The Common-Scaling Social Cost-of-Living Index.

Thomas F. Crossley and Krishna Pendakur

May 27, 2009

Abstract

If preferences or budgets are heterogeneous across people (as they clearly are), then individual cost-of-living indices are also heterogeneous. Thus, any social cost-of-living index faces an aggregation problem. In this paper, we provide a solution to this problem which we call a 'common-scaling' social cost-of-living index (CS-SCOLI). In addition, we describe nonparametric methods for estimating such social cost-of-living indices. As an application, we consider changes in the social cost of living in the U.S. between 1988 and 2000.

JEL Classifications: D11, D12, D63, E31

Keywords: Inflation, Social cost-of-living, Demand, CPI, Average Derivatives

1 Introduction

"How has the cost-of-living changed?" is among the first questions that policy makers and the public ask of economists. One reason is that a vast amount of public expenditure is tied to measured changes in the cost-of-living. For example, many public pensions are indexed to measures of the overall or "social" cost-of-living. While economists have a well developed theory of the cost-of-living for a person, they do not have similarly well developed theory for the cost-of-living for a society. If preferences and budgets are identical across people, then the cost-of-living index is identical across people, and there is no problem in identifying the social cost-of-living index. However, if preferences or budgets are heterogeneous across people (as they clearly are), then different people experience different changes in the cost-of-living which must be somehow aggregated into a cost of living index for the society as a whole. In this paper we present a new class of social cost-of-living indices. These indices aggregate the cost-of-living indices of heterogeneous individuals. In addition, we describe nonparametric methods for estimating these and other social cost-of-living indices.

Most social cost-of-living indices in use—such as the Consumer Price Index (CPI)—can be understood as aggregator functions of approximations of household cost-of-living indices (see, e.g., Prais 1958 or Nicholson 1975 and, especially, Diewert's 1998 overview). The CPI is the expenditure-weighted average of first-order approximations of each household's cost-of-living index (COLI). It is troubling for at least three reasons. First, it has not been shown to have a welfare economic foundation for the case where agents are heterogeneous. Second, the CPI uses an expenditure-weighted average which downweights the experience of poor households relative to rich households (and thus is sometimes called a "plutocratic" index). Finally, it uses only first-order approximations of each individual's cost-of-living index, and thus ignores substitution effects.

Many researchers have used an alternative, called the "democratic index", equal to the arithmetic mean of household COLIs. In practice, a first order approximation to this index is implemented (using only first-order approximations of each individual's cost-of-living index.) Recent work includes Kokoski (2000), Crawford and Smith (2002) and Ley (2002, 2005). This alternative approach addresses our second concern, but not the other two. Pollak (1981) offers a social cost-of-living index which is explicitly grounded in a welfare economic problem. His solution is elegant but, as we shall elaborate below, it can be difficult to compute and interpret. In particular, the thought experiment corresponding to his index involves different income adjustments for different individuals. However, in practice, we use social cost-of-living indices to make equiproportionate adjustments to incomes, for example, in adjusting public pension levels.

We provide a simple, easy-to-interpret, easy-to-estimate solution to the problem of aggregating different changes in the cost-of-living into a social cost-of-living index. For an individual, the change in the cost-of-living is the answer to a question: what scaling of expenditure would hold utility constant over a price change? Our new social cost-of-living index is the answer to the following question. What single scaling to everyone's expenditure would hold social welfare constant over a price change? We call this the "common-scaling" social cost-of-living index (CS-SCOLI). It has the following advantages. First, the "common-scaling" aspect of our index is attractive both because it is analogous to the individual cost-of-living index and because it corresponds directly to the feasible policy uses of a social cost-of-living index (i.e., equiproportionate income adjustments). Second, our social cost-of-living index has social welfare foundations, and allows the investigator to easily choose the 'welfare weight' placed on rich and poor households. Third, nonparametric first-order approximations of the CS-SCOLI are easy-to-implement with commonly available commodity price and consumer expenditure data. Fourth, it is possible to compute nonparametric second-order approximations which capture substitution effects. Finally, the CS-SCOLI nests the plutocratic index and an object which is quite similar to the democratic index.

The rest of the paper proceeds as follows. In the next section we formally define the CS-SCOLI and compare it to other social cost-of-living indices. In Section Three we derive the first-order approximation to the CS-SCOLI, and compare its first-order approximation to that of other social-cost-of living indices. Section Four develops second order approximations to the CS-SCOLI, the plutocratic SCOLI and the democratic SCOLI. It also describes nonparametric methods for estimating them. Our method relies on nonparametric estimates of average derivatives, and is similar in spirit to that proposed by Deaton and Ng (1998). In the presence of unobserved preference heterogeneity, estimation of the second order terms in the approximations to plutocratic SCOLI and the democratic SCOLI faces a problem posed by Lewbel (2001). Another contribution of this paper is that we offer a solution to this problem. Section Five contains a small Monte Carlo study, in which we show that both welfare weights and second-order effects might matter. In Section 6, we present an empirical illustration which considers changes in the social cost-of-living in the U.S. between 1988 and 2000. We find that both the weighting of rich and poor households and the incorporation of second order effects have only modest impacts on our assessment of changes in the social cost-of-living. Section 7 concludes.

2 Theory

2.1 Individual Cost-of-Living Index

The standard theory of the cost-of-living for a person is as follows. Let $u = V(\mathbf{p}, x, \mathbf{z})$ be the indirect utility function that gives the utility level for an individual living in a household with a T-vector of demographic or other characteristics z, total expenditure x and facing the price vector $\mathbf{p} = [p^1, ..., p^M]$. Let $x = C(\mathbf{p}, u, \mathbf{z})$ be the cost function, which is the inverse of V over x. Let $i = 1, ..., N$ index individuals, each of whom lives in a household with one or more members. For each individual, the number n_i gives the number of members in that person's household. Each individual has an expenditure level x_i equal to total expenditure of that individual's household.

Many calculations are done at the household, rather than the individual, level. For household-level calculations, let $h = 1, ..., H$ index households, let x_h be the total expenditure, n_h be the number of members, and $\mathbf{z}_h$ be the characteristics of household h. Note that $x_h = x_i$ (because x_i is the total expenditure of household to which individual i belongs), so that, for example $\sum_{h=1}^{H} x_h = \sum_{i=1}^{N} x_i/n_i$. For simplicity, we consider environments where expenditure levels and characteristics vary across households, but not within households, and where price vectors are common across all individuals/households at a point in time. Thus, we assume that all members of a given household attain the same utility level, and consequently have the same cost-of-living index. Addressing intrahousehold variation in expenditure, and hence welfare, would be an interesting extension but is beyond the scope of the current paper.

We define the individual's cost-of-living index (COLI), $\pi(\mathbf{p}, \overline{\mathbf{p}}, x, \mathbf{z})$, as the scaling to expenditure x which equates utility between a reference price vector, $\overline{\mathbf{p}}$, and a different price vector, $\mathbf{p}$. Formally,

$$V(\overline{\mathbf{p}}, x, \mathbf{z}) = V(\mathbf{p}, \pi(\mathbf{p}, \overline{\mathbf{p}}, x, \mathbf{z})x, \mathbf{z}) \tag{1}$$

for π. This individual COLI function over the pair of price vectors $\overline{\mathbf{p}}$ and $\mathbf{p}$ is defined for any expenditure level, x, including the actual expenditure of the household when facing $\overline{\mathbf{p}}$ or $\mathbf{p}$.

Let $\overline{x}_i$ be the expenditure level of person i when facing the reference price vector (typically observed in the data). Let $\overline{u}_i$ be the utility level of person i when facing the reference price vector: $\overline{u}_i = V(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i)$. We denote the individual's COLI evaluated at reference expenditures, $\overline{x}_i$, for the price change from reference prices, $\overline{\mathbf{p}}$, to new prices, $\mathbf{p}$, as

$$\pi_i = \pi(\mathbf{p}, \overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i) = C(\mathbf{p}, \overline{u}_i, \mathbf{z}_i) / \overline{x_i} \tag{2}$$

For a household-level calculation, we note that $\pi_i = \pi_h$ for all i in household h. Although most previous work is motivated with household-level calculations, the welfarist framework that we employ below necessitates an individual-level analysis. Since all household members are identical, and thus have the same COLI, moving between these levels of analysis is straightforward, and amounts to reweighting.

2.2 The Common-Scaling Social Cost-of-Living Index

We propose a social cost-of-living index (SCOLI) that is similar in spirit to the individual COLI defined by (1). Let the direct social welfare function $s = W(u_1, ..., u_N)$ give the level of social welfare, s, corresponding to a vector of utilities, $u_1, ..., u_N$. We define the common-scaling social cost-of-living index (CS-SCOLI), Π^*, as the single scaling of *all* expenditures that equates social welfare at the two different price vectors. This definition requires reference expenditure vector which we take to be the actual expenditure vector when facing the reference price vector. Denoting $\overline{x}_1, .., \overline{x}_N$ as the reference expenditure vector, we solve

$$W\left(V(\overline{\mathbf{p}}, \overline{x}_1, \mathbf{z}_1), ..., V(\overline{\mathbf{p}}, \overline{x}_N, \mathbf{z}_N)\right) = W\left(V(\mathbf{p}, \Pi^* \overline{x}_1, \mathbf{z}_1), ..., V(\mathbf{p}, \Pi^* \overline{x}_N, \mathbf{z}_N)\right) \tag{3}$$

for $\Pi^* = \Pi^*\left(\mathbf{p}; \overline{\mathbf{p}}, \overline{x}_1, .., \overline{x}_N, \mathbf{z}_1, ..., \mathbf{z}_N\right)$. Just as a person's cost-of-living index is the scaling to her expenditure that holds her utility constant over a price change, the CS-SCOLI is the scaling to *everyone's* expenditure that holds social welfare constant over a price change. Clearly $\Pi^* = 1$ at reference prices. However, at other prices it can be a complex implicit functional of the direct welfare function and indirect utility functions. Below, we will show cases in which the CS-SCOLI has an explicit representation, and we show how to approximate it when it does not.

2.3 Previous Approaches to the Social Cost-of-Living

Since the COLI is different for individuals with different x and $\mathbf{z}$, a social cost-of-living index must somehow aggregate these individual COLIs. The most commonly used SCOLI is the so-called plutocratic SCOLI, Π^P, which may be defined as a weighted average of individual COLIs given by

$$\Pi^P = \frac{1}{\sum_{h=1}^{H} \overline{x}_h} \sum_{h=1}^{H} \overline{x}_h \pi_h = \frac{1}{\sum_{i=1}^{N} \frac{\overline{x}_i}{n_i}} \sum_{i=1}^{N} \frac{\overline{x}_i}{n_i} \pi_i. \tag{4}$$

This SCOLI assigns the reference household expenditure weight to each household-specific COLI, or, equivalently, assigns the reference household per-capita expenditure weight to each person-specific COLI.

An alternative is the democratic SCOLI, Π^D, which uses the unweighted average of household COLIs instead of the expenditure-weighted average as follows:

$$\Pi^D = \frac{1}{H}\sum_{h=1}^{H}\pi_h = \frac{1}{\sum_{i=1}^{N}\frac{1}{n_i}}\sum_{i=1}^{N}\frac{1}{n_i}\pi_i. \tag{5}$$

Here, individual COLIs are weighted by the reciprocal of the number of household members. Both the plutocratic and democratic SCOLIs are weighted averages of individual COLIs. The avoidance of expenditure weights in the weighted average is the great advantage of the democratic SCOLI (see, e.g., Ley 2005). We will show below in our discussion of first-order approximations that the CS-SCOLI is approximately a weighted average of approximate individual COLIs, where the weights are determined by the curvature of the social welfare function in expenditures.

Although the plutocratic and democratic SCOLIs are social aggregator functions, neither has a welfare-economic basis. Pollak (1981) offers a SCOLI which is explicitly grounded in a welfare economic problem. Define the indirect social cost function

$$\begin{aligned} M(\mathbf{p}, s, \mathbf{z}_1, ..., \mathbf{z}_N) &\equiv \min_{x_1,...,x_N}\sum_{h=1}^{H} x_h = \sum_{i=1}^{N} x_i/n_i \\ st \;\; W\left(V(\mathbf{p}, x_1, \mathbf{z}_1), ..., V(\mathbf{p}, x_N, \mathbf{z}_N)\right) &\geq s \\ x_i &= x_h \; \forall i \epsilon h \end{aligned}$$

as the minimum total (across households) expenditure required to attain the level of social welfare s for a population with characteristics $\mathbf{z}_1, ..., \mathbf{z}_N$ facing prices $\mathbf{p}$. Pollak's proposal for a SCOLI is

$$\Pi^M\left(\mathbf{p}, \overline{\mathbf{p}}, s, \mathbf{z}_1, ..., \mathbf{z}_N\right) = \frac{M(\mathbf{p}, s, \mathbf{z}_1, ..., \mathbf{z}_N)}{M(\overline{\mathbf{p}}, s, \mathbf{z}_1, ..., \mathbf{z}_N)}$$

where s equals initial social welfare, new social welfare, or some other social welfare level. Here, the numerator is equal to the minimum total expenditure across all households required to get a welfare level of s when facing prices $\mathbf{p}$, and the denominator is the minimum total expenditure when facing reference prices $\overline{\mathbf{p}}$.

Pollak's is a very elegant solution to the aggregation problem, and it has been implemented by Jorgenson and Slesnick (1983), Jorgenson (1990) and Slesnick (2001). Note that this procedure requires an optimization step in which the investigator determines the optimal distribution of expenditure (income) in each price regime. With heterogeneous preferences across households, this optimization can be hard. Moreover, if actual expenditure distributions are not optimal, comparisons of optimal expenditure distributions may not be very compelling. Finally, individual adjustments to expenditure may not correspond to feasible uses of a social cost-of-living index.

Let $\overline{s} = W\left(V(\overline{\mathbf{p}}, \overline{x}_1, \mathbf{z}_1), ..., V(\overline{\mathbf{p}}, \overline{x}_N, \mathbf{z}_N)\right)$ be the reference level of social welfare attained if households have reference expenditures $\overline{x}_1, ..., \overline{x}_N$, characteristics $\mathbf{z}_1, ..., \mathbf{z}_N$ and face the reference price vector $\overline{\mathbf{p}}$. If the program defining Pollak's index is evaluated at this level of welfare, then the CS-SCOLI can be understood as the solution to Pollak's program subject to the additional constraint that expenditures must be proportional to the reference expenditure vector $\overline{x}_1, .., \overline{x}_N$:

$$
\begin{aligned}
M(\mathbf{p}, s, \mathbf{z}_1, ..., \mathbf{z}_N) &\equiv \min_{x_1,...,x_N} \sum_{i=1}^{N} x_i/n_i \\
st \;\; W\left(V(\mathbf{p}, x_1, \mathbf{z}_1), ..., V(\mathbf{p}, x_N, \mathbf{z}_N)\right) &\geq \overline{s}, \\
x_i &= \lambda \overline{x}_i \;\forall i, \\
x_i &= x_h \;\forall i \epsilon h
\end{aligned}
$$

One can see by inspection that the CS-SCOLI is equal to the factor of proportionality: $\Pi^* = \lambda$. This restriction to Pollak's program was independently proposed by Fisher (2005), and we discuss his results below. The "common-scaling" restriction is attractive. It corresponds directly to the feasible policy uses of a social cost-of-living index (i.e., equiproportionate income adjustments). In addition, it constrains us to work with actual, rather than optimal, income distributions.

The relationship between the CS-SCOLI and Pollak's program is illustrated in Figure 1. The axes measure total expenditure of two individuals. The straight lines running diagonally down and to the right are social isocost lines. (If all households are of the same size, these have slope -1.) The point C indicates the initial expenditure distribution.

The curved dotted lines are level sets of the social welfare function (SWF). At the initial expenditure distribution ($\overline{X}$) indicated by point C and at the initial (reference) prices ($\overline{P}$) , social welfare $\overline{W}$ results. This level of social welfare can be achieved at minimum social cost at point A (at a point of tangency between a social isocost line and the level set of the SWF corresponding to .) Thus A represents the 'optimal' expenditure distribution.

A change in prices (from $\overline{P}$ to P) shifts the level set of the SWF in expenditure space. $\overline{W}$ is now achieved with minimum social cost at point B. The Pollak SCOLI is the ratio of the social cost of B to

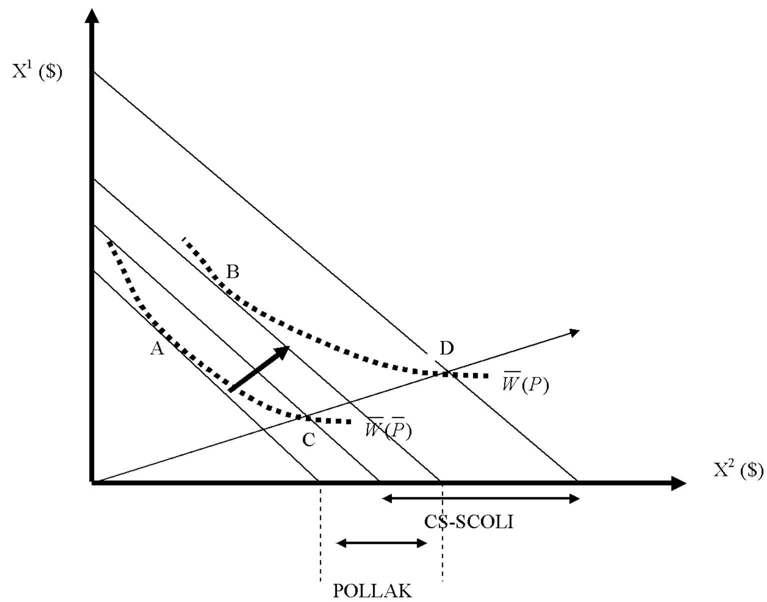

Figure 1: The CS-SCOLI and Pollak's SCOLI

the social cost of A.

The ray from the origin indicates equiproportionate scalings of the original expenditure distribution. The point D indicates where $\overline{W}$ can be achieved (given new prices P) with an equiproportionate scaling of budgets. (This is the point where the ray from the origin intersects the level set labelled $\overline{W}(P)$). The CS-SCOLI is the ratio of the social cost of D to the social cost of C.

The comparison of A and B invoked by Pollak's SCOLI is a comparison of 2 unobserved, optimal, distributions. The income adjustments relating them are individual-specific. In contrast, the comparison of C and D invoked by the CS-SCOLI is a comparison of an observed distribution, which may not be optimal, and a different distribution yielding the same welfare, which also may not be optimal. The income adjustments relating them, though, are common to all people. The equiproportionate nature of the CS-SCOLI makes it a reasonable candidate for policy purposes: most public policy applies the same cost-of-living adjustments to a large group of people.

2.4 Some Special Cases of the CS-SCOLI

The CS-SCOLI given by equation (3) is expressed as an implicit function. In some cases, the CS-SCOLI may be expressed explicitly. The proposition below shows the explicit representation of the CS-SCOLI for a few interesting cases. Our first case puts no restrictions on the welfare function, but requires identically homothetic utility. The second case allows for nonhomothetic utility via the PIGL structure (see Muellbauer 1976 for discussion) but is restricted to utilitarian welfare. The third case allows for non-utilitarian welfare via the S-Gini structure (see Donaldson and Weymark 1980) and invokes a somewhat restricted form of nonhomothetic PIGL utility. Throughout we assume that individuals with the same $\mathbf{z}$ have same preferences.

Proposition 1 *Let $h(t)$ be monotonically increasing in t. Let $a(\mathbf{p}, \mathbf{z})$ be homogeneous of degree 0 in $\mathbf{p}$, and $b(\mathbf{p}, \mathbf{z})$ and $c(\mathbf{p})$ be homogeneous of degree 1 in $\mathbf{p}$. Let $d(\mathbf{z})$ be an arbitrary function of $\mathbf{z}$. Consider the PIGL indirect utility function defined by $V(\mathbf{p}, x, \mathbf{z}) = a(\mathbf{p}, \mathbf{z}) + \left(\frac{x}{b(\mathbf{p},\mathbf{z})}\right)^{\theta}/\theta$ for $\theta \neq 0$ and $V(\mathbf{p}, x, \mathbf{z}) = a(\mathbf{p}, \mathbf{z})\left[\ln x - \ln b(\mathbf{p}, \mathbf{z})\right]$ for $\theta = 0$, with associated cost function $C(\mathbf{p}, u, \mathbf{z}) = b(\mathbf{p}, \mathbf{z})\theta^{1/\theta}\left(u - a(\mathbf{p}, \mathbf{z})\right)^{1/\theta}$ for $\theta \neq 0$ and $\ln C(\mathbf{p}, u, \mathbf{z}) = \frac{u}{a(\mathbf{p},\mathbf{z})} + \ln b(\mathbf{p}, \mathbf{z})$ for $\theta = 0$.*

1. *If indirect utility is identically homothetic with $V(\mathbf{p}, x, \mathbf{z}) = h\left(\frac{x}{b(\mathbf{p})d(\mathbf{z})}\right)$ and the direct welfare function depends only on the utility vector (i.e.„ is 'welfarist'), then the CS-SCOLI is given by*

$$\Pi^*(\mathbf{p};\cdot) = \frac{b(\mathbf{p})}{b(\overline{\mathbf{p}})}.$$

2. *If indirect utility is PIGL and the direct welfare function is utilitarian, where $W(\cdot) = \sum_{i=1}^{N} u_i$, then the CS-SCOLI is given by*

$$\begin{aligned} \Pi^*(\mathbf{p};\cdot) &= \left[\frac{\sum_{i=1}^{N}\left(a(\overline{\mathbf{p}}, \mathbf{z}_i) - a(\mathbf{p}, \mathbf{z}_i) + \left(\frac{x_i}{b(\overline{\mathbf{p}},\mathbf{z}_i)}\right)^{\theta}/\theta\right)}{\sum_{i=1}^{N}\left(\frac{x_i}{b(\mathbf{p},\mathbf{z}_i)}\right)^{\theta}/\theta}\right]^{1/\theta}, \text{ for } \theta \neq 0 \\ &= \exp\left(\frac{\sum_{i=1}^{N}\left[\left(a(\overline{\mathbf{p}}, \mathbf{z}_i) - a(\mathbf{p}, \mathbf{z}_i)\right)\ln x_i + a(\mathbf{p}, \mathbf{z}_i)\ln b(\mathbf{p}, \mathbf{z}_i) + a(\overline{\mathbf{p}}, \mathbf{z}_i)\ln b(\overline{\mathbf{p}}, \mathbf{z}_i)\right]}{\sum_{j=1}^{N} a(\mathbf{p}, \mathbf{z}_j)}\right), \text{ for } \theta = 0 \end{aligned}$$

3. *If indirect utility is PIGL with a independent of z ($a(\mathbf{p}, \mathbf{z}) = a(\mathbf{p})$) and b multiplicatively separable with $b(\mathbf{p}, \mathbf{z}) = c(\mathbf{p})d(\mathbf{z})$, and the direct welfare function is a weighted sum of utilities, where $W(\cdot) = \sum_{i=1}^{N} g_i u_i$, with S-Gini weights g_i which sum to 1 and depend only on person i's position*

in the distribution of utilities, then the CS-SCOLI is given by

$$\Pi^*(\mathbf{p};\cdot) = \left[c(\mathbf{p})^\theta \frac{a(\overline{\mathbf{p}}) - a(\mathbf{p})}{\sum_{i=1}^N g_i \left(\frac{x_i}{d(\mathbf{z}_i)}\right)^\theta / \theta} + \left(\frac{c(\mathbf{p})}{c(\overline{\mathbf{p}})}\right)^\theta \right]^{1/\theta}, \textit{ for } \theta \neq 0$$

$$= \exp\left[\frac{[a(\overline{\mathbf{p}}) - a(\mathbf{p})]}{a(\mathbf{p})} \sum_{i=1}^N g_i \left[\ln x_i - \ln d(\mathbf{z}_i)\right] + \ln c(\mathbf{p}) - \ln c(\overline{\mathbf{p}}) \frac{a(\overline{\mathbf{p}})}{a(\mathbf{p})} \right], \textit{ for } \theta = 0$$

Proof. Case 1: Solving (1) with $V(\mathbf{p}, x, \mathbf{z}) = h\left(\frac{x}{b(\mathbf{p})d(\mathbf{z})}\right)$ yields $\pi_i = \frac{b(\mathbf{p})}{b(\overline{\mathbf{p}})}$ for all $i = 1, .., N$. Thus, setting $\Pi^* = \frac{b(\mathbf{p})}{b(\overline{\mathbf{p}})}$, we have $V(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i) = V(\mathbf{p}, \Pi^*\overline{x}_i, \mathbf{z}_i)$ for all $i = 1, ..., N$. Since welfare depends only on the vector of utilities, $\Pi^* = \frac{b(\mathbf{p})}{b(\overline{\mathbf{p}})}$ solves (3). Since W is monotonic in u_i and h is monotonic in x_i, this is a unique solution. Cases 2 and 3: Substitute $W(\cdot)$ and $V(\mathbf{p}, x, \mathbf{z})$ into (3) and re-arrange to get explicit formulae for Π^*. Note that in Case 3, the restrictions on a and b ensure that, given the distribution of expenditures, the distribution of utilities—and therefore the set of weights g_i—does not depend on prices. ■

In cases 2 and 3, the direct welfare function is (piecewise) linear and indirect utility is linear in a function of expenditure at each price vector. This ensures that both sides of the implicit definition of the CS-SCOLI given by equation (3) are linear in a function of Π^*, so that solving for Π^* is straightforward. The linearity of the solutions (combined with the concavity of individual cost in prices) also implies that the CS-SCOLI is concave in prices for these cases.

In case 1, the direct welfare function may be nonlinear, but since preferences are homothetic, this nonlinearity drops out of the expression for Π^*. That is, since this homothetic indirect utility function implies individual COLIs are independent of expenditure and demographics, the CS-SCOLI is equal to this (common) individual COLI regardless of the inequality-aversion of the social welfare function.

We note that the PIGL functional form for indirect utility exploited in the proposition is also used by Muellbauer (1974) in his exploration of the political economy of price indices. However, his focus was on how a common price index might be redistributive, whereas ours is on the welfare basis of a potential common price index.

Two examples of Case 2 of Proposition 1 are of specific interest. Given PIGL utility, one may substitute the individual COLI, π_i, into the expression for the CS-SCOLI. If $\theta = 1$, then the marginal utility of money is constant and preferences are quasihomothetic so that commodity demands are linear in expenditure. If b is multiplicatively separable with $b(\mathbf{p}, \mathbf{z}) = c(\mathbf{p})d(\mathbf{z})$ and $d(\mathbf{z}) = n_i$, then costs are proportional to the number of household members and independent of all other demographic characteristics. With these restrictions, some (tedious) algebra reveals that the CS-SCOLI is given by the plutocratic SCOLI:

$$\Pi^*\left(\mathbf{p};\overline{\mathbf{p}},\overline{x}_1,..,\overline{x}_N,\mathbf{z}_1,...,\mathbf{z}_N\right)=\frac{\sum_{i=1}^{N}\frac{\overline{x}_i\pi_i}{n_i}}{\sum_{i=1}^{N}\frac{\overline{x}_i}{n_i}}.$$

Thus, with quasihomothetic utility and costs depending linearly on household size, the CS-SCOLI coincides with the plutocratic SCOLI.

If $\theta=0$, then the marginal utility of log-money is a constant and preferences are PIGLOG with commodity budget shares that are linear in the log of expenditure. If b is is multiplicatively separable with $b(\mathbf{p},\mathbf{z})=c(\mathbf{p})d(\mathbf{z})$, then the CS-SCOLI is given by the *unitary SCOLI* defined as the geometric mean of individual COLIs:

$$\Pi^*\left(\mathbf{p};\overline{\mathbf{p}},\overline{x}_1,..,\overline{x}_N,\mathbf{z}_1,...,\mathbf{z}_N\right)=\exp\left(\frac{1}{N}\sum_{i=1}^{N}\ln\pi_i\right).$$

The unitary social cost-of-living index gives the same weight to everyone's individual cost-of-living-index, so it is similar in spirit to the democratic SCOLI. However, the unitary SCOLI is a geometric rather than an arithmetic mean.

3 First-Order Approximation

3.1 First-Order Approximation of the CS-SCOLI

Consider the approximation of the cost-of-living index for an individual household. Let $\mathbf{w}$ be the $M-$vector of budget-shares, with a subscript for household or individual. Define the compensated (Hicksian) budget-share vector-function, $\boldsymbol{\omega}(\mathbf{p},u,\mathbf{z})$, to give the budget-share vector for a person with utility u and characteristics $\mathbf{z}$ facing prices $\mathbf{p}$. Note that, by Shephard's Lemma, $\mathbf{w}=\boldsymbol{\omega}(\mathbf{p},u,\mathbf{z})=\boldsymbol{\nabla}_{\ln\mathbf{p}}\ln C(\mathbf{p},u,\mathbf{z})$. We typically invoke Shepphard's Lemma at the reference price vector and utility value. Define the uncompensated (Marshallian) budget-share vector-functions, $\mathbf{w}(\mathbf{p},x,\mathbf{z})$, to give the budget-share vector for a person with household expenditure x and characteristics $\mathbf{z}$ facing prices $\mathbf{p}$.

We assume knowledge of the reference, or initial, price vector, $\overline{\mathbf{p}}$, and of micro-data on budget-shares, expenditures and demographics, $\{\overline{\mathbf{w}}_i,\overline{x}_i,\mathbf{z}_i\}_{i=1}^{N}$, for individuals facing this initial price vector. We assume knowledge only of price data, $\mathbf{p}$, in the non-reference, or final, period. We note that $\mathbf{z}_i$

does not have an overbar because it is assumed to be invariant for an individual. In this paper, the reference price vector is taken to be the initial price vector, so that our approach is similar in spirit to the Laspeyres index for an individual COLI. It is possible to reformulate our concept with the reference price vector given by the final price vector, which is similar in spirit to the Paasche formulation. Since our approximation strategies require information on budget-shares in the reference period, we feel that the Laspeyres approach is more useful. Donaldson and Pendakur (2009) consider the question of when these two indices are equivalent.

The first-order approximation of π_i around $\overline{\mathbf{p}}$ for a person with expenditure $\overline{x}_i$ and demographics $\mathbf{z}_i$ is the well-known Laspeyres index for the individual, π_i^L,

$$\pi_i \approx \pi_i^L \equiv 1 + d\mathbf{p}'\overline{\mathbf{w}}_i \tag{6}$$

where $d\mathbf{p} \equiv \left[\frac{p^1 - \overline{p}^1}{\overline{p}^1}, ..., \frac{p^M - \overline{p}^M}{\overline{p}^M}\right]$ is the $M-$vector of proportionate price changes. This type of approximation requires micro-level information only from the reference period, so it is easily implemented with real-world data. For example, Crawford and Smith (2002) evaluate this approximation to the household-level COLI for a sample of UK households between 1976 and 2000 and find great heterogeneity in (approximate) COLIs across households.

A first-order approximation to the plutocratic SCOLI is commonly computed by statistical agencies for several reasons. For example, the linearity of the approximation makes it decomposable by groups. In addition, it may be interpreted in terms of an 'aggregate consumer' whose behavior is described by that of the economy as a whole. From our point of view, its key feature is that it is computable using only aggregate data. In particular, because Π^P is linear in π_i, we may substitute (6) into (4) to obtain a first-order approximation of Π^P:

$$\Pi^P \approx \frac{1}{\sum_{i=1}^N \frac{\overline{x}_i}{n_i}} \sum_{i=1}^N \frac{\overline{x}_i}{n_i} \pi_i^L = 1 + d\mathbf{p}'\widetilde{\mathbf{w}}^P, \tag{7}$$

where $\widetilde{\mathbf{w}}^P \equiv \frac{1}{\sum_{i=1}^N \frac{\overline{x}_i}{n_i}} \sum_{i=1}^N \frac{\overline{x}_i}{n_i} \overline{\mathbf{w}}_i = \frac{1}{\sum_{h=1}^H \overline{x}_h} \sum_{h=1}^H \overline{x}_h \overline{\mathbf{w}}_i$ is the aggregate reference budget share vector for the population. This methodology is used by the Bureau of Labor Statistics to compute the CPI.

Crawford and Smith (2002) estimate the first-order approximation of Π^D,

$$\Pi^D \approx \frac{1}{\sum_{i=1}^N \frac{1}{n_i}} \sum_{i=1}^N \frac{1}{n_i} \pi_i^L = 1 + d\mathbf{p}'\widetilde{\mathbf{w}}^D, \tag{8}$$

where $\widetilde{\mathbf{w}}^D \equiv \frac{1}{\sum_{i=1}^N \frac{1}{n_i}} \sum_{i=1}^N \frac{1}{n_i} \overline{\mathbf{w}}_i = \frac{1}{H} \sum_{h=1}^H \overline{\mathbf{w}}_h$ is the household-level average budget share vector (rather than the aggregate budget share vector). They find in their study of UK data that the approximate democratic SCOLI shows slightly less inflation than does the approximate plutocratic SCOLI.

Defining the weights

$$\phi_i^P = \frac{\frac{\overline{x}_i}{n_i}}{\sum_{i=1}^N \frac{\overline{x}_i}{n_i}}, \tag{9}$$

$$\phi_i^D = \frac{\frac{1}{n_i}}{\sum_{i=1}^N \frac{1}{n_i}}, \tag{10}$$

we may write the first-order approximations of the plutocratic and democratic indices as

$$\Pi^m \approx \sum_{i=1}^N \phi_i^m \pi_i^L$$

for $m = P, D$. It turns out that the first-order approximation of the CS-SCOLI also takes the form of a weighted average of individual Laspeyres indices.

An approximation of the CS-SCOLI, Π^*, around $\overline{\mathbf{p}}$ may be obtained via the implicit function theorem. Let $(\cdot)$ denote the reference values of function arguments, and let overbars denote the values of functions evaluated at their reference arguments. Thus, we denote reference utility levels and functions as $\overline{u}_i = V(\cdot) \equiv V(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i)$, and the reference welfare level and function as $\overline{s} = W(\cdot) \equiv W(\overline{u}_1, ..., \overline{u}_N) = W\left(V(\overline{\mathbf{p}}, \overline{x}_1, \mathbf{z}_1), ..., V(\overline{\mathbf{p}}, \overline{x}_N, \mathbf{z}_N)\right)$. The equation defining the CS-SCOLI may be rewritten as

$$\overline{s} = W\left(V(\mathbf{p}, \Pi^*(\mathbf{p})\, \overline{x}_1, \mathbf{z}_1), ..., V(\mathbf{p}, \Pi^*(\mathbf{p})\, \overline{x}_N, \mathbf{z}_N)\right).$$

where $\Pi^*(\mathbf{p}) = \Pi^*(\mathbf{p}; \overline{\mathbf{p}}, \overline{x}_1, .., \overline{x}_N, \mathbf{z}_1, ..., \mathbf{z}_N)$ suppresses the dependence of the CS-SCOLI on the reference price and expenditure vectors and demographic characteristics vectors. Here, we emphasize that $\Pi^*(\mathbf{p})$ depends on $\mathbf{p}$ as a variable. Application of the implicit function theorem yields

$$\nabla_{\mathbf{p}} \Pi^*(\mathbf{p}) = -\frac{\sum_{i=1}^N \nabla_{u_i} W(\cdot) \nabla_{\mathbf{p}} V(\overline{\mathbf{p}}, \Pi^*(\mathbf{p})\, \overline{x}_i, \mathbf{z}_i)}{\sum_{i=1}^N \nabla_{u_i} W(\cdot) \nabla_{x_i} V(\overline{\mathbf{p}}, \Pi^*(\mathbf{p})\, \overline{x}_i, \mathbf{z}_i) \overline{x}_i}. \tag{11}$$

Define the normalized proportionate welfare weight for person i's household expenditure as

$$\phi_i(\mathbf{p}) = \phi_i(\mathbf{p}; \overline{\mathbf{p}}, \overline{x}_1, .., \overline{x}_N, \mathbf{z}_1, ..., \mathbf{z}_N) \equiv \frac{\nabla_{u_i} W(\cdot) \nabla_{x_i} V(\mathbf{p}, \Pi^*(\mathbf{p})\, \overline{x}_i, \mathbf{z}_i) \overline{x}_i}{\sum_{i=1}^N \nabla_{u_i} W(\cdot) \nabla_{x_i} V(\mathbf{p}, \Pi^*(\mathbf{p})\, \overline{x}_i, \mathbf{z}_i) \overline{x}_i}, \tag{12}$$

and let

$$\overline{\phi}_i = \phi_i(\overline{\mathbf{p}}) = \frac{\nabla_{u_i} W(\cdot) \nabla_{x_i} V(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i) \overline{x}_i}{\sum_{i=1}^N \nabla_{u_i} W(\cdot) \nabla_{x_i} V(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i) \overline{x}_i} \tag{13}$$

$$= \frac{\nabla_{u_i} W(\cdot) \nabla_{\ln x_i} V(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i)}{\sum_{i=1}^N \nabla_{u_i} W(\cdot) \nabla_{\ln x_i} V(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i)} \tag{14}$$

be the reference value of the normalized proportionate welfare weight for person i's household expenditure. Since $\overline{\phi}_i = \phi_i(\overline{\mathbf{p}})$ is evaluated at reference prices, and since $\Pi^*(\overline{\mathbf{p}}) = 1$ at reference prices, $\Pi^*(\mathbf{p})$ drops out of the expression for $\overline{\phi}_i$. Again, we suppress the dependence of ϕ_i on $(\overline{\mathbf{p}}, \overline{x}_1, .., \overline{x}_N, \mathbf{z}_1, ..., \mathbf{z}_N)$ because $\phi_i(\mathbf{p})$ only depends on $\mathbf{p}$ as a variable. A welfare weight is usually defined as the response of social welfare to individual expenditure, $\nabla_{u_i} W(u_i, ..., u_N) \nabla_{x_i} V(\mathbf{p}, x_i, \mathbf{z}_i)$. For convenience, we use a *proportionate* welfare weight, which multiplies this quantity by expenditure, and gives the response of welfare to a proportionate change in person i's household expenditure. It is *normalized* to sum to 1. Finally, it is a *reference* weight because it evaluated at the reference price and expenditure vectors. The following proposition shows the first-order approximation of Π^*.

Proposition 2 *Given the normalized reference proportionate welfare weights $\overline{\phi}_i$ given by equations (12) and (13) and reference budget shares $\overline{\mathbf{w}}_i$, a first-order approximation of the CS-SCOLI Π^* is*

$$\Pi^* \approx \sum_{i=1}^N \overline{\phi}_i \pi_i^L, \tag{15}$$

or, equivalently,

$$\Pi^* \approx 1 + d\mathbf{p}'\widetilde{\mathbf{w}} \tag{16}$$

where $d\mathbf{p} \equiv \left[\frac{p^1 - \overline{p}^1}{\overline{p}^1}, ..., \frac{p^M - \overline{p}^M}{\overline{p}^M}\right]$ is the M-vector of proportionate price changes and $\widetilde{\mathbf{w}} \equiv \sum_{i=1}^N \overline{\phi}_i \overline{\mathbf{w}}_i$ is the weighted-average reference budget-share vector.

Proof. *Let $\Pi^*(\cdot)$ denote $\Pi^*(\mathbf{p}; \overline{\mathbf{p}}, \overline{x}_1, .., \overline{x}_N, \mathbf{z}_1, ..., \mathbf{z}_N)$. Using the first-derivative of $\Pi^*(\cdot)$ given by (11) and the fact that $\Pi^*(\overline{\mathbf{p}}; \overline{\mathbf{p}}, \overline{x}_1, .., \overline{x}_N, \mathbf{z}_1, ..., \mathbf{z}_N) = 1$, a first-order approximation of $\Pi^*(\mathbf{p}; \overline{\mathbf{p}}, \overline{x}_1, .., \overline{x}_N, \mathbf{z}_1, ..., \mathbf{z}_N)$ is given by*

$$\begin{aligned}
\Pi^*(\mathbf{p}; \overline{\mathbf{p}}, \overline{x}_1, .., \overline{x}_N, \mathbf{z}_1, ..., \mathbf{z}_N) &= 1 + (\mathbf{p} - \overline{\mathbf{p}})' \nabla_{\mathbf{p}} \Pi^*(\cdot) \\
&= 1 - (\mathbf{p} - \overline{\mathbf{p}})' \frac{\sum_{i=1}^N \nabla_{u_i} W(\cdot) \nabla_{\mathbf{p}} V(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i)}{\sum_{i=1}^N \nabla_{u_i} W(\cdot) \nabla_{x_i} V(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i) \overline{x}_i}.
\end{aligned}$$

Substituting in $d\mathbf{p}$ *and the logarithmic form of Roy's Identity evaluated at reference prices and expenditures,* $\overline{\mathbf{w}}_i = -\nabla_{\ln \mathbf{p}} V(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i)/\nabla_{\ln x} V(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i)$*, gives the approximation in terms of reference budget shares* $\overline{\mathbf{w}}_i$*:*

$$\Pi^* \approx 1 + d\mathbf{p}' \left(\frac{\sum_{i=1}^{N} \nabla_{u_i} W(\cdot) \nabla_{x_i} V(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i) \overline{x}_i \overline{\mathbf{w}}_i}{\sum_{i=1}^{N} \nabla_{u_i} W(\cdot) \nabla_{x_i} V(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i) \overline{x}_i} \right).$$

Substituting in the welfare weights (12) yields (16). ■

3.2 Relationship to first-order approximations to other SCOLIs

In Section 2.4 we gave conditions under which the CS-SCOLI and the plutocratic SCOLI coincide. Of course, when the exact indices coincide, their first order approximations coincide as well. The first order approximation of the CS-SCOLI, (15), coincides with that of the democratic SCOLI if and only if

$$\overline{\phi}_i = \frac{1/n_i}{\sum_{i=1}^{N} 1/n_i} \tag{17}$$

for all $i = 1, ..., N$. This would obtain if welfare were utilitarian and indirect utility were PIGLOG with $V(\mathbf{p}, x, \mathbf{z}) = a(\mathbf{p}) + \frac{b(\mathbf{p})}{n_i} \ln x_i$. This indirect utility function corresponds to the widely-used Almost Ideal demand system combined with constant marginal utility of log-money. Almost Ideal utility implies budget-shares that are linear in the log of expenditure, a restriction which is not typically satisfied in real-world data (see, e.g., Banks, Blundell and Lewbel 1997). Here, the marginal social value of utility is a constant, and the marginal utility of money for an individual is declining. (We note that this incorporation of household size is hard to interpret: the more common strategy in applied consumer demand or inequality measurement would be to deflate x_i by a function of a weakly concave function of n_i, rather than to deflate $\ln x_i$ by n_i.)

Computationally, the expression in terms of weighted average budget shares is convenient. Of course, the welfare-weights, $\overline{\phi}_i$, are themselves functions of the price, expenditures and demographics (evaluated at reference prices and expenditures, $\overline{\mathbf{p}}, \overline{x}_1, ..., \overline{x}_N$), so implementation requires this information and requires the full knowledge of the welfare weight function. Further, since welfare weights must be calculated by the researcher, they must not depend on any unobserved factors. Such knowledge is routinely assumed in empirical investigations of inequality and poverty, and is required for an empirical investigation of the CS-SCOLI as well.

First-order approximations are exact for infinitesimal price changes. Fisher (2005) shows the CS-

SCOLI is equivalent to Pollack's SCOLI (Π^M defined above) for an infinitesimal price change if the initial distribution is optimal. This is because, due to an envelope result applied to the social optimization, the already optimal income distribution need not be re-adjusted in the face of a tiny price change. Fisher, Pollack and Diewert all note that if the initial distribution is optimal, then the first-order approximations of Pollack's SCOLI and the plutocratic SCOLI coincide. This is due to an envelope result applied to the individual cost minimizations.

The assumption that the initial distribution is socially optimal also allows one to bound Pollak's SCOLI (Π^M)for large price changes. Pollak (1980) defines the 'Scitovsky-Laspeyres' SCOLI as the proportionate increase total expenditure required to keep all households on their reference indifference curve, and shows that if the reference distribution of expenditure is socially optimal, then the Scitovsky-Laspeyres SCOLI is an upper bound on Π^M. A first-order approximation to the 'Scitovsky-Laspeyres' SCOLI is the aggregate Laspeyres index (7). Diewert (2001) considers the analogous 'Scitovsky-Paasche' SCOLI defined as the proportionate increase total expenditure required to keep all households on their final indifference curve, and shows that if the final distribution is socially optimal, then the Scitovsky-Paasche SCOLI is a lower bound on Π^M. A first-order approximation of this SCOLI is given by the aggregate Paasche index.

First-order approximations of bounds to Pollak's SCOLI are thus easily implemented, and while these bounds are in principle very useful, they apply only to the case where the initial (or final) distribution of household expenditure is socially optimal. In contrast, the first-order approximation to the CS-SCOLI given in Proposition 2 approximates the index itself, rather than bounds for it, and applies to expenditure distributions that may not be socially optimal.

As noted above, if the CS-SCOLI is given by one of the explicit representations given in Proposition 1, it is concave in prices. Consequently, the first-order approximation to the CS-SCOLI provides an upper bound to its value. This bounding result contrasts with the bounding results of Pollak and Diewert concerning Pollak's SCOLI. Their bounds apply only when the initial (or final) distribution of expenditure is socially optimal. Our bound applies regardless of the optimality of the expenditure distribution if welfare and indirect utility satisfy the requirements of one of the cases.

4 Second-Order Approximation

In this section we derive second-order approximations to the plutocratic and democratic SCOLIs and to the CS-SCOLI. We also discuss estimation of these approximations. The second-order terms in the approximations to the plutocratic and democratic SCOLIs involve weighted averages of compensated price effects (on demand). Compensated price effects are defined by the Slutsky equation, and involve products of levels and expenditure derivatives of demand equations. Lewbel (2001) shows that such compensated price effects may be difficult to estimate in the presence of unobserved preference heterogeneity. Below, we provide a solution to this problem, so that second-order approximations of the Plutocratic and Democratic SCOLIs may be estimated nonparametrically.

The second-order term in an approximation to the CS-SCOLI involves average *un*compensated price effects, which do not involve products of levels and derivatives of demand equations. These average uncompensated price effects may be estimated directly from data on budget shares, prices, total expenditure and demographics using standard nonparametric methods (following, for example, Deaton and Ng, 1998). However, the second-order approximation to the CS-SCOLI brings a different problem. As we shall show below, it involves a term which is the response of the welfare weight functions to price changes. This cannot be estimated from demand data. We provide a proposition showing conditions under which this trailing term is exactly zero, and in the following section, we present a Monte Carlo study, one goal of which is to assess the likely size of this trailing term under a range of different assumptions. We begin with the plutocratic and democratic SCOLIs.

4.1 Plutocratic and Democratic SCOLIs

Define $\overline{\mathbf{W}}_i \equiv \overline{\mathbf{w}}_i\overline{\mathbf{w}}_i'$, and $diag(\overline{\mathbf{w}}_i)$ as a diagonal matrix with $\overline{\mathbf{w}}_i$ on the main diagonal. Define the matrix of compensated budget-share semi-elasticities as $\mathbf{\Gamma}(\mathbf{p}, u, \mathbf{z}) \equiv \nabla_{\ln \mathbf{p}'}\boldsymbol{\omega}(\mathbf{p}, u, \mathbf{z}) = \nabla_{\ln \mathbf{p}'}\mathbf{w}(\mathbf{p}, x, \mathbf{z}) + \nabla_{\ln x}\mathbf{w}(\mathbf{p}, x, \mathbf{z})\mathbf{w}(\mathbf{p}, x, \mathbf{z})'$ by the Slutsky Theorem (that is, the chain rule). The second-order approximation of an individual COLI evaluated at reference prices and expenditure is given by

$$\pi_i \approx \pi_i^L + \pi_i^S \tag{18}$$

where

$$\begin{aligned}
\pi_i^S &= \frac{1}{2}\mathbf{dp}' \left[\overline{\mathbf{W}}_i - diag\left(\overline{\mathbf{w}}_i\right) + \nabla_{\ln \mathbf{p}'}\mathbf{w}(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i) + \nabla_{\ln x}\mathbf{w}(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i)\mathbf{w}(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i)'\right] \mathbf{dp} \\
&= \frac{1}{2}\mathbf{dp}' \left[\overline{\mathbf{W}}_i - diag\left(\overline{\mathbf{w}}_i\right) + \overline{\mathbf{\Gamma}}_i\right] \mathbf{dp},
\end{aligned}$$

where

$$\overline{\boldsymbol{\Gamma}}_i \equiv \nabla_{\ln \mathbf{p}'} \boldsymbol{\omega}(\overline{\mathbf{p}}, \overline{u}_i, \mathbf{z}_i) = \nabla_{\ln \mathbf{p}'} \mathbf{w}(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i) + \nabla_{\ln x} \mathbf{w}(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i) \mathbf{w}(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i)'$$

is the matrix of compensated budget-share price semi-elasticities evaluated at reference prices and utility (expenditure). The second-order term π_i^S is a quadratic form in the Slutsky matrix of substitution terms (expressed in budget-share form) and captures substitution effects in demand and savings from substitution in the cost-of-living. von Haefen (2000) shows how the budget-share form is related to the more familiar quantity form of the Slutsky matrix.

Before considering the CS-SCOLI, which is defined implicitly, it is instructive to consider the Plutocratic and Democratic indices. Since both the plutocratic and democratic indices are weighted averages of individual COLIs (given by (4) and (5), respectively), the second-order approximations of these indices are the corresponding weighted averages of (18):

$$\Pi^m \approx \sum_{i=1}^{N} \phi_i^m \left(\pi_i^L + \pi_i^S \right),$$

for $m = P, D$ and where ϕ_i^P and ϕ_i^D are given by equations (9) and (10), respectively.

Estimates of these second-order approximations use weights ϕ_i^m (which depend on reference expenditures, x_i and household sizes, n_i) plus budget shares, $\overline{\mathbf{w}}_i$, and compensated semi-elasticities, $\overline{\boldsymbol{\Gamma}}_i$. All but the last are observable, and, in the absence of unobserved preference heterogeneity, the compensated semi-elasticities may be estimated nonparametrically using data on characteristics, budget shares, prices and expenditures. Of course, consumption panel data would be appropriate, but repeated cross-section data will also suffice. The nonparametric estimator of the compensated semi-elasticity matrix for an individual facing $(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i)$ is

$$\widehat{\overline{\boldsymbol{\Gamma}}}_i = \widehat{\nabla_{\ln \mathbf{p}'}} \mathbf{w}(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i) + \widehat{\nabla_{\ln x}} \mathbf{w}(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i) \overline{\mathbf{w}}_{\mathbf{i}},' \tag{19}$$

where $\widehat{\nabla_{\ln \mathbf{p}'}} \mathbf{w}(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i)$ and $\widehat{\nabla_{\ln x}} \mathbf{w}(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i)$ are nonparametric estimates of $\nabla_{\ln \mathbf{p}'} \mathbf{w}(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i)$ and $\nabla_{\ln x} \mathbf{w}(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i)$, respectively (see Lewbel 2001). Since the convergence rate of this object is dominated by the derivatives, the estimators for Π^P and Π^D behave like (weighted) average derivative estimators.

In the presence of unobserved preference heterogeneity, Lewbel (2001) shows that the above estimator for $\overline{\boldsymbol{\Gamma}}_i$ may be inconsistent. The basic problem is that unobserved preference parameters may affect both

$\nabla_{\ln x}\mathbf{w}(\overline{\mathbf{p}},\overline{x}_i,\mathbf{z}_i)$ and $\overline{\mathbf{w}}_i$, which could induce covariance between them. In this case, $\widehat{\nabla_{\ln x}}\mathbf{w}(\overline{\mathbf{p}},\overline{x}_i,\mathbf{z}_i)\overline{\mathbf{w}}_i'$ would not equal the expectation of $\nabla_{\ln x}\mathbf{w}(\overline{\mathbf{p}},\overline{x}_i,\mathbf{z}_i)\overline{\mathbf{w}}_i'$, but would instead have a matrix of bias terms equal to the unobserved covariance between $\nabla_{\ln x}\mathbf{w}(\overline{\mathbf{p}},\overline{x}_i,\mathbf{z}_i)$ and $\overline{\mathbf{w}}_i$. The following proposition gives an estimator for $\overline{\mathbf{\Gamma}}_i$ that exploits the restriction of Slutsky symmetry to get around Lewbel's problem.

Proposition 3 *Taking Slutsky symmetry as a maintained assumption for individual demands, and assuming that unobserved preference heterogeneity is independent of* $\mathbf{p},x,\mathbf{z}$ *(which implies conditional exogeneity), then* $\overline{\mathbf{\Gamma}}_i$ *may be estimated consistently with*

$$\widehat{\overline{\Gamma}}_i \equiv \frac{1}{2}\left(\widehat{\nabla_{\ln \mathbf{p}'}}\mathbf{w}(\overline{\mathbf{p}},\overline{x}_i,\mathbf{z}_i) + \left(\widehat{\nabla_{\ln \mathbf{p}'}}\mathbf{w}(\overline{\mathbf{p}},\overline{x}_i,\mathbf{z}_i)\right)' + \widehat{\nabla_{\ln x}}\left[\mathbf{w}(\overline{\mathbf{p}},\overline{x}_i,\mathbf{z}_i)\mathbf{w}(\overline{\mathbf{p}},\overline{x}_i,\mathbf{z}_i)'\right]\right). \qquad (20)$$

where $\widehat{\nabla_{\ln \mathbf{p}'}}\mathbf{w}(\overline{\mathbf{p}},\overline{x}_i,\mathbf{z}_i)$ *is a consistent estimator for the uncompensated price semi-elasticity matrix of budget shares and* $\widehat{\nabla_{\ln x}}\left[\mathbf{w}(\overline{\mathbf{p}},\overline{x}_i,\mathbf{z}_i)\mathbf{w}(\overline{\mathbf{p}},\overline{x}_i,\mathbf{z}_i)'\right]$ *is a consistent matrix-valued estimator for the the expenditure semi-elasticity of the outer product of budget-share vector-functions. Any of the standard nonparametric estimators (including, for example, local linear derivative estimators) provides a consistent estimator for these objects.*

Proof. Given Slutsky symmetry, $\mathbf{\Gamma}(\mathbf{p},u,\mathbf{z}) = \mathbf{\Gamma}(\mathbf{p},u,\mathbf{z})'$, so that

$$\nabla_{\ln \mathbf{p}'}\mathbf{w}(\mathbf{p},x,\mathbf{z}) + \nabla_{\ln x}\mathbf{w}(\mathbf{p},x,\mathbf{z})\mathbf{w}(\mathbf{p},x,\mathbf{z})' = \left(\nabla_{\ln \mathbf{p}'}\mathbf{w}(\mathbf{p},x,\mathbf{z})\right)' + \mathbf{w}(\mathbf{p},x,\mathbf{z})\nabla_{\ln x}\mathbf{w}(\mathbf{p},x,\mathbf{z})',$$

which implies that

$$2\mathbf{\Gamma}(\mathbf{p},u,\mathbf{z}) = \nabla_{\ln \mathbf{p}'}\mathbf{w}(\mathbf{p},x,\mathbf{z}) + \nabla_{\ln x}\mathbf{w}(\mathbf{p},x,\mathbf{z})\mathbf{w}(\mathbf{p},x,\mathbf{z})' + \left(\nabla_{\ln \mathbf{p}'}\mathbf{w}(\mathbf{p},x,\mathbf{z})\right)' + \mathbf{w}(\mathbf{p},x,\mathbf{z})\nabla_{\ln x}\mathbf{w}(\mathbf{p},x,\mathbf{z})'.$$

Thus, we have

$$\mathbf{\Gamma}(\mathbf{p},u,\mathbf{z}) = \frac{1}{2}\left(\nabla_{\ln \mathbf{p}'}\mathbf{w}(\mathbf{p},x,\mathbf{z}) + \nabla_{\ln x}\mathbf{w}(\mathbf{p},x,\mathbf{z})\mathbf{w}(\mathbf{p},x,\mathbf{z})' + \left(\nabla_{\ln \mathbf{p}'}\mathbf{w}(\mathbf{p},x,\mathbf{z})\right)' + \mathbf{w}(\mathbf{p},x,\mathbf{z})\nabla_{\ln x}\mathbf{w}(\mathbf{p},x,\mathbf{z})'\right),$$

and the symmetry-restricted estimated matrix of compensated semi-elasticities for a person i, may be written as in (20) ■

The intuition here is as follows. The troublesome term is $\widehat{\nabla_{\ln x}}\mathbf{w}(\overline{\mathbf{p}},\overline{x}_i,\mathbf{z}_i)\overline{\mathbf{w}}_i'$. Even if unobserved preference heterogeneity parameters are independent of $\mathbf{p},x,\mathbf{z}$, there may be correlation between how such heterogeneity affects the slope and level of budget-shares. Since we are interested in the product of those, such a correlation would manifest as bias. In contrast, under the maintained restriction of Slutsky symmetry, the expenditure derivative of the outer product of budget shares has the appropriate product inside it, so we can estimate it directly. Thus, it is the added restriction of Slutsky symmetry that allows us to avoid Lewbel's problem.

Denoting the estimated value of $\overline{\mathbf{\Gamma}}_i$ as $\widehat{\overline{\mathbf{\Gamma}}}_i$ (using either of the estimators given by equations (19) or (20)), define the plutocratic and democratic weighted average budget-shares, outer-product of budget-shares and compensated semi-elasticity matrices as

$$\begin{aligned}\widetilde{\mathbf{w}}^m &= \sum_{i=1}^{N} \phi_i^m \overline{\mathbf{w}}_i, \\ \widetilde{\mathbf{W}}^m &= \sum_{i=1}^{N} \phi_i^m \overline{\mathbf{W}}_i, \\ \widetilde{\mathbf{\Gamma}}^m &= \sum_{i=1}^{N} \phi_i^m \widehat{\overline{\mathbf{\Gamma}}}_i,\end{aligned}$$

for $m = P, D$. Given these objects, we can write the second-order approximations of the plutocratic and democratic indices as

$$\Pi^P \approx 1 + \mathbf{dp}'\widetilde{\mathbf{w}}^P + \frac{1}{2}\mathbf{dp}'\left[\widetilde{\mathbf{W}}^P - diag\left(\widetilde{\mathbf{w}}^P\right) + \widetilde{\mathbf{\Gamma}}^P\right] d\mathbf{p}, \tag{21}$$

$$\Pi^D \approx 1 + \mathbf{dp}'\widetilde{\mathbf{w}}^D + \frac{1}{2}\mathbf{dp}'\left[\widetilde{\mathbf{W}}^D - diag\left(\widetilde{\mathbf{w}}^D\right) + \widetilde{\mathbf{\Gamma}}^D\right] d\mathbf{p}. \tag{22}$$

4.2 CS-SCOLI

As with the Plutocratic and Democratic SCOLIs, second order approximation of the CS-SCOLI allows the researcher to account for substitution effects. However, unlike those SCOLIs, the CS-SCOLI is not generally linear the individual COLIs, which makes its approximation slightly more complex. In particular, the second-order approximation of the CS-SCOLI has an additional term, which accounts for changes in the welfare weights as prices change. The following proposition establishes the second-order approximation of the CS-SCOLI:

Proposition 4 *Given the normalized reference proportionate welfare weights, $\overline{\phi}_i$, their local price responses, $\nabla_{\ln \mathbf{p}'}\phi_i\left(\overline{\mathbf{p}}, \overline{x}_1, .., \overline{x}_N, \mathbf{z}_1, ..., \mathbf{z}_N\right)$, reference budget shares, $\overline{\mathbf{w}}_i$, and their price and expenditure derivatives, the second-order approximation of the CS-SCOLI Π^* is*

$$\Pi^* \approx 1 + d\mathbf{p}'\widetilde{\mathbf{w}} + \frac{1}{2}d\mathbf{p}'\left[\widetilde{\mathbf{w}}\ \widetilde{\mathbf{w}} - diag\left(\widetilde{\mathbf{w}}\right) + \widetilde{\nabla_{\ln \mathbf{p}'}\mathbf{w}} + \widetilde{\nabla_{\ln x}\mathbf{w}}\ \widetilde{\mathbf{w}}' + \widetilde{\mathbf{\Phi}}\right] d\mathbf{p} \tag{23}$$

where

$$\widetilde{\mathbf{w}} \equiv \sum_{i=1}^{N} \overline{\phi}_i \overline{\mathbf{w}}_i,$$
$$\widetilde{\nabla_{\ln \mathbf{p}'} \mathbf{w}} \equiv \sum_{i=1}^{N} \overline{\phi}_i \nabla_{\ln \mathbf{p}'} \mathbf{w}\left(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i\right),$$
$$\widetilde{\nabla_{\ln x} \mathbf{w}} = \sum_{i=1}^{N} \overline{\phi}_i \nabla_{\ln x} \mathbf{w}(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i),$$

and

$$\widetilde{\mathbf{\Phi}} \equiv \sum_{i=1}^{N} \overline{\phi}_i \overline{\mathbf{w}}_i \nabla_{\ln \mathbf{p}'} \ln \phi_i \left(\overline{\mathbf{p}}, \overline{x}_1, .., \overline{x}_N, \mathbf{z}_1, ..., \mathbf{z}_N\right).$$

Here, tildas (with no subscript i, and no superscript) denote welfare-weighted averages.

Proof. See Appendix B. ■

The second-order approximation of the CS-SCOLI differs from a weighted average of second-order approximations of individual COLIs in two ways. First, it contains products of weighted averages rather than weighted averages of products, for example $\widetilde{\mathbf{w}}\ \widetilde{\mathbf{w}}$ rather than $\widetilde{\mathbf{W}}$. Second, it contains $\widetilde{\mathbf{\Phi}}$, which captures the response of the welfare weight functions to price changes. We note that if the CS-SCOLI is equal to the plutocratic SCOLI (for example, if welfare is utilitarian, the marginal utility of money is a constant and preferences are quasihomothetic), then the second-order approximation of the CS-SCOLI equals that given for the plutocratic SCOLI in equation (21). In that case, $\widetilde{\mathbf{w}}\ \widetilde{\mathbf{w}} + \widetilde{\nabla_{\ln x} \mathbf{w}}\ \widetilde{\mathbf{w}}' = \widetilde{\mathbf{W}} + \widetilde{\nabla_{\ln x} \mathbf{w}}\ \mathbf{w}'$ and $\widetilde{\mathbf{\Phi}} = \mathbf{0}$ (the first follows from the constant marginal propensities to consume under quasihomothetic preferences, and we discuss the latter condition in detail below).

For purposes of calculation and estimation, equation (23) is convenient. The first three terms in this approximation are easily constructed from data given the welfare weights $\overline{\phi}_i$. The weighted average budget shares, $\widetilde{\mathbf{w}}$, are computed directly from the data and the welfare weights. The next two terms may be computed using the estimated weighted average derivatives, $\widetilde{\nabla_{\ln \mathbf{p}'} \mathbf{w}}$ and $\widetilde{\nabla_{\ln x} \mathbf{w}}$. These objects may be estimated at efficient convergence rates using any of the standard average derivative estimators, including a weighted version of the score-type estimator for the average of $\nabla_{\ln \mathbf{p}'} \mathbf{w}$ proposed by Deaton and Ng (1998).

The trailing term of the approximation is driven by the price elasticity of the normalized proportionate reference welfare weights. If, at reference arguments, these welfare weights are highly elastic with respect to prices, then this term may potentially be large. Unfortunately, although applied researchers studying inequality have been very willing to assume knowledge of welfare weights, our guess is that assuming knowledge of the elasticity of these weights with respect to prices will be less popular.

We propose explore two ways to get a sense of the size of this trailing term. First, in the proposition below, we provide some sets of sufficient conditions for this term to be exactly zero. Note that the trailing term may be written as $\widetilde{\boldsymbol{\Phi}} = \sum_{i=1}^{N} \overline{\mathbf{w}}_i \nabla_{\ln \mathbf{p}'} \phi_i \left(\overline{\mathbf{p}}, \overline{x}_1, .., \overline{x}_N, \mathbf{z}_1, ..., \mathbf{z}_N\right)$, which is the empirical covariance between budget-shares and the price semi-elasticities of the welfare weight functions (it is a covariance matrix, rather than just a moment matrix, because the sum of $\nabla_{\ln \mathbf{p}'} \phi_i$ is zero by construction). The sufficient conditions we give in our proposition are sufficient conditions for uncorrelatedness, either by making $\overline{\mathbf{w}}_i$ the same for all $i = 1, ..., N$ or by making $\nabla'_{\ln \mathbf{p}} \ln \phi_i \left(\overline{\mathbf{p}}\right) = 0$ for all $i = 1, ..., N$. Identically homothetic preferences implies the former case. A welfare weight function which is independent of prices implies the latter. The cases in which we show that the trailing term is zero are very similar to the cases in which we show above that the CS-SCOLI has an explicit representation. A second way we can assess the importance of the trailing term is via a Monte Carlo exercise. We show in the next section that in a plausible setting, the trailing term is very small.

Proposition 5 *Define $a(\mathbf{p}, \mathbf{z}), b(\mathbf{p}, \mathbf{z}), c(\mathbf{p}), d(\mathbf{z})$ and $h(t)$ as in Proposition 1. Define PIGL indirect utility as in Proposition 1: $V(\mathbf{p}, x, \mathbf{z}) = a(\mathbf{p}, \mathbf{z}) + \left(\frac{x}{b(\mathbf{p},\mathbf{z})}\right)^{\theta} / \theta$ for $\theta \neq 0$ and $V(\mathbf{p}, x, \mathbf{z}) = a(\mathbf{p}, \mathbf{z}) \left[\ln x - \ln b(\mathbf{p}, \mathbf{z})\right]$ for $\theta = 0$. The trailing term in (23) is exactly zero if any of the following sets of conditions hold:*

1. *indirect utility is identically homothetic with $V(\mathbf{p}, x, \mathbf{z}) = h\left(\frac{x}{c(\mathbf{p})d(\mathbf{z})}\right)$, and the direct welfare function depends only on the utility vector (i.e.„ is 'welfarist');*
2. *indirect utility is PIGL with $\theta \neq 0$, and b is multiplicatively separable where $b(\mathbf{p}, \mathbf{z}) = c(\mathbf{p})d(\mathbf{z})$, and the direct welfare function is utilitarian where $W(\cdot) = \sum_{i=1}^{N} u_i$;*
3. *indirect utility is PIGL with $\theta = 0$ (aka PIGLOG), a is multiplicatively separable where $a(\mathbf{p}, \mathbf{z}) = c(\mathbf{p})d(\mathbf{z})$, and the direct welfare function is utilitarian where $W(\cdot) = \sum_{i=1}^{N} u_i$;*
4. *indirect utility is PIGL, a is independent of z ($a(\mathbf{p}, \mathbf{z}) = a(\mathbf{p})$), b is multiplicatively separable where $b(\mathbf{p}, \mathbf{z}) = c(\mathbf{p})d(\mathbf{z})$, and the direct welfare function is a weighted sum of utilities where $W(\cdot) = \sum_{i=1}^{N} g_i u_i$, with S-Gini weights g_i which sum to 1 and depend only on person i's position in the distribution.*

Proof. Case 1: With identically homothetic preferences, π_i (and therefore $d\pi_i$) is identical for all $i = 1, ..., N$, and so is uncorrelated with $d\overline{\phi}_i$. Since $\sum_{i=1}^{N} d\overline{\phi}_i = 0$, we have $\sum_{i=1}^{N} d\pi_i^L d\overline{\phi}_i = 0$. Cases 2, 3 and 4: See Appendix B. ■

These cases include reasonable combinations of welfare and indirect utility functions. For example, Case 2 includes the case where the direct welfare function is utilitarian with $W = \sum_{i=1}^{N} u_i$ and indirect utility is PIGLOG (corresponding to the Almost Ideal demand system (Deaton and Muellbauer 1980)) with $V(\mathbf{p}, x, \mathbf{z}) = a(\mathbf{p})[\ln x - \ln b(\mathbf{p}, \mathbf{z})]$. In this case, $\nabla_{u_i} W(\cdot) = 1$ and $\nabla_{\ln x} V(\mathbf{p}, x, \mathbf{z}) = a(\mathbf{p})$. Consequently, we have $\phi_i(\mathbf{p}, x_1, .., x_N, \mathbf{z}_1, ..., \mathbf{z}_N) = 1/N$, which is obviously invariant to prices.

The uncorrelatedness required for the trailing term to be zero is clearly restrictive. However, we note two important features of our sufficient conditions. First, cases 2 and 3 have PIGL indirect utility wherein preferences may be nonhomothetic. The PIGL class contains as cases both quasihomothetic preferences and PIGLOG preferences, both of which are commonly used rank-2 demand systems. This result is slightly counter-intuitive because many have shown (e.g., Banks, Blundell and Lewbel 1997) that standard welfare weights defined as $\nabla_{u_i} W(u_i, ..., u_N)\nabla_{x_i} V(\mathbf{p}, x_i, \mathbf{z}_i)$ are independent of prices if and only if V is quasihomothetic. This is because $\nabla_{x_i} V$ is independent of prices if and only if V is affine in (a function of) expenditures. Quasihomotheticity of indirect utility implies marginal propensities to consume goods which are independent of expenditure for all goods, and is consequently an undesirable restriction. However, since our normalized welfare weights have the population sum of welfare weights in the denominator, we require a weaker restriction on V, namely that $\nabla_{x_i} V$ is multiplicatively separable into a function of prices only and a function of expenditure and demographics.

Second, cases 1 and 3 allow for a direct welfare function which averse to inequality in utilities. Although case 1 requires identically homothetic preferences, case 3, which covers the S-Gini class of inequality-averse direct welfare functions (see Donaldson and Weymark 1980; Barrett and Pendakur 1995), does not require homothetic preferences.

This section develops a strategy to estimate second-order approximations of various social cost-of-living indices. As in the first-order approximation, our method does not require a parametric model of consumer demand. Another similarity is that our nonparametric estimation strategy is robust to independent preference heterogeneity. However, in contrast to the first-order approximation case, our nonparametric estimation strategy may suffer from bias (endogeneity) if unobserved preference heterogeneity is correlated with observables such as expenditure or demographics. Consequently, a second-order approximation may not be better than a first-order approximation.

5 Monte Carlo Experiment

Here, we set up an experiment to establish two features the CS-SCOLI in plausible environments: (1) the trailing term in the 2nd-Order approximation is small; and (2) both the welfare weights and the second-order approximation may affect the SCOLI in sizeable ways. To assess these for a given price vector, we need to specify the distributions of expenditures and demographic characteristics, and the structure of consumer preferences. For these, we use the data and empirical specification developed in Pendakur (2002). In addition, we need to specify the social welfare function and consumer marginal utility of money. For these, we use utilitarian and inequality-averse social welfare functions plus utility linear either in expenditure or the log of expenditure.

To mimic the data used by Pendakur (2002), we use a single demographic characteristic, the number of household members, and draw 1000 values of log-expenditure and the number of household members from independent standard normals with means of 4.54 and 2.17, respectively, and standard deviations of 0.65 and 1.31, respectively. The number of household members is discretized to $1, 2, 3....$

For consumer preferences, we use the 9-good parametric demand system estimated by Pendakur (2002) in which all households have single demographic characteristic, z, equal to the number of household members and have Quadratic Almost Ideal (QAI) indirect utility (see Banks, Blundell and Lewbel 1997) given by:

$$V(\mathbf{p}, x, z) = k + h\left(\frac{\ln x - \ln a(\mathbf{p}, z)}{b(\mathbf{p}) + q(\mathbf{p})\left(\ln x - \ln a(\mathbf{p}, z)\right)}\right)$$

where h is monotonically increasing and $k = 4$ is chosen to keep V positive in the simulation. We use the following functional forms for a, b, and q:

$$\begin{aligned}
\ln a(\mathbf{p}, z) &= d_0 + dz + \ln \mathbf{p}'\mathbf{a} + \ln \mathbf{p}'\mathbf{D}z + \frac{1}{2}\ln \mathbf{p}'\mathbf{A}_0 \ln \mathbf{p} + \frac{1}{2}\ln \mathbf{p}'\mathbf{A}_z \ln \mathbf{p}z, \\
\ln b(\mathbf{p}) &= \ln \mathbf{p}'\mathbf{b}, \\
q(\mathbf{p}) &= \ln \mathbf{p}'\mathbf{q},
\end{aligned}$$

where all parameter values are taken from Pendakur (2002, Table 3) and $\boldsymbol{\iota}'\mathbf{a} = 1$, $\boldsymbol{\iota}'\mathbf{b} = \boldsymbol{\iota}'\mathbf{q} = 0$, $\boldsymbol{\iota}'\mathbf{D} = \boldsymbol{\iota}'\mathbf{A}_0 = \boldsymbol{\iota}'\mathbf{A}_z = \mathbf{0}_M$, $\mathbf{A}_0 = \mathbf{A}_0'$ and $\mathbf{A}_z = \mathbf{A}_z'$. For convenience, let the reference price vector be $\overline{\mathbf{p}} = \mathbf{1}_M \cdot 100$ and let $d_0 = -\ln 100$. Thus, $\ln a(\overline{\mathbf{p}}, z) = \ln b(\overline{\mathbf{p}}) = q(\overline{\mathbf{p}}) = 0$. Engel curves are quadratic in the log of expenditure. However, if $\mathbf{q}$ is set to $\mathbf{0}$, then Engel curves are linear in the log of expenditure.

For this indirect utility function, utility at the reference price vector is given by

$$\overline{u}_i = V(\overline{\mathbf{p}}, x_i, z_i) = k + h(\ln x_i - dz_i).$$

Consequently, the function h determines the marginal utility of money. If $h(t) = \exp(t)$, then $V(\overline{\mathbf{p}}, x, z) = k + x/\exp(dz_i)$, so that at a vector of unit prices, the marginal utility of money is constant $(1/\exp(dz_i))$. We refer to this case as "money-metric utility". If, in contrast, $h(t) = t$, then $V(\overline{\mathbf{p}}, x, z) = k + \ln x - dz_i$ and, at a vector of unit prices, the marginal utility of a proportionate increase in money is constant. We refer to this case as "log-money-metric utility".

For the social welfare function, we use the generalized utilitarian form with $W(u_1, ..., u_N) = \sum_{i=1}^{N} g(u_i)$. This yields utilitarianism, which is neutral to inequality of utilities, if $g(u_i) = u_i$. It yields social welfare which is somewhat averse to inequality in utilities if $g(u_i) = \ln u_i$ and strongly averse if $g(u_i) = -(u_i)^{-1}$. These welfare functions correspond to the *Mean-of-Order r*, or *Atkinson*, class of social welfare functions with social welfare ordinally equivalent to the arithmetic mean of utility (if $g(u_i) = u_i$), the geometric mean of utility (if $g(u_i) = \ln u_i$) or the harmonic mean of utility (if $g(u_i) = -(u_i)^{-1}$). The parameter k allows us to keep utility positive, so that we can use this power function for social welfare. It also implies that $g(u_i) = u_i$ combined with $h(t) = t$ is not equivalent to $g(u_i) = \ln u_i$ combined with $h(t) = \exp(t)$.

Given g and h, the normalized reference proportionate welfare weights, $\overline{\phi}_i$, are given by

$$\overline{\phi}_i = \frac{\partial g(u_i)/\partial u \cdot \partial h(\ln x_i - dz_i)/\partial x_i \cdot x_i}{\sum_{i=1}^{N} \partial g(u_i)/\partial u \cdot \partial h(\ln x_i - dz_i)/\partial x_i \cdot x_i}.$$

Here, if $g(u) = u_i$ and $h(t) = t$, we have identical welfare weights $\overline{\phi}_i = 1$, $i = 1, ..., N$. However, for some choices of g and h, the welfare weights vary a lot: for example, if $g(u) = u_i^{-1}$ and $h(t) = t$, the largest welfare weight is approximately three times as large as the smallest.

Case 2 of Proposition 4 is satisfied, and thus the trailing term is exactly zero, if $q(\mathbf{p}) = 0$ (preferences are almost ideal), $h(t) = t$ (utility is log-money metric), and $g(u_i) = u_i$ (social welfare is utilitarian). Using the estimated parametric demand system from Pendakur (2002), either with the estimated $\mathbf{q}$ for the QAI case, or with $\mathbf{q}$ set to $\mathbf{0}$ for the Almost Ideal (AI) case, and the various choices for g and h above, we can assess the size of the trailing term when they lie outside the conditions given in Proposition 4.

We may rewrite the second-order approximation of Π^* given in equation (23) as

$$\Pi^* \approx 1 + d\mathbf{p}'\widetilde{\mathbf{w}} + \frac{1}{2}d\mathbf{p}'\widetilde{\mathbf{S}}d\mathbf{p} + \frac{1}{2}d\mathbf{p}'\widetilde{\mathbf{\Phi}}d\mathbf{p}$$

where

$$\widetilde{\mathbf{S}} = \widetilde{\mathbf{w}}\widetilde{\mathbf{w}}' - diag\left(\widetilde{\mathbf{w}}\right) + \widetilde{\nabla_{\ln \mathbf{p}'}\mathbf{w}} + \widetilde{\nabla_{\ln x}\mathbf{w}\mathbf{w}'}.$$

To assess the cost of ignoring the trailing term, we compare the size of $\widetilde{\mathbf{\Phi}}$ to the size of $\widetilde{\mathbf{S}}$.

Since the second-order approximation of the CS-SCOLI uses quadratic forms, $d\mathbf{p}'\widetilde{\mathbf{S}}d\mathbf{p}$ and $d\mathbf{p}'\widetilde{\mathbf{\Phi}}d\mathbf{p}$, we compare the sizes of these quadratic forms over many random draws of $d\mathbf{p}$ for a single population of households (randomly drawn as described above). Since the population of households is constant across the randomizations, $\widetilde{\mathbf{S}}$ and $\widetilde{\mathbf{\Phi}}$ are constant across the randomizations. We compute $\widetilde{\mathbf{S}}$ directly from the QAI structure. We compute $\widetilde{\mathbf{\Phi}}$ via numerical differentiation.

We draw 1000 observations of the vector $d\mathbf{p}$ from independent random uniform distributions, each with minimum -0.45 and maximum 0.55. The uniform distribution chosen has small but positive mean in order to reflect underlying price growth. The independence of prices from each other allows for lots of relative price variation. For each draw, $d\mathbf{p}_t$, $t = 1, ..., 1000$, we compute the ratio $abs\left(d\mathbf{p}_t'\widetilde{\mathbf{\Phi}}d\mathbf{p}_t / d\mathbf{p}_t'\widetilde{\mathbf{S}}d\mathbf{p}_t\right)$, which captures the relative size of the two parts of the second-order approximation. We use the absolute value to account for the fact that this ratio may be positive or negative for different price vectors, because the numerator and denominator need not have the same sign. Table 1 presents the average over the 1000 draws of $abs\left(d\mathbf{p}_t'\widetilde{\mathbf{\Phi}}d\mathbf{p}_t / d\mathbf{p}_t'\widetilde{\mathbf{S}}d\mathbf{p}_t\right)$ for cases with varying marginal utility, social welfare functions and preference structures (the set of price vectors used is held constant across rows).

Table 1: How Big is the Trailing Term?

Social Welfare Function	Marginal Utility	Preferences	Mean	5th p'tile	95th p'tile
$g(u_i) = u_i$	$h(t) = \exp(t)$	QAI	0.062	0.002	0.141
(*Utilitarian*)	(*money-metric*)	AI	0.045	0.003	0.090
	$h(t) = t$	QAI	0.022	0.001	0.066
	(*log-money-metric*)	AI	0	0	0
$g(u_i) = \ln u_i$	$h(t) = \exp(t)$	QAI	0.084	0.003	0.175
(*Inequality-Averse*)	(*money-metric*)	AI	0.052	0.002	0.101
	$h(t) = t$	QAI	0.023	0.001	0.063
	(*log-money-metric*)	AI	0.022	0.001	0.043
$g(u_i) = (u_i)^{-1}$	$h(t) = \exp(t)$	QAI	0.044	0.001	0.091
(*Strongly Inequality-Averse*)	(*money-metric*)	AI	0.030	0.001	0.079
	$h(t) = t$	QAI	0.070	0.001	0.256
	(*log-money-metric*)	AI	0.054	0.002	0.125

The main lesson from Table 1 is that the trailing term is relatively small under plausible circumstances. The average of the 1000 values of the $abs\left(d\mathbf{p}_t'\widetilde{\mathbf{\Phi}}d\mathbf{p}_t/d\mathbf{p}_t'\widetilde{\mathbf{S}}d\mathbf{p}_t\right)$ takes its largest value of 0.084 for the CS-SCOLI with $g(u_i) = \ln u_i$ and $h(t) = \exp(t)$ combined with QAI preferences. Even in this worst case, the substitution term is an order of magnitude more important than the trailing term. The average of $abs\left(d\mathbf{p}_t'\widetilde{\mathbf{\Phi}}d\mathbf{p}_t/d\mathbf{p}_t'\widetilde{\mathbf{S}}d\mathbf{p}_t\right)$ is less than $1/10th$ in all of the cases. Even the 95th percentile of $abs\left(d\mathbf{p}_t'\widetilde{\mathbf{\Phi}}d\mathbf{p}_t/d\mathbf{p}_t'\widetilde{\mathbf{S}}d\mathbf{p}_t\right)$ is less than about 1/4 in all cases. This means that in the range economic environments we consider, there are very few vectors of price changes for which the trailing term is important.

When $g(u_i) = u_i$, $h(t) = \exp(t)$ and preferences are QAI, the average value of $abs\left(d\mathbf{p}_t'\widetilde{\mathbf{\Phi}}d\mathbf{p}_t/d\mathbf{p}_t'\widetilde{\mathbf{S}}d\mathbf{p}_t\right)$ is less than $1/20th$. This case corresponds to the plutocratic weights plus quadratic Engel curves. When $g(u_i) = u_i$, $h(t) = t$ and preferences are AI, we have from Proposition 7 that $\widetilde{\mathbf{\Phi}} = 0$, so $d\mathbf{p}_t'\widetilde{\mathbf{\Phi}}d\mathbf{p}_t/d\mathbf{p}_t'\widetilde{\mathbf{S}}d\mathbf{p}_t$ is also zero. This case has unitary weights (similar to democratic weights) plus linear Engel curves. These results suggest that ignoring the trailing term may not be too damaging in empirical work.

Now, we consider whether or not the welfare weights can matter in the assessment of the social cost of living. Rather than average over many vectors of prices changes (as in Table 1), here we examine a particular price change. Consider a 20% increase in the price of food purchased from stores. We note that the food share is interesting because the estimated parameter values from Pendakur (2002) show a large, negative and statistically significant own-price semi-elasticity for this budget share. That is, if the price of food goes up, the (uncompensated) food share goes down. Thus, this is a case where second-order effects should matter. In addition, because the share of food—a necessity—is strongly decreasing in expenditure, the welfare weights should also play a role. Using the QAI parameter values, artificial data, welfare functions, marginal utilities and random number seed from above, we provide first- and second-order approximations of the percentage increase in the CS-SCOLI for this increase in the price of food purchased from stores. In addition, we provide the second-order approximation both excluding and including the trailing term.

Table 2: Can Welfare Weights Matter?

Marginal Utility	Social Welfare Function		
	$g(u_i) = u_i$	$g(u_i) = \ln u_i$	$g(u_i) = (u_i)^{-1}$
	first-order approximation		
$h(t) = \exp(t)$	3.89	4.14	4.37
$h(t) = t$	5.16	5.59	6.07
	second-order approximation: no trailing term		
$h(t) = \exp(t)$	3.17	3.38	3.58
$h(t) = t$	4.28	4.65	5.09
	second-order approximation: with trailing term		
$h(t) = \exp(t)$	3.15	3.36	3.56
$h(t) = t$	4.29	4.66	5.10

Comparing (horizontally) within panels shows how the welfare weights affect the CS-SCOLI for this experiment. Considering the first-order approximation, we see that increasing the inequality-aversion of the welfare function increases the magnitude of the CS-SCOLI by 0.48 percentage points if marginal utility is constant ($h(t) = \exp(t)$) and by 0.93 percentage points if marginal utility is declining ($h(t) = t$). In both of these cases, the change in the CS-SCOLI is increased by a factor of about one-eighth between utilitarian welfare and strongly inequality-averse welfare.

Comparing the top and middle panels of Table 2, we see that the second-order part of the approximation can make an appreciable difference. Since substitution effects are always negative, it is comforting that the difference between the first- and second-order estimates is always negative. The magnitude of the second-order term is smallest when social welfare is least inequality-averse and marginal utility is constant. In this case, substitution effects reduce the social cost-of-living impact by 0.72

percentage points. In contrast, when social welfare is inequality-averse and marginal utility is declining, substitution effects can reduce the social cost-of-living impact by almost a full percentage point.

Comparing the middle and bottom panels of Table 2, we see that the trailing term has almost no impact on the estimated social cost-of-living increase. In particular, with constant marginal utility, inclusion of the trailing term decreases the social cost-of-living change by 0.02 percentage points, and, with declining marginal utility, inclusion of the trailing term increases the change by 0.01 percentage points.

The bottom line from this experiment is that the welfare weights can matter in the cost of living, that the second-order part of the approximation (which captures substitution effects) can matter, and that ignoring the trailing term in that second-order approximation (which captures the response of welfare weights to price changes) may be acceptable.

6 Illustration

As an empirical illustration of the CS-SCOLI, we consider changes in the social cost-of-living in the U.S. between 1988 and 2000. The approximate CS-SCOLIs described above all employ weighted averages of expenditure shares and weighted average derivatives of expenditures shares. Our analysis is done at the level of individuals, which are assumed identical within households. Thus, we replicate budget-share vectors $\mathbf{w}_h$ for each individual in each household to generate $\mathbf{w}_i$. For estimation of $\overline{\Gamma}_i$, we use nonparametric estimation of $\overline{\Gamma}_h$ on the basis of H household-level observations, and replicate these estimates for each individual in the household to generate $\widehat{\overline{\Gamma}}_i$, from which we compute $\widetilde{\Gamma}$.

The estimation of weighted average derivatives may be implemented by various empirical strategies. For example, Deaton and Ng (1998) use Hardle and Stoker's (1989) estimator which does not use estimates of derivatives of demands for any particular observations, but rather recovers the average derivative by multiplying the derivative of the density function with the level of demand. We use a more direct approach: we use a high dimensional nonparametric kernel estimator to generate an estimate of the symmetry-restricted matrix of compensated price semi-elasticities, $\widehat{\overline{\Gamma}}_i$ defined in equation (20), for all $i = 1, ..., N$. Because the data are household-level, each person in household h is assigned the estimated value for that household. Then, we compute $\widetilde{\mathbf{\Gamma}}$ as the weighted average of $\widehat{\overline{\Gamma}}_i$ which can be plugged into the second-order approximation given by equation (23). With these household-level data, welfare-weighted sample means and weighted average derivatives both converge at $\sqrt{H}$, where H is the number of households (observations) facing the reference price vector, since the welfare weights are strictly positive and bounded by construction.

We use household-level microdata on expenditures from the American Consumer Expenditure Surveys (CES), 1980 to 1998, and aggregate commodity price data from 1980 to 2000, both of which are publicly available from the Bureau of Labor Statistics (BLS). These are part of the data which underlie the Consumer Price Index produced by the BLS. As described above, the CPI is a first-order approximation to the plutocratic SCOLI, Π^P. Unfortunately, the publicly available data do not allow us to take into account any cross-sectional (i.e., regional) variation in prices.

We estimate our model using household expenditures in 19 distinct price regimes representing annual commodity prices for 9 goods for each year 1980 to 1998. The CES microdata are available at the monthly and quarterly level, but since our commodity price data are annual over calendar years, we use only households for which a full year of expenditure is available, with the full year starting in December, January or February. Since the rental flows from owned accommodation are difficult to impute and commodity prices are available only for urban residents, we use only rental-tenure urban residents of the continental USA. There remain 4705 households in our restricted sample, with approximately 300 observations in each year from 1980 to 1998. Following Harris and Sabelhaus (2000), we reweight all household data to reflect these sample restrictions. These weights are used in the constructing sample weighted averages and weighted average derivatives, but not in the kernel estimation step. Summary statistics for the sample are presented in Appendix A.

The nine commodities are: food at home; food out; rent; household furnishing and operation; clothing; motor fuel; public transportation; alcohol; and tobacco products. These commodities account for approximately 3/4 of household consumption for households in the sample. We condition on two household demographic characteristics: the number of household members; and the age of the head of the household.

We present results for the plutocratic and democratic SCOLIs and for various CS-SCOLIs. The plutocratic SCOLI uses weights $\overline{\phi}_i^P = x_i n_i^{-1} / \sum_i x_i n_i^{-1}$ and the democratic SCOLI uses weights $\overline{\phi}_i^D = n_i^{-1} / \sum_i n_i^{-1}$, respectively. Recall from Proposition 1 that the CS-SCOLI is equal to the plutocratic SCOLI if $V(\mathbf{p}, x_i, \mathbf{z}_i) = a(\mathbf{p}, \mathbf{z}_i) + \frac{b(\mathbf{p})}{n_i} x_i$ and welfare is utilitarian. We consider CS-SCOLIs that nest the plutocratic SCOLI by keeping this form for indirect utility and adding curvature to the welfare function. We use the same three social welfare functions as in the calibration experiment: $W(\cdot) = \sum_i u_i$ is utilitarian yielding plutocratic weights $\overline{\phi}_i^P$; $W(\cdot) = \sum_i \ln(u_i)$ is somewhat inequality-averse yielding *unitary* weights $\overline{\phi}_i = 1$; $W(\cdot) = \sum_i -(u_i)^{-1}$ is strongly inequality-averse yielding *reciprocal* weights

$\overline{\phi}_i = (x_i/n_i)^{-1} / \sum_i (x_i/n_i)^{-1}$. As noted above, these welfare functions are ordinally equivalent to the arithmetic, geometric and harmonic means, respectively, of utility. These correspondences are not unique: for example, unitary weights also result if indirect utility is Almost Ideal with $V(\mathbf{p}, x_i, \mathbf{z}_i) = a(\mathbf{p}, \mathbf{z}_i) + b(\mathbf{p}) \ln (x_i/d(\mathbf{z}_i))$ and welfare is utilitarian with $W(\cdot) = \sum_i u_i$.

Table 3 gives approximations of SCOLIs using plutocratic, democratic, unitary and reciprocal weights. The left-hand side gives first-order estimates and the right-hand side gives second-order estimates. We do not provide standard errors in the table because the variance in the estimates due to the variance of the estimate of $\widetilde{\mathbf{\Gamma}}$ is very small (bootstrapped standard errors are less than 0.05 percentage points for all estimates shown). As noted by Ley (2002, 2005), the variance induced by the variance of $\widetilde{\mathbf{\Gamma}}$ and $\widetilde{\mathbf{w}}$ is likely dwarfed by the variance induced by measurement error in price changes, $d\mathbf{p}$.

We present illustrative results for 2 periods: 1988 to 1998 and 1999 to 2000. The former period is chosen because the BLS used 1982-4 expenditure weights for calculating the CPI over that entire period. We use 1983 expenditure weights for that period. (Since the late 1990s, the BLS has updated the weights used in the CPI about every 2 years.) The latter period is chosen because although most prices were fairly stable over 1999 to 2000, the price of motor fuel rose by 30% over this year. Since motor fuel represents a comparatively large expenditure share for the bottom half of households, we may expect the distributional weights to matter over such a price change. We use 1998 expenditure weights to assess this price change.

Table 3: Estimated CS-SCOLIs, percentage changes

	first-order				second-order			
	Pluto	Demo	Unitary	Reciprocal	Pluto	Demo	Unitary	Reciprocal
1988-1998	32.5	33.5	33.6	34.2	31.9	32.7	32.8	33.4
1999-2000	4.8	4.9	5.0	5.1	4.8	4.9	4.9	5.0

Over 1988 to 1998 and 1999 to 2000, the BLS reports that the CPI rose by 35.6% and 3.3%, respectively, which is different from our reported plutocratic SCOLI. However, we do not expect the CPI and our plutocratic SCOLI to be numerically identical because the CPI is computed from a much larger and finer set of commodities. Consider first-order approximations over the long period 1988 to 1998 reported in Table 3. The Plutocratic SCOLI rose by 32.5%, but, because price changes favoured rich households over poor households during this period, the democratic SCOLI which up-weights the experience of poor households, rose by 33.5%. Similarly, the index with unitary weights shows an increase of 33.6%. The index with reciprocal per-capita expenditure weights, which upweights the experience of poor households even more, shows an increase of 34.2%. Thus, over this 10-year period, we see that different plausible weighting structures in the CS-SCOLI yield different pictures of the path for the social cost of living. In particular, the index which emphasizes the experience of poor households shows nearly 2 percentage points more inflation than that which emphasizes the experience of rich households. These results are consistent with other studies showing variation in the cost-of-living across income classes (see, e.g., Pendakur 2002, Ley 2002, Chiru 2005a,b).

Turning to the one-year price change for 1999 to 2000, we see that the large increase in the relative price of motor fuel had a noticeable distributional effect. The first-order approximations to the plutocratic- and democratic-weighted CS-SCOLIs rose 4.8% and 4.9%, respectively. The first-order approximation to the CS-SCOLI using reciprocal per-capita expenditure weights rose by 5.1%. This is due to the fact that motor fuel is a necessity whose price increases affects poor households more than rich households. Given that this is only a 1-year price change, these are large differences across the various weighting structures.

The right-hand panel of Table 3 presents estimates of second-order approximations of the same CS-SCOLIs for the same years. In these approximations, we ignore the contribution of the trailing term in the approximation. Here, we may illustrate the importance of accounting for substitution effects in the assessment of the social cost-of-living. If the CS-SCOLI is concave in prices (which it is for the cases given in Proposition 1), then accounting for substitution effects must (weakly) reduce the estimated SCOLI. For all indices and price changes shown, this is the case.

For the period 1988 to 1998, accounting for substitution effects reduces our estimate of the increase in plutocratic CS-SCOLI from 32.5% to 31.9%, a difference of 0.6 percentage points. For all three other CS-SCOLI's shown, the difference between first- and second-order approximations is 0.8 percentage points. This magnitude for substitution effects is plausible given the high level of commodity aggregation in our illustration, though somewhat smaller than the magnitudes identified in the Boskin Report (1996).

It is natural to expect that substitution effects will matter more if expenditure weights are updated infrequently or with long lags. This is because if expenditure weights are updated continuously and instantly, a fine sequence of first-order approximations will capture the behavioral responses that the substitution effects 'predict' (see Vartia 1983). Since the BLS has substantially reduced delays and increased the frequency of expenditure updates, it may be important to assess the size of second-order effects over short periods. Turning to the one-year price change from 1999 to 2000 which uses 1998 expenditure weights, we still see substitution effects of noticeable magnitude. During this period, the price of motor fuel rose by 30%, which is large enough in principle to induce changes in behavior to reduce the cost of the price change. The bottom row of the right-hand panel of Table 2 suggests that this effect was small. For all the plutocratic and democratic SCOLIs, the first-order approximation is larger than the second-order approximation, but by less than 0.05 percentage points. For the CS-SCOLIs with unitary and reciprocal weights, the first-order approximations are 0.08 and 0.09 percentage points larger, respectively, than the second-order approximations.

7 Conclusion

For an individual, the change in the cost-of-living is the scaling of expenditure required to hold utility constant over a price change. Because preferences and resources differ across people, for any price change, there is heterogeneity across individuals in their cost-of-living changes. Thus, a social cost-of-living approach to the measurement of price change faces a formidable aggregation problem. Whose cost-of-living should we be measuring?

The common-scaling social cost-of-living index (CS-SCOLI) developed in this paper answers the following question. What single scaling to everyone's expenditure would hold social welfare constant across a price change? The common-scaling feature of this index corresponds directly to the feasible policy uses of a social cost-of-living index—equiproportionate income adjustments, for example, as in adjusting the payments from public pensions or benefit programs. The CS-SCOLI has social welfare foundations, and allows the investigator to easily choose the weight placed on rich and poor households. It is easy to implement, and we have provided methods for estimating both first- and second-order approximations to the index. The latter capture substitution effects. Finally, the CS-SCOLI has as special cases objects that are either identical or very similar to all the commonly used social cost-of-living indices, and in particular, the plutocratic and democratic SCOLIs. This is important. In our framework there is a social welfare function and utility function which lead to the CPI. Thus the CPI is given an explicit social welfare foundation. Moreover, an investigator who finds the social welfare and utility function corresponding to the CPI unpalatable can easily generate a SCOLI more to her tastes.

8 Acknowledgements

We acknowledge the comments and thoughts of our friends and colleagues (Ian Crawford, David Donaldson, Greg Dow, Geoff Dunbar, Arthur Lewbel and Steeve Mongrain) as well as the helpful suggestions of the editor, associate editor and two anonymous referees. We gratefully acknowledge the financial support of the Social Sciences and Humanities Research Council of Canada.

Appendix A: Summary Statistics

Table A1 gives summary statistics for the data we use in our analysis.

Table A1: The Data		Min	Max	Mean	Std Dev
expenditure shares	food-in	0	0.85	0.26	0.13
	food-out	0	0.63	0.08	0.07
	rent	0	0.94	0.41	0.15
	hh furn/equip	0	0.45	0.04	0.05
	clothing	0	0.41	0.06	0.05
	motor fuel	0	0.43	0.07	0.06
	public trans	0	0.39	0.09	0.04
	alcohol	0	0.54	0.03	0.04
	tobacco	0	0.26	0.03	0.04
log-expenditure		6.66	10.76	9.05	0.55
log household size		0	2.56	0.65	0.59
age of head (less 40)		-24	24	2.9	11

Appendix B: Proofs

Proof. (Proposition 4) CS-SCOLI is defined by

$$\begin{aligned}\overline{s} &= W\left(V(\mathbf{p}, \Pi^*\left(\mathbf{p}; \overline{\mathbf{p}}, \overline{x}_1, .., \overline{x}_N, \mathbf{z}_1, ..., \mathbf{z}_N\right)\overline{x}_1, \mathbf{z}_1), ..., V(\mathbf{p}, \Pi^*\left(\mathbf{p}; \overline{\mathbf{p}}, \overline{x}_1, .., \overline{x}_N, \mathbf{z}_1, ..., \mathbf{z}_N\right)\overline{x}_N, \mathbf{z}_N)\right) \\ &= W\left(V(\mathbf{p}, \Pi^*\left(\mathbf{p}\right)\overline{x}_1, \mathbf{z}_1), ..., V(\mathbf{p}, \Pi^*\left(\mathbf{p}\right)\overline{x}_N, \mathbf{z}_N)\right)\end{aligned}$$

where $\Pi^*\left(\mathbf{p}\right) = \Pi^*\left(\mathbf{p}; \overline{\mathbf{p}}, \overline{x}_1, .., \overline{x}_N, \mathbf{z}_1, ..., \mathbf{z}_N\right)$, because Π^* depends only on $\mathbf{p}$ as a variable. The implicit function theorem gives

$$\begin{aligned}\nabla_{\mathbf{p}}\Pi^*\left(\mathbf{p}\right) &= -\frac{\sum_{i=1}^{N}\nabla_{u_i}W\left(V(\mathbf{p}, \Pi^*\left(\mathbf{p}\right)\overline{x}_1, \mathbf{z}_1), ..., V(\mathbf{p}, \Pi^*\left(\mathbf{p}\right)\overline{x}_N, \mathbf{z}_N)\right)\nabla_{\mathbf{p}}V(\mathbf{p}, \Pi^*\left(\mathbf{p}\right)\overline{x}_i, \mathbf{z}_i)}{\sum_{i=1}^{N}\nabla_{u_i}W\left(V(\mathbf{p}, \Pi^*\left(\mathbf{p}\right)\overline{x}_1, \mathbf{z}_1), ..., V(\mathbf{p}, \Pi^*\left(\mathbf{p}\right)\overline{x}_N, \mathbf{z}_N)\right)\nabla_{x_i}V(\mathbf{p}, \Pi^*\left(\mathbf{p}\right)\overline{x}_i, \mathbf{z}_i)\overline{x}_i} \\ &= \sum_{i=1}^{N}\phi_i\left(\mathbf{p}\right)\frac{\mathbf{q}(\mathbf{p}, \Pi^*\left(\mathbf{p}\right)\overline{x}_i, \mathbf{z}_i)}{\overline{x}_i}\end{aligned}$$

where $\mathbf{q}(\mathbf{p}, \Pi^*(\mathbf{p})\overline{x}_i, \mathbf{z}_i) = \nabla_{\mathbf{p}} V(\mathbf{p}, \Pi^*(\mathbf{p})\overline{x}_i, \mathbf{z}_i) / \nabla_{x_i} V(\mathbf{p}, \Pi^*(\mathbf{p})\overline{x}_i, \mathbf{z}_i)$ by Roy's Identity, and

$$\phi_i(\mathbf{p}) = \frac{\nabla_{u_i} W\left(V(\mathbf{p}, \Pi^*(\mathbf{p})\overline{x}_1, \mathbf{z}_1), ..., V(\mathbf{p}, \Pi^*(\mathbf{p})\overline{x}_N, \mathbf{z}_N)\right) \nabla_{x_i} V(\mathbf{p}, \Pi^*(\mathbf{p})\overline{x}_i, \mathbf{z}_i)\overline{x}_i}{\sum_{i=1}^N \nabla_{u_i} W\left(V(\mathbf{p}, \Pi^*(\mathbf{p})\overline{x}_1, \mathbf{z}_1), ..., V(\mathbf{p}, \Pi^*(\mathbf{p})\overline{x}_N, \mathbf{z}_N)\right) \nabla_{x_i} V(\mathbf{p}, \Pi^*(\mathbf{p})\overline{x}_i, \mathbf{z}_i)\overline{x}_i}.$$

Here, $\phi_i(\mathbf{p})$ depends on $\Pi^*(\mathbf{p})$.

Differentiate $\nabla_{\mathbf{p}}\Pi^*(\mathbf{p})$ to get

$$\begin{aligned}
\nabla_{\mathbf{pp}'}\Pi^*(\mathbf{p}) &= \sum_{i=1}^N \left(\frac{\mathbf{q}(\mathbf{p}, \Pi^*(\mathbf{p})\overline{x}_i, \mathbf{z}_i)}{\overline{x}_i}\right) \nabla_{\mathbf{p}'}\phi_i(\mathbf{p}) + \quad \sum_{i=1}^N \phi_i(\mathbf{p}) \nabla_{\mathbf{p}'}\left(\frac{\mathbf{q}(\mathbf{p}, \Pi^*(\mathbf{p})\overline{x}_i, \mathbf{z}_i)}{\overline{x}_i}\right) \\
&= \sum_{i=1}^N \left(\frac{\mathbf{q}(\mathbf{p}, \Pi^*(\mathbf{p})\overline{x}_i, \mathbf{z}_i)}{\overline{x}_i}\right) \nabla_{\mathbf{p}'}\phi_i(\mathbf{p}) + \quad \sum_{i=1}^N \phi_i(\mathbf{p}) \left(\frac{\nabla_{\mathbf{p}'}\mathbf{q}(\mathbf{p}, \Pi^*(\mathbf{p})\overline{x}_i, \mathbf{z}_i)}{\overline{x}_i}\right) + \\
&\qquad \sum_{i=1}^N \phi_i(\mathbf{p}) \left(\frac{\nabla_{x_i}\mathbf{q}(\mathbf{p}, \Pi^*(\mathbf{p})\overline{x}_i, \mathbf{z}_i)\overline{x}_i \nabla_{\mathbf{p}}\Pi^*(\mathbf{p})}{\overline{x}_i}\right).
\end{aligned}$$

This can be written as

$$\begin{aligned}
\nabla_{\mathbf{pp}'}\Pi^*(\mathbf{p}) \quad = \quad & \sum_{i=1}^N \left(\frac{\mathbf{q}(\mathbf{p}, \Pi^*(\mathbf{p})\overline{x}_i, \mathbf{z}_i)}{\overline{x}_i}\right) \nabla'_{\mathbf{p}}\phi_i(\mathbf{p}) + \\
& \sum_{i=1}^N \phi_i(\mathbf{p}) \left(\frac{\nabla'_{\mathbf{p}}\mathbf{q}(\mathbf{p}, \Pi^*(\mathbf{p})\overline{x}_i, \mathbf{z}_i)}{\overline{x}_i}\right) + \\
& \sum_{i=1}^N \phi_i(\mathbf{p}) \left(\nabla_{x_i}\mathbf{q}(\mathbf{p}, \Pi^*(\mathbf{p})\overline{x}_i, \mathbf{z}_i)\right) \sum_{j=1}^N \phi_j(\mathbf{p}) \frac{\mathbf{q}(\mathbf{p}, \Pi^*(\mathbf{p})\overline{x}_j, \mathbf{z}_j)'}{\overline{x}_j}.
\end{aligned}$$

Evaluated at $\overline{\mathbf{p}}$, we have $\Pi^*(\overline{\mathbf{p}}) = 1$, so that

$$\nabla_{\mathbf{pp}'}\Pi^*(\overline{\mathbf{p}}) = \sum_{i=1}^N \frac{\overline{\mathbf{q}}_i}{\overline{x}_i} \nabla'_{\mathbf{p}}\phi_i(\overline{\mathbf{p}}) + \sum_{i=1}^N \overline{\phi}_i \left(\frac{\nabla'_{\mathbf{p}}\mathbf{q}(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i)}{\overline{x}_i}\right) + \sum_{i=1}^N \overline{\phi}_i \left(\nabla_{x_i}\mathbf{q}(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i)\right) \sum_{i=1}^N \overline{\phi}_j \frac{\overline{\mathbf{q}}'_j}{\overline{x}_j}.$$

where $\overline{\mathbf{q}}_i = \mathbf{q}(\overline{\mathbf{p}}, \overline{x}_i, \mathbf{z}_i)$ and $\overline{\phi}_i = \phi_i(\overline{\mathbf{p}})$.

To convert to budget-share form, note that:

$$\begin{aligned}
\frac{\mathbf{P}\nabla_{\mathbf{p}}\mathbf{q}(\mathbf{p}, x, \mathbf{z})\mathbf{P}}{x} \quad &= \quad \nabla_{\ln \mathbf{p}'}\mathbf{w}(\mathbf{p}, x, \mathbf{z}) - diag\left[\mathbf{w}(\mathbf{p}, x, \mathbf{z})\right], \text{ and} \\
\mathbf{P}\nabla_{x_i}\mathbf{q}(\mathbf{p}, x, \mathbf{z}) \quad &= \quad \nabla_{\ln x}\mathbf{w}(\mathbf{p}, x, \mathbf{z}) + \mathbf{w}(\mathbf{p}, x, \mathbf{z}).
\end{aligned}$$

The second-order term pre-and post-multiplies this by $\overline{\mathbf{P}}$ (recall $d\mathbf{p}'\overline{\mathbf{P}}\nabla_{\mathbf{pp}'}\Pi^*(\overline{\mathbf{p}})\overline{\mathbf{P}}d\mathbf{p} = (\mathbf{p}-\overline{\mathbf{p}})'\nabla_{\mathbf{pp}'}\Pi^*(\overline{\mathbf{p}})(\mathbf{p}-\overline{\mathbf{p}})$):

$$\begin{aligned}
\overline{\mathbf{P}}\nabla_{\mathbf{pp}'}\Pi^*(\overline{\mathbf{p}})\overline{\mathbf{P}} &= \sum_{i=1}^{N}\overline{\phi}_i\mathbf{w}_i\nabla'_{\ln\mathbf{p}}\ln\phi_i(\overline{\mathbf{p}}) + \sum_{i=1}^{N}\overline{\phi}_i\left(\nabla_{\ln\mathbf{p}'}\mathbf{w}(\overline{\mathbf{p}},\overline{x}_i,\mathbf{z}_i) - diag\left[\mathbf{w}(\overline{\mathbf{p}},\overline{x}_i,\mathbf{z}_i)\right]\right) + \\
&\quad \sum_{i=1}^{N}\overline{\phi}_i\left(\nabla_{\ln x}\mathbf{w}(\overline{\mathbf{p}},\overline{x}_i,\mathbf{z}_i) + \mathbf{w}_i\right)\sum_{i=1}^{N}\overline{\phi}_j\mathbf{w}'_j \\
&= \sum_{i=1}^{N}\overline{\phi}_i\mathbf{w}_i\nabla'_{\ln\mathbf{p}}\ln\phi_i(\overline{\mathbf{p}}) + \sum_{i=1}^{N}\overline{\phi}_i\nabla_{\ln\mathbf{p}'}\mathbf{w}(\overline{\mathbf{p}},\overline{x}_i,\mathbf{z}_i) + \sum_{i=1}^{N}\overline{\phi}_i\nabla_{\ln x}\mathbf{w}(\overline{\mathbf{p}},\overline{x}_i,\mathbf{z}_i)\sum_{i=1}^{N}\overline{\phi}_j\mathbf{w}'_j \\
&\quad + \sum_{i=1}^{N}\overline{\phi}_i\mathbf{w}_i\sum_{i=1}^{N}\overline{\phi}_j\mathbf{w}'_j - \sum_{i=1}^{N}\overline{\phi}_i diag\left[\mathbf{w}_i\right]
\end{aligned}$$

Thus, in our tilda notation:

$$\overline{\mathbf{P}}\nabla_{\mathbf{pp}'}\Pi^*(\overline{\mathbf{p}})\overline{\mathbf{P}} = \sum_{i=1}^{N}\overline{\phi}_i\mathbf{w}_i\nabla'_{\ln\mathbf{p}}\ln\phi_i(\overline{\mathbf{p}}) + \widetilde{\nabla_{\ln\mathbf{p}'}\mathbf{w}} + \widetilde{\nabla_{\ln x}\mathbf{w}}\cdot\widetilde{\mathbf{w}}' + \widetilde{\mathbf{w}}\widetilde{\mathbf{w}}' - diag\left[\widetilde{\mathbf{w}}\right]$$

■

Proof. (Proposition 5) Case 2. The welfare function is $W(u_1,...,u_N) = \sum_{i=1}^{N}u_i$, implying that $\nabla_{u_i}W(u_1,...,u_N) = 1$. Indirect utility is $V(\mathbf{p},x,\mathbf{z}) = a(\mathbf{p},\mathbf{z}) + \left(\frac{x}{b(\mathbf{p},\mathbf{z})}\right)^{\theta}/\theta$, so that marginal utility is given by $\nabla_{x_i}V(\mathbf{p},x_i,\mathbf{z}_i) = \left(\frac{x}{b(\mathbf{p},\mathbf{z})}\right)^{\theta-1}\frac{1}{b(\mathbf{p},\mathbf{z})}$. Imposing multiplicative separability on b implies $\nabla_{x_i}V(\mathbf{p},x_i,\mathbf{z}_i) = \left(\frac{x}{c(\mathbf{p})d(\mathbf{z})}\right)^{\theta-1}\frac{1}{c(\mathbf{p})d(\mathbf{z})}$, so that

$$\phi_i(\cdot) = \frac{\left(\frac{x_i}{c(\mathbf{p})d(\mathbf{z}_i)}\right)^{\theta-1}\frac{1}{c(\mathbf{p})d(\mathbf{z})}x_i}{\sum_{i=1}^{N}\left(\frac{x_i}{c(\mathbf{p})d(\mathbf{z}_i)}\right)^{\theta-1}\frac{1}{c(\mathbf{p})d(\mathbf{z})}x_i} = \frac{c(\mathbf{p})^{-\theta}\left(\frac{x_i}{d(\mathbf{z}_i)}\right)^{\theta}}{c(\mathbf{p})^{-\theta}\sum_{i=1}^{N}\left(\frac{x_i}{d(\mathbf{z}_i)}\right)^{\theta}} = \frac{\left(\frac{x_i}{d(\mathbf{z}_i)}\right)^{\theta}}{\sum_{i=1}^{N}\left(\frac{x_i}{d(\mathbf{z}_i)}\right)^{\theta}},\ \text{for } \theta \neq 0,$$

which is independent of prices.

Case 3, $V(\mathbf{p},x,\mathbf{z}) = a(\mathbf{p},\mathbf{z})\left[\ln x - \ln b(\mathbf{p},\mathbf{z})\right]$ and $\nabla_{x_i}V(\mathbf{p},x_i,\mathbf{z}_i) = a(\mathbf{p},\mathbf{z})/x$. Imposing multiplicative separability on a implies $\nabla_{x_i}V(\mathbf{p},x_i,\mathbf{z}_i) = c(\mathbf{p})d(\mathbf{z})/x$ so that

$$\phi_i(\cdot) = \frac{c(\mathbf{p})d(\mathbf{z}_i)}{\sum_{i=1}^{N}c(\mathbf{p})d(\mathbf{z}_i)} = \frac{c(\mathbf{p})d(\mathbf{z}_i)}{c(\mathbf{p})\sum_{i=1}^{N}d(\mathbf{z}_i)} = \frac{d(\mathbf{z}_i)}{\sum_{i=1}^{N}d(\mathbf{z}_i)},$$

which is independent of prices.

Case 4. The welfare function is $W(u_1,...,u_N) = \sum_{i=1}^{N}g_iu_i$, implying that $\nabla_{u_i}W(u_1,...,u_N) = g_i$, and the indirect utility function is $V(\mathbf{p},x,\mathbf{z}) = a(\mathbf{p}) + \left(\frac{x}{b(\mathbf{p})d(\mathbf{z})}\right)^{\theta}/\theta$, so that marginal utility is given by $\nabla_{x_i}V(\mathbf{p},x_i,\mathbf{z}_i) = \left(\frac{x}{b(\mathbf{p})d(\mathbf{z})}\right)^{\theta-1}\frac{1}{b(\mathbf{p})d(\mathbf{z})}$. Since g_i depends only on the rank of individual i in the distribution of utilities, if rank is independent of prices, then the g_i are invariant to prices. Since both

a and b are independent of $\mathbf{z}$ and x, changes in $\mathbf{p}$ do not change the rankings of $u_i = V(\mathbf{p}, x_i, \mathbf{z}_i)$, so the S-Gini weights g_i are the same at every price vector. The normalized proportional welfare weight function is thus

$$\phi_i(\mathbf{p}, x_1, .., x_N, \mathbf{z}_1, ..., \mathbf{z}_N) = \frac{g_i \left(\frac{x}{b(\mathbf{p})d(\mathbf{z})}\right)^{\theta-1} \frac{1}{b(\mathbf{p})d(\mathbf{z})} x_i}{\sum_{i=1}^N g_i \left(\frac{x}{b(\mathbf{p})d(\mathbf{z})}\right)^{\theta-1} \frac{1}{b(\mathbf{p})d(\mathbf{z})} x_i} = \frac{g_i \left(\frac{x_i}{d(\mathbf{z}_i)}\right)^{\theta}}{\sum_{i=1}^N g_i \left(\frac{x_i}{d(\mathbf{z}_i)}\right)^{\theta}},$$

which is independent of prices. ■

References

[1] Banks, J., R. Blundell and A. Lewbel, 1996, "Tax Reform and Welfare Measurement: Do We Need Demand System Estimation," *Economic Journal,* 106:1227-1241.

[2] Banks, J., R. Blundell and A. Lewbel, 1997, "Quadratic Engel Curves and Consumer Demand," *Review of Economics and Statistics*, 79(4):527-539.

[3] Barrett, Garry and Krishna Pendakur, 1995, "Asymptotic Distributions for the Generalised Gini Inequality Indices" Canadian Journal of Economics, 28(4b), pp 1042- 55.

[4] Boskin et al., 1996, "Towards a More Accurate Measure of the Cost of Living," Final report to the US Senate Finance Committee, Washington D.C., December.

[5] Chiru, R., 2005a, "Is Inflation Higher for Seniors?" Statistics Canada Analytical Paper, 11-621-MIE – No. 027.

[6] Chiru, R., 2005b, "Does Inflation Vary with Income?" Statistics Canada Analytical Paper, 11-621-MIE – No. 030.

[7] Crawford, I. and Z. Smith, 2002, "Distributional Aspects of Inflation," The Institute for Fiscal Studies, Commentary 90.

[8] Deaton, A., and S. Ng, 1998, "Parametric and Nonparametric Approaches to Price and Tax Reform," *Journal of the American Statistical Association*, 93(443): 900-909.

[9] Deaton, A., and J. Mullbauer, 1980, "An Almost Ideal Demand System," *American Economic Review*, 70(3):312-326.

[10] Diewert, W., E., 1998, "Index Number Issues in the Consumer Price Index," *Journal of Economic Perspectives*, 12(1):47-58.

[11] Diewert, W.E., 2001, "The Consumer Price Index and Index Number Purpose," *Journal of Economic and Social Measurement*, 27, pp 167-248.

[12] Donaldson, D. and J. Weymark, 1980, "A Single-Parameter Generalisation of the Gini Index of Inequality," *Journal of Economic Theory* 22, pp 67-86.

[13] Donaldson, D. and K. Pendakur, 2009, "Reversability of the CS-SCOLI," unpublished working paper.

[14] Fisher, F.M., 2005, "Price Index Aggregation: Plutocratic Weights, Democratic Weights, and Value Judgements," *Annales d'Economie et de Statistique* 79/80, special issue in memory of Zvi Griliches.

[15] von Haefen, R., 2000, "A Complete Characterization of the Implications of Slutzky Symmetry for the Linear, Log-Linear, and Semi-Log Incomplete Demand System Models," U.S. Bureau of Labor Statistics Working Paper 332, September 2000

[16] Hardle, W. and T. Stoker, 1989, "Investigating Smooth Multiple Regression by the Method of Average Derivatives," *Journal of the American Statistical Association*, 84:986-995.

[17] Harris, E. and J. Sabelhaus, 2000, "Consumer Expenditure Survey Family Level Extracts," National Bureau of Economic Research, http://www.nber.org/data/ces_cbo.html.

[18] Jorgenson, D.W., 1990, "Aggregate Consumer Behaviour and the Measurement of Social Welfare," *Econometrica*, Vol. 58, No. 5 (September, 1990), 1007-1040.

[19] Jorgenson, D.W. and D. Slesnick, 1983, "Individual and Social Cost-of-Living Indexes", in W.E. Diewert and C. Montmarquette (eds), *Price Level Measurement*, Ottawa: Statistics Canada, 1983, pp 241-336.

[20] Kokoski, M.F., 2000, "Alternative CPI Aggregations: Two Approaches," *BLS Monthly Labor Review*, 123:31–39.

[21] Lewbel, A., 2001, "Demand Systems With and Without Error Terms," *American Economic Review*, 91(3):611-618.

[22] Ley, E., 2005, "Whose Inflation? A Characterization of the CPI Plutocratic Gap," *Oxford Economic Papers* 57(4):634-646.

[23] Ley, E., 2002, "On Plutocratic and Democratic CPIs," *Economics Bulletin*, 4(3):1-5.

[24] Muellbauer, J., 1974, "The Political Economy of Price Indices," Birbeck Discussion Paper no 22.

[25] Muellbauer, J., 1976, "Community Preference Fields and the Representative Consumer," Econometrica, 44(5): 979-999.

[26] Nicholson, J.L., 1975, "Whose Cost of Living?" *Journal of the Royal Statistical Society, Series A*, 138:540–2.

[27] Pendakur, K., 2002, "Taking Prices Seriously in the Measurement of Inequality," *Journal of Public Economics*, 86(1):47-69.

[28] Pollak, R., 1980, "Group Cost-of-Living Indexes," *American Economic Review*, May 1980, pp 273-278.

[29] Pollak, R., 1981, "The Social Cost-of-Living Index," *Journal of Public Economics*, 15:311–36.

[30] Prais, S., 1958, "Whose Cost of Living?" *Review of Economic Studies*, 26:126–34.

[31] Slesnick, D.T., 2001, *Consumption and Social Welfare: Living Standards and Their Distribution in the United States*, Cambridge, UK: Cambridge University Press.

[32] Vartia, Yrjo, 1983, "Efficient Methods of Measuring Welfare Change and Compensated Income in Terms of Ordinary Demand Functions," *Econometrica*, 51(1): pp. 79-98.