymbol{p}_{t+1}, \overline{\boldsymbol{q}}_{t+1}; \boldsymbol{p}_t, \overline{\boldsymbol{q}}_t\right)}\right), \qquad t < 1984, \tag{34}$$

To find the estimated value of the nonhomotheticity correction function $\widehat{\Lambda}_t^{1984}\left(\overline{c}_{t+1}\right)$, we apply Algorithm 1 over a single period between 1984 and any given year t, using the 1984 cross-section of households and the cumulative inflation in prices from 1984 and that year. We alternatively consider expressing real consumption in constant 2019 prices, in which case we apply Algorithm 1 to chain the observed cross sections between 2019 and every year after 1984 and then use the approximation above to extend the analysis to the years before 1984. Due to the absence of cross-sectional data in the intervening period, we use a linear approximation of the cross-sections of price indices as our baseline analysis, to avoid overfitting the composition of consumption in 1984 to the previous decades.

Figure 11a shows how the evolution of the gaps in the cumulative geometric price index formulas across different quintiles over the entire time horizon from 1954 to 2019. As explained above, in the absence of cross-sectional data before 1954, we use the 1984 cross-section to compute the index over the entire period. For the post-1984 period, the figure compares this unchained index against the chained geometric index that uses the cross-sectional data between 1984 and 2019.

Over this period, we find that the two yield similar pictures for the inflation-income relationship: the overall inflation gap between the highest and the lowest quintile based on the unchained index is within 1% of the value of this gap based on the chained index.

Figure 11b shows the growth in aggregate real consumption as measured by the uncorrected measures of real consumption from 1954 to 2019. The growth in the first three decades is more rapid, reaching to 91% by 1984. The growth in the remainder of the period is slower, such that the cumulative growth by 2019 reaches 142% only.

Figure 12a shows the evolution of the annual $\overline{\lambda}_t^b$ and cumulative $\overline{\lambda}_{C,(t_0,t)}^b$ growth bias in reduced-form measures of aggregate real consumption growth based on our nonhomotheticity correction scheme. First, let us consider the case with constant 1984 prices as base. We choose this year as the base due to the availability of the earliest cross-section in this year. However, this year also happens to be roughly in the middle of the entire 75-year horizon of our data. The figure compares the values of annual growth bias under two approaches: first, when we use the 1984 cross-section for the entire period, and second, when we use this cross-section for all periods before and up to 1984 and use available data on the following cross-sections for all periods between 1984 and 2019. We find it reassuring for our pre-1984 analysis that the bias (and the resulting correction) are similar for both approaches for the post-1984 period.

Recall that the covariance between inflation and income elasticities are negative both before and after 1984 (Figure 11a). As a result, when compared to real consumption exprssed in 1984 prices, the bias in the uncorrected measure of growth changes sign before and after 1984. Figure 12a shows that the annual (cumulative) bias falls from +4.3% (+3.0%) of the uncorrected measured growth in 1954 to around -5.7% (-2.7%) by 2019. In contrast, when we use the prices in 2019 as our base, the annual (cumulative) bias of measured growth is positive over the entire period, reaching to 9.6% (6.9%) by 1954.

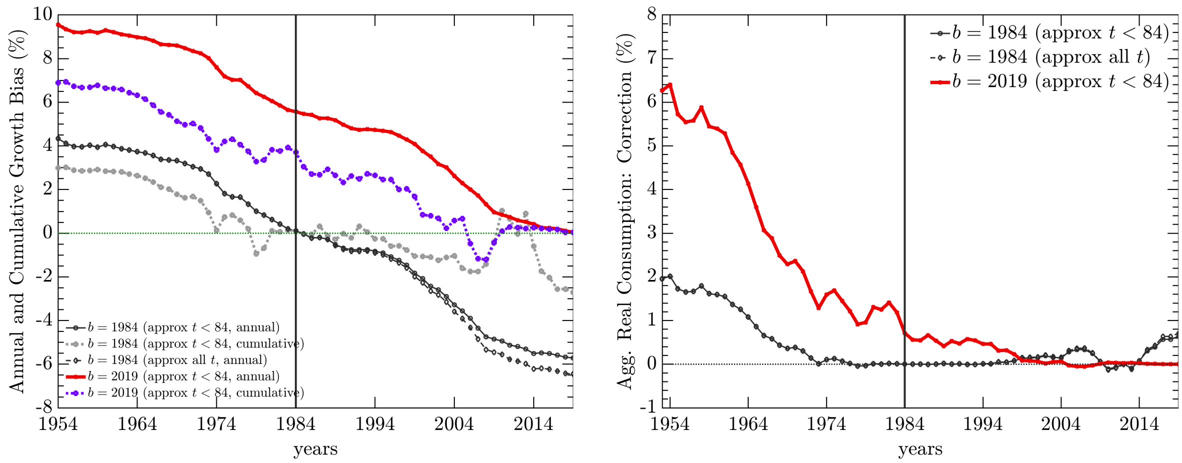
Figure 12b shows the corresponding evolution of the corrected *level* of aggregate real consumption relative to the uncorrected measure.[28] When expressed in constant 1984 prices, standard indices of real consumption underestimate aggregate real consumption in 1954 and 2019 by around 2.0% and 0.6%, respectively. In contrast, when we express real consumption in constant 2019 prices, the uncorrected indices overestimate the level of real consumption in 1954 by around 6.3%.

Expressed in terms of real consumption *growth*, from the perspective of prevailing prices in 2019, the standard measures based on geometric price index formulas overestimate (per capita) real consumption growth over the past 75 years by around 14.3 percentage points, or 22 basis points annually. The majority of this overestimation comes from the first half of the data (1954-

[28]Figures A9a and A9b in Appendix C presents the same graph using a quadratic approximation for fitting the cross-sectional variations in price indices

Figure 12: CEX and BLS Data: Correction Over Longer Horizon

(a) Nonhomotheticity Growth Bias (b) Real Consumption Correction



Note: Panels (a) and (b) show the evolution of the nonhomotheticity correction and the corrected relative to uncorrected measure of real consumption using aggregate consumption expenditure shares.

1984) that contains faster growth rates in the early postwar era. The cumulative overestimation over this 30-year period is around 10 percentage points, or 32 basis points annually.

The corrected and uncorrected measures of growth appear much closer if real consumption is expressed from the perspective of 1984 prices, in which case the bias before and after 1984 changes sign, leading to on overestimation of only 3.1 percentage points over the entire 75-year period (or 5 basis points annually). However, the uncorrected measures contain more substantial error if we use the same constant vector of 1984 prices to compare real consumption between other periods, e.g., between 1954 and 1984. Over this shorter 30-year period, the overestimation of growth is even larger than that over the entire 75-year period, reaching 3.7 percentage points, or 12 basis points annually. These results highlight that the uncorrected measures do not allow us to perform consistent quantitative comparisons of welfare between arbitrary periods.

4.3 Robustness

Using the three alternative datasets previously described, we show that the previous results are robust to data construction and aggregation choices: the size of the nonhomotheticity correction is relatively small with modest inflation and income growth, in particular over the past three decades in the US, but grows substantially when we consider large inflation and growth regimes, as is the case when we extend the analysis to include the entire postwar US experience. All the figures are included in Appendix C.

Official CPI Expenditure Weights First, we repeat the analysis using the aggregate category-level consumption weights used by the Bureau of Labor Statistics in the official CPI index. Using this dataset, the relationship relationship between household income is negative (Figure A10a), as in our baseline dataset 6a. We find that the higher level of aggregation in this alternative dataset attenuates both measured real consumption growth (FigureA10b) and the cross-sectional variations in the price index formulas across household groups, but they still remain decreasing in household income. Figures A11a-A12b show the evolution of the nonhomotheticity bias and the corrected measures of real consumption across different quintiles over the period 1984-2019 with the robustness dataset, which are somewhat smaller than those we found in Figures 7a-9b with the baseline dataset. When we compute the evolution of average real consumption over the period, we find smaller corrections for both initial (1984) and final (2019) base periods (Figures A13a-A13b). Despite the smaller magnitudes due to the more aggregate nature of the BLS data, the overall patterns are broadly similar between the two datasets.

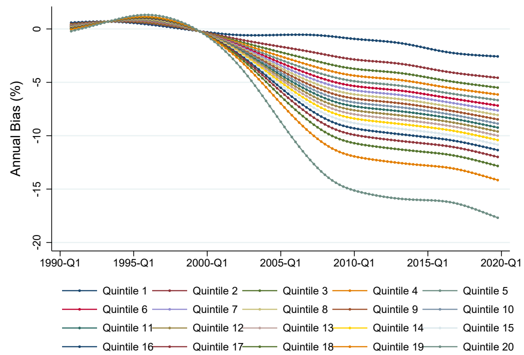
When we consider the longer horizon of the postwar US experience in the robustness dataset, we again find sizable corrections (Figures A14a and A14a). Depending on whether we express real consumption in constant 1984 or 2019 prices, the conventional measures overestimate annual rates of growth in real consumption by 11.3 and 12.4 basis points, respectively.

Household-level CEX Data linked to CPI inflation rates Second, we repeat the analysis using household-level expenditure data from the Consumer Expenditure Survey (CEX) from 1990 to 2019, which we link to inflation data from the CPI across 61 product categories. We implement our algorithm with household-level price indices to compute the correction for nonhomotheticity non-parametrically. The results are reported in Figures 13a-13d, holding prices fixed in the baseline period.[29] Figure 13a reports the results without controlling for household characteristics, and shows that the nonhomotheticity correction is sizable for the largest income vingtiles: after 29 years, the annual bias in the conventional measure of real consumption growth reaches 18%. The bias is smaller for less affluent households. Figure 13b shows the biases for the cumulative growth in real consumption. The bias is meaningful for the richest households. The conventional measure of real consumption growth, computed with a geometric price index, indicates that real consumption was 50% for the top 5% of the expenditure distribution, while growth reached 55% according to the corrected measure accounting for nonhomotheticities. For lower quantiles, the corrected growth measures tend to be lower than the conventional measure. In Figures 13c and 13d, we conduct the same analysis with household-level controls for age, education and family size. The patterns are similar.

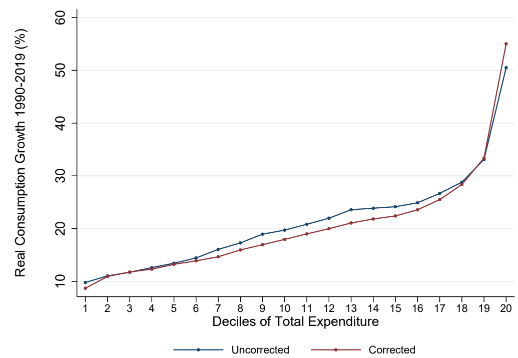
[29] We use a HP-filter to smooth the data; the results obtained without HP-filter are similar.

Figure 13: CEX and BLS Data: Household-Level CEX Data

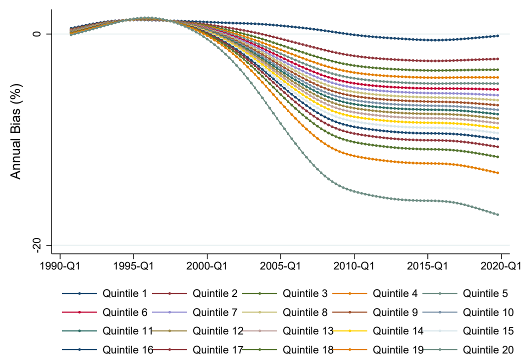
(a) Annual bias, no covariates



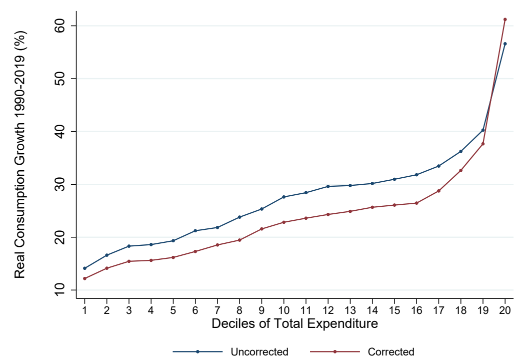
(b) Corrected and uncorrected growth, no covariates



(c) Annual bias, with covariates



(d) Corrected and uncorrected growth, with covariates



Note: Panels (a) and (b) show the annual bias in real consumption growth and the cumulative growth between 1990 and 2019 for different vingtiles of income, without using household characteristics (age, family size, education) as covariates. Panels (c) and (d) show the same results after controlling for these household characteristics.

Nielsen Scanner Data Third, we consider the Nielsen data as another robustness check to examine the nonhomotheticity correction in a setting with more disaggregated consumption data. The data covers a much shorter time horizon (2004-2014), but the annual level of inflation inequality is stronger, with substantially lower levels of inflation for higher income households (Figure A15). The magnitude of the annual correction reaches 4% of the uncorrected measures after only a decade (Figures A16a-A17b). Overall, these results indicate the robustness of the findings obtained with our baseline dataset.

5 Conclusion

In this paper, we extended insights from classical demand theory to the cases where the composition of demand depends on income (nonhomotheticity) and other consumer characteristics. We obtained a procedure for nonparametric measurement of consumer welfare, providing a theoretically consistent measure of real consumption while imposing minimal restrictions on the underlying preferences. This approach remains valid under any observable heterogeneity and requires only data on spending patterns in a cross-section of consumers. We showed the practical relevance of the correction for nonhomotheticities when computing long-run growth in consumer welfare. With our correction taking 2019 prices as base, growth in consumer welfare is significantly attenuated in the United States in the post-war era, due to the combination of fast growth and lower inflation for income-elastic product categories. Extending this analysis to other countries and time periods, as well as to the measurement of purchasing power parity (PPP) indices across countries with preference heterogeneity, is a promising direction for future research.

References

Argente, David and Munseob Lee, "Cost of Living Inequality during the Great Recession," *Journal of the European Economic Association*, 2021, *19* (2), 913–952. 2, 7, 33

Atkin, David, Benjamin Faber, Thibault Fally, and Marco Gonzalez-Navarro, "Measuring Welfare and Inequality with Incomplete Price Information," 2020. 6, 13

Baqaee, David and Ariel Burstein, "Welfare and Output with Income Effects and Demand Instability," 2021. 2, 6, 13

Battistin, Erich, Richard Blundell, and Arthur Lewbel, "Why Is Consumption More Log Normal than Income? Gibrat's Law Revisited," *Journal of Political Economy*, 2009, *117* (6), 1140–1154. 24

Comin, Diego, Danial Lashkari, and Marti Mestieri, "Structural change with long-run income and price effects," *Econometrica*, 2021, *89* (1), 311–374. 4, 24

Deaton, Angus and John Muellbauer, "An almost ideal demand system," *American Economic Revieweview*, 1980, *70* (3), 312–326. 6

Diewert, W Erwin, "Exact and superlative index numbers," *Journal of Econometrics*, 1976, *4* (2), 115–145. 6, 19, 20

— , "Superlative Index Numbers and Consistency in Aggregation," *Econometrica*, 1978, *46* (4), 883–900. 19

Diewert, Walter E, "The economic theory of index numbers: a survey," in Walter E Diewert and Alice O Nakamura, eds., *Essays in Index Number Theory*, Emerald Group Publishing Ltd, 1993, chapter 7, pp. 163–208. 6, 10, 19

Engel, Ernst, "Die Produktions- und Consumtionsverhältnisse des Königreichs Sachsen," *Zeitschrift des Statistischen Büreaus des Königlich Sächischen Ministeriums des Innern*, 1857, *8-9*, 1–54. 2

Fajgelbaum, Pablo D and Amit K Khandelwal, "Measuring the Unequal Gains from Trade," *Quarterly Journal of Economics*, 2016, *131* (3), 1113–1180. 6, 13

Hobijn, Bart and David Lagakos, "Inflation inequality in the United States," *Review of Income and Wealth*, 2005, *51* (4), 581–606. 7

Hulten, Charles R, "Divisia Index Numbers," *Econometria*, 1973, *41* (6), 1017–1025. 9

Jaravel, Xavier, "The unequal gains from product innovations: evidence from the US retail sector," *Quarterly Journal of Economics*, 2019, *134* (2), 715–783. 2, 7, 32, 33

Kaplan, Greg and Sam Schulhofer-Wohl, "Inflation at the household level," *Journal of Monetary Economics*, 2017, *91*, 19–38. 2, 7, 33

Klick, Josh and Anya Stockburger, "Experimental CPI for lower and higher income households," 2021. 2

McGranahan, Leslie and Anna Paulson, "Federal Reserve Bank of Chicago Demographic Group : 1983-2005 Inflation Experiences by Demographic Group : 1983-2005 Federal Reserve Bank of Chicago," 2006. 7

Newey, Whitney K., "Convergence rates and asymptotic normality for series estimators," *Journal of Econometrics*, 1997, *79* (1), 147–168. 20, A18

Pollak, Robert A, "The theory of the cost-of-living index," in W Erwin Diewert, ed., *Price Level Measurement*, Elsevier North Holland Inc., 1990, pp. 5–77. 6

Redding, Stephen J. and David E. Weinstein, "Measuring Aggregate Price Indices with Taste Shocks: Theory and Evidence for CES Preferences," *Quarterly Journal of Economics*, 2020, *135* (1), 503–560. 6

Samuelson, Paul A and S Swamy, "Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis," *American Economic Review*, 1974, *64* (4), 566–593. 2, 6, 10, 20

A Theory Appendix

A.1 Second-Order Algorithm

Algorithm 2. *Let $\widehat{c}_b^n \equiv y_b^n$, define function $\widehat{\mathscr{P}}_{b,b}(c) \equiv 1$, and consider sequence of log-power functions $\{f_k(z) \equiv z^k\}_{k=0}^{K_N}$ where K_N grows with N, the number of consumers in the cross-section. For each $t \geq b$, apply the following steps:*

1. *Initialize the values of the real consumption $\widehat{c}_{t+1}^{n,(0)}$ for each consumer at $t+1$ using Equations (23)–(27) as in Algorithm 1.*

2. *For each $t \geq b$, iterate over the following steps over $\tau \in \{0, 1, \cdots\}$ until convergence for some tolerance $\epsilon \ll 1$:*

   (a) *Solve for the coefficients $\left(\widehat{\alpha}_{k,t}^{\dagger}\right)_{k=0}^{K}$ in the following problem:*

$$\min_{\left(\alpha_{k,t}^{\dagger}\right)_{k=0}^{K}} \sum_{n=1}^{N} \left(\pi_t^n - \sum_{k=0}^{K} \alpha_{k,t}^{\dagger} f_k \left(\log \widehat{c}_{t+1}^{n,(\tau)} \right) \right)^2, \tag{A1}$$

   where $\pi_t^n \equiv \log \mathbb{P}_I \left(\boldsymbol{p}_{t+1}, \boldsymbol{q}_{t+1}^n; \boldsymbol{p}_t, \boldsymbol{q}_t^n \right)$ for $I \in \{G, L, P\}$..

   (b) *Update the next period function $\widehat{\mathscr{P}}_{b,t+1}(\cdot)$:*

$$\log \widehat{\mathscr{P}}_{b,t+1}(c) \equiv \log \widehat{\mathscr{P}}_{b,t}(c) + \sum_{k=0}^{K} \widehat{\beta}_{k,t} f_k(\log c), \tag{A2}$$

   where the coefficients $\left(\widehat{\beta}_{k,t}\right)_{k=0}^{K}$ solve the following problem:

$$\min_{\left(\widehat{\alpha}_{k,t}\right)_{k=0}^{K}} \sum_{n=1}^{N} \left(\pi_t^{*,n} + \rho_t^{n,(\tau)} - \sum_{k=0}^{K} \widehat{\beta}_{k,t} f_k(\log \widehat{c}_t^n) \right)^2, \tag{A3}$$

   with $\lambda_t^{n,(\tau)}$ is defined as:

$$\rho_t^{n,(\tau)} \equiv \frac{1}{4} \sum_{k=0}^{K} \widehat{\alpha}_{k,t}^{\dagger} \left[f_k'(\log \widehat{c}_t^n) + f_k' \left(\log \widehat{c}_{t+1}^{n,(\tau)} \right) \right] \log \left(\frac{\widehat{c}_{t+1}^{n,(\tau)}}{\widehat{c}_t^n} \right). \tag{A4}$$

   and where $\pi_t^{*,n}$ is the value of a second-order price index for consumer n:

$$\pi_t^{*,n} \equiv \log \mathbb{P}_I \left(\boldsymbol{p}_t, \boldsymbol{q}_t^n; \boldsymbol{p}_{t+1}, \boldsymbol{q}_{t+1}^n \right), \qquad I \in \{T, F, S\}.$$

(c) *Update the real consumption in the next period for each consumer:*

$$\log \widehat{c}_{t+1}^{n,(\tau+1)} = \log \widehat{c}_t^{n,(\tau)} + \frac{1}{1+\frac{1}{2}\left[\widehat{\Lambda}_t\left(\widehat{c}_t^n\right)+\widehat{\Lambda}_{t+1}\left(\widehat{c}_{t+1}^{n,(\tau)}\right)\right]}\left(\log\left(\frac{y_{t+1}^n}{y_t^n}\right)-\pi_t^{*,n}\right), \quad \text{(A5)}$$

where we have defined the approximate nonhomothetic correction function as:

$$\widehat{\Lambda}_{b,t+1}(c) \equiv \sum_{k=0}^{K}\left(\sum_{\tau=b+1}^{t+1}\widehat{\beta}_{k,\tau}\right) f_k'(\log c). \quad \text{(A6)}$$

(d) *Stop if* $\max_n \left|\widehat{c}_{t+1}^{n,(\tau+1)} - \widehat{c}_{t+1}^{n,(\tau)}\right| < \epsilon$ *and set* $\widehat{c}_{t+1}^n \equiv \widehat{c}_{t+1}^{n,(\tau+1)}$.

Function $\widehat{\mathscr{P}}_{b,t+1}(c)$ defined in Equation (A2) provides a second-order approximation for the true price index function $\mathscr{P}_{b,t+1}^b(c)$ defined in Equation (12). Equation (A5) then updates our current guess $\widehat{c}_{t+1}^{n,(\tau)}$ about the next-period real consumption. The following proposition then establishes that this iterative process yields a second-order approximation to the true price index between any periods t and $t+1$.

Proposition 2. *Assume that* $\log E(\cdot;\cdot)$ *is continuously differentiable of order* $m \geq 5$. *If Assumption 1 holds, if* $\Delta \equiv \Delta_p + \Delta_y < 1$ *with* Δ_p *defined as in Equation (20) and* $\Delta_y \equiv \max_{n,i}\left\{\log(y_{t+1}^n/y_t^n)\right\}$, *then the sequences of real consumptions* $\widehat{c}_t^n$ *constructed by Algorithm 2 satisfy:*

$$\log\left(\frac{c_{t+1}^n}{c_t^n}\right) = \log\left(\frac{\widehat{c}_{t+1}^n}{\widehat{c}_t^n}\right) + O\left(\Delta^3+\epsilon\right) + O_p\left(K_N^3\left(\sqrt{\frac{K_N}{N}}\left(\Delta^3+K_N^{4-m}\right)^2+K_N^{1-m}\right)\Delta\right). \quad \text{(A7)}$$

Proof. See Appendix ??. □

A.2 Approximating Welfare Change with Observed Heterogeneity

Next, we provide an algorithm that allows us to approximate changes in real consumption growth to the first order of approximation. This algorithm closely parallels Algorithm 1 in the previous section: we use the cross-sectional variations in the reduced-form price indices to nonparametrically approximate the correction functions for nonhomotheticity and all covariates.

Algorithm 3. *Let* $\widehat{c}_b^n \equiv c_b^n \equiv c_b^n$ *and define a function* $\widehat{\mathscr{P}}_{b,b}(c;\boldsymbol{x}) \equiv 1$, *and consider a sequence* $\{f_k(c,\boldsymbol{x})\}_{k=0}^{K_N}$ *of log-power functions of* c *and* $\boldsymbol{x}$ *where* K_N *depends on* N, *the number of consumers in the cross-section. For each* $t \geq b$, *apply the following steps:*

1. *Compute the next period function* $\widehat{\mathscr{P}}_{b,t+1}(\cdot;\cdot)$*:*

$$\log \widehat{\mathscr{P}}_{b,t+1}(c;\boldsymbol{x}) \equiv \log \widehat{\mathscr{P}}_{b,t}(c;\boldsymbol{x}) + \sum_{k=0}^{K} \widehat{\alpha}_{k,t} f_k(c,\boldsymbol{x}), \tag{A8}$$

where the coefficients $(\widehat{\alpha}_{k,t})_{k=0}^{K_N}$ *solve the following problem:*

$$\min_{(\alpha_{k,t})_{k=0}^{K}} \sum_{n=1}^{N} \left(\pi_t^n - \sum_{k=0}^{K_N} \alpha_{k,t} f_k(\widehat{c}_t^n, \boldsymbol{x}_t^n) \right)^2, \tag{A9}$$

where $\pi_t^n \equiv \log \mathbb{P}_G(\boldsymbol{p}_t, \boldsymbol{q}_t^n; \boldsymbol{p}_{t+1}, \boldsymbol{q}_{t+1}^n)$.

2. *Compute the real consumption in the next period for each household:*

$$\log \widehat{c}_{t+1}^n = \log \widehat{c}_t^n + \frac{1}{1+\widehat{\Lambda}_{t+1}(\widehat{c}_t^n;\boldsymbol{x}_t^n)} \left[\log(y_{t+1}^n/y_t^n) - \pi_t^n - \sum_{d=1}^{D} \widehat{\Gamma}_{d,t+1}(c_t^n;\boldsymbol{x}_t^n) \cdot \log\left(\frac{x_{d,t+1}}{x_{d,t}}\right) \right] \tag{A10}$$

where we have defined the approximate nonhomotheticity correction function as:

$$\widehat{\Lambda}_{t+1}(c;\boldsymbol{x}) = \sum_{k=0}^{K_N} \left(\sum_{\tau=b+1}^{t+1} \widehat{\alpha}_{k,\tau} \right) \frac{\partial f_k(c,\boldsymbol{x})}{\partial \log c}, \tag{A11}$$

and the following approximation for the characteristic-d correction function:

$$\widehat{\Gamma}_{d,t+1}(c;\boldsymbol{x}) \equiv \frac{\partial}{\partial \log x_d} \log \widehat{\mathscr{P}}_{b,t+1}(c;\boldsymbol{x}) = \sum_{k=0}^{K_N} \left(\sum_{\tau=b+1}^{t+1} \widehat{\alpha}_{k,\tau} \right) \frac{\partial f_k(c,\boldsymbol{x})}{\partial \log x_d}.$$

Proposition 3 establishes bounds on the approximation error of the sequences of real consumption growth found by Algorithm 3. The main additional requirement, compared to Proposition 1, is that we now require the expenditure function to be infinitely differentiable.

Proposition 3. *Assume that the expenditure function* $\log E(\cdot;\cdot)$ *is an analytical function, that* $\Delta_p + \Delta_y < 1$ *with* Δ_p *defined as in Equation (20) and* $\Delta_y \equiv \max_{n,i} \left\{ \log(y_{t+1}^n/y_t^n) \right\}$*, and that the maximum change in the logarithm of the characteristics across consumers in bounded above by a constant* $\Delta_x < 1$ *such that*

$$\max_{d,n} \left| \log\left(\frac{x_{d,t+1}^n}{x_{d,t}^n} \right) \right| \le \Delta_x. \tag{A12}$$

Then, the sequences of real consumptions constructed by Algorithm 3 satisfy:

$$\log\left(\frac{c_{t+1}^n}{c_t^n}\right) = \log\left(\frac{\widehat{c}_{t+1}^n}{\widehat{c}_t^n}\right) + O\left(\Delta^2\right) + O_p\left(K_N^3\left(\sqrt{\frac{K_N}{N}}\cdot\Delta^4 + K_N^{-m}\right)\Delta\right), \tag{A13}$$

where $\Delta \equiv \max\left\{\Delta_p + \Delta_y, \Delta_x\right\}$ *and* m *is any positive number.*

Proof. See Appendix ??. □

With the stronger assumption imposed by Proposition 3 on the differentiability of the expenditure function, we find a tighter bound in Equation (A13). Again, so long as $K_N \to \infty$ and $K_N^7/N \to 0$, the error in our approximation converges to zero.

Below, we further provide a generalization of the second-order approximation for the case of observed heterogeneity. Algorithm 4 and Proposition 4 provide generalizations of Algorithm 2 and Proposition 2, respectively, to the cases involving observed heterogeneity.

Algorithm 4. *Let* $\widehat{c}_b^n \equiv c_b^n \equiv y_b^n$ *and* $\widehat{\mathscr{P}}_{b,b}(c;\boldsymbol{x}) \equiv 1$, *consider a sequence* $\{g_k(q,\boldsymbol{x})\}_{k=0}^{K_N}$ *of power functions of* q *and* $\boldsymbol{x}$ *where* N *is the number of households in the cross-section. For each* $t \geq b$, *apply the following steps:*

1. *Initialize the values of the real consumption* $\widehat{c}_{t+1}^{n,(0)}$ *for each household at* $t+1$ *using Equations (23)–(27) as in Algorithm 1.*

2. *For each* $t \geq b$, *iterate over the following steps over* $\tau \in \{0, 1, \cdots\}$ *until convergance for some* $\epsilon \ll 1$*:*

   (a) *Solve for the coefficients* $\left(\widehat{\alpha}_{k,t}^{\dagger}\right)_{k=0}^{K}$ *in the following problem:*

$$\min_{\left(\alpha_{k,t}^{\dagger}\right)_{k=0}^{K}} \sum_{n=1}^{N}\left(\pi_t^n - \sum_{k=0}^{K}\alpha_{k,t}^{\dagger} f_k\left(\widehat{c}_{t+1}^{n,(\tau)}, \boldsymbol{x}_{t+1}^n\right)\right)^2, \tag{A14}$$

   where $\pi_t^n \equiv \log\mathbb{P}_G\left(\boldsymbol{p}_{t+1}, \boldsymbol{q}_{t+1}^n; \boldsymbol{p}_t, \boldsymbol{q}_t^n\right)$.

   (b) *Update the next period function* $\widehat{\mathscr{P}}_{b,t+1}(\cdot,\cdot)$*:*

$$\log\widehat{\mathscr{P}}_{b,t+1}(c;\boldsymbol{x}) \equiv \log\widehat{\mathscr{P}}_{b,t}(c;\boldsymbol{x}) + \sum_{k=0}^{K}\widehat{\beta}_{k,t} f_k(c,\boldsymbol{x}), \tag{A15}$$

   where the coefficients $\left(\widehat{\beta}_{k,t}\right)_{k=0}^{K}$ *solve the following problem:*

$$\min_{\left(\widehat{\alpha}_{k,t}\right)_{k=0}^{K}} \sum_{n=1}^{N}\left(\pi_t^{n,*} + \rho_t^{n,(\tau)} - \sum_{k=0}^{K}\widehat{\beta}_{k,t} f_k\left(\widehat{c}_t^n, \boldsymbol{x}_t^n\right)\right)^2, \tag{A16}$$

where $\pi_t^{n,*} \equiv \mathbb{P}_T\left(\boldsymbol{p}_t, \boldsymbol{q}_t^n; \boldsymbol{p}_{t+1}, \boldsymbol{q}_{t+1}^n\right)$ *and where* $\rho_t^{n,(\tau)}$ *is defined as:*

$$\rho_t^{n,(\tau)} \equiv \frac{1}{4}\sum_{k=0}^{K}\widehat{\alpha}_{k,t}^{\dagger}\left[\frac{\partial f_k\left(\widehat{q}_t^n, \boldsymbol{x}_t^n\right)}{\partial \log c} + \frac{\partial f_k\left(\widehat{c}_{t+1}^{n,(\tau)}, \boldsymbol{x}_{t+1}^n\right)}{\partial \log c}\right]\log\left(\frac{\widehat{q}_{t+1}^{n,(\tau)}}{\widehat{q}_t^n}\right) \tag{A17}$$

$$+\frac{1}{4}\sum_{d=1}^{D}\sum_{k=0}^{K}\widehat{\alpha}_{k,t}^{\dagger}\left[\frac{\partial f_k\left(\widehat{c}_t^n, \boldsymbol{x}_t^n\right)}{\partial \log x_d} + \frac{\partial f_k\left(\widehat{q}_{t+1}^{n,(\tau)}, \boldsymbol{x}_{t+1}^n\right)}{\partial \log x_d}\right]\log\left(\frac{x_{d,t+1}}{x_{d,t}}\right). \tag{A18}$$

(c) *Update the real consumption in the next period for each household:*

$$\log \widehat{c}_{t+1}^{n,(\tau+1)} = \log \widehat{c}_t^{n,(\tau)} + \frac{1}{1+\frac{1}{2}\left[\widehat{\Lambda}_t\left(\widehat{c}_t^n; \boldsymbol{x}_t^n\right) + \widehat{\Lambda}_{t+1}\left(\widehat{c}_{t+1}^{n,(\tau)}; \boldsymbol{x}_{t+1}^n\right)\right]} \tag{A19}$$

$$\times\left[\log\left(\frac{y_{t+1}^n}{y_t^n}\right) - \frac{1}{2}\sum_{d=1}^{D}\left[\widehat{\Gamma}_{d,t}\left(c_t^n; \boldsymbol{x}_t^n\right) + \widehat{\Gamma}_{d,t+1}\left(c_{t+1}^n; \boldsymbol{x}_{t+1}^n\right)\right]\cdot \log\left(\frac{x_{d,t+1}}{x_{d,t}}\right)\right], \tag{A20}$$

where we have defined the approximate correction functions as:

$$\widehat{\Lambda}_{t+1}(c; \boldsymbol{x}) = \sum_{k=0}^{K}\left(\sum_{\tau=b+1}^{t+1}\widehat{\beta}_{k,\tau}\right)\frac{\partial f_k(c, \boldsymbol{x})}{\partial \log c}, \tag{A21}$$

$$\widehat{\Gamma}_{d,t}(c; \boldsymbol{x}) = \sum_{k=0}^{K_N}\left(\sum_{\tau=b+1}^{t+1}\widehat{\alpha}_{k,\tau}\right)\frac{\partial f_k(c, \boldsymbol{x})}{\partial \log x_d}, \tag{A22}$$

(d) *Stop if* $\max_n\left|\widehat{c}_{t+1}^{n,(\tau+1)} - \widehat{c}_{t+1}^{n,(\tau)}\right| < \epsilon$ *and set* $\widehat{c}_{t+1}^n \equiv \widehat{c}_{t+1}^{n,(\tau+1)}$.

Proposition 4. *Assume that the expenditure function* $\log E(\cdot;\cdot)$ *is an analytical function, , that* $\Delta_p + \Delta_y < 1$ *with* Δ_p *defined as in Equation (20) and* $\Delta_y \equiv \max_{n,i}\left\{\log(y_{t+1}^n / y_t^n)\right\}$*, and that the maximum change in the logarithm of the characteristics across consumers in bounded above* $\Delta_x < 1$ *following (A12). Then, the sequences of real consumptions constructed by Algorithm 4 satisfy:*

$$\log\left(\frac{c_{t+1}^n}{c_t^n}\right) = \log\left(\frac{\widehat{c}_{t+1}^n}{\widehat{c}_t^n}\right) + O\left(\Delta^2\right) + O_p\left(K_N^3\left(\sqrt{\frac{K_N}{N}}\cdot\Delta^4 + K_N^{-m}\right)\Delta\right), \tag{A23}$$

where $\Delta \equiv \max\left\{\Delta_p + \Delta_y, \Delta_x\right\}$ *and* m *is any positive number.*

Proof. See Appendix ??. □

A.3 Proofs and Derivations

A.3.1 Proofs of the Lemmas and Propositions

Proof of Equation (8). From Definitions 1 and 4, we have:

$$\frac{c_t^b}{c_{t_0}^b} = \frac{E(u_t;\boldsymbol{p}_b)}{E\left(u_{t_0};\boldsymbol{p}_b\right)} = \frac{E(u_t;\boldsymbol{p}_b)}{E\left(u_{t_0};\boldsymbol{p}_b\right)} \times \frac{E(u_t;\boldsymbol{p}_b)}{E(u_t;\boldsymbol{p}_t)} \times \frac{E\left(u_{t_0};\boldsymbol{p}_{t_0}\right)}{E\left(u_{t_0};\boldsymbol{p}_b\right)},$$
$$= \frac{y_t}{y_{t_0}} \times \frac{1}{\mathscr{P}_{b,t}^b(c_t)} \times \frac{1}{\mathscr{P}_{t_0,t}^b\left(c_{t_0}\right)}.$$

□

Proof of Lemma 2. From Lemma 1, we have:

$$\frac{d\log c_t^{b_2}}{d\log c_t^{b_1}} = \frac{1+\Lambda_t^{b_1}\left(c_t^{b_1}\right)}{1+\Lambda_t^{b_2}\left(c_t^{b_2}\right)} = \frac{\left.\frac{\partial \log \widetilde{E}^{b_1}(c;\boldsymbol{p}_t)}{dc}\right|_{c=c_t^{b_1}}}{\left.\frac{\partial \log \widetilde{E}^{b_2}(c;\boldsymbol{p}_t)}{dc}\right|_{c=c_t^{b_1}}} = \frac{\left.\left(\frac{\frac{\partial \log E(u;\boldsymbol{p}_t)}{du}}{\frac{\partial \log E\left(u;\boldsymbol{p}_{b_1}\right)}{du}}\right)\right|_{u=u_t}}{\left.\left(\frac{\frac{\partial \log E(u;\boldsymbol{p}_t)}{du}}{\frac{\partial \log E\left(u;\boldsymbol{p}_{b_2}\right)}{du}}\right)\right|_{u=u_t}} = \left.\left(\frac{\frac{\partial \log E\left(u;\boldsymbol{p}_{b_2}\right)}{du}}{\frac{\partial \log E\left(u;\boldsymbol{p}_{b_1}\right)}{du}}\right)\right|_{u=u_t},$$

leading to the desired result. □

Proof of Lemma 3. We can write the growth in the consumer expenditure as

$$\frac{d\log \widetilde{E}(c_t;\boldsymbol{p}_t,\boldsymbol{x}_t)}{dt} = \sum_i \frac{\partial \log \widetilde{E}(c_t;\boldsymbol{p}_t,\boldsymbol{x}_t)}{d\log p_{it}}\frac{d\log p_{it}}{dt} + \sum_d \frac{\partial \log \widetilde{E}(c_t;\boldsymbol{p}_t,\boldsymbol{x}_t)}{d\log x_{dt}}\frac{d\log x_{dt}}{dt}$$
$$+ \frac{\partial \log \widetilde{E}(c_t;\boldsymbol{p}_t,\boldsymbol{x}_t)}{dc_t}\frac{d\log c_t}{dt},$$

where the left hand side equals $\frac{d\log y_{it}}{dt}$ and where we have suppressed the base period superscripts b to simplify the expression. The desired result follows from the observation that $c = \widetilde{E}(c;\boldsymbol{p}_b,\boldsymbol{x})$ for all $\boldsymbol{x}$. □

Proof of Lemma 4. Performing a Taylor series expansion of the expenditure function around the vector of prices $\boldsymbol{p}_{t_0}$, we find:

$$\log \mathscr{P}_{t_0,t}^b(c_\tau) = \log \frac{\widetilde{E}^b(c_\tau;\boldsymbol{p}_t)}{\widetilde{E}^b\left(c_\tau;\boldsymbol{p}_{t_0}\right)} = \sum_i \widetilde{\Omega}_i^b\left(c_\tau;\boldsymbol{p}_{t_0}\right)\log\left(\frac{p_{i,t}}{p_{i,t_0}}\right) + O\left(\Delta_p^2\right),$$

where $\Delta_p \equiv \max_i \left\{\log\left(p_{it+1}/p_{it}\right)\right\}$, and where we have used Shephard's lemma in the second equality. If the preferences are homotheic, we have $s_{it_0} = \widetilde{\Omega}_i^b\left(c_{t_0};\boldsymbol{p}_{t_0}\right) = \widetilde{\Omega}_i^b\left(c_\tau;\boldsymbol{p}_{t_0}\right)$ and the desired result follows. Otherwise, performing of a Taylor series expansion of the expenditure function around the vector of prices $\boldsymbol{p}_{t_0}$, we find:

$$\widetilde{\Omega}_i^b\left(c_\tau;\boldsymbol{p}_{t_0}\right) = s_{it_0} + \frac{\partial \widetilde{\Omega}_i^b\left(c_{t_0};\boldsymbol{p}_{t_0}\right)}{\partial \log c_{t_0}} \cdot \log\left(\frac{c_\tau}{c_{t_0}}\right) + O\left(\left(\log\left(\tfrac{c_\tau}{c_{t_0}}\right)\right)^2\right),$$

where we have substituted $s_{it_0} = \widetilde{\Omega}_i^b\left(c_{t_0};\boldsymbol{p}_{t_0}\right)$ on the right hand side. We now apply Lemma 6 in Section A.3.2 below to bound the size of $\log\left(\frac{c_\tau}{c_{t_0}}\right)$, and we find: □

$$\log \mathscr{P}_{t_0,t}^b\left(c_\tau\right) = \log \mathbb{P}_G\left(\boldsymbol{p}_{t_0},\boldsymbol{q}_{t_0};\boldsymbol{p}_t,\boldsymbol{q}_t\right) + \log\left(\frac{c_\tau}{c_{t_0}}\right) \cdot \sum_{i=1}^{I} \frac{\partial \widetilde{\Omega}_i^b\left(c_{t_0};\boldsymbol{p}_{t_0}\right)}{\partial \log c_{t_0}} \log\left(\frac{p_{i,t}}{p_{i,t_0}}\right) + O\left(\Delta^2\right),$$

where we have let $\left|\log\left(\frac{c_\tau}{c_{t_0}}\right)\right| = O\left(\Delta_p + \Delta_y\right)$. Since $\widetilde{E}^b\left(c;\boldsymbol{p}\right)$ is second order continuously differentiable, the second term above is of the order $O\left(\Delta^2\right)$ as well, and we reach Equation (21).

Applying Lemma 5 in Section A.3.2 to the expenditure function we find:

$$\begin{aligned}
\log \mathscr{P}_{t_0,t}^b(c) &= \log \frac{\widetilde{E}^b\left(c;\boldsymbol{p}_t\right)}{\widetilde{E}^b\left(c;\boldsymbol{p}_{t_0}\right)}, \\
&= \frac{1}{2}\sum_{i=1}^{I}\left[\widetilde{\Omega}_i^b\left(c;\boldsymbol{p}_{t_0}\right) + \widetilde{\Omega}_i^b\left(c;\boldsymbol{p}_t\right)\right]\log\left(\frac{p_{i,t}}{p_{i,t_0}}\right) + O\left(\Delta_p^3\right).
\end{aligned} \tag{A24}$$

Once again assuming homotheticity, we have that $s_{it} = \widetilde{\Omega}_i^b\left(c;\boldsymbol{p}_t\right)$ for all c and the desired result follows. Otherwise, using Lemma 5 in Section A.3.2 on the Hicksian expenditure share function we find:

$$\widetilde{\Omega}_i^b\left(c;\boldsymbol{p}_{t_0}\right) = s_{it_0} + \frac{1}{2}\left[\frac{\partial \widetilde{\Omega}_i^b\left(c_{t_0};\boldsymbol{p}_{t_0}\right)}{\partial \log c_{t_0}} + \frac{\partial \widetilde{\Omega}_i^b\left(c;\boldsymbol{p}_{t_0}\right)}{\partial \log c}\right] \cdot \log\left(\frac{c}{c_{t_0}}\right) + O\left(\left(\log\left(\tfrac{c_\tau}{c_{t_0}}\right)\right)^3\right),$$

$$\widetilde{\Omega}_i^b\left(c;\boldsymbol{p}_t\right) = s_{it} + \frac{1}{2}\left[\frac{\partial \widetilde{\Omega}_i^b\left(c_t;\boldsymbol{p}_t\right)}{\partial \log c_t} + \frac{\partial \widetilde{\Omega}_i^b\left(c;\boldsymbol{p}_{t_0}\right)}{\partial \log c}\right] \cdot \log\left(\frac{c}{c_t}\right) + O\left(\left(\log\left(\tfrac{c_\tau}{c_{t_0}}\right)\right)^3\right),$$

Substituting this expression in Equation (A24), invoking Lemma 6 in Section A.3.2 below, we find:

$$\log \mathscr{P}^{b}_{t_0,t}(c) = \log \mathbb{P}_T\left(\boldsymbol{p}_{t_0}, \boldsymbol{q}_{t_0}; \boldsymbol{p}_t, \boldsymbol{q}_t\right)$$
$$+\frac{1}{2}\cdot \log\left(\frac{c}{c_{t_0}}\right)\cdot \sum_{i=1}^{I}\left[\frac{\partial \widetilde{\Omega}^b_i\left(c_{t_0}; \boldsymbol{p}_{t_0}\right)}{\partial \log c_{t_0}} + \frac{\partial \widetilde{\Omega}^b_i\left(c; \boldsymbol{p}_{t_0}\right)}{\partial \log c}\right] \log\left(\frac{c}{c_{t_0}}\right)\log\left(\frac{p_{i,t}}{p_{i,t_0}}\right)$$
$$+\frac{1}{2}\cdot \log\left(\frac{c}{c_t}\right)\cdot \sum_{i=1}^{I}\left[\frac{\partial \widetilde{\Omega}^b_i\left(c_t; \boldsymbol{p}_t\right)}{\partial \log c_t} + \frac{\partial \widetilde{\Omega}^b_i\left(c; \boldsymbol{p}_{t_0}\right)}{\partial \log c}\right] \log\left(\frac{c}{c_t}\right)\log\left(\frac{p_{i,t}}{p_{i,t_0}}\right) + O\left(\Delta^3\right), \tag{A25}$$

where $\Delta \equiv \Delta_p + \Delta_y$ were Δ_p and Δ_y are given by Equation (A35) in Section A.3.2 below. Now, we use the third-order continuously differentiable property of the expenditure function to find

$$\frac{\partial \widetilde{\Omega}^b_i\left(c_{t'}; \boldsymbol{p}_{t''}\right)}{\partial \log c_{t'}} = \frac{\partial \widetilde{\Omega}^b_i\left(c_{t_0}; \boldsymbol{p}_{t_0}\right)}{\partial \log c_{t_0}} + O(\Delta), \qquad t', t'' \in [t_0, t],$$

and to substitute for the expressions within the square brackets in Equation (A25). This leads to the following result

$$\log \mathscr{P}_{CL}\left(\boldsymbol{p}_{t_0}, \boldsymbol{p}_t; \boldsymbol{q}_b\right) = \log \mathbb{P}_T\left(\boldsymbol{p}_{t_0}, \boldsymbol{q}_{t_0}; \boldsymbol{p}_t, \boldsymbol{q}_t\right)$$
$$+\log\left[\frac{(c)^2}{c_{t_0}\cdot c_t}\right]\cdot \sum_{i=1}^{I}\frac{\partial \widetilde{\Omega}^b_i\left(c_{t_0}; \boldsymbol{p}_{t_0}\right)}{\partial \log c_{t_0}} \log\left(\frac{p_{i,t}}{p_{i,t_0}}\right) + O\left(\Delta^3\right).$$

Thus, letting $c = \sqrt{c_{t_0}\cdot c_t}$ the second term on the right hand side vanishes and we reach Equation (22).

Finally, Lemma 7 below shows that we can replace the geometric index with Laspeyres or Paasche and the Törqvist with Fisher or Sato-Vartia formulas and reach similar error bounds.

Proof for Proposition 1. First, we will establish a bound on the error corresponding to the approximation of the nonhomothetic correction function $\Lambda_{t+1}(c)$ with the nonparametric estimation $\widehat{\Lambda}_{t+1}(c)$. By definition of the nonhomotheticity correction function, we have

$$\Lambda_{t+1}(c) = \frac{\partial}{\partial \log c}\log\left(\frac{\widetilde{E}^b\left(c; \boldsymbol{p}_{t+1}\right)}{\widetilde{E}^b\left(c; \boldsymbol{p}_b\right)}\right) = \sum_{\tau=b}^{t}\frac{\partial}{\partial \log c}\log\left(\frac{\widetilde{E}^b\left(c; \boldsymbol{p}_{\tau+1}\right)}{\widetilde{E}^b\left(c; \boldsymbol{p}_\tau\right)}\right).$$

Applying Lemma 4, from Equation (21), we have:

$$\log\left(\frac{\widetilde{E}^b\left(c_t^n;\boldsymbol{p}_{t+1}\right)}{\widetilde{E}^b\left(c_t^n;\boldsymbol{p}_t\right)}\right)=\log\mathbb{P}_G\left(\boldsymbol{p}_{t_0},\boldsymbol{q}_{t_0}^n;\boldsymbol{p}_t,\boldsymbol{q}_t^n\right)+O\left(\Delta^2\right).$$

Let us now begin with the first period away from the base period where we observe the real consumption index $c_b^n \equiv y_b^n$. We apply Lemma 8 below for $y^n \equiv \pi_t^n$, $x^n \equiv \log\widehat{c}_b^n$, $z^n \equiv \log c_b^n$, $v^n \equiv \delta_v \equiv 0$, and $\Delta_\varepsilon \equiv \Delta_p^2$ to find:

$$\frac{\partial}{\partial \log c}\log\left(\frac{\widetilde{E}^b\left(c;\boldsymbol{p}_{b+1}\right)}{\widetilde{E}^b\left(c;\boldsymbol{p}_b\right)}\right)=\sum_{k=0}^{K_N}\widehat{\alpha}_{k,b}f_k'(\log c)+O_p\left(K_N^3\left(\sqrt{\frac{K_N}{N}}\cdot\Delta^4+K_N^{-(m-1)}\right)\right),$$

where $\left(\widehat{\alpha}_{k,b}\right)_{k=0}^{K_N}$ solve Equation (24) for $t=b$. Therefore, we have:

$$\Lambda_{b+1}(c)=\widehat{\Lambda}_{b+1}(c)+O_p\left(K_N^3\left(\sqrt{\frac{K_N}{N}}\cdot\Delta_p^4+K_N^{-(m-1)}\right)\right).$$

Applying Lemma 9, the desired result follows for the first period.

Next, we recursively apply Lemma 8 below for $y^n \equiv \log\mathbb{P}_G\left(\boldsymbol{p}_t,\boldsymbol{p}_{t+1};\boldsymbol{q}_t^n,\boldsymbol{q}_{t+1}^n\right)$, $x^n \equiv \log\widehat{c}_b^n$, $z^n \equiv \log c_b^n$, with δ_v denoting the error from the previous period approximation error in the ReC index and $\Delta_\varepsilon \equiv \Delta_p^2$ to find:

$$\frac{\partial}{\partial \log c}\log\left(\frac{\widetilde{E}^b\left(c;\boldsymbol{p}_{t+1}\right)}{\widetilde{E}^b\left(c;\boldsymbol{p}_t\right)}\right)=\sum_{k=0}^{K_N}\widehat{\alpha}_{k,t}g_k'(c)+O_p\left(K_N^3\left(\sqrt{\frac{K_N}{N}}\cdot\Delta_p^4+K_N^{-(m-1)}\right)\right),$$

where $\left(\widehat{\alpha}_{k,t}\right)_{k=0}^{K_N}$ solve Equation (24). Note that the term $O(\delta_v)$ is of the same order as the error term on the right hand side of the equation above and is therefore absorbed in that error. Therefore, we have for all t:

$$\Lambda_{t+1}(c)=\widehat{\Lambda}_{t+1}(c)+O_p\left(K_N^3\left(\sqrt{\frac{K_N}{N}}\cdot\Delta_p^4+K_N^{-(m-1)}\right)\right).$$

Applying Lemma 9, the desired result follows for all t. Using Lemma 7 allows us to replace the geometric index with Laspeyres or Paasche indices. □

Proof for Proposition 2. As in the proof of Proposition 1, we will first establish a bound on the error corresponding to the approximation of the nonhomothetic correction function $\Lambda_t(c)$ with the nonparametric estimation $\widehat{\Lambda}_t(c)$. By definition, we have:

$$\Lambda_t(c) = \frac{\partial}{\partial \log c} \log\left(\frac{\widetilde{E}^b(c;\boldsymbol{p}_{t+1})}{\widetilde{E}^b(c;\boldsymbol{p}_b)}\right) = \sum_{\tau=b}^{t} \frac{\partial}{\partial \log c} \log\left(\frac{\widetilde{E}^b(c;\boldsymbol{p}_{\tau+1})}{\widetilde{E}^b(c;\boldsymbol{p}_\tau)}\right).$$

To approximate this function, we first note that

$$\begin{aligned}
\log\left(\frac{\widetilde{E}^b(c_t^n;\boldsymbol{p}_{t+1})}{\widetilde{E}^b(c_t^n;\boldsymbol{p}_t)}\right) &= \frac{1}{2}\sum_{i=1}^{I}\left[\widetilde{\Omega}_i^b(c_t^n;\boldsymbol{p}_t) + \widetilde{\Omega}_i^b(c_t^n;\boldsymbol{p}_{t+1})\right]\log\left(\frac{p_{i,t+1}}{p_{i,t}}\right), \\
&= \mathbb{P}_T(\boldsymbol{p}_t, \boldsymbol{q}_t^n; \boldsymbol{p}_{t+1}, \boldsymbol{q}_{t+1}^n) \\
&\quad - \frac{1}{4}\sum_{i=1}^{I}\left(\frac{\partial \widetilde{\Omega}_i^b(c_t^n;\boldsymbol{p}_t)}{\partial \log c} + \frac{\partial \widetilde{\Omega}_i^b(c_t^n;\boldsymbol{p}_{t+1})}{\partial \log c}\right)\log\left(\frac{p_{i,t+1}}{p_{i,t}}\right),
\end{aligned} \tag{A26}$$

where we have used

$$\widetilde{\Omega}_i^b(c_t^n;\boldsymbol{p}_{t+1}) = \widetilde{\Omega}_i^b(c_{t+1}^n;\boldsymbol{p}_{t+1}) - \frac{1}{2}\left(\frac{\partial \widetilde{\Omega}_i^b(c_t^n;\boldsymbol{p}_{t+1})}{\partial \log c} + \frac{\partial \widetilde{\Omega}_i^b(c_{t+1}^n;\boldsymbol{p}_{t+1})}{\partial \log c}\right)\log\left(\frac{c_{t+1}^n}{c_t^n}\right) + O(\Delta^3).$$

Defining

$$\mathscr{P}_t^\dagger(c) \equiv \frac{\partial}{\partial \log c}\left[\sum_{i=1}^{I}\widetilde{\Omega}_i^b(c;\boldsymbol{p}_{t+1}) \cdot \log\left(\frac{p_{i,t+1}}{p_{i,t}}\right)\right],$$

we can now rewrite Equation (A26) as:

$$\log\left(\frac{\widetilde{E}^b(c_t^n;\boldsymbol{p}_{t+1})}{\widetilde{E}^b(c_t^n;\boldsymbol{p}_t)}\right) = \mathbb{P}_T(\boldsymbol{p}_t, \boldsymbol{q}_t^n; \boldsymbol{p}_{t+1}, \boldsymbol{q}_{t+1}^n) - \frac{1}{4}\left[\mathscr{P}_t^\dagger(c_t^n) + \mathscr{P}_t^\dagger(c_{t+1}^n)\right]\log\left(\frac{c_{t+1}^n}{c_t^n}\right) + O(\Delta^3). \tag{A27}$$

The key observation is to note that, through the definition of the geometric index, we have:

$$\sum_{i=1}^{I}\widetilde{\Omega}_i^b(c_{t+1}^n;\boldsymbol{p}_{t+1}) \cdot \log\left(\frac{p_{i,t+1}}{p_{i,t}}\right) = -\log \mathbb{P}_G(\boldsymbol{p}_{t+1}, \boldsymbol{q}_{t+1}^n; \boldsymbol{p}_t, \boldsymbol{q}_t^n).$$

We now apply Lemma 8 for $y^n \equiv \mathbb{P}_G(\boldsymbol{p}_{t+1}, \boldsymbol{q}_{t+1}^n; \boldsymbol{p}_t, \boldsymbol{q}_t^n)$, $x^n \equiv \log \widehat{c}_b^n$, $z^n \equiv \log c_b^n$, $v^n \equiv \delta_v \equiv 0$, and $\Delta_\varepsilon \equiv \Delta^2$ to find that for ρ_t^n defined by Equations (A1) and (A4), we have:

$$-\frac{1}{4}\left[\mathscr{P}_t^\dagger(c_t^n) + \mathscr{P}_t^\dagger(c_{t+1}^n)\right] = \rho_t^n + O(\epsilon) + O_p(K_N^{4-m}).$$

Therefore, we can now rewrite Equation (A27) as:

$$\log\left(\frac{\widetilde{E}^b\left(c_t^n;\boldsymbol{p}_{t+1}\right)}{\widetilde{E}^b\left(c_t^n;\boldsymbol{p}_t\right)}\right)=\mathbb{P}_T\left(\boldsymbol{p}_t,\boldsymbol{q}_t^n;\boldsymbol{p}_{t+1},\boldsymbol{q}_{t+1}^n\right)+\rho_t^n+O(\epsilon)+O_p\left(K_N^{4-m}\right)+O\left(\Delta_p^3\right). \tag{A28}$$

This allows us to again apply Lemma 8 for $y^n\equiv\mathbb{P}_T\left(\boldsymbol{p}_t,\boldsymbol{q}_t^n;\boldsymbol{p}_{t+1},\boldsymbol{q}_{t+1}^n\right)+\rho_t^n$, $x^n\equiv\log\widehat{c}_t^n$, $z^n\equiv\log c_t^n$, and $\Delta_\varepsilon\equiv O(\epsilon)+O_p\left(K_N^{4-m}\right)+O\left(\Delta^3\right)$ (similar to that in the proof of Proposition 1) shows that Equation (A6) indeed approximates $\Lambda_{t+1}(c)$. Applying Lemma 10 in Section A.3.2 below leads to Equation (A7). Again, Using Lemma 7 allows us to replace the geometric index with Laspeyres or Paasche indices, and the Törqvist index with Fisher or Sato-Vartia. □

Proof for Proposition 3. We need to establish bounds on the error corresponding to the approximations of the correction functions $\Lambda_{t+1}(c;\boldsymbol{x})$ and $\Gamma_{d,t+1}(c;\boldsymbol{x})$and with the by $\widehat{\Lambda}_{t+1}(c;\boldsymbol{x})$ and $\widehat{\Gamma}_{d,t+1}(c;\boldsymbol{x})$. The former follows closely that in the proof of Proposition 1, except that we now invoke the multi-dimensional case of Lemma 8, requiring the CoL index to be a infinitely differentiable. This leads us to:

$$\Lambda_{t+1}(c;\boldsymbol{x})=\widehat{\Lambda}_{t+1}(c,\boldsymbol{x})+O_p\left(K_N^3\left(\sqrt{\frac{K_N}{N}}\cdot\Delta^4+K_N^{-m}\right)\right),$$
$$\Gamma_{d,t+1}(c;\boldsymbol{x})=\widehat{\Gamma}_{d,t+1}(c,\boldsymbol{x})+O_p\left(K_N^3\left(\sqrt{\frac{K_N}{N}}\cdot\Delta^4+K_N^{-m}\right)\right),$$

where m is any positive number.

We next show that

$$\log\left(\frac{c_{t+1}^n}{c_t^n}\right)=\frac{1}{1+\Lambda_{t+1}\left(c_t^n;\boldsymbol{x}_{t+1}^n\right)}\left[\log\left(\frac{y_{t+1}^n}{y_t^n}\right)-\log\mathscr{P}_{t,t+1}\left(c_t^n;\boldsymbol{x}_t^n\right)\right. \tag{A29}$$
$$\left.-\sum_{d=1}^{D}\Gamma_{b,d}\left(c;\boldsymbol{x}_t^n\right)\cdot\log\left(\frac{x_{d,t+1}^n}{x_{d,t}^n}\right)\right]+O\left(\Delta^2\right), \tag{A30}$$

which, when combined with the above result, establishes the proposition. To derive Equation (A30), we perform a first-order Taylor expansion of the left-hand-side of the equation above in terms of q_{t+1}^n, assuming that $\log\left(\frac{q_{t+1}^n}{q_t^n}\right)<\Delta_q$:

$$\log\left(\frac{y_{t+1}^n}{y_t^n}\right)=\log\frac{\widetilde{E}\left(c_t^n;\boldsymbol{p}_{t+1},\boldsymbol{x}_t^n\right)}{\widetilde{E}\left(c_t^n;\boldsymbol{p}_t,\boldsymbol{x}_t^n\right)}+\sum_{d=1}^{D}\frac{\partial\log\widetilde{E}(c;\boldsymbol{p},\boldsymbol{x})}{\partial\log x_d}\bigg|_{(c;\boldsymbol{p},\boldsymbol{x})\equiv\left(c_t^n;\boldsymbol{p}_{t+1},\boldsymbol{x}_t^n\right)}\cdot\log\left(\frac{x_{d,t+1}^n}{x_{d,t}^n}\right)$$

$$+\left.\frac{\partial \log \widetilde{E}(c;\boldsymbol{p},\boldsymbol{x})}{\partial \log c}\right|_{(c;\boldsymbol{p},\boldsymbol{x})\equiv\left(c_t^n;\boldsymbol{p}_{t+1},\boldsymbol{x}_t^n\right)}\cdot\log\left(\frac{c_{t+1}^n}{c_t^n}\right)+O\left(\Delta^2\right).$$

This allows us to write:

$$\log\left(\frac{c_{t+1}^n}{c_t^n}\right)=\frac{\log\left(\frac{y_{t+1}^n}{y_t^n}\right)-\sum_{d=1}^D\Gamma_{d,t+1}(c;\boldsymbol{x}_t^n)\cdot\log\left(\frac{x_{d,t+1}^n}{x_{d,t}^n}\right)}{1+\Lambda_{t+1}(c;\boldsymbol{x}_t^n)},$$

where we have defined:

$$\begin{aligned}\Lambda_{t+1}(c;\boldsymbol{x}_t^n)&=\frac{\partial \log \widetilde{E}\left(c;\boldsymbol{p}_{t+1},\boldsymbol{x}_t^n\right)}{\partial \log c}-\frac{\partial \log \widetilde{E}(c;\boldsymbol{p}_b,\boldsymbol{x}_t^n)}{\partial \log c},\\&=\frac{\partial}{\partial \log c}\log\left(\frac{\widetilde{E}\left(c;\boldsymbol{p}_{t+1},\boldsymbol{x}_t^n\right)}{\widetilde{E}(c;\boldsymbol{p}_b,\boldsymbol{x}_t^n)}\right),\end{aligned}\tag{A31}$$

noting that $\frac{\partial \log \widetilde{E}(c;\boldsymbol{p}_b,\boldsymbol{x}_t^n)}{\partial \log c}\equiv 1$ for all $\boldsymbol{x}_t^n$, and:

$$\begin{aligned}\Gamma_{d,t+1}(q;\boldsymbol{x}_t^n)&=\frac{\partial \log \widetilde{E}\left(c;\boldsymbol{p}_{t+1},\boldsymbol{x}_t^n\right)}{\partial \log x_d}-\frac{\partial \log \widetilde{E}(c;\boldsymbol{p}_b,\boldsymbol{x}_n^n)}{\partial \log x_d},\\&=\frac{\partial}{\partial \log x_d}\log\left(\frac{\widetilde{E}\left(c;\boldsymbol{p}_{t+1},\boldsymbol{x}_t^n\right)}{\widetilde{E}(c;\boldsymbol{p}_b,\boldsymbol{x}_t^n)}\right),\end{aligned}\tag{A32}$$

noting that $\frac{\partial \log \widetilde{E}(c;\boldsymbol{p}_b,\boldsymbol{x}_t^n)}{\partial \log x_d}\equiv 0$ for all $\boldsymbol{x}_t^n$. □

Proof for Proposition 4. First, we establish the following result:

$$\begin{aligned}\log\left(\frac{c_{t+1}^n}{c_t^n}\right)=&\frac{1}{1+\frac{1}{2}\left[\Lambda_t(c_t^n;\boldsymbol{x}_t^n)+\Lambda_{t+1}\left(c_{t+1}^n;\boldsymbol{x}_{t+1}^n\right)\right]}\\&\times\left[\log\left(\frac{y_{t+1}^n}{y_t^n}\right)-\pi_t^{n,*}-\frac{1}{2}\sum_{d=1}^D\left[\Gamma_{d,t}(c_t^n;\boldsymbol{x}_t^n)+\Gamma_{d,t+1}\left(c_{t+1}^n;\boldsymbol{x}_{t+1}^n\right)\right]\cdot\log\left(\frac{x_{d,t+1}}{x_{d,t}}\right)\right],\end{aligned}\tag{A33}$$

where $\pi_t^{n,*}\equiv\log\mathbb{P}_T\left(\boldsymbol{p}_t,\boldsymbol{q}_t^n;\boldsymbol{p}_{t+1},\boldsymbol{q}_{t+1}^n\right)$. To show this, we again rely on Lemma 5 to find:

$$\log\left(\frac{y_{t+1}^n}{y_t^n}\right)=\log\frac{\widetilde{E}\left(c_{t+1}^n;\boldsymbol{p}_{t+1},\boldsymbol{x}_{t+1}^n\right)}{\widetilde{E}(c_t^n;\boldsymbol{p}_t,\boldsymbol{x}_t^n)},$$

$$
\begin{aligned}
&= \frac{1}{2}\sum_{i=1}^{I}\left[\left.\frac{\partial \log \widetilde{E}(c;\boldsymbol{p},\boldsymbol{x})}{\partial \log p_i}\right|_{(c;\boldsymbol{p},\boldsymbol{x})\equiv(c_t^n;\boldsymbol{p}_t,\boldsymbol{x}_t^n)} + \left.\frac{\partial \log \widetilde{E}(c;\boldsymbol{p},\boldsymbol{x})}{\partial \log p_i}\right|_{(c;\boldsymbol{p},\boldsymbol{x})\equiv\left(c_{t+1}^n;\boldsymbol{p}_{t+1},\boldsymbol{x}_{t+1}^n\right)}\right]\cdot \log\left(\frac{p_{i,t+1}}{p_{i,t}}\right) \\
&+ \frac{1}{2}\sum_{d=1}^{D}\left[\left.\frac{\partial \log \widetilde{E}(c;\boldsymbol{p},\boldsymbol{x})}{\partial \log x_d}\right|_{(c;\boldsymbol{p},\boldsymbol{x})\equiv(c_t^n;\boldsymbol{p}_t,\boldsymbol{x}_t^n)} + \left.\frac{\partial \log \widetilde{E}(c;\boldsymbol{p},\boldsymbol{x})}{\partial \log x_d}\right|_{(c;\boldsymbol{p},\boldsymbol{x})\equiv\left(c_{t+1}^n;\boldsymbol{p}_{t+1},\boldsymbol{x}_{t+1}^n\right)}\right]\cdot \log\left(\frac{x_{d,t+1}}{x_{d,t}}\right) \\
&+ \frac{1}{2}\left[\left.\frac{\partial \log \widetilde{E}(c;\boldsymbol{p},\boldsymbol{x})}{\partial \log c}\right|_{(c;\boldsymbol{p},\boldsymbol{x})\equiv(c_t^n;\boldsymbol{p}_t,\boldsymbol{x}_t^n)} + \left.\frac{\partial \log \widetilde{E}(c;\boldsymbol{p},\boldsymbol{x})}{\partial \log c}\right|_{(c;\boldsymbol{p},\boldsymbol{x})\equiv\left(c_{t+1}^n;\boldsymbol{p}_{t+1},\boldsymbol{x}_{t+1}^n\right)}\right]\cdot \log\left(\frac{c_{t+1}^n}{c_t^n}\right) \\
&\quad + O\left(\Delta^3\right), \\
&= \frac{1}{2}\sum_{i=1}^{I}\left[\widetilde{\Omega}_i\left(c_t^n;\boldsymbol{p}_t,\boldsymbol{x}_t^n\right) + \widetilde{\Omega}_i\left(c_{t+1}^n;\boldsymbol{p}_{t+1},\boldsymbol{x}_{t+1}^n\right)\right]\log\left(\frac{p_{i,t+1}}{p_{i,t}}\right) \\
&+ \frac{1}{2}\sum_{d=1}^{D}\left[\Gamma_{d,t}\left(c_t^n;\boldsymbol{x}_t^n\right) + \Gamma_{d,t+1}\left(c_{t+1}^n;\boldsymbol{x}_{t+1}^n\right)\right]\cdot\log\left(\frac{x_{d,t+1}}{x_{d,t}}\right) \\
&\quad + \left(1 + \frac{1}{2}\left[\Lambda_t\left(c_t^n;\boldsymbol{x}_t^n\right) + \Lambda_{t+1}\left(c_{t+1}^n;\boldsymbol{x}_{t+1}^n\right)\right]\right)\cdot\log\left(\frac{c_{t+1}^n}{c_t^n}\right) + O\left(\Delta^3\right),
\end{aligned}
$$

where in the second equality we have again used Shephard's lemma, as well as the definition of the first-order nonhomotheticity correction function.

Next, we need to first find an approximation to $\mathscr{P}_{t,t+1}(c_t^n;\boldsymbol{x}_t^n)$. Applying Lemma 5, we have

$$
\begin{aligned}
\log \mathscr{P}_{t,t+1}\left(c_t^n;\boldsymbol{x}_t^n\right) &= \log \frac{\widetilde{E}\left(c_t^n;\boldsymbol{p}_{t+1},\boldsymbol{x}_t^n\right)}{\widetilde{E}\left(c_t^n;\boldsymbol{p}_t,\boldsymbol{x}_t^n\right)}, \\
&= \frac{1}{2}\sum_{i=1}^{I}\left[\widetilde{\Omega}_i\left(c_t^n;\boldsymbol{p}_t,\boldsymbol{x}_t^n\right) + \widetilde{\Omega}_i\left(c_t^n;\boldsymbol{p}_{t+1},\boldsymbol{x}_t^n\right)\right]\log\left(\frac{p_{i,t+1}}{p_{i,t}}\right) + O\left(\Delta_p^3\right).
\end{aligned}
$$

For the second term inside the square bracket above, using Lemma 5 again we have:

$$
\begin{aligned}
\widetilde{\Omega}_i\left(c_{t+1}^n;\boldsymbol{p}_{t+1},\boldsymbol{x}_{t+1}^n\right) = \widetilde{\Omega}_i\left(c_t^n;\boldsymbol{p}_{t+1},\boldsymbol{x}_t^n\right) &+ \frac{1}{2}\left(\left.\frac{\partial \widetilde{\Omega}_i\left(c;\boldsymbol{p}_{t+1},\boldsymbol{x}_t^n\right)}{\partial \log c}\right|_{c=c_t^n} + \left.\frac{\partial \Omega_i\left(c;\boldsymbol{p}_{t+1},\boldsymbol{x}_{t+1}^n\right)}{\partial \log c}\right|_{c=c_{t+1}^n}\right)\cdot\log\left(\frac{c_{t+1}^n}{c_t^n}\right) \\
&+ \frac{1}{2}\sum_{d=1}^{D}\left(\left.\frac{\partial \widetilde{\Omega}_i\left(c_t^n;\boldsymbol{p}_{t+1},\boldsymbol{x}\right)}{\partial \log x_d}\right|_{\boldsymbol{x}=\boldsymbol{x}_t^n} + \left.\frac{\partial \widetilde{\Omega}_i\left(c_{t+1}^n;\boldsymbol{p}_{t+1},\boldsymbol{x}\right)}{\partial \log x_d}\right|_{\boldsymbol{x}=\boldsymbol{x}_{t+1}^n}\right)\cdot\log\left(\frac{x_{d,t+1}^n}{x_{d,t}^n}\right) \\
&\quad + O\left(\Delta^3\right).
\end{aligned}
$$

Thus, we find:

$$\log\left(\frac{\widetilde{E}\left(c_t^n;\boldsymbol{p}_{t+1},\boldsymbol{x}_t^n\right)}{\widetilde{E}\left(c_t^n;\boldsymbol{p}_t,\boldsymbol{x}_t^n\right)}\right)=\mathbb{P}_T\left(\boldsymbol{p}_t,\boldsymbol{q}_t^n;\boldsymbol{p}_{t+1},\boldsymbol{q}_{t+1}^n\right) \tag{A34}$$

$$-\frac{1}{4}\left[\mathscr{P}_{t,t+1}^{\dagger}\left(c_t^n;\boldsymbol{x}_t^n\right)+\mathscr{P}_{t,t+1}^{\dagger}\left(c_{t+1}^n;\boldsymbol{x}_{t+1}^n\right)\right]\cdot\log\left(\frac{c_{t+1}^n}{c_t^n}\right)$$

$$-\frac{1}{4}\sum_{d=1}^{D}\left[\mathscr{P}_{d,t,t+1}^{\dagger}\left(c_t^n;\boldsymbol{x}_t^n\right)+\mathscr{P}_{d,t,t+1}^{\dagger}\left(c_{t+1}^n;\boldsymbol{x}_{t+1}^n\right)\right]\cdot\log\left(\frac{x_{d,t+1}^n}{x_{d,t}^n}\right)+O\left(\Delta^3\right),$$

where we have defined:

$$\mathscr{P}_{t,t+1}^{\dagger}\left(c_t^n;\boldsymbol{x}\right)\equiv\frac{\partial}{\partial\log c}\left[\sum_i\widetilde{\Omega}_i\left(c;\boldsymbol{p}_{t+1},\boldsymbol{x}\right)\cdot\log\left(\frac{p_{i,t+1}}{p_{i,t}}\right)\right],$$

$$\mathscr{P}_{d,t,t+1}^{\dagger}\left(c_t^n;\boldsymbol{x}\right)\equiv\frac{\partial}{\partial\log x_d}\left[\sum_i\widetilde{\Omega}_i\left(q;\boldsymbol{p}_{t+1},\boldsymbol{x}\right)\cdot\log\left(\frac{p_{i,t+1}}{p_{i,t}}\right)\right].$$

The remainder of the proof follows along the structure of the proof of Proposition 2. □

A.3.2 Additional Lemmas and Propositions

Lemma 5. *Consider a function $f(\boldsymbol{x})$ defined in the space of $\boldsymbol{x}\in\mathbb{R}^I$. To the second order of approximation, we have:*

$$f(\boldsymbol{y})-f(\boldsymbol{x})=\frac{1}{2}\sum_{i=1}^{I}\left[\frac{\partial f(\boldsymbol{y})}{\partial y_i}+\frac{\partial f(\boldsymbol{x})}{\partial x_i}\right](y_i-x_i).$$

Proof. Using Taylor's expansion, up to the second order in $\boldsymbol{y}-\boldsymbol{x}$, we have:

$$f(\boldsymbol{y})=f(\boldsymbol{x})+\sum_{i=1}^{I}\frac{\partial f(\boldsymbol{x})}{\partial x_i}(y_i-x_i)+\frac{1}{2}\sum_{i,j=1}^{I}\frac{\partial^2 f(\boldsymbol{x})}{\partial x_i\partial x_j}(y_i-x_i)\left(y_j-x_j\right),$$

$$f(\boldsymbol{x})=f(\boldsymbol{y})+\sum_{i=1}^{I}\frac{\partial f(\boldsymbol{y})}{\partial x_i}(x_i-y_i)+\frac{1}{2}\sum_{i,j=1}^{I}\frac{\partial^2 f(\boldsymbol{y})}{\partial x_i\partial x_j}(y_i-x_i)\left(y_j-x_j\right).$$

Together, the two equations imply:

$$f(\boldsymbol{y})=f(\boldsymbol{x})+\frac{1}{2}\sum_{i=1}^{I}\left[\frac{\partial f(\boldsymbol{y})}{\partial y_i}+\frac{\partial f(\boldsymbol{x})}{\partial x_i}\right](y_i-x_i)+\frac{1}{4}\sum_{i,j=1}^{I}\left[\frac{\partial^2 f(\boldsymbol{x})}{\partial x_i\partial x_j}-\frac{\partial^2 f(\boldsymbol{y})}{\partial x_i\partial x_j}\right](y_i-x_i)\left(y_j-x_j\right).$$

This gives us the desired result since:

$$\frac{\partial^2 f(\boldsymbol{y})}{\partial x_i \partial x_j} - \frac{\partial^2 f(\boldsymbol{x})}{\partial x_i \partial x_j} = \sum_k \frac{\partial^3 f(\boldsymbol{x})}{\partial x_k \partial x_i \partial x_j} (y_k - x_k).$$

□

Lemma 6. *Define:*

$$\Delta_y \equiv \left| \log \left(\frac{y_t}{y_{t_0}} \right) \right|, \qquad \Delta_p \equiv \max_i \left| \log \left(\frac{p_{i,t}}{p_{i,t_0}} \right) \right|, \tag{A35}$$

If the expenditure function is continuously differentiable, we have:

$$\left| \log \left(\frac{c_t^b}{c_{t_0}^b} \right) \right| = O\left(\Delta_p + \Delta_y \right). \tag{A36}$$

Proof. Define:

$$\Delta_y \equiv \left| \log \left(\frac{y_t}{y_{t_0}} \right) \right|, \qquad \Delta_p \equiv \max_i \left| \log \left(\frac{p_{i,t}}{p_{i,t_0}} \right) \right|, \qquad \Delta_c \equiv \left| \log \left(\frac{c_t^b}{c_{t_0}^b} \right) \right|,$$

and assume that $\Delta \equiv \Delta_y + \Delta_p < 1$. Using a Taylor expansion of $\widetilde{E}^b\left(c; \boldsymbol{p}_{t_0}\right)$ around $c = c_{t_0}^b$ and $\boldsymbol{p}_{t_0}$, and since the expenditure function $\widetilde{E}^b(c; \boldsymbol{p})$ is continuously differentiable we have

$$\begin{aligned} \log \left(\frac{y_t}{y_{t_0}} \right) &= \log \left(\frac{\widetilde{E}^b\left(c_t^b; \boldsymbol{p}_t\right)}{\widetilde{E}^b\left(c_{t_0}^b; \boldsymbol{p}_{t_0}\right)} \right), \\ &= \sum_i h_i\left(c_t^b, \boldsymbol{p}_t\right) \log \left(\frac{p_{i,t}}{p_{i,t_0}} \right) + h^c\left(c_t^b, \boldsymbol{p}_t\right) \log \left(\frac{c_t^b}{c_{t_0}^b} \right), \end{aligned}$$

for some bounded functions $h_i\left(c_t^b, \boldsymbol{p}_t\right)$ and $h^c\left(c_t^b, \boldsymbol{p}_t\right)$. We thus have:

$$\left| \log \left(\frac{c_t^b}{c_{t_0}^b} \right) \right| < M\left(\Delta_c + \Delta_p \right),$$

for some $M > 0$. □

Lemma 7. *Assume that the expenditure function $E(\cdot;\cdot)$ is second-order continuously differentiable in all its arguments. It follows that:*

$$\log \mathbb{P}_G\left(\boldsymbol{p}_{t_0}, \boldsymbol{q}_{t_0}; \boldsymbol{p}_t, \boldsymbol{q}_t\right) = \log \mathbb{P}_L\left(\boldsymbol{p}_{t_0}, \boldsymbol{q}_{t_0}; \boldsymbol{p}_t, \boldsymbol{q}_t\right) + O\left(\Delta^2\right) = \log \mathbb{P}_L\left(\boldsymbol{p}_{t_0}, \boldsymbol{q}_{t_0}; \boldsymbol{p}_t, \boldsymbol{q}_t\right) + O\left(\Delta^2\right),$$

$$\log \mathbb{P}_T\left(\boldsymbol{p}_{t_0},\boldsymbol{q}_{t_0};\boldsymbol{p}_t,\boldsymbol{q}_t\right)=\log \mathbb{P}_F\left(\boldsymbol{p}_{t_0},\boldsymbol{q}_{t_0};\boldsymbol{p}_t,\boldsymbol{q}_t\right)+O\left(\Delta^3\right)=\log \mathbb{P}_S\left(\boldsymbol{p}_{t_0},\boldsymbol{q}_{t_0};\boldsymbol{p}_t,\boldsymbol{q}_t\right)+O\left(\Delta^3\right),$$

where $\Delta \equiv \Delta_y + \Delta_p$ with Δ_y and Δ_p defined as in Equation (20), and where Laspeyres $\mathbb{P}_L$ and Paasche $\mathbb{P}_P$ indices are given by Eqution (18), Fisher by $\mathbb{P}_F \equiv (\mathbb{P}_P \cdot \mathbb{P}_L)^{\frac{1}{2}}$, and Sato-Vartia $\mathbb{P}_S = \prod_i \left(\frac{p_{i,t}}{p_{i,t_0}}\right)^{\omega_i}$ where Sato-Vartia weights are proportional to $\omega_i \propto s_{it}/s_{it_0}/\log\left(s_{it}/s_{it_0}\right)$ and sum to 1.

Proof. First, note that given the fact that the expenditure function is second-order continuously differentiable, we have

$$\log\left(\frac{s_{it+1}}{s_{it}}\right)=\sum_i \frac{\partial \log \widetilde{\Omega}_i\left(c_t;\boldsymbol{p}_{it}\right)}{\partial \log p_{it}}\log\left(\frac{p_{it+1}}{p_{it}}\right)+\frac{\partial \log \widetilde{\Omega}_i\left(c_t;\boldsymbol{p}_{it}\right)}{\partial \log y_t}\log\left(\frac{c_{t+1}}{c_t}\right)+O\left(\Delta^2\right), \quad \text{(A37)}$$

where $\Delta \equiv \Delta_p$ if the preferences are homothetic and $\Delta \equiv \Delta_p + \Delta_y$ from Lemma 6 above.

For the Laspeyres price index formula, we have:

$$\begin{aligned}
\log P_L &= \log\left(s_{it}\exp\left(\log\left(\tfrac{p_{it+1}}{p_{it}}\right)\right)\right),\\
&=\log\left(1+\sum_i s_{it}\log\left(\tfrac{p_{it+1}}{p_{it}}\right)+\tfrac{1}{2}\sum_i s_{it}\left(\log\left(\tfrac{p_{it+1}}{p_{it}}\right)\right)^2+O\left(\Delta^3\right)\right),\\
&=\sum_i s_{it}\log\left(\tfrac{p_{it+1}}{p_{it}}\right)+\tfrac{1}{2}\sum_i s_{it}\left(\log\left(\tfrac{p_{it+1}}{p_{it}}\right)\right)^2-\frac{1}{2}\left(\sum_i s_{it}\log\left(\tfrac{p_{it+1}}{p_{it}}\right)+\tfrac{1}{2}\sum_i s_{it}\left(\log\left(\tfrac{p_{it+1}}{p_{it}}\right)\right)^2\right)^2+O\left(\Delta^2\right),\\
&=\log P_G+O\left(\Delta^2\right),
\end{aligned}$$

where in the second and the third equality we have used the Taylor series expansion for $\exp(x)$ and $\log(1+x)$, respectively. For the Paasche price index formula, we find:

$$\begin{aligned}
\log P_P &= -\log\left(s_{it+1}\exp\left(-\log\left(\tfrac{p_{it+1}}{p_{it}}\right)\right)\right),\\
&=-\log\left(1-\sum_i s_{it+1}\log\left(\tfrac{p_{it+1}}{p_{it}}\right)+\tfrac{1}{2}\sum_i s_{it+1}\left(\log\left(\tfrac{p_{it+1}}{p_{it}}\right)\right)^2+O\left(\Delta^3\right)\right),\\
&=\sum_i s_{it+1}\log\left(\tfrac{p_{it+1}}{p_{it}}\right)-\tfrac{1}{2}\sum_i s_{it}\left(\log\left(\tfrac{p_{it+1}}{p_{it}}\right)\right)^2+\frac{1}{2}\left(\sum_i s_{it}\log\left(\tfrac{p_{it+1}}{p_{it}}\right)+\tfrac{1}{2}\sum_i s_{it}\left(\log\left(\tfrac{p_{it+1}}{p_{it}}\right)\right)^2\right)^2+O\left(\Delta^2\right),\\
&=\sum_i s_{it}\exp\left(\log\left(\tfrac{s_{it+1}}{s_{it}}\right)\right)\times\log\left(\tfrac{p_{it+1}}{p_{it}}\right)+O\left(\Delta^2\right),\\
&=\log P_G+O\left(\Delta^2\right),
\end{aligned}$$

where in the last equality, we have used the fact that $\log\left(\frac{s_{it+1}}{s_{it}}\right)=O(\Delta)$ from Equation (A37) above.

For the Fisher price index formula, we have:Fisher:

$$
\begin{aligned}
\log \mathbb{P}_F &= \frac{1}{2}\log \mathbb{P}_L + \frac{1}{2}\log \mathbb{P}_P, \\
&= \frac{1}{2}\log\left(s_{it}\exp\left(\log\left(\tfrac{p_{it+1}}{p_{it}}\right)\right)\right) - \frac{1}{2}\log\left(s_{it+1}\exp\left(-\log\left(\tfrac{p_{it+1}}{p_{it}}\right)\right)\right), \\
&= \frac{1}{2}\log\left(1 + \sum_i s_{it}\log\left(\tfrac{p_{it+1}}{p_{it}}\right) + \tfrac{1}{2}\sum_i s_{it}\left(\log\left(\tfrac{p_{it+1}}{p_{it}}\right)\right)^2\right) \\
&\quad - \frac{1}{2}\log\left(1 - \sum_i s_{it+1}\log\left(\tfrac{p_{it+1}}{p_{it}}\right) + \tfrac{1}{2}\sum_i s_{it+1}\left(\log\left(\tfrac{p_{it+1}}{p_{it}}\right)\right)^2\right) + O\left(\Delta^3\right), \\
&= \frac{1}{2}\sum_i (s_{it}+s_{it+1})\log\left(\tfrac{p_{it+1}}{p_{it}}\right) + \frac{1}{4}\sum_i s_{it}\left(\log\left(\tfrac{p_{it+1}}{p_{it}}\right)\right)^2 - \frac{1}{4}\sum_i s_{it+1}\left(\log\left(\tfrac{p_{it+1}}{p_{it}}\right)\right)^2 \\
&\quad - \frac{1}{4}\left(\sum_i s_{it}\log\left(\tfrac{p_{it+1}}{p_{it}}\right) + \frac{1}{2}\sum_i s_{it}\left(\log\left(\tfrac{p_{it+1}}{p_{it}}\right)\right)^2\right)^2 \\
&\quad + \frac{1}{4}\left(\sum_i s_{it+1}\log\left(\tfrac{p_{it+1}}{p_{it}}\right) - \frac{1}{2}\sum_i s_{it+1}\left(\log\left(\tfrac{p_{it+1}}{p_{it}}\right)\right)^2\right)^2 + O\left(\Delta^3\right), \\
&= \log \mathbb{P}_T + \frac{1}{4}\left(\sum_i s_{it}\left(\log\left(\tfrac{p_{it+1}}{p_{it}}\right)\right)^2 - \sum_i s_{it+1}\left(\log\left(\tfrac{p_{it+1}}{p_{it}}\right)\right)^2\right) + O\left(\Delta^3\right) \\
&+ \frac{1}{4}\left[\sum_i s_{it+1}\log\left(\tfrac{p_{it+1}}{p_{it}}\right) - \sum_i s_{it}\log\left(\tfrac{p_{it+1}}{p_{it}}\right) - \tfrac{1}{2}\sum_i s_{it+1}\left(\log\left(\tfrac{p_{it+1}}{p_{it}}\right)\right)^2 - \tfrac{1}{2}\sum_i s_{it}\left(\log\left(\tfrac{p_{it+1}}{p_{it}}\right)\right)^2\right] \\
&\quad \times\left[\sum_i s_{it+1}\log\left(\tfrac{p_{it+1}}{p_{it}}\right) - \tfrac{1}{2}\sum_i s_{it+1}\left(\log\left(\tfrac{p_{it+1}}{p_{it}}\right)\right)^2 + \sum_i s_{it}\log\left(\tfrac{p_{it+1}}{p_{it}}\right) + \tfrac{1}{2}\sum_i s_{it}\left(\log\left(\tfrac{p_{it+1}}{p_{it}}\right)\right)^2\right], \\
&= \log \mathbb{P}_T - \frac{1}{4}\sum_i s_{it}\left(\tfrac{s_{it+1}-s_{it}}{s_{it}}\right)\left(\log\left(\tfrac{p_{it+1}}{p_{it}}\right)\right)^2 \\
&\quad + \frac{1}{2}\left[\sum_i s_{it}\left(\tfrac{s_{it+1}-s_{it}}{s_{it}}\right)\log\left(\tfrac{p_{it+1}}{p_{it}}\right) - \sum_i \bar{s}_{it}\left(\log\left(\tfrac{p_{it+1}}{p_{it}}\right)\right)^2\right] \\
&\quad \times\left[\sum_i \bar{s}_{it}\log\left(\tfrac{p_{it+1}}{p_{it}}\right) - \frac{1}{4}\sum_i s_{it}\left(\tfrac{s_{it+1}-s_{it}}{s_{it}}\right)\left(\log\left(\tfrac{p_{it+1}}{p_{it}}\right)\right)^2\right] + O\left(\Delta^3\right), \\
&= \log \mathbb{P}_T + O\left(\Delta^3\right).
\end{aligned}
$$

where we have let $\bar{s}_{it} \equiv \frac{1}{2}\left(s_{it}+s_{it+1}\right)$ and have used Equation (A37) and the following approximations:

$$
\begin{aligned}
\tfrac{s_{it+1}-s_{it}}{s_{it}} &= \exp\left(\log\left(\tfrac{s_{it+1}}{s_{it}}\right)\right) - 1 = O(\Delta), \\
\sum_i \bar{s}_{it}\left(\log\left(\tfrac{p_{it+1}}{p_{it}}\right)\right)^m &= O(\Delta^m), \qquad 1 \le m.
\end{aligned}
$$

Finally, for Sato-Vartia, we begin by the following approximation:

$$\begin{aligned}
\frac{s_{it+1}-s_{it}}{\log\left(\frac{s_{it+1}}{s_{it}}\right)} &= \frac{\frac{1}{2}s_{it}\left(\exp\left(\log\left(\frac{s_{it+1}}{s_{it}}\right)\right)-1\right)+\frac{1}{2}s_{it+1}\left(1-\exp\left(-\log\left(\frac{s_{it+1}}{s_{it}}\right)\right)\right)}{\log\left(\frac{s_{it+1}}{s_{it}}\right)}, \\
&= \frac{1}{2}s_{it}\left(1+\frac{1}{2}\log\left(\frac{s_{it+1}}{s_{it}}\right)+\frac{1}{6}\left(\log\left(\frac{s_{it+1}}{s_{it}}\right)\right)^2\right) \\
&\quad +\frac{1}{2}s_{it+1}\left(1-\frac{1}{2}\log\left(\frac{s_{it+1}}{s_{it}}\right)+\frac{1}{6}\left(\log\left(\frac{s_{it+1}}{s_{it}}\right)\right)^2\right)+O\left(\Delta^3\right), \\
&= \overline{s}_{it}\left(1+\tfrac{1}{6}\left(\log\left(\frac{s_{it+1}}{s_{it}}\right)\right)^2\right)-\frac{1}{4}\left(s_{it+1}-s_{it}\right)\log\left(\frac{s_{it+1}}{s_{it}}\right), \\
&= \overline{s}_{it}\left(1+\tfrac{1}{6}\left(\log\left(\frac{s_{it+1}}{s_{it}}\right)\right)^2\right)-\frac{s_{it}}{4}\left(\exp\left(\log\left(\frac{s_{it+1}}{s_{it}}\right)\right)-1\right)\log\left(\frac{s_{it+1}}{s_{it}}\right)+O\left(\Delta^3\right), \\
&= \overline{s}_{it}\left(1+\tfrac{1}{6}\left(\log\left(\frac{s_{it+1}}{s_{it}}\right)\right)^2\right)-\frac{s_{it}}{4}\left(\log\left(\frac{s_{it+1}}{s_{it}}\right)\right)^2+O\left(\Delta^3\right),
\end{aligned}$$

where we have again used Equation (A37). Substituting this result in the definition of the price index formula, we find

$$\begin{aligned}
\log\mathbb{P}_S &\equiv \frac{\log\mathbb{P}_T+O\left(\Delta^3\right)}{1+O\left(\Delta_s^3\right)}, \\
&= \log\mathbb{P}_T+O\left(\Delta_s^3\right),
\end{aligned}$$

where we have used the fact that $\sum_i \overline{s}_{it}=1$. □

Lemma 8. *Assume that we observe $(y^n,x^n)_{n=1}^N$ such that*

$$\begin{aligned}
y^n &= f\left(z^n\right)+\varepsilon^n, \\
x^n &= z^n+v^n,
\end{aligned}$$

with $y^n,z^n,x^n\in\mathbb{R}$ where the following conditions are satisfied:

1. *z^n is distributed according to a probability distribution function that is bounded away from zero over the interval $[\underline{z},\overline{z}]$*
2. *Function $f(\cdot)$ is continuouly differentiable of order m over the interval $[\underline{z},\overline{z}]$*
3. *$|\varepsilon^n|<\Delta_\varepsilon$ and $|v^n|<\delta_v$.*

Let coefficients $(\widehat{\alpha}_k)_{k=0}^{K_N}$ solve the following problem:

$$\min_{(\alpha_k)_{k=0}^{K_N}}\sum_{n=1}^{N}\left(y^n-\sum_{k=0}^{K_N}\alpha_k g_k\left(x^n\right)\right)^2, \tag{A38}$$

where $g_k(q)$ is the Legendre polynomial of order k. Then, we have

$$f'(q) = \sum_{k=0}^{K_N} \widehat{\alpha}_k g_k'(q) + O(\delta_v) + O_p\left(K_N^3\left(\sqrt{\frac{K_N}{N}} \cdot \Delta_\varepsilon^2 + K_N^{-(m-1)}\right)\right). \tag{A39}$$

If $z^n, x^n \in \mathbb{R}^D$ for $D \geq 2$ and z^n belongs to a Cartesian product of compact connected intervals, and its probability distribution is bounded away from zero over this set, Equation A39 holds for any arbitrary integer m if function $f(\cdot)$ is analytical.

Proof. The proof closely follows the proof of Theorem 1 of Newey (1997). Define $\boldsymbol{g}(z) \equiv \left[g_0(z), \cdots, g_{K_N}(z)\right]^t$ where superscript t stands for the transpose of the matrix, and let

$$\begin{aligned} \boldsymbol{G}^* &\equiv \left[\boldsymbol{g}(z^1), \cdots, \boldsymbol{g}(z^n)\right]^t, \\ \boldsymbol{G} &\equiv \left[\boldsymbol{g}(x^1), \cdots, \boldsymbol{g}(x^n)\right]^t. \end{aligned}$$

Define:

$$\begin{aligned} \widehat{\boldsymbol{\alpha}} &\equiv (\boldsymbol{G}^t \boldsymbol{G})^{-1} \boldsymbol{G}^t \boldsymbol{y}, \\ \boldsymbol{\alpha}^* &\equiv \left((\boldsymbol{G}^*)^t \boldsymbol{G}^*\right)^{-1} (\boldsymbol{G}^*)^t \boldsymbol{y}. \end{aligned}$$

First, assumptions (i) and (ii) correspond to Assumptions 8 and 9 in Newey (1997). This implies that Assumption 3 of Newey (1997) is satisfied for first derivative function such that:

$$\sup_{z \in [\underline{z}, \overline{z}]} \left| f'(z) - \sum_{k=0}^{K_N} \alpha_k^* g_k'(z) \right| = O\left(K_N^{-(m-1)}\right).$$

Using the results of Section 5, we additionally find that:

$$\xi_1(K_N) \equiv \sup_{z \in [\underline{z}, \overline{z}]} \left\| \left(g_0'(z), \cdots, g_{K_N}'(z)\right) \right\| = O\left(K_N^3\right),$$

where $||\cdots||$ corresponds to the Euclidean norm. It follows from the same steps as in the proof of Theorem 1 of Newey (1997) that:

$$f'(z) = \sum_{k=0}^{K_N} \alpha_k^* g_k'(z) + O_p\left(K_N^3\left(\sqrt{\frac{K_N}{N}} \cdot \Delta_\varepsilon^2 + K_N^{-(m-1)}\right)\right).$$

with the only difference being the fact that here $\mathbb{E}[\varepsilon_n \varepsilon_{n'}]$ is not a constant, but rather we have $\mathbb{E}[\varepsilon_n \varepsilon_{n'}] = O\left(\Delta_\varepsilon^2\right)$.

Define $\boldsymbol{g}(z) \equiv \left[g_0'(z), \cdots, g_{K_N}'(z)\right]^t$ and note that:

$$\boldsymbol{G} = \boldsymbol{G}^* + \left[\boldsymbol{g}'\left(x^1\right) \cdot v^1, \cdots, \boldsymbol{g}'\left(x^n\right) \cdot v^n\right]^t + O\left(\delta_v^2\right),$$

which implies:

$$\widehat{\boldsymbol{\alpha}} = \boldsymbol{\alpha}^* + O\left(\delta_v\right).$$

Equation (A39) then follows from the observation that:

$$\sum_{k=0}^{K_N} \alpha_k^* g_k'(z) - \sum_{k=0}^{K_N} \widehat{\alpha}_k g_k'(z) = O\left(\delta_v\right).$$

□

Lemma 9. *Assume that the expenditure function $E(\cdot;\cdot)$ is second-order continuously differentiable. Then the growth in real consumption between periods t_0 and t satisfies*

$$\log\left(\frac{c_t^b}{c_{t_0}^b}\right) = \frac{1}{1+\Lambda_t^b\left(c_{t_0}^b\right)} \log\left(\frac{y_t/y_{t_0}}{\mathscr{P}_{t_0,t}^b\left(c_{t_0}^b\right)}\right) + O\left(\Delta^2\right), \tag{A40}$$

where $\Delta \equiv \Delta_p + \Delta_y$ with Δ_y and Δ_p defined as in Equation (20).

Proof. First, note that we have:

$$\log\left(\frac{y_t}{y_{t_0}}\right) = \log \frac{\widetilde{E}^b\left(c_t^b; \boldsymbol{p}_t\right)}{\widetilde{E}^b\left(c_{t_0}^b; \boldsymbol{p}_{t_0}\right)}.$$

We can do a first-order Taylor expansion of the left-hand-side of the equation above in terms of c_t^b, and use Lemma 6 to find:

$$\log\left(\frac{y_t}{y_{t_0}}\right) = \log \frac{\widetilde{E}^b\left(c_{t_0}^b; \boldsymbol{p}_t\right)}{\widetilde{E}^b\left(c_{t_0}^b; \boldsymbol{p}_{t_0}\right)} + \left.\frac{\partial \log \widetilde{E}^b\left(c; \boldsymbol{p}_t\right)}{\partial \log c}\right|_{c \equiv c_{t_0}^b} \cdot \log\left(\frac{c_t^b}{c_{t_0}^b}\right) + O\left(\Delta^2\right).$$

□

Lemma 10. *If the expenditure function $\log \widetilde{E}(\cdot;\cdot)$ is continuously differentiable of order at least 3, then we have*

$$\log\left(\frac{c_t^b}{c_{t_0}^b}\right) = \frac{1}{1+\frac{1}{2}\left[\Lambda_{t_0}^b\left(c_{t_0}^b\right) + \Lambda_t^b\left(c_t^b\right)\right]} \log\left(\frac{y_t/y_{t_0}}{\mathbb{P}_T\left(\boldsymbol{p}_{t_0}, \boldsymbol{q}_{t_0}; \boldsymbol{p}_t, \boldsymbol{q}_t\right)}\right) + O\left(\Delta^3\right), \tag{A41}$$

where $\Delta \equiv \Delta_p + \Delta_y$ with Δ_y and Δ_p defined as in Equation (20).

‘

Proof. Once again, we start with:

$$\log\left(\frac{y_t}{y_{t_0}}\right) = \log \frac{\widetilde{E}^b\left(c_t; \boldsymbol{p}_t\right)}{\widetilde{E}^b\left(c_{t_0}; \boldsymbol{p}_{t_0}\right)},$$

and rely on Lemma 5 for variables $\boldsymbol{x} \equiv (c; \boldsymbol{p})$ to find:

$$\begin{aligned}
\log\left(\frac{y_t^n}{y_{t_0}^n}\right) = & \frac{1}{2}\sum_{i=1}^{I}\left[\left.\frac{\partial \log \widetilde{E}^b(c;\boldsymbol{p})}{\partial \log p_i}\right|_{(c;\boldsymbol{p})\equiv\left(c_{t_0};\boldsymbol{p}_{t_0}\right)} + \left.\frac{\partial \log \widetilde{E}^b(c;\boldsymbol{p})}{\partial \log p_i}\right|_{(c;\boldsymbol{p})\equiv\left(c_{t+1};\boldsymbol{p}_{t+1}\right)}\right] \cdot \log\left(\frac{p_{i,t+1}}{p_{i,t}}\right) \\
& + \frac{1}{2}\left[\left.\frac{\partial \log \widetilde{E}^b(c;\boldsymbol{p})}{\partial \log c}\right|_{(c;\boldsymbol{p})\equiv\left(c_{t_0};\boldsymbol{p}_{t_0}\right)} + \left.\frac{\partial \log \widetilde{E}^b(c;\boldsymbol{p})}{\partial \log c}\right|_{(c;\boldsymbol{p})\equiv\left(c_{t+1};\boldsymbol{p}_{t+1}\right)}\right] \cdot \log\left(\frac{c_{t+1}}{c_t}\right) \\
& + O\left(\Delta^3\right), \\
= & \frac{1}{2}\sum_{i=1}^{I}\left[\widetilde{\Omega}_i^b\left(c_{t_0};\boldsymbol{p}_{t_0}\right) + \widetilde{\Omega}_i^b\left(c_t;\boldsymbol{p}_t\right)\right]\log\left(\frac{p_{i,t+1}}{p_{i,t}}\right) \\
& + \left(1 + \frac{1}{2}\left[\Lambda_{t_0}\left(c_{t_0};\boldsymbol{p}_{t_0}\right) + \Lambda_t\left(c_t;\boldsymbol{p}_t\right)\right]\right) \cdot \log\left(\frac{c_t}{c_{t_0}}\right) + O\left(\Delta^3\right),
\end{aligned}$$

where in the second equality we have again used Shephard's lemma, as well as the definition of the first-order nonhomotheticity correction function. □

B Data Appendix

CPI dataset. The majority of our datasets are linked to the Consumer Price Index (CPI) data series, which contain monthly or quarterly price indexes for over 200 detailed product categories, with various time frames for availability.[30] To obtain a balanced panel of inflation series derived from the CPI price indexes, whenever a category is missing we use a more aggregate series in the product hierarchy as proxy, since higher-level series usually has longer time coverage. If broad categories also have limited data availability, but one or more immediate sub-categories are populated, then we take the simple average of sub-categories as a proxy.

Main linked CEX-CPI dataset. The preferred dataset contains 19 expenditure product categories that collectively cover the full consumption basket from 1953 to 2019.[31] These product

[30] The data is available at https://download.bls.gov/pub/time.series/cu.

[31] The 19 categories are: alcoholic beverages; apparel and services; entertainement; food at home; food away from

categories are defined in the annual summary tables of expenditures from the Consumer Expenditure Survey (CEX) published by the U.S. Bureau of Labor Statistics, which include expenditure shares across all items by quintiles of income before taxes from 1984 to 2019. For prior year, we extrapolate expenditure shares based on the available data in 1984. Using a crosswalk we build by hand based on matching of product description and aggregation level, these 19 categories are each mapped to one or more inflation series from the CPI price data. Table 1 below reports a sample of the crosswalk we build to link price and expenditure categories.

Robustness dataset #1. The first alternative dataset for robustness check uses the official consumption weights used by the Bureau of Labor Statistics for calculating the U.S. overall price index.[32] This dataset is limited in terms of the numbers of product categories but has the benefit of an extended time frame. The ten broad product categories included in this dataset are: food and beverages, shelter, fuels and utilities, household furnishings and operations, apparel, transportation, medical care, recreation, education and communication, other goods and services. Due to the evolution of product categories and product hierarchy over the years, some sub-categories are reassigned by BLS from one broad category to another over time. For example, BLS places "Telephone services" under housing until 1997, then under "Education and communication." To address this issue, we adjust the placement of certain sub-categories and their allocated weights so that the composition of broad categories remains consistent from 1953 to 2020. In addition to the aggregate consumption weights, our linked dataset uses aggregate expenditures across income quintiles from the Consumer Expenditure Survey summary tables published by the BLS, which is available from 1984 onwards, as in the main dataset. Prior to 1984, we assume the expenditure levels to be constant and identical to 1984. We use the expenditure shares of each income quintile to distribute aggregate consumption across income groups, so that we obtain a linked dataset with consumption patterns by income groups while keeping aggregate consumption weights identical to the official weights of the BLS.

Data on consumption growth by income groups. The CEX annual summary tables, which we use to obtain expenditures and expenditure share by income groups, also contain information on total average annual expenditure as well as average income before and after tax for each quintile. This data is used to measure the nominal growth rates of consumption by income quintiles from 1984 to 2000. Prior to 1984, we assign a common growth rate to all income groups using the growth rate of per capita personal consumption expenditures, as measured by the BEA.

---

home; gasoline, other fuels, and motor oil; healthcare; households furnishings and equipment; household operations; housekeeping supplies; miscellaneous; public and other transportation; vehicle purchases; other vehicle expenses; personal care products and services; reading; shelter; tobacco products and smoking supplies; utilities, fuels, and public services.

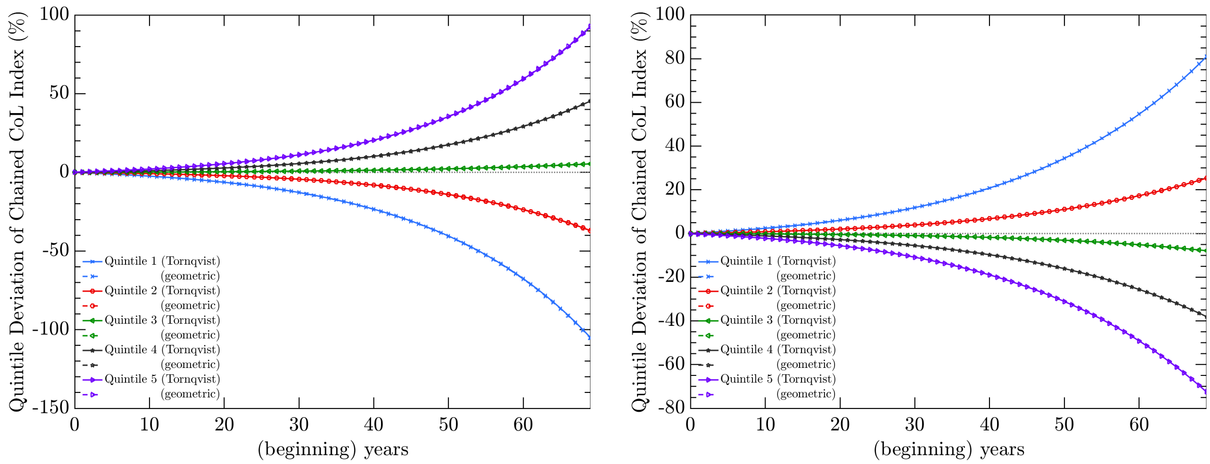
[32]The official consumption weights are available at https://www.bls.gov/cpi/tables/relative-importance/home.htm.

Table A1: Crosswalk Between CEX Product Category and CPI Price Series

Product Category in CEX Data	Linked CPI Series
Alcoholic beverages	CUSR0000SAF116 – Alcoholic beverages
Apparel and services	CUSR0000SAA – Apparel
	CUUR0000SEGD03 – Laundry and dry cleaning services
	CUUR0000SEGD04 – Apparel services other than laundry and dry cleaning
Entertainment	CUUR0000SERA – Video and audio; CUSR0000SERC – Sporting goods
	CUUR0000SERD – Photography; CUSR0000SERE – Other recreational goods
	CUSR0000SERF – Other recreation services
	CUUR0000SEEE03 – Internet services and electronic information providers
Food at home	CUSR0000SAF11 – Food at home
Food away from home	CUSR0000SEFV – Food away from home
Gasoline, other fuels, and motor oil	CUSR0000SETB – Motor fuel; CUUR0000SS47021 – Motor oil, coolant, and fluids
Healthcare	CUSR0000SAM – Medical care
Household furnishings and equipment	CUUR0000SEHH – Window and floor coverings and other linens
	CUSR0000SEHJ – Furniture and bedding
	CUUR0000SEHK – Appliances
	CUUR0000SEHL – Other household equipment and furnishings
	CUUR0000SEHM – Tools, hardware, outdoor equipment and supplies
	CUUR0000SEEE01 – Computers, peripherals, and smart home assistants
Household operations	CUUR0000SEHP – Household operations
Housekeeping supplies	CUSR0000SEHN – Housekeeping supplies; CUUR0000SEEC01 – Postage
Miscellaneous	CUUR0000SEGD – Miscellaneous personal services; CUUR0000SEGE – Miscellaneous personal goods
Public and other transportation	CUSR0000SETG – Public transportation
Vehicle purchases	CUSR0000SETA – New and used motor vehicles
Other vehicle expenses	CUUR0000SETC – Motor vehicle parts and equipment
	CUSR0000SETD – Motor vehicle maintenance and repair
	CUSR0000SETE – Motor vehicle insurance; CUUR0000SETF – Motor vehicle fees
Personal care products and services	CUUR0000SEGB – Personal care products; CUUR0000SEGC – Personal care services
Reading	CUSR0000SERG – Recreational reading materials
Shelter	CUSR0000SAH1 – Shelter
Tobacco products and smoking supplies	CUSR0000SEGA – Tobacco and smoking products
Utilities, fuels, and public services	CUSR0000SAH2 – Fuels and utilities; CUUR0000SEED – Telephone services

Figure A1: Example: Chained Cumulative Price Index Formulas by Quintile

(a) Positive Covariance (b) Negative Covariance



Note: Panels (a) and (b) show the deviation from average of the first and second order cumulative price index formulas for households at different quntiles of initial total expenditure over the period for the cases with positive and negative covariances of price inflations and income elasticities, respectively.

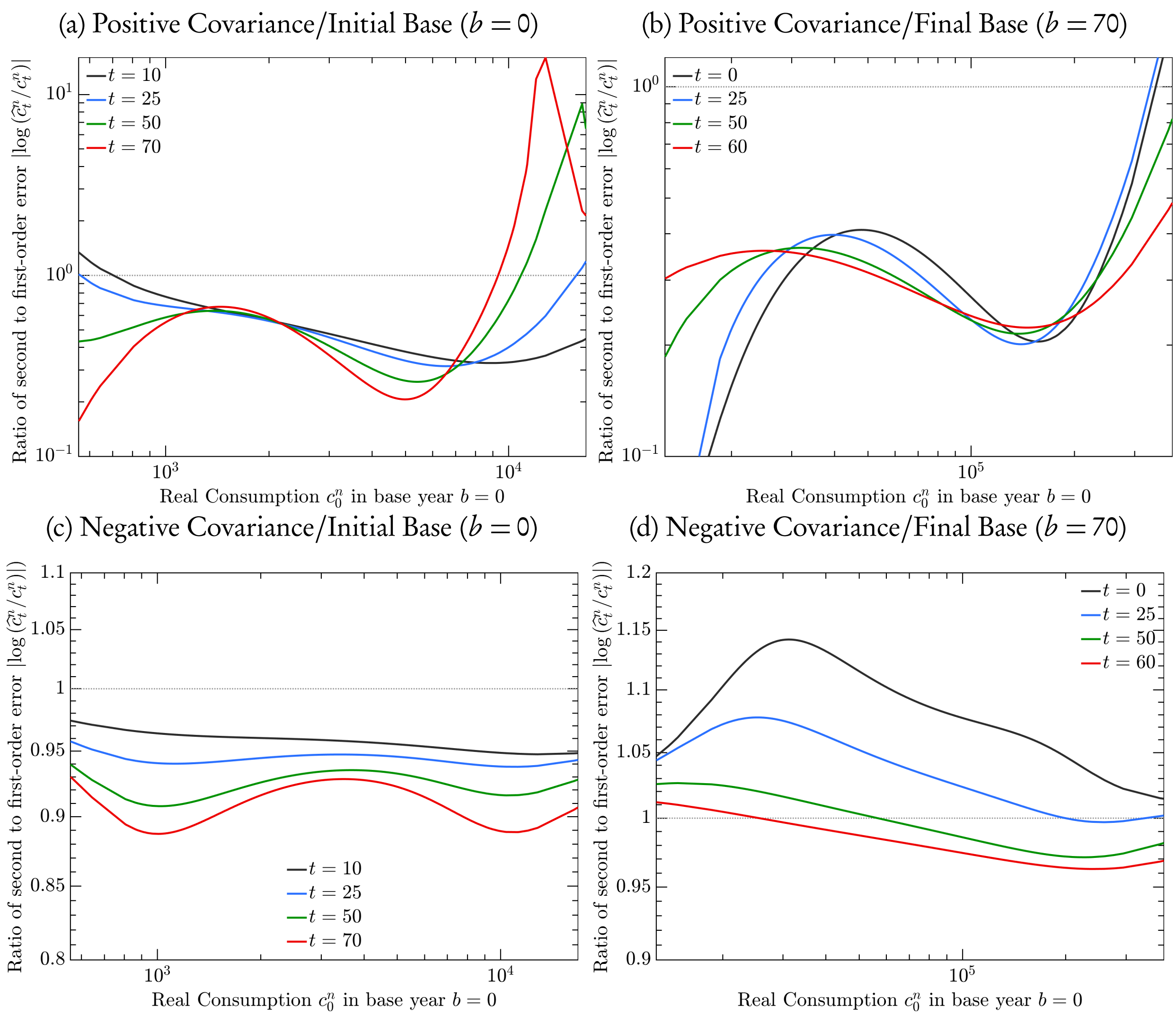
Nielsen data. As a third robustness exercise, we study product-level data covering consumer packaged goods. The Nielsen Homescan Consumer Panel (HMS) records consumption from 2004 to 2015 for a rotating panel of about 50,000 households, who are instructed to scan any product they purchase that has a barcode. These products are typically found in department stores, grocery stores, drug stores, convenience stores, and other similar retail outlets. Price and expenditure data are observed for each barcode, along with the socio-demographic characteristics of the participants, including household income.

C Additional Results

C.1 Further Details on the Illustrative Example

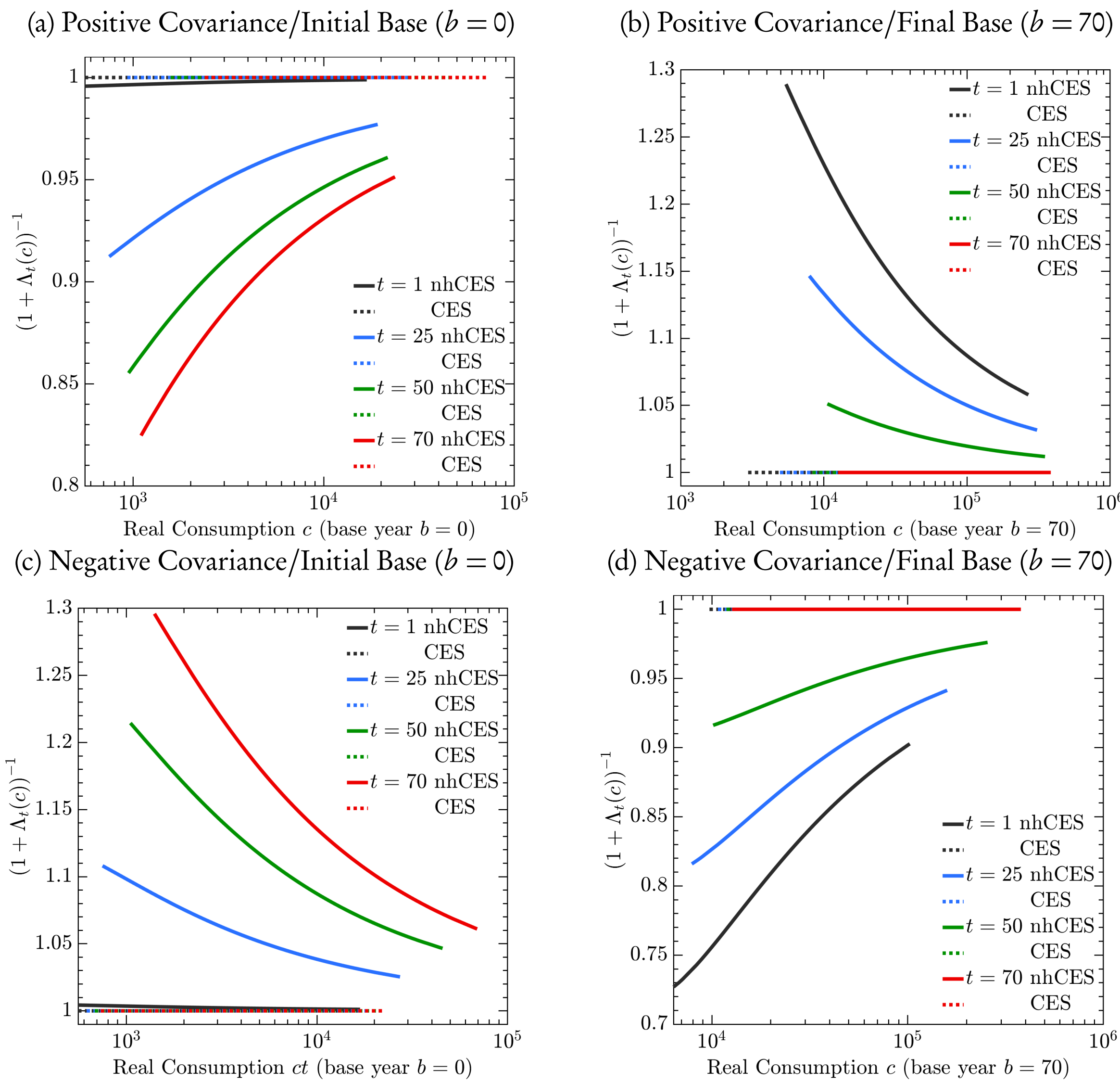
C.2 Additional Empirical Results

Figure A2: nhCES Example: Second vs. First-order Correction



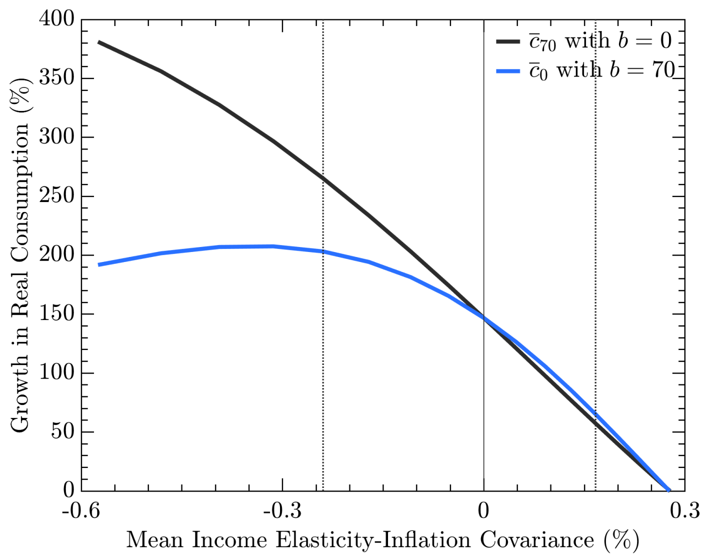
Note: The figures compare the error in the approximate value of real consumption between the the first-order and second-order algorithms. The correct value of real consumption is calculated based on the underlying parameters of the nhCES preferences. The panels show the error for the choices of base period (a) $b = 0$ and (b) $b = 70$ with the positive income elasticity-inflation covariance and (c) $b = 0$ and (d) $b = 70$ with the negative covariance.

Figure A3: Example: The Nonhomotheticity Correction



Note: The figure shows the ratios of the ReC and CoL quantity indices across households as a function of the current real consumption for the choices of base period (a) $b = 0$ and (b) $b = 70$ with the positive income elasticity-inflation covariance and (c) $b = 0$ and (d) $b = 70$ with the negative covariance. The underlying data is drawn based on preferences defined in Equation (33) with parameters corresponding to a nonhomothetic CES $(\sigma, \varepsilon_a, \varepsilon_m, \varepsilon_s) = (0.3, 0.2, 1, 1.7)$ (nhCES) and homothetic CES $(\sigma, \varepsilon_a, \varepsilon_m, \varepsilon_s) = (0.3, 1, 1, 1)$ functions.

Figure A4: Example: Real Consumption Growth and Income Elasticity-Inflation Covariance

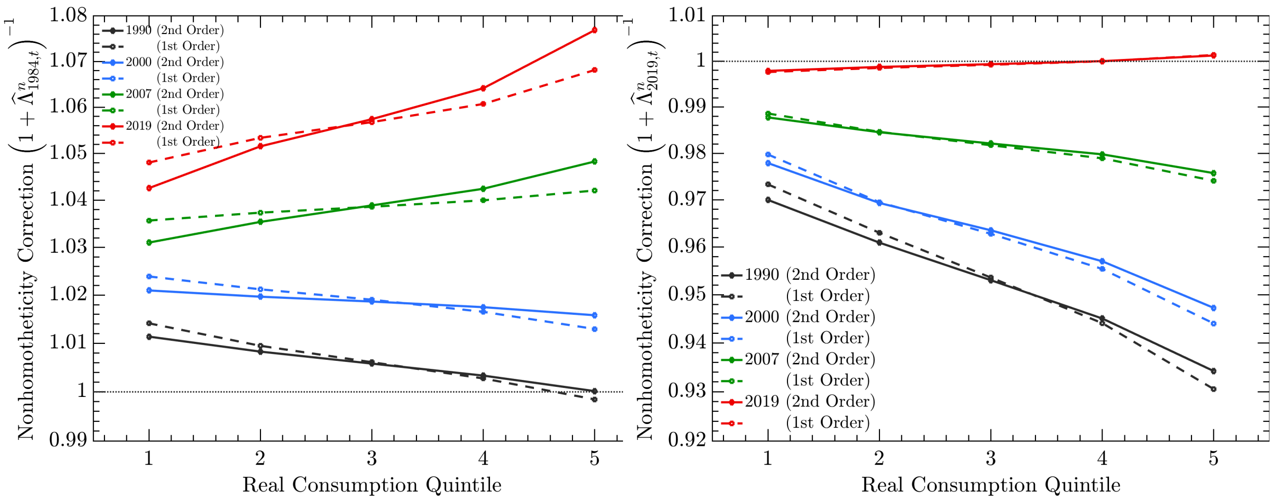


Note: The figure compares the growth in average real consumption for the initial and final periods as base, respectively, as a function of the mean covariance between price inflations and expenditure elasticities over the period.

Figure A5: CEX and BLS Data: Nonhomotheticity Correction

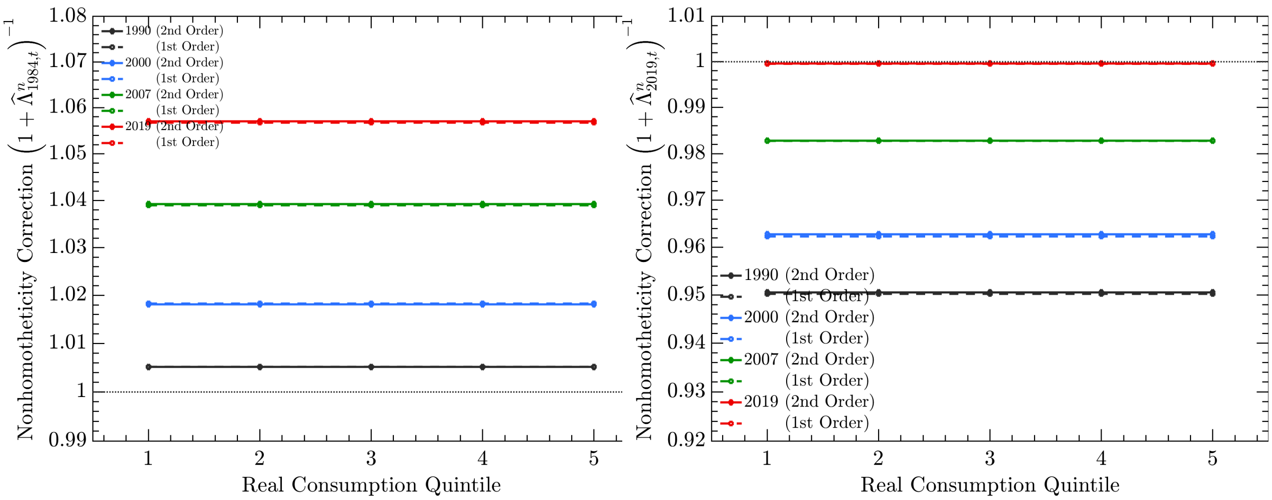
(a) Initial Period as Base ($b = 1984$)

(b) Final Period as Base ($b = 2019$)



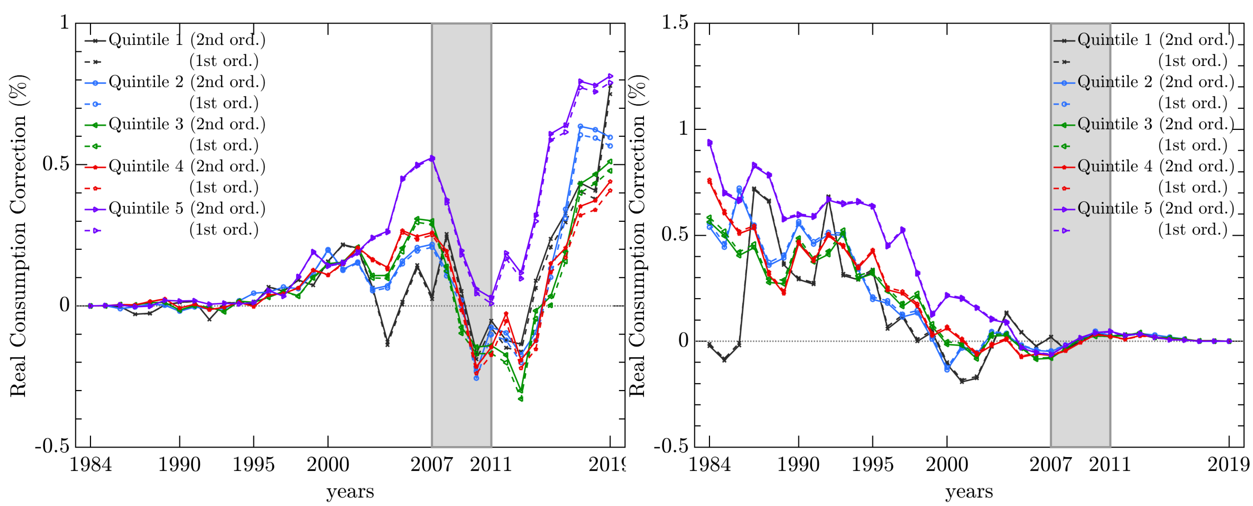
Note: Panels (a) and (b) show how the "Nonhomotheticity Correction" varies over time for different quintiles of income for the initial and final years as the base period.

Figure A6: CEX and BLS Data: Nonhomotheticity Correction ($K = 1$)



Note: Panels (a) and (b) show how the "Nonhomotheticity Correction" varies over time for different quintiles of income for the initial and final years as the base period with $K = 1$.

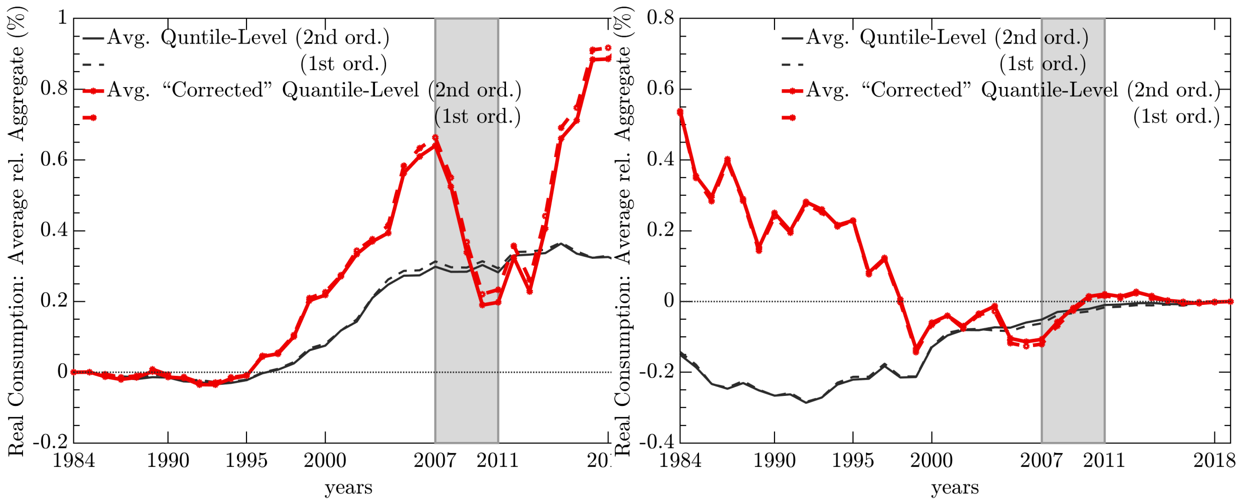
Figure A7: CEX and BLS Data: Corrected relative to Uncorrected Real Consumption ($K = 1$)



Note: Panels (a) and (b) show how the ratio of the corrected to the uncorrected measures of real consumption vary over time for each quintile of initial real consumption for the initial and final years as the base period, respectively, with $K = 1$. The great recession has been indicated in grey background.

Figure A8: CEX and BLS Data: Average Growth in Real Consumption ($K = 1$)

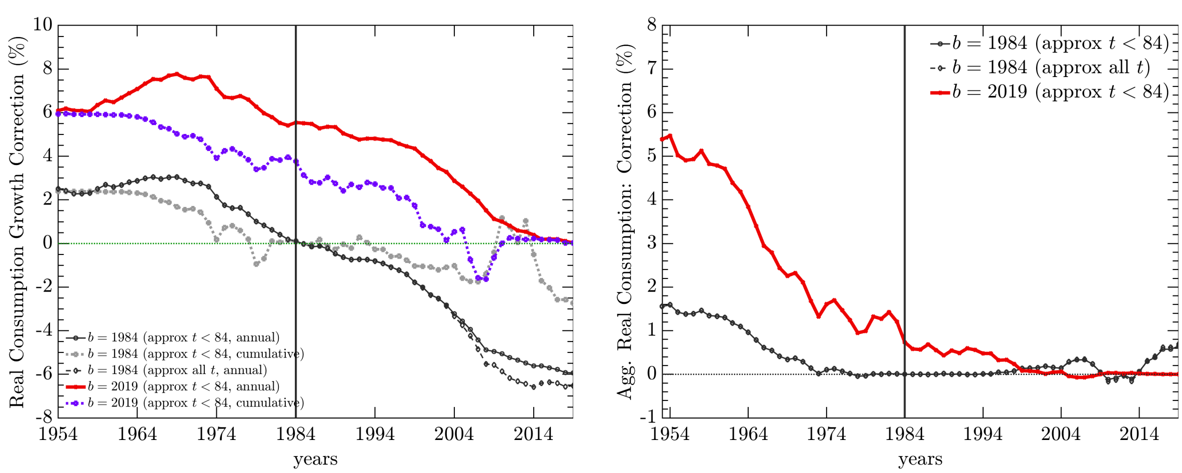
(a) Initial Period as Base ($b = 1984$) (b) Final Period as Base ($b = 2019$)



Note: The figure shows the evolution the average corrected and uncorrected measures of real consumption across quintiles relative to the measure of aggregate real consumption that ignores income heterogeneity. The latter defines the reduced-form index of real consumption using aggregate consumption expenditure shares. Panels (a) and (b) show the correction for the initial and final years as the base period, respectively, with $K = 1$. The reduced-form prices indices used for the 2nd and 1st order approximations are geometric and Tornqvist indices, respectively.

Figure A9: CEX and BLS Data: Correction Over Longer Horizon ($K = 2$)

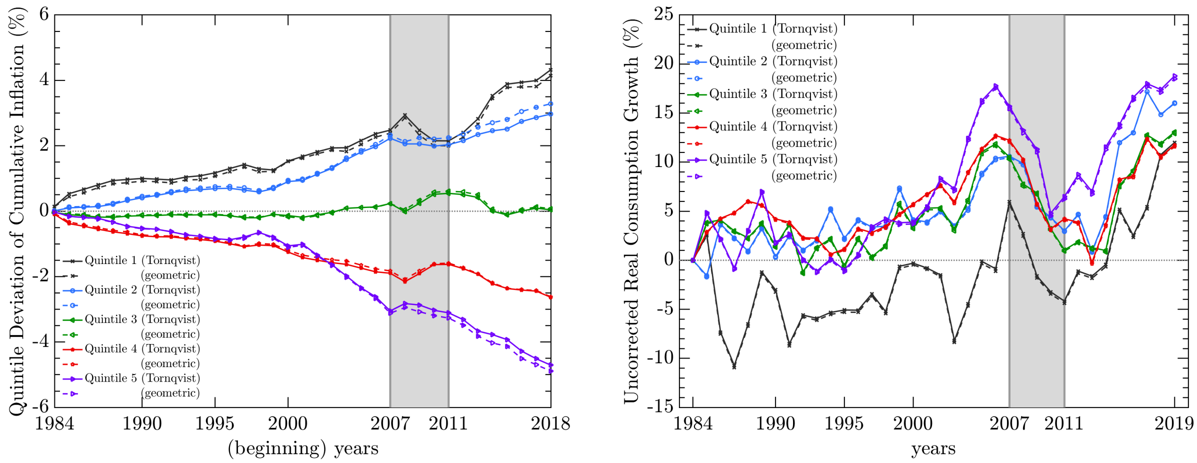
(a) Nonhomotheticity Growth Bias (b) Real Consumption Correction



Note: Panels (a) and (b) show the evolution of the nonhomotheticity correction and the corrected relative to uncorrected index of real consumption using aggregate consumption expenditure shares.

Figure A10: BLS Data: Conventional Price and Quantity Indices

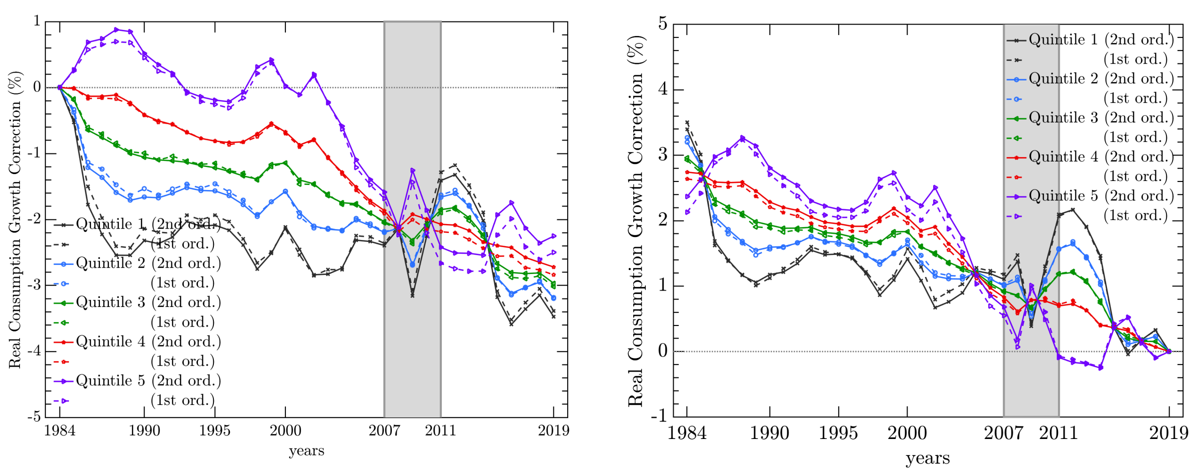
(a) Quintile Deviation of Price Indices (b) Quintile Quantity Indices



Note: Panels (a) and (b) show the evolution of cumulative reduced-form price and quantity indices for each quintile of income over the period, respectively. The great recession has been indicated in grey background.

Figure A11: BLS Data: Bias in Reduced-form Real Consumption Growth ($K = 2$)

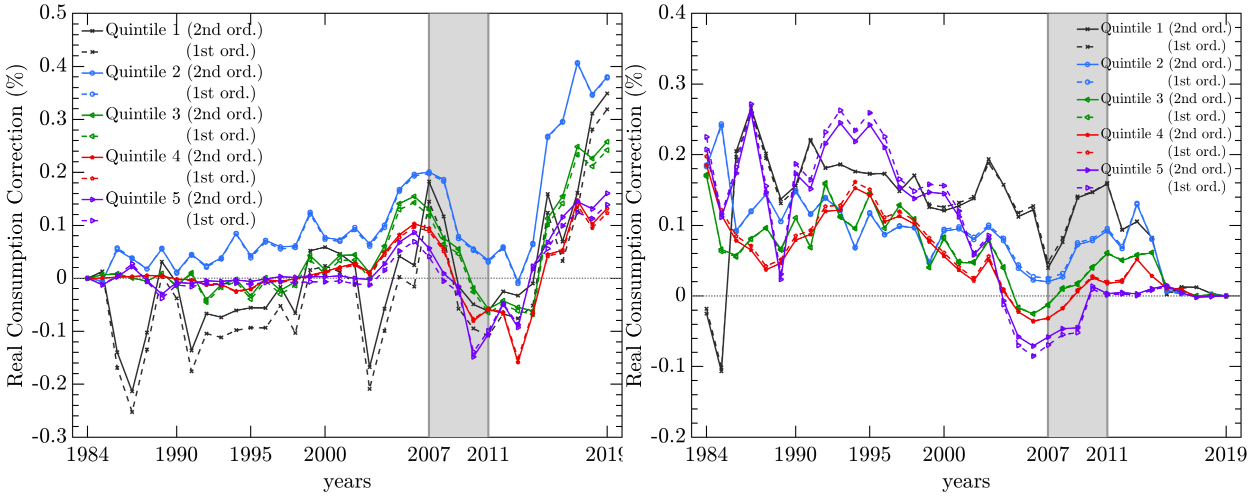
(a) Initial Period as Base ($b = 1984$) (b) Final Period as Base ($b = 2019$)



Note: Panels (a) and (b) show how the evolution of the annual bias in the reduced-form measures of real consumption growth $\lambda^n_{b,t}$, defined in Equation (29), for different quintiles of income for the initial and final years as the base period.

Figure A12: BLS Data: Corrected relative to Uncorrected Real Consumption ($K = 2$)

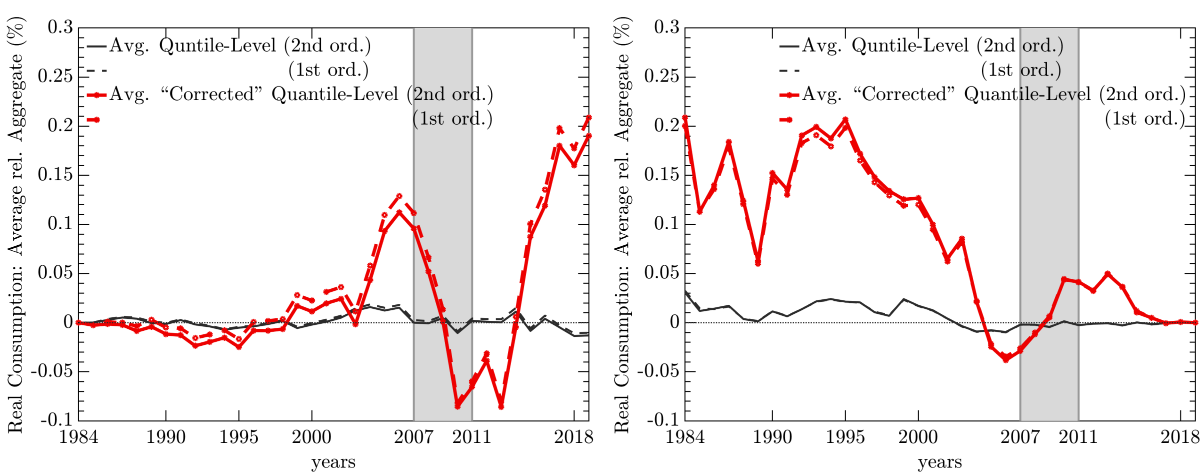
(a) Initial Period as Base ($b = 1984$) (b) Final Period as Base ($b = 2019$)



Note: Panels (a) and (b) show how the ratio of the corrected to the uncorrected measures of real consumption vary over time for each quintile of initial real consumption for the initial and final years as the base period, respectively. The great recession has been indicated in grey background.

Figure A13: BLS Data: Average Growth in Real Consumption ($K = 2$)

(a) Initial Period as Base ($b = 1984$) (b) Final Period as Base ($b = 2019$)

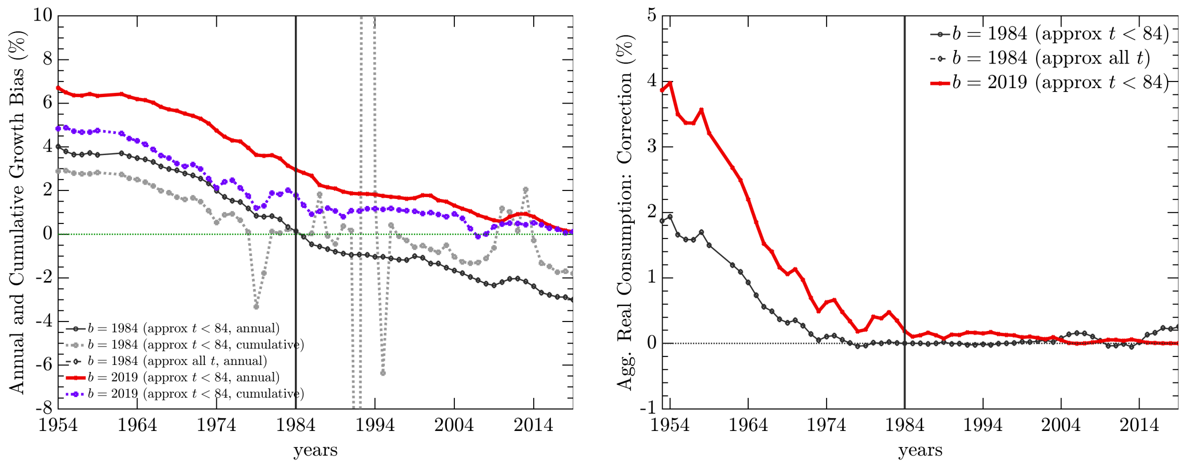


Note: The figure shows the evolution the average corrected and uncorrected measures of real consumption across quintiles relative to the measure of aggregate real consumption that ignores income heterogeneity. The latter defines the reduced-form index of real consumption using aggregate consumption expenditure shares. Panels (a) and (b) show the correction for the initial and final years as the base period, respectively. The reduced-form prices indices used for the 2nd and 1st order approximations are geometric and Tornqvist indices, respectively.

Figure A14: BLS Data: Correction Over Longer Horizon ($K = 1$)

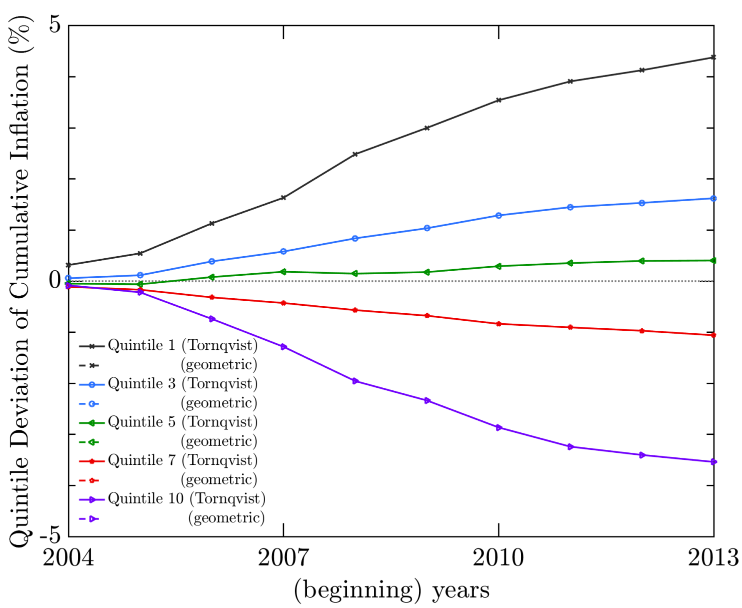
(a) Nonhomotheticity Growth Bias

(b) Real Consumption Correction



Note: Panels (a) and (b) show the evolution of the nonhomotheticity correction and the corrected relative to uncorrected index of real consumption using aggregate consumption expenditure shares.

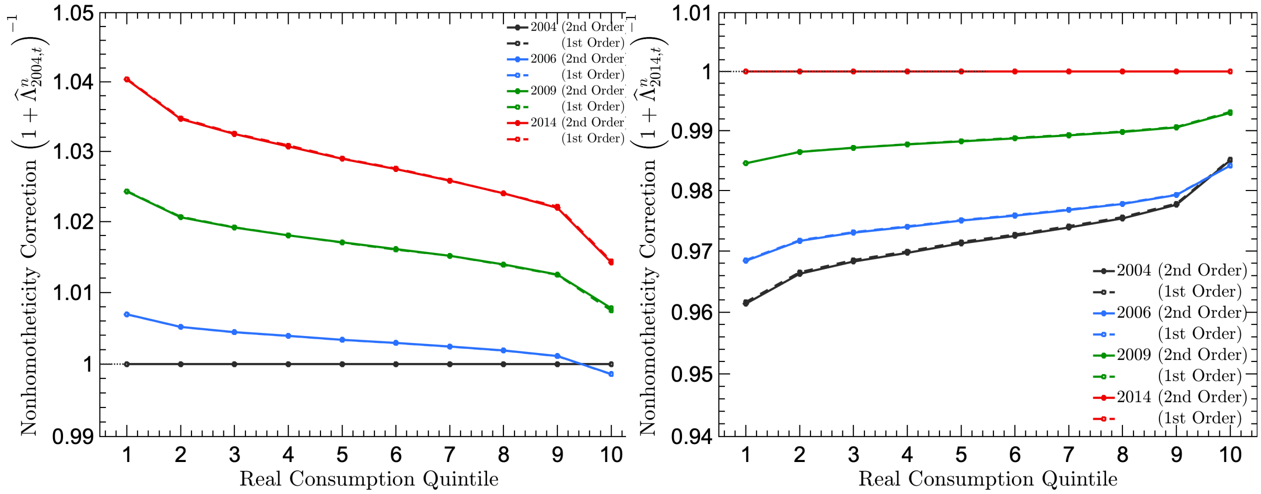
Figure A15: Nielsen Data: Price Indices



Note: The figure shows the evolution of cumulative reduced-form price indices for each quintile of income over the period. The great recession has been indicated in grey background.

Figure A16: Nielsen Data: Nonhomotheticity Correction ($K = 2$)

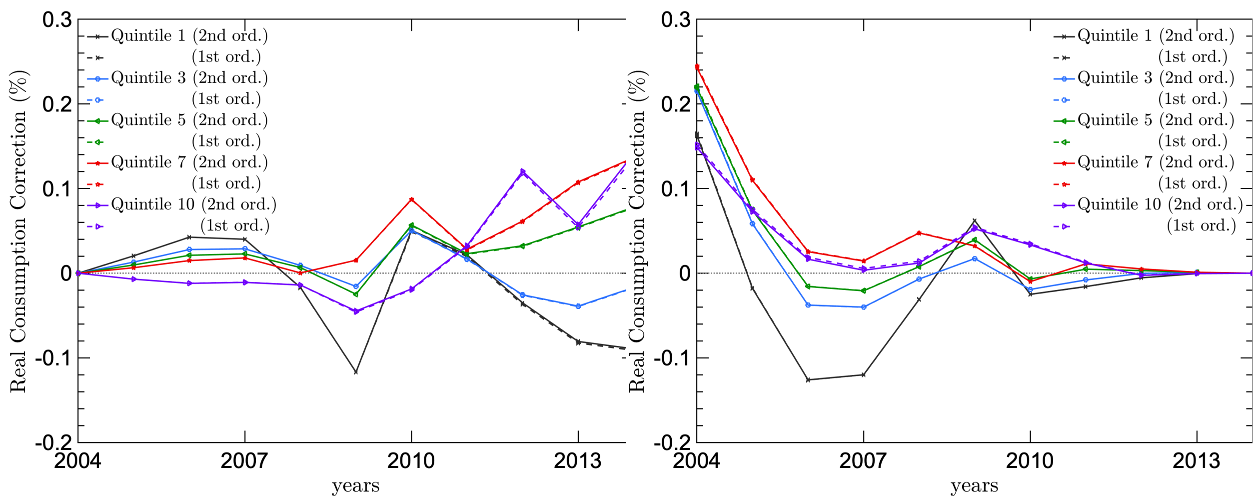
(a) Initial Period as Base ($b = 1984$) (b) Final Period as Base ($b = 2019$)



Note: Panels (a) and (b) show how the "Nonhomotheticity Correction" varies over time for different quintiles of income for the initial and final years as the base period.

Figure A17: Nielsen Data: Corrected relative to Uncorrected Real Consumption ($K = 2$)

(a) Initial Period as Base ($b = 1984$) (b) Final Period as Base ($b = 2019$)



Note: Panels (a) and (b) show how the ratio of the corrected to the uncorrected measures of real consumption vary over time for each quintile of initial real consumption for the initial and final years as the base period, respectively. The great recession has been indicated in grey background.