

Whose Inflation? A Characterization of the CPI Plutocratic Bias

Eduardo Ley
IMF Institute

Version: June 7, 2001

Abstract. Prais (1958) showed that the standard CPI computed by most statistical agencies can be interpreted as a weighed average of household price indexes, where the weight of each household is determined by its total expenditures. In this paper, we decompose the difference between the standard CPI and a democratically weighed index (*i.e.*, the so-called plutocratic bias) as the product of average income, income inequality and the sample covariance between the elementary individual price indexes and a parameter which is a function of the income elasticity of each good. This decomposition allows us to interpret variations in the size and sign of the plutocratic bias, and also to discuss issues pertaining to group indexes.

Keywords. Consumer price index, plutocratic index, democratic index, group index, aggregation, equivalence scales, inflation.

JEL Classification System. C43, D31, D63

Address. Eduardo Ley, IS-3-1343, International Monetary Fund, 700 19th Street NW, Washington DC 20431. Email: eley@imf.org. Vox: +1 (202) 623 6107. eFax: +1 (202) 589 6107.

1. Introduction

The Hong Kong Census and Statistics Department routinely reports three consumer price indexes, by income bracket, along with an overall consumer price index (CPI).¹ CPI-A is based on the expenditure patterns of the bottom 50 percent of the population, CPI-B uses the next 30 percent, and CPI-C is for the next 10 percent. The composite CPI takes into account the expenditure patterns of all these households taken together —which cover 90 percent of the population. For the year 2000, while inflation measured by the overall index was -3.7 percent, inflation rates of the group indexes were, respectively: -2.8 , -3.8 , and -4.5 percent. Thus, differences in inflation rates by income group can be quite substantial. In most countries, however, a single CPI is reported. Even when more than one index is made available by the statistical agency, a single one is often used as an inflation gauge from a macroeconomic policy perspective. How representative is, in general, the official inflation rate, as measured by the CPI?

It is known since Prais (1958)² that the CPI computed by statistical agencies can be interpreted as a weighed average of household price indexes. The weight of each household is given by its total expenditure, hence the term ‘plutocratic index.’ Alternatively, we could construct a democratically-weighed index, where each household weighs the same. We shall define the CPI *plutocratic bias* as the difference between the plutocratic index and the democratic one. Whether price behavior in a given period hurts relatively more the better-off or the worse-off households can be expressed in terms of this single scalar (Fry and Pashardes, 1985).

This paper investigates the sources of possible discrepancies between plutocratic and democratic indexes. We show that the plutocratic bias can be expressed as the product of mean income, a measure of variation of household expenditures, and the sample covariance between the elementary individual price indexes and the corresponding good’s expenditure-share regression coefficient on household income. This coefficient, in turn, is a function of the income elasticity of each good. Consequently, because the decomposition is multiplicative, three elements are required for the bias to exist. First, there must be some dispersion in the distribution of income across households. Second, there must be differences in behavior among households at different income brackets. And, third, there must be differences in behavior among different prices. The bias decomposition allows us to interpret the empirical results obtained on the size and the sign of the plutocratic bias, and suggests that averaging the bias over long time periods may be misleading.

This paper is organized as follows: Section 2 presents analytical results regarding the

I thank Javier Ruiz-Castillo for numerous conversations and stimulating discussions on the subject of this paper. I also thank Woon Gyu Choi, Nuri Erbas, Roberto García-Saltos, Josh Greene, Mario Izquierdo, Mohsin Khan, Rafa Repullo, Hal Varian and Shlomo Yitzhaki for useful comments.

¹ Prior to June 1999, these indexes were computed and reported by Hang Seng Bank, a private entity. Methodological information on the elaboration of CPIs (and other economic data) can be obtained from the IMF Data Dissemination Standard website, <http://dsbb.imf.org/>.

² J.L. Nicholson derived similar results about the same time, which were later published as Nicholson (1975). See also Pollak (1980).

plutocratic and democratic budget shares, and relates these new results with the approximation in Prais (1958). Section 3 derives a characterization of the CPI plutocratic bias, interprets the empirical evidence under this decomposition, and discusses issues related to group indexes. Section 4 presents an alternative approach based on weighing each household proportionately to the number of its members, and Section 5 concludes.

2. Plutocratic and Democratic CPI budget shares

Define the *plutocratic* budget shares for good i in the aggregate CPI by

$$\tilde{s}_i^P = \frac{1}{X} \sum_h x^h s_i^h, \quad (1)$$

where x^h denotes household h total expenditures, x_i^h is the expenditure on good i , so that household h budget share for good i is given by $s_i^h = x_i^h/x^h$. Total aggregate expenditure is given by $X = \sum x^h$.³ (See Table 1 for a summary of the notation used.) The CPI (at time t) is given by⁴

$$CPI^P = \sum_i \tilde{s}_i^P \mathcal{I}_i, \quad (2)$$

where $\mathcal{I}_i = (p_{it}/p_{i0})$ are elementary price indexes. Noting that household h individual index is given by $cp_i^h = \sum_i s_i^h \mathcal{I}_i$, the CPI^P in (2) may be interpreted as a ‘representative’ CPI.⁵ It is natural to ask then what is the household better represented by the CPI^P . Muellbauer (1974) searched for the household whose budget shares were closest to the \tilde{s}_i^P aggregate weights in the U.K. CPI, and found it to be at the 71 percentile in the household expenditures distribution. For the U.S. in 1990, Deaton (1998) estimates that this consumer occupies the 75 percentile. Thus, the ‘representative’ consumer embedded in (2) is biased towards upper-income households.

Alternatively, we could use *democratic* budget shares,

$$\tilde{s}_i^D = \frac{1}{H} \sum_h s_i^h, \quad (3)$$

³ All expenditure magnitudes shall refer to the household-survey performed at the base period, 0. We shall assume that the statistical agency observes prices for all goods at 0. See Ruiz-Castillo *et al.* (2000) for issues regarding modified Laspeyres indexes when prices become available only some time after the household survey, as it is the case in most practical situations. See BLS (1997), chapter 17, for a good description about the computation of the U.S. CPI.

⁴ Prices are typically sampled over J geographical areas, obtaining the elementary price indexes, $\mathcal{I}_{ij} = (p_{ijt}/p_{ij0})$. The official CPI is given by $CPI^P = \sum_j \sum_i S_{ij} \mathcal{I}_{ij}$, where the aggregate shares are computed as $S_{ij} = (1/X) \sum_{h \in j} x_i^h$. Noting that $x_i^h = x^h s_i^h$, we have that $CPI^P = \sum_h \sum_i (x^h/X) s_i^h \mathcal{I}_{ih}$, where \mathcal{I}_{ih} denotes \mathcal{I}_{ij} for the region j where household h is located.

⁵ It can be established that $CPI^P = \sum p_{it} \bar{q}_i / \sum p_{i0} \bar{q}_i$, where \bar{q}_i is the average consumption of good i . Consequently, at least in this sense, the CPI^P is indeed the CPI of an average consumer. Similarly, by simply multiplying the average quantities by H , we obtain that the CPI^P is also the CPI of an aggregate consumer.

Table 1. Notation

$i = 1, \dots, N$	Good subscript
$h = 1, \dots, H$	Household superscript
$j = 1, \dots, J$	Geographical area subscript
x_i^h	Expenditure on good i by household h at 0
n^h	Number of people in household h
$m^h = (n^h)^\theta$	Equivalent number of adults in household h
$x^h = \sum_i x_i^h$	Total expenditure by household h at 0
$\mathbf{x} = (x^1, \dots, x^H)$	Distribution of household expenditures, x^h
$X = \sum_h x^h$	Total aggregate expenditure in all goods at 0
$s_i^h = x_i^h/x^h$	Household h 's budget share for good i at 0
p_{it}	Price of good i at t
$q_i^h = x_i^h/p_{it}$	Quantity of good i purchased by household h at 0
p_{ijt}	Price of good i in geographical area j at t
$\mathcal{I}_i = p_{it}/p_{i0}$	Good i elementary price index between t and 0
$\mathcal{I}_{ij} = p_{ijt}/p_{ij0}$	Good i elementary price index between t and 0 in area j

where H denotes the number of households, to construct a democratically-weighted index,

$$CPI^D = \sum_i \tilde{s}_i^D \mathcal{I}_i. \quad (4)$$

(Other possibility, explored later, consists in weighing each household proportionately to the number of its members using an equivalence-scale approach.)

From equations (1) and (3), the difference between good i plutocratic and democratic shares in the CPI is given by

$$(\tilde{s}_i^P - \tilde{s}_i^D) = \frac{1}{\bar{x}H} \sum_h (x^h - \bar{x})s_i^h = \frac{1}{\bar{x}} \hat{\sigma}(x, s_i), \quad (5)$$

where $\bar{x} = X/H$ is the sample mean of total expenditures, and $\hat{\sigma}(x, s_i)$ is the sample covariance, across households, of the budget share of good i , s_i^h , and total expenditure. We can now rescale the covariance term in (5) and convert it into a regression coefficient by simply dividing it by the variance of any of the two variables involved (which would then become the independent variable). Thus, multiplying and dividing the right-hand side of expression (5) by the sample variance of household total expenditures, $\hat{\sigma}^2 = \hat{\sigma}(x, x)$, we obtain:

$$(\tilde{s}_i^P - \tilde{s}_i^D) = \bar{x} \hat{\zeta} \hat{\beta}_i, \quad (6)$$

where $\hat{\zeta}$ is the square of the sample coefficient of variation of household total expenditures, $\hat{\zeta} = (\hat{\sigma}/\bar{x})^2$, and $\hat{\beta}_i$ denotes the OLS estimate in the regression given by

$$(s_i^h - \tilde{s}_i^D) = \beta_i(x^h - \bar{x}) + \varepsilon_i^h. \quad (7)$$

Equation (6) indicates that the difference in good i 's plutocratic and democratic CPI shares depends on the product of: (i) average income, \bar{x} ; (ii) a measure of inequality of

household expenditure, $\hat{\zeta}$; and (iii) a measure of how good i 's budget share varies with total expenditure in the household sample, $\hat{\beta}_i$. Since the decomposition is multiplicative, the shares must coincide when there is no income inequality or when expenditure shares are not affected by income differences.

It is important to note that no distributional or behavioral assumptions are needed to obtain $\hat{\beta}_i$, because we can always estimate the regression coefficient in equation (7). Of course, if assumptions are made, then different interpretations could be given to the parameters involved. For now, however, we simply want to stress that the decomposition in (6) holds because of algebraic identities, and does not rely on any assumptions on consumer behavior or household-expenditures distribution.

Note that $\hat{\zeta} = 2I_2(\mathbf{x})$, where $I_2(\mathbf{x})$ corresponds to the Generalized Entropy inequality measure, $I_c(\mathbf{x})$, for $c = 2$. The parameter c summarizes the sensitivity of \mathcal{I}_c in different parts of the household total expenditures distribution: the more positive (negative) c is, the more sensitive \mathcal{I}_c is to differences at the top (bottom) of the distribution (Cowell and Kuga, 1981). The Generalized Entropy family of inequality indexes are the only measures of relative inequality that satisfy the usual normative properties required for an inequality index and, in addition, are decomposable by inequality subgroups (Shorrocks, 1984). Finally, using the fact that $\sum_i s_i^h = 1$, it follows from equation (6) that $\sum \hat{\beta}_i = 0$ (so that if, for some i , we have $\hat{\beta}_i > 0$ we must also have $\hat{\beta}_j < 0$ for some j).

2.1. Income Elasticity

For any observed variable, y^h , we can compute its distance from the population average, $\Delta y^h = (y^h - \bar{y})$, where $\bar{y} = \frac{1}{H} \sum y^h$ is the sample mean of y^h . Noting that all households (in the same geographical area) face the same prices, we can compute a sample analogue of the elasticity of the budget share of good i with respect to income, by computing the ratio of percent deviations across households from average quantities:⁶

$$\frac{\Delta s_i^h \bar{x}}{\Delta x^h \bar{s}_i} = (\eta_i^h - 1), \quad (8)$$

where:

$$\eta_i^h = \frac{\Delta q_i^h \bar{x}}{\Delta x^h \bar{q}_i}, \quad (9)$$

is a sample analogue of good i 's income elasticity —showing the percent deviation of household h 's consumption of good i from the average consumption, Δq_i^h , divided by household h 's percent deviation from average income (expenditure), Δx^h .

As discussed in Olkin and Yitzhaki (1992), we can rewrite the standard OLS estimator in (7) as a weighed average of slopes:

$$\hat{\beta}_i = \sum_h \omega^h \frac{\Delta s_i^h}{\Delta x^h} = \frac{\bar{s}_i}{\bar{x}} \sum_h \omega^h (\eta_i^h - 1), \quad (10)$$

⁶ The differential counterpart can be easily obtained by noting that $\log(s_i^h) = \log(p_i) + \log(q_i^h) - \log(x^h)$ and differentiating with respect to $\log(x^h)$ in order to obtain the i budget share income elasticity.

where the normalized weights are given by⁷ $\omega^h = (\Delta x^h)^2 / H\hat{\sigma}^2$. Equation (10) suggests the following estimate of the income elasticity along the income distribution:

$$\hat{\eta}_i = \sum_h \omega^h \eta_i^h = \frac{\bar{x}}{\bar{q}_i} \sum_h \omega^h \frac{\Delta q_i^h}{\Delta x^h} = \frac{\bar{x}}{\bar{q}_i} \hat{\theta}_i, \quad (11)$$

where $\hat{\theta}_i$ is the OLS coefficient of the regression of Δq_i^h on Δx^h . Equation (11) allows us to rewrite equation (10) as:

$$\hat{\beta}_i = \frac{\bar{s}_i}{\bar{x}} (\hat{\eta}_i - 1), \quad (12)$$

which implies that:

$$\hat{\beta}_i > 0 \iff \hat{\eta}_i > 1. \quad (13)$$

Thus, the sign of $\hat{\beta}_i$ is determined by the empirical estimate of good i 's income elasticity $\hat{\eta}_i$. Superior goods, displaying an income elasticity larger than 1, will be associated with positive coefficients. Necessities, on the other hand, will have negative regression coefficients. Using (12), we can now express equation (6) as:

$$(\tilde{s}_i^P - \tilde{s}_i^D) = \bar{s}_i \hat{\zeta} (\hat{\eta}_i - 1). \quad (14)$$

If $\eta_i^h = \eta_i$ for all households—which is guaranteed by constant-elasticity Engel curves, as assumed in Prais (1958)—then, using (14) we find that the percentage difference between plutocratic and democratic shares is given by:

$$\frac{\tilde{s}_i^P - \tilde{s}_i^D}{\tilde{s}_i^D} = \hat{\zeta} (\eta_i - 1) = \frac{\hat{\sigma}^2}{\bar{x}^2} (\eta_i - 1), \quad (15)$$

which is related to the expression obtained by Prais (1958) (his equation (8) on pages 127 and 134):

$$\frac{s_i^P - s_i^D}{s_i^P} \approx \sigma^2 (\eta_i - 1). \quad (16)$$

Note, however, that while (16) is an approximation, equation (15) is exact. The approximation in (16) requires assuming that x^h is lognormally distributed (in which case: $\sigma^2 = \log(1 + \zeta)$, where the symbols without hats refer to *population* parameters), and relies on the fact that if ϵ is small, then $\log(1 + \epsilon) \approx \epsilon$ —see the Appendix in Prais (1958) for details.⁸ Our analysis shows that no distributional assumption on household expenditures is necessary to obtain (15).

⁷ Note that this weighing scheme is plutocratic squared! See Yitzhaki (1996) for a discussion of issues involving regressions on income in the context of welfare economics.

⁸ Under this assumptions, we have that $\sigma^2 = \log(1 + \zeta) \approx \zeta$, and using (15) we can write: $((s_i^P - s_i^D)/s_i^D) \approx \hat{\sigma}^2 (\eta_i - 1)$.

3. The CPI Plutocratic Bias

As discussed before, we shall define the plutocratic bias, \mathcal{B} , as:

$$\mathcal{B} \equiv \frac{\Pi^P - \Pi^D}{100} = (CPI^P - CPI^D), \quad (17)$$

where $\Pi = (CPI - 1) \times 100$ is the inflation rate between 0 and t (in percent); using (5) we find:

$$\mathcal{B} = \sum_i (\tilde{s}_i^P - \tilde{s}_i^D) \mathcal{I}_i = \bar{x} \hat{\zeta} \sum_i \hat{\beta}_i \mathcal{I}_i = \bar{x} \hat{\zeta} \sum_i \hat{\beta}_i (\mathcal{I}_i - \bar{\mathcal{I}}), \quad (18)$$

where $\bar{\mathcal{I}}$ is a simple average, *i.e.*, $\bar{\mathcal{I}} = \frac{1}{N} \sum \mathcal{I}_i$, which represents average inflation. Equation (18) may be rewritten as:

$$\mathcal{B} = \bar{x} \hat{\zeta} N \hat{\sigma}(\hat{\beta}, \mathcal{I}), \quad (19)$$

where $\hat{\sigma}(\hat{\beta}, \mathcal{I})$ refers to the sample covariance of $\hat{\beta}_i$ and \mathcal{I}_i , this time over goods instead that over households.

Equation (19) is our fundamental result. It shows that the plutocratic bias is determined by average income (expenditure), the dispersion of household expenditure, measured by $\hat{\zeta}$, and the sample covariance between $\hat{\beta}_i$ and \mathcal{I}_i . The sign of the plutocratic bias is determined by the covariance term. A positive covariance term means that the goods favored by the richer households experience higher than average inflation and necessities a lower than average inflation. Similarly, a negative covariance implies that necessities experience higher than average inflation while superior or luxury goods experience lower than average inflation. These effects are also scaled by the magnitude of the inequality of household expenditures, as measured by $\hat{\zeta}$, as well as average income (expenditure). *Ceteris paribus*, the higher the dispersion in household expenditures, the higher the size of the plutocratic bias. Similarly for average income, the higher the average income, the higher the bias.

Inspection of equation (19) indicates that three elements are required for the plutocratic bias to be different from zero: (a) there must be some dispersion in the distribution of household expenditures (reflected by $\hat{\zeta} \neq 0$); (b) there must be some observed behavioral differences among households with different total expenditures (reflected by $\hat{\beta}_i \neq 0$ for some i); and (c) there must be some differences in price behavior across some goods which display behavioral differences across households (reflected by $\mathcal{I}_i \neq \bar{\mathcal{I}}$ for some i which has $\hat{\beta}_i \neq 0$).

3.1. Geographical Price Variation

In practical situations, however, not all households face the same prices. For each good i , the statistical agency collects j prices, p_{ijt} , one for each geographical area j . This can be readily accommodated within the present analysis in, at least, two ways.

One approach is to think of item i in different areas as different goods. We would expand the good space to include $N \times J$ goods. As a result, s_i^h will be always zero for the ‘goods’

outside the geographical area where household h resides. All the previous analysis applies without further changes.

An alternative approach, closer perhaps to the actual computation of the CPI, would be to simply apply (18) within a geographical area j where all the households face the same prices p_{ijt} :

$$\mathcal{B}_j = \bar{x}_j \hat{\zeta}_j N \hat{\sigma}(\hat{\beta}_{ij}, \mathcal{I}_{ij}), \quad (20)$$

where $\hat{\sigma}(\hat{\beta}_{ij}, \mathcal{I}_{ij})$ now refers to the covariance between $\hat{\beta}_{ij}$, the OLS estimate from the regression (7) in area j , and \mathcal{I}_{ij} are the set of i prices in area j , \mathcal{I}_{ij} . Then the overall bias becomes:

$$\mathcal{B} = \sum_j \mathcal{B}_j = N \sum_j \bar{x}_j \hat{\zeta}_j \hat{\sigma}(\hat{\beta}_{ij}, \mathcal{I}_{ij}). \quad (21)$$

3.2. Empirical Estimates of the CPI Bias

Table 2 summarizes the main findings of various empirical studies that have estimated the plutocratic bias for different countries during various time periods. As we shall discuss below, the bias for shorter periods is typically larger, so that when averaging over long periods positive and negative biases may tend to cancel off, resulting in an apparently smaller overall bias. In addition, the size of the bias is likely to be underestimated because none of the studies uses the most disaggregated data, known only to the statistical offices.

Given a household survey, $\hat{\zeta}$ and the $\hat{\beta}_i$'s are then fixed, and any source of variation in the sign and size of the bias for, *e.g.*, each year must be solely explained by the price behavior reflected by the \mathcal{I}_i 's. The movements in the \mathcal{I}_i 's may cause $\hat{\sigma}(\hat{\beta}, \mathcal{I})$ to change sign from one year to another (Table 3 below). Thus, as noted, looking at the overall \mathcal{B} , simply averaging over a long period may be misleading.

Because of data limitations, most of the results in Table 2 are based on a smaller number of goods than the number for which prices were collected by the statistical agencies. For instance, Yahav and Yitzhaki (1991) are based on 10 or 28 aggregate goods, depending on the period. Similarly, Ruiz-Castillo *et al.* (1999), uses 2,042 goods for the 1990s (see below), but only 58 and 57 for the 1980s and 1970s.⁹ In these instances, working with highly aggregated goods causes an underestimation of the true plutocratic bias for two reasons. First, price aggregates already embody a plutocratic bias. Second, income elasticities revert to the mean (*i.e.*, to one) as we aggregate goods. As a result, the true size of the plutocratic bias is underestimated.¹⁰

⁹ In addition, some studies do not have information on geographical price variation. For example, Kokosky (2000) assumes that the same U.S. national average CPI prices apply to all households in the sample, and focuses on the effect of expenditure shares variability across households.

¹⁰ Algebraically, in equation (18), the first effect implies that the \mathcal{I}_i 's are artificially close to CPI^P . The second effect implies that the $\hat{\beta}_i$'s are artificially close to zero. The end result is that the size of $\sigma(\hat{\beta}, \mathcal{I})$ shrinks towards zero, producing an underestimation of the true plutocratic bias, \mathcal{B} .

Table 2. Empirical Studies of the CPI Plutocratic Bias

	Country	Time Period	Range (percentage points per year)	
			Π^P	\mathcal{B}
Carruthers <i>et al.</i> (1980)	U.K.	1975–78	8.2 to 24.2	–0.1
Fry and Pashardes (1985)	<i>id.</i>	1974–82	8.2 to 24.2	negative
Deaton and Muellbauer (1980)	<i>id.</i>	1975/76	16	–2.0
Crawford (1996)	<i>id.</i>	1979–92	3.4 to 18.0	+0.16
Newberry (1994)	<i>id.</i>	1980s	3.4 to 18.0	negligible
<i>ibid.</i>	Hungary	1980s	4.5 to 16.9	negligible
Kokoski (1987)	U.S.	1972–80	3.3 to 13.5	–0.1 to –0.3
Erbas and Sayers (1998)	<i>id.</i>	1986–95	1.9 to 5.4	negative
Garner <i>et al.</i> (1999)	<i>id.</i>	1980s	1.9 to 13.5	slightly negative
Kokoski (2000) ^a	<i>id.</i>	1987–97	2.0 to 5.25	–0.28 to +0.56
Lódola <i>et al.</i> (2000)	Argentina	1989–91	220 to 10,781	+2.3 to +663.4
<i>ibid.</i>	<i>id.</i>	1991–93	11.2 to 20.0	–0.66 to –0.78
<i>ibid.</i>	<i>id.</i>	1993–98	1.2 to 3.3	–0.48 to +0.65
Yahav and Yitzhaki (1991)	Israel	1960–71	1.99 to 12.06	–0.12 to +0.25
<i>ibid.</i>	<i>id.</i>	1981–86	19.9 to 373.8	–1.7 to +6.3
Ruiz-Castillo <i>et al.</i> (1999)	Spain	1973–81	14.54 to 23.02	–0.04 to +0.53
<i>ibid.</i>	<i>id.</i>	1981–91	4.59 to 9.48	–0.19 to +0.30
<i>ibid.</i>	<i>id.</i>	1991–98	2.49 to 6.99	–0.08 to +0.15

Source: Studies cited, IMF Government Financial Statistics and author’s calculations.

^a This paper has a typo in its Table 2: the column headings “Democratic” and “Plutocratic” should be switched.

For Spain during the 1990s, Ruiz-Castillo *et al.* (1999) estimate that the average plutocratic bias in Spain amounts to 0.055 percent per year.¹¹ However, as shown in Table 3, annual biases are typically larger, and price movements significantly change the sign and magnitude of the annual bias. The results in Table 3 are based on $21 \times 18 + 32 \times 52 = 2,042$ different \mathcal{I}_{ij} ’s; 21 food goods in 18 autonomous communities and 32 non-food goods in 52 provinces.¹²

Thus, as discussed above, the sign and magnitude of the bias may vary significantly year after year, even when using the same budget survey. As a result, finding the bias

¹¹ The results in Table 3 are based on the Spanish household budget survey collected by the Spanish statistical agency 1990–91. This is a household budget survey of 21,155 household sample points, representative of a population of approximately 11 million households and 38 million persons occupying residential housing in all of Spain. The survey was collected from April 1990 to March 1991.

¹² The statistical agency collects elementary price indexes for a commodity basket consisting of 471 items in 130 municipalities spread over the 52 Spanish provinces under the CPI present system, based in 1992. Approximately 150,000 prices are collected each month from approximately 29,000 establishments. Price information at this disaggregated level is not publicly available. Prices are generally collected once a month at each establishment, except for perishable items which are collected three times a month from all establishments.

Table 3. Decomposition of the CPI Plutocratic Bias: Spain 1993–97
(All values in percentage points)

$N = 2,042, \bar{x} = 2.56\text{E}+6$					
	Π^P	Π^D	$\hat{\zeta}$	$\bar{x}N\hat{\sigma}(\hat{\beta}, \mathcal{I})$	\mathcal{B}
1993	5.271	5.165	4.96E-01	0.2115	0.105
1994	4.621	4.701	4.96E-01	-0.1612	-0.080
1995	4.079	4.130	4.96E-01	-0.1007	-0.050
1996	3.180	3.090	4.96E-01	0.1813	0.090
1997	2.494	2.369	4.96E-01	0.2518	0.125

small during one particular period has little bearing over its size and sign at other time when prices may behave differently. For different household surveys, not only the price dynamics may change, but also income inequality may be different (*e.g.*, $\hat{\zeta}$ was 2% larger for Spain in 1980–81). As a result, findings for one country may have little implications for other countries with larger income inequality and different price dynamics. For instance, income inequality in Latin America is very large, IDB (1998) reports that countries in the region experience the largest income inequality in the World. It is very likely then that the CPI plutocratic bias be of a larger significance in Latin America, especially in countries with double-digit inflation that may have more differentiated price dynamics.¹³

3.3. Group Indexes

The approach presented here may also be used for ‘group’ indexes —*i.e.*, indexes referred to a group or population of households (Pollak, 1980). Equation (6) implies that the plutocratic shares for the whole population are given by $\tilde{s}_i^P = \bar{s}_i + \bar{x} \hat{\zeta} \hat{\beta}_i$, while plutocratic shares for a subgroup G

$$\tilde{s}_i^G = \bar{s}_i^G + \bar{x}^G \hat{\zeta}^G \hat{\beta}_i^G, \quad (22)$$

where the G superscript means that the quantities have been computed for $h \in G$. Then, we have that:

$$CPI^P - CPI^G = \sum_i \left((\bar{s}_i - \bar{s}_i^G) + (\bar{x} \hat{\zeta} \hat{\beta}_i - \bar{x}^G \hat{\zeta}^G \hat{\beta}_i^G) \right) \mathcal{I}_i. \quad (23)$$

The term $(\bar{s}_i - \bar{s}_i^G)$ picks up differences in average spending patterns. The term $(\bar{x} \hat{\zeta} \hat{\beta}_i - \bar{x}^G \hat{\zeta}^G \hat{\beta}_i^G)$ reflects differences in income inequality and behavioral responses to income. If

¹³ Interestingly, the CPI weights in Ecuador are computed excluding the richer households because of the ‘large dispersion in their consumption patterns’ —see <http://www.inec.gov.ec/pry15.htm>. (In the U.K., the index weight calculation excludes the top 4% of the population by income and also pensioners mainly dependent on state benefits.) Nonetheless, in Ecuador, as is typically done in many countries (including the U.S.), the household survey on which the CPI is based is restricted to urban areas.

the group is very homogeneous in income, so that $\hat{\zeta}^G \approx 0$, then:

$$CPI^P - CPI^G = \sum_i \left((\bar{s}_i - \bar{s}_i^G) + \bar{x} \hat{\zeta} \hat{\beta}_i \right) \mathcal{I}_i = \mathcal{B} + \sum_i (\bar{s}_i - \bar{s}_i^G) \mathcal{I}_i. \quad (24)$$

Thus, the difference between the overall CPI and the group index would depend on the difference of average spending patterns and the plutocratic bias term.

There are two important issues regarding group indexes that fall outside the framework developed in this paper which implicitly assumes that goods are well defined. In reality the prices for all existing commodities and services cannot be sampled, and statistical agencies must choose specific items to represent categories of goods. Thus, a fundamental issue pertains to the selection of items that represent particular goods. Take, for instance, the ‘alcoholic beverages’ expenditure category; the statistical agency may chose to follow the price of domestic beer, imported beer, or, say, champagne (Pollak, 1998). Each choice may be more relevant for a particular population group, so the decision of which goods to follow has implications for whose inflation is being measured. A related issue concerns the location where the prices are sampled; as long as distinct population groups may shop at different places, again the selection of outlets may implicitly shape whose inflation is being targeted.

While the Hong Kong indexes mentioned in the introduction are obtained by simple reweighing,¹⁴ there are instances where statistical agencies make an effort to adequately address these two issues. The Indian Ministry of Statistics and Programme Implementation computes four different CPIs: CPI-IW, CPI-RL, CPI-AL, and CPI-UNME. The CPI-IW covers households headed by industrial workers in 70 industrial centers following 260 items and sampling approximately 160,000 retail price quotes from 16,545 outlets and selected open markets. The CPI-RL covers all rural households and it is compiled for 20 states and for all India, covering 85 to 106 items; 61,005 monthly price quotes are collected from retail outlets in 600 villages. The CPI-AL, similar to the RL, except that it covers only households headed by agricultural laborers. Finally, the CPI-UNME covers households headed by non-manual workers in 59 urban centers sampling 1022 price quotes on a varying number of items depending on the center, from 146 in Imphal to 345 in Delhi. Inflation rates may differ significantly for these indexes. For instance, in January 2001, the annual inflation rates were 3.3, -2.0, -1.6, and 5.8 percent, respectively. (See Deaton and Tarozzi (2000) for an independent set of calculations of Indian price indexes using data for two year-long periods, 1987–88, and 1993–94.)

4. An Equivalence-Scale Approach

Instead of weighing equally each household, an alternative approach is to consider explicitly the number of members in each household using an equivalence-scale approach. Equivalence scales are used in empirical studies of consumption behavior to take into account economies of scale in household composition. Following Buhmann *et al.* (1988) we

¹⁴ France also publishes a CPI for blue-collar employees which is computed by reweighing.

adopt an equivalence scale model in which scale economies in consumption depend only on household size. Let n^h be the number of members of household h ; then $m(\theta)^h = (n^h)^\theta$, with $\theta \in [0, 1]$, can be interpreted as the number of ‘equivalent adults’ in a household of size n^h . Making $\theta = 0$ we have $m^h = 1$, which would weigh each household equally as in our democratic index. At the other extreme, with $\theta = 1$, m^h would simply represent the number of members for a super-democratic index.

The budget shares, now a function of θ , would be given by

$$\tilde{s}_i^E(\theta) = \frac{1}{H\bar{m}(\theta)} \sum_h m(\theta)^h s_i^h, \quad (25)$$

which implies

$$(\tilde{s}_i^E(\theta) - \tilde{s}_i^D) = \frac{\hat{\sigma}(m(\theta), s_i)}{\bar{m}(\theta)} = \bar{m}(\theta) \hat{\xi}(\theta) \hat{\delta}_i(\theta), \quad (26)$$

where $\hat{\xi}$ is the inequality index applied to $m(\theta)^h$, and $\hat{\delta}_i(\theta)$ is the OLS estimate of the regression $(\tilde{s}_i^E(\theta) - \tilde{s}_i) = \delta_i(\theta)(m(\theta)^h - \bar{m}(\theta)) + u_i$. The inequality of the number of household members is typically smaller than expenditure inequality. For the Spanish 1990–91 household survey we have that $\hat{\zeta} = 0.496$, and $\hat{\xi} = 0.213$.¹⁵ For the previous survey, 1980–81, we have $\hat{\zeta} = 0.506$, and $\hat{\xi} = 0.225$, which are 2 percent and 5 percent higher, respectively.

The difference between the equivalence-scale index and the democratic index is given by:

$$CPI^E(\theta) - CPI^D = N \bar{m}(\theta) \hat{\xi}(\theta) \hat{\sigma}(\hat{\delta}(\theta), \mathcal{I}). \quad (27)$$

As noted before, this difference is zero for $\theta = 0$ because, by definition, $CPI^E(0) = CPI^D$.

Define $\mathcal{B}(\theta) = (CPI^P - CPI^E(\theta))$, using equations (19) and (27), we have that:

$$\begin{aligned} \mathcal{B}(\theta) &= N \{ \bar{x} \hat{\zeta} \hat{\sigma}(\hat{\beta}, \mathcal{I}) - \bar{m}(\theta) \hat{\xi}(\theta) \hat{\sigma}(\hat{\delta}(\theta), \mathcal{I}) \} \\ &= \sum_i \left\{ \frac{\hat{\sigma}(x, s_i)}{\bar{x}} - \frac{\hat{\sigma}(m(\theta), s_i)}{\bar{m}(\theta)} \right\} \mathcal{I}_i. \end{aligned} \quad (28)$$

Note that equation (28) is a generalization of (19), since $\mathcal{B}(0) = \mathcal{B}$.¹⁶

Table 4 shows $\mathcal{B}(\theta)$ for different values of θ , for Spain in the 1990s. The magnitude of the bias decreases with θ . Moreover, there is even a sign reversal, for 1995, with $\mathcal{B}(1) > 0$ while for all other values of θ the bias, $\mathcal{B}(\theta)$, is negative. The results in Table 4 are a consequence of the fact that, as it is to be expected, for Spain, in that time period, household size and total expenditure are correlated. Nonetheless, even for $\theta = 1$, the size of the biases in Table 4 is not negligible.

¹⁵ $\bar{x} = 2,563,502$ pesetas, and $\hat{\sigma} = 1,806,178$; for number of household members, $\bar{n} = 3.4070$ and $\hat{\sigma}_n = 1.5719$.

¹⁶ Muellabuer (1974) proposes a class of *homogeneous social price indexes* parameterized by a measure of aversion to inequality, $\delta \in [0, 1]$. The shares are defined by $s_i(\delta) \propto \sum (x^h/X)^{1-\delta} s_i^h$, which reduces to democratic shares when $\delta = 1$, and plutocratic shares when $\delta = 0$. The indexes using these shares range from the standard Laspeyres to the democratic Laspeyres, depending on δ .

Table 4. CPI Plutocratic Bias for different values of θ : Spain 1993–97
(All values in percentage points)

	Π^P	$\mathcal{B}(0)$	$\mathcal{B}(0.25)$	$\mathcal{B}(0.50)$	$\mathcal{B}(0.75)$	$\mathcal{B}(1)$
1993	5.271	0.105	0.098	0.090	0.083	0.077
1994	4.621	-0.080	-0.065	-0.052	-0.042	-0.034
1995	4.079	-0.050	-0.033	-0.018	-0.004	0.009
1996	3.180	0.090	0.088	0.087	0.086	0.085
1997	2.494	0.125	0.121	0.117	0.113	0.110

5. Concluding Remarks

What is the appropriate inflation gauge from a macroeconomic perspective? How should we adjust, *e.g.*, public pensions, or social programs transfers annually?¹⁷ BLS (1997, p. 172) warns “CPI users should understand that the CPI may not be applicable to all questions about price movements for all population groups.” Nevertheless, in most places and in most times, these quantities are invariably revised according to a plutocratic CPI. Thus, a dollar-weight logic prevails over a household-weight logic.

Indexing by the current CPI may result in over- or under-compensation relative to a democratic index during different time periods. These deviations may tend to cancel off over longer periods. There is, however, an important perversity emphasized by Ruiz-Castillo *et al.* (1999). While inflation itself affects household welfare, the plutocratic bias in the CPI often accentuates the welfare change rather than smoothing it. Thus, the worse-off households are hurt by the CPI adjustments during periods when inflation is more harmful to them —*i.e.*, when they can least afford it. In periods where the plutocratic bias is negative, when prices behave in an anti-poor way, then social programs, which primarily benefit the poor, are revised less than what would be the case with a democratic group index. Similarly, when price movements are less detrimental to the poorer households —*i.e.*, when the plutocratic bias is positive— indexed social transfers grow more than cost-of-living adjustments would dictate. Thus, the plutocratic CPI adjustments involving the poorer households display harmful ‘procyclical’ features.¹⁸

Nonetheless, the plutocratic CPI has its own merits. It naturally arises when computing the aggregate Laspeyres price index, and it is consistent with macroeconomic-aggregate deflators arising from the national accounts. It also provides an upper bound for the

¹⁷ See Triplett (1983), Fry and Pashardes (1985), Griliches (1995), and Pollak (1998).

¹⁸ For the richer households, however, since the plutocratic index is always closer to their true index than a democratic one, this problem is unlikely. In particular, possibly the most important CPI adjustment involving the richer households involves the revision of income-tax brackets. In this case, when inflation is more detrimental to the richer households, the positive plutocratic bias ensures that the brackets are appropriately revised upwards so that tax payments are lower —*i.e.*, there is a countercyclical effect that smoothes changes in welfare.

theoretical aggregate compensating variation (Hicks, 1940) —*i.e.*, by how much would monetary national income need increase to compensate for a price variation.

While different indexes could be easily computed for different uses, Prais (1958, p. 131) asked: Can more than one index numbers be tolerated without confusion?¹⁹ There is a crucial tradeoff between the simplicity of the current prevailing one-size-fits-all approach and the conceptual superiority of a piecemeal menu approach to index numbers. The best resolution may well vary in different places and at different times. This paper shows that the larger the income inequality, the more different the consumption patterns by income group, and the larger the variance in individual price behavior, the less appealing is a single plutocratic CPI as the only policy adjuster. Finally, if a single index number is to be computed, then as Prais (1958, p. 126) asked: Whose cost of living should one have in mind?

References

- BLS (1997), *Handbook of Methods*, Washington DC: U.S. Department of Labor.
- Buhmann, B., L. Rainwater, G. Schmauss and T. Smeeding (1988), “Equivalence Scales, Well-Being, Inequality and Poverty: Sensitivity Estimates Across Ten Countries Using the Luxembourg Income Study Database,” *Review of Income and Wealth*, **34**: 115–142.
- Carruthers, A., D. Sellwood and P. Ward (1980), “Recent Developments in the Retail Price Index,” *The Statistician*, 29: 1–32.
- Crawford, Ian (1996), “UK Household Cost-of-Living indexes, 1979-92,” in J. Hills (ed), *New Inequalities: the Changing Distribution of Income and Wealth in the United Kingdom*, Cambridge: Cambridge University Press.
- Cowell, F.A. and K. Kuga (1981), “Additivity and the Entropy Concept: An Axiomatic Approach to Inequality Measurement,” *Journal of Economic Theory*, 25, 131–143.
- Deaton, A. (1998), “Getting Prices Right: What Should Be Done?,” *Journal of Economic Perspectives*, 12: 37-46.
- , and J. Muellbauer (1980), *Economics and Consumer Behavior*, New York: Cambridge University Press.
- , and Alessandro Tarozzi (2000), “Prices and poverty in India,” mimeo, Research Program in Development Studies, Princeton University.
- Erbas, N. and C.L. Sayers (1998), “Is the United States CPI biased across income and age groups?,” Washington DC: IMF Working Paper 98-136. <http://www.imf.org>
- Fry, V. And P. Pashardes (1985), *The RPI and the Cost of Living*, Report Series No. 22, London: Institute for Fiscal Studies.
- Garner, T, J. Ruiz-Castillo and M. Sastre (1999), “The Influence of Demographics and Household Specific Price indexes on Expenditure Based inequality and Welfare: A

¹⁹ Referring to the Indian experience, where the government normally uses the WPI, wholesale price index, as a gauge of inflation rather than any of the four indexes mentioned above, Rajani X. Desai remarks: “Different measures of inflation ought to be used depending on one’s purpose. Actually, the Government deliberately creates some confusion on this score.” *Research Unit for Political Economy*, 26/7, November 1998. Available from the Maoist documentation project website: <http://www.maoism.org/misc/india/rupe>.

- Comparison of Spain and the United States,” Working Paper 9963 Economic Series 25, Universidad Carlos III, Madrid. <http://eco.uc3m.es>
- Griliches, Z. (1995), “Prepared Statement for the U.S. Senate Hearings on the Consumer Price Index,” Senate Hearing 104-69, Washington D.C.: U.S. Printing Office.
- Hicks, J. (1940), “The Valuation of Social Income,” *Econometrica*, 7, 108–124.
- IDB (1998), *Facing Up to Inequality in Latin America. Economic and Social Progress in Latin America, 1998–99 Report*; Washington DC: Inter-American Development Bank.
- Kokoski, M.F. (1987), “Consumer Price indexes by Demographic Group,” BLS Working Paper 167 (April).
- , (2000), “Alternative CPI Aggregations: two approaches,” *BLS Monthly Labor Review*, vol. 123, 31–39.
- Lódola, A., M. Busso and F. Cerimedo (2000), “Sesgos en el índice de precios al consumidor: el sesgo plutocrático en Argentina,” working paper, Universidad Nacional de La Plata.
- Muellbauer, J. (1974), “The Political Economy of Price indices,” Birbeck Discussion Paper no 22.
- Newberry, David M. (1994), “The Distributional Impact of Price Changes in Hungary and the UK,” University of Cambridge Discussion Papers on Economic Transition: DPET 9402.
- Nicholson, J.L. (1975), “Whose Cost of Living?,” *Journal of the Royal Statistical Society, Series A*, 138:4, 540–2.
- Olkin, I. and S. Yitzhaki (1992), “Gini Regression Analysis,” *International Statistical Review*, 66(2), 185–196.
- Pollak, R. (1980), “Group Cost-of-Living Indexes,” *American Economic Review*, 70:2, 273–78.
- , R. (1998), “The Consumer Price Index: A Research Agenda and Three Proposals,” *Journal of Economic Perspectives*, 12: 69–78.
- Prais, S. (1958), “Whose Cost of Living?,” *Review of Economic Studies*, 26: 126–134.
- Ruiz-Castillo, J., E. Ley and M. Izquierdo (1999), “The Plutocratic Bias in the CPI: Evidence from Spain”, Madrid: FEDEA Working Paper 99/15. <http://www.fedea.es>
- , (2000), “The Laspeyres Bias in the Spanish CPI”, Madrid: FEDEA Working Paper 2000–05. <http://www.fedea.es>
- Shorrocks, A.F. (1984), “Inequality Decomposition by Population Subgroups,” *Econometrica*, 52(6): 1369–85.
- Triplett, J. (1983), “Escalation Measures: What Is the Answer, What Is the Question?,” in W. E. Diewert and C. Montmarquette (eds.), *Price Level Measurement*, Ottawa: Statistics Canada.
- Yahav, J. and S. Yitzhaki (1991), “On the Appropriate Index for Cost-of-Living Adjustment,” in E. Helpman and Y. Nathan (eds) *Studies in Israel’s Economy, 1989*, Jerusalem: The Israel Economic Association.
- Yitzhaki, S. (1996), “On Using Linear Regressions in Welfare Economics,” *Journal of Business and Economic Statistics*, 14(4), 478–486.