The Laspeyres bias in the Spanish consumer price index

J. RUIZ-CASTILLO, E. LEY* and M. IZQUIERDO‡

*Dept. de Economía, Universidad Carlos III de Madrid, Spain, *IMF Institute, Washington DC, USA and ‡FEDEA and Bank of Spain, Madrid, Spain*

The CPI compares the cost of acquiring a reference quantity vector at current and base prices. Such reference vector is the vector of mean quantities actually bought by a reference population, whose consumption patterns are investigated during a period τ prior to the index base period 0. This paper shows that unless the price change between these two dates is taken into account, the CPI ceases to be a proper statistical price index of the Laspeyres type. Among several negative consequences, the most important is that this omission produces a bias in the measurement of inflation: the 'Laspeyres bias'. Using Spanish data, the size of the Laspeyres bias is estimated at -0.061% per year, during 1992–1998. The Laspeyres bias in shorter time periods reached -0.122% per year in 1992, and -0.108 in 1997.

I. INTRODUCTION

A true Cost of Living Index (or COLI for short) for an individual consumer compares a vector of current prices with a given vector of base prices while maintaining constant the consumer's living standard or utility level. Whilst the consumer's preferences would need to be known in order to estimate a COLI, a Statistical Price Index (SPI for short) serves the same purpose but maintains constant a reference quantity vector which, in principle, can be directly observed. When the utility level or the quantity vector correspond to the base period, we say that the COLI or the SPI are of the Laspeyres type. In this case, the (observable) Laspeyres SPI provides an upper bound to the (conceptually appealing) Laspeyres COLI (Konüs, 1924).

In practice, official statistical agencies are concerned with group indexes which are meant to be representative of a certain population of households (individuals or consumers). Consumers' behaviour is investigated by means of a household budget survey. On the basis of this information, a Consumer Price Index (CPI) – which is an aggregate SPI where the reference the vector is the one of mean quantities actually bought by the reference population –

can then be constructed. However, because the survey's collection period typically precedes the base period, statistical offices must properly take into account the price change between these two dates. In those countries where this is done, the CPI becomes what has been called a modified Laspeyres SPI (Moulton, 1996). Such a CPI is a weighted average of a set of household-specific CPIs in which each household's reference quantity vector is the one she acquired during the survey period. The connection between each individual's CPI and a modified Laspeyres COLI provides the basis for a normative rationale for the CPI.

This paper is concerned with those countries where statistical offices do not make any adjustment in the CPI weights for the price change between the survey's collection period and the base period. It is important to emphasize that those countries which conduct a yearly household survey in order to construct a chained-Laspeyres price index can also be affected by this problem. As pointed out by Fry and Pashardes (1986), the weights used in the UK Retail Price Index correspond to expenditures incurred, on average, 12 months prior to the base change which takes place in January of every year. Since, at the time, the price changes occurred during these 12 months were not taken into account to express the weights at base prices, the UK chained-price index was not a group index of the Laspeyres type.

The US, Germany, France, UK and many OECD countries presently correct for this problem. However, together with Spain, among the countries which change the base after a number of years, Argentina and Austria (for detailed products) suffer the problem. Among the countries with a yearly chained price index, in Norway the base period has a six month lag relative to the aggregate weights which are estimated as a three year average. On the other hand, many non-OECD countries – like Colombia, Chile, Denmark and Perú – have recently made the necessary methodological changes to correct the problem.

Those countries which do not take fully into account the price change between the survey's collection period and the base period face three difficulties: (1) The nexus between an individual CPI and a COLI breaks down and, with it, the basis for a normative justification of the aggregate CPI. (2) The individual CPIs are no longer valid for expressing household total expenditures in the survey period at the constant prices of some other period. (3) Using a wrong group index of this type, rather than a true modified Laspeyres SPI, creates what we call a Laspeyres bias in the measurement of inflation. The sign of this bias is empirically related to the plutocratic gap, namely, the bias in the measurement of inflation which appears when we use the present plutocratic CPI rather than a democratic one in which all households receive the same weight – see Izquierdo *et al.* (2002) and Ley (2002).

In Spain, the household budget surveys which serve to estimate the weights (or the reference quantity vector) of the official CPI are the EPFs (*Encuestas de Presupuestos Familiares*), conducted by the INE (*Instituto Nacional de Estadística*). This paper constructs a series of modified Laspeyres indexes for each household interviewed in each of the two latest EPFs. These surveys, gathered in 1980–1981 and 1990–1991, have been used in the CPI systems based in 1983 and 1992, respectively. The main empirical results are that for the two periods 1983 to December 1992, and 1992 to January 1998, the Laspeyres bias is equal to -0.026 and -0.061, respectively. Thus, during the last 12 years the official Spanish CPI has been (slightly) underestimating the inflation which would have been observed if the true modified Laspeyres price indexes has been used.

The rest of the paper is organized into three Sections. Section II introduces the notation for individual and group indexes. Section III presents the empirical results, while Section IV concludes. Some data details are relegated to an Appendix.

II. INDIVIDUAL AND GROUP PRICE INDEXES

Let there be N goods and H households indexed by $i = 1, \ldots, N$ and $h = 1, \ldots, H$, respectively, and let $\mathbf{q} = (q_1, \ldots, q_N)$ be a commodity vector.[1] Each household h is characterized by her total expenditures, x^h, and her preferences represented by a utility function $U^h(\mathbf{q})$. Assume that all households have the same preferences, so that $u = U^h(\mathbf{q}) = U(\mathbf{q})$ for all h, and let $c(u, \mathbf{p})$, be the cost or expenditure function, which gives the minimum cost of achieving the utility level u at prices $\mathbf{p}$. Under general conditions, it is known that $x^h = c(U(\mathbf{q}^h), \mathbf{p})$, where $\mathbf{q}^h$ is the utility maximizing commodity vector at prices $\mathbf{p}$ when household expenditures are x^h.

Consider two price vectors $\mathbf{p}_0$ and $\mathbf{p}_t$ in periods 0 and t. A true or a Konüs COLI for each household which takes as reference the utility level u^h, is defined as the ratio of the minimum cost of achieving that utility level at prices $\mathbf{p}_t$ and $\mathbf{p}_0$: $\kappa(\mathbf{p}_t, \mathbf{p}_0; u^h) = c(\mathbf{p}_t, u^h)/c(\mathbf{p}_0, u^h)$. When the reference utility is the utility maximizing level at prices $\mathbf{p}_0$, denoted by u_0^h, it is said that the COLI $\kappa(\mathbf{p}_t, \mathbf{p}_0; u_0^h)$ is a Laspeyres type index. Given a reference commodity vector, $\mathbf{q}^h$, a SPI is defined as the ratio of the cost of acquiring $\mathbf{q}^h$ at prices $\mathbf{p}_t$ and $\mathbf{p}_0$:[2] $\ell(\mathbf{p}_t, \mathbf{p}_0; \mathbf{q}^h) = \mathbf{p}_t \cdot \mathbf{q}^h / \mathbf{p}_0 \cdot \mathbf{q}^h$. When $\mathbf{q}^h = \mathbf{q}_0^h$, the utility maximizing vector at prices $\mathbf{p}_0$, it is said that the SPI $\ell(\mathbf{p}_t, \mathbf{p}_0; \mathbf{q}_0^h)$ is a Laspeyres type index.

A fundamental theorem in Konüs (1924) establishes that, under general assumptions, the Laspeyres SPI provides an upper bound to the Laspeyres COLI, i.e., if $u_0^h = U(\mathbf{q}_0^h)$, then $\kappa(\mathbf{p}_t, \mathbf{p}_0; u_0^h) \leq \ell(\mathbf{p}_t, \mathbf{p}_0; \mathbf{q}_0^h)$. Equality is obtained when preferences are of the Leontief type, i.e., when there is no substitution between goods.

The modified Laspeyres CPI

Define the vector of mean quantities $\bar{\mathbf{q}}_0 = (\bar{q}_{10}, \ldots, \bar{q}_{I0})$, where $\bar{q}_{i0} = (1/H) \sum_h q_{i0}^h$. The aggregate Laspeyres SPI is defined as follows:

$$\ell(\mathbf{p}_t, \mathbf{p}_0; \bar{\mathbf{q}}_0) = \frac{\mathbf{p}_t \cdot \bar{\mathbf{q}}_0}{\mathbf{p}_0 \cdot \bar{\mathbf{q}}_0} \tag{1}$$

However, the CPI actually computed by statistical agencies is not exactly an aggregate Laspeyres SPI of the type defined in Equation 1. The reason is that individual behaviour is typically investigated by means of a household budget survey conducted in a period τ prior to the index base period, say period 0.

---

[1] A few words on the notation; superscripts will be used for households, and subscripts for goods and time. Boldface symbols will be used to denote vectors, and the "·" operator will indicate a vector inner product: $\mathbf{p} \cdot \mathbf{q} = \sum_i p_i q_i$.
[2] Defining the budget shares $w_i^h = p_{i0} q_i^h / \sum_i p_{i0} q_i^h$, an SPI can also be conveniently expressed as a weighted sum of individual price changes: $\ell(\mathbf{p}_t, \mathbf{p}_0; \mathbf{q}^h) = \sum_i w_i^h (p_{it}/p_{i0})$.

Household budget surveys provide information on individual expenditures in each good, $x_{i\tau}^h$, and on total expenditures, x_τ^h, not on individual prices and quantities – which are often hard to define. Under the assumption that all households living in the same area face the same prices, observable household expenditures on item i by household h living in area j can also be viewed as the product of a price, $p_{ij\tau}$, and a quantity, $q_{i\tau}^h$, – i.e., $x_{i\tau}^h = p_{ij\tau} q_{i\tau}^h$. When holding information on the prices at τ, $p_{ij\tau}$, quantities purchased can then be recovered for each good, $q_{i\tau}^h = x_{i\tau}^h / p_{ij\tau}$, and $\bar{\mathbf{q}}_\tau$ used in the index construction instead of $\bar{\mathbf{q}}_0$.[3] In this setting, the CPI based in period 0 is an aggregate SPI defined as:

$$CPI_t \equiv \ell(\mathbf{p}_t, \mathbf{p}_0; \bar{\mathbf{q}}_\tau) = \frac{\mathbf{p}_t \cdot \bar{\mathbf{q}}_\tau}{\mathbf{p}_0 \cdot \bar{\mathbf{q}}_\tau} = \frac{\ell(\mathbf{p}_t, \mathbf{p}_\tau; \bar{\mathbf{q}}_\tau)}{\ell(\mathbf{p}_0, \mathbf{p}_\tau; \bar{\mathbf{q}}_\tau)} \qquad (2)$$

This is what the BLS calls a modified Laspeyres aggregate price index (Moulton, 1996), with base period 0 and reference consumption patterns surveyed at τ.

For each household h:

$$cpi_t^h \equiv \ell(\mathbf{p}_t, \mathbf{p}_0; \mathbf{q}_\tau^h) = \frac{\mathbf{p}_t \cdot \mathbf{q}_\tau^h}{\mathbf{p}_0 \cdot \mathbf{q}_\tau^h} = \frac{\ell(\mathbf{p}_t, \mathbf{p}_\tau; \mathbf{q}_\tau^h)}{\ell(\mathbf{p}_0, \mathbf{p}_\tau; \mathbf{q}_\tau^h)}$$

Define, for each household h the plutocratic weight:

$$\phi^h = \frac{1}{H} \frac{\mathbf{p}_0 \cdot \mathbf{q}_\tau^h}{\mathbf{p}_0 \cdot \bar{\mathbf{q}}_\tau}$$

Then, as known from Prais (1958):

$$\sum_h \phi^h cpi_t^h = \frac{1}{H} \sum_h \frac{\mathbf{p}_0 \cdot \mathbf{q}_\tau^h}{\mathbf{p}_0 \cdot \bar{\mathbf{q}}_\tau} \frac{\mathbf{p}_t \cdot \mathbf{q}_\tau^h}{\mathbf{p}_0 \cdot \mathbf{q}_\tau^h} = \frac{(1/H) \sum_h \mathbf{p}_t \cdot \mathbf{q}_\tau^h}{\mathbf{p}_0 \cdot \bar{\mathbf{q}}_\tau} = \frac{\mathbf{p}_t \cdot \bar{\mathbf{q}}_\tau}{\mathbf{p}_0 \cdot \bar{\mathbf{q}}_\tau}$$

$$= CPI_t \qquad (3)$$

Thus, on one hand, the CPI is an SPI which serves to compare the price vector in any period t with the price vector in the base period 0, while maintaining constant the aggregate vector $\bar{\mathbf{q}}_\tau$ of mean quantities actually purchased during the survey period τ – see Equation 2. On the other hand, the CPI is the plutocratic weighted mean of a set of household-specific modified Laspeyres price indexes – see Equation 3.

The question is, what is the normative basis for such a construction? To answer this question it is necessary to define a set of household-specific modified Laspeyres COLIs. For each h, let $u_\tau^h = U(\mathbf{q}_\tau^h)$. It is easy to see that the ratio of the corresponding Laspeyres COLIs leads to what is called a modified Laspeyres COLI:

$$\frac{\kappa(\mathbf{p}_t, \mathbf{p}_\tau; u_\tau^h)}{\kappa(\mathbf{p}_0, \mathbf{p}_\tau; u_\tau^h)} = \frac{c(\mathbf{p}_t, u_\tau^h)}{c(\mathbf{p}_0, u_\tau^h)} = \kappa(\mathbf{p}_t, \mathbf{p}_0; u_\tau^h)$$

Konüs theorem ensures that, for each h, $\ell(\mathbf{p}_s, \mathbf{p}_\tau; \mathbf{q}_\tau^h) - \kappa(\mathbf{p}_s, \mathbf{p}_\tau; u_\tau^h) \geq 0$ for $s = 0, t$, but it says nothing about the ratio of the Laspeyres indexes which give rise to an individual CPI. However, the household budget survey collection period τ is typically not far apart from the base year 0 of the CPI system. Thus, under the assumption that the substitution bias $\ell(\mathbf{p}_0, \mathbf{p}_\tau; \mathbf{q}_\tau^h) - \kappa(\mathbf{p}_0, \mathbf{p}_\tau; u_\tau^h)$ is smaller than $\ell(\mathbf{p}_t, \mathbf{p}_\tau; \mathbf{q}_\tau^h) - \kappa(\mathbf{p}_t, \mathbf{p}_\tau; u_\tau^h)$, a household-specific CPI provides an upper bound to a modified Laspeyres COLI. In view of Equation 3, it is seen that $CPI_t \geq \sum_h \phi^h \kappa(\mathbf{p}_t, \mathbf{p}_0; \mathbf{q}_\tau^h)$. Thus, only under the assumption that for a sufficiently large number of households:

$$cpi_t^h = \frac{\ell(\mathbf{p}_t, \mathbf{p}_\tau; \mathbf{q}_\tau^h)}{\ell(\mathbf{p}_0, \mathbf{p}_\tau; \mathbf{q}_\tau^h)} \geq \frac{\kappa(\mathbf{p}_t, \mathbf{p}_\tau; u_\tau^h)}{\kappa(\mathbf{p}_0, \mathbf{p}_\tau; u_\tau^h)} = \kappa(\mathbf{p}_t, \mathbf{p}_0; u_\tau^h)$$

the aggregate CPI provides an upper bound to a plutocratic weighted mean of modified Laspeyres COLIs.

The Spanish CPI

In Spain, the INE does not use information on the prices $p_{ij\tau}$ when constructing its price index. Therefore, only an IPC (*Indice de Precios de Consumo*) can be defined. At the individual level, for any household h, the individual IPC is defined as:

$$ipc_t^h \equiv \sum_i w_{i\tau}^h \frac{p_{it}}{p_{i0}} = \sum_i \frac{p_{i\tau} q_{i\tau}^h}{\mathbf{p}_\tau \cdot \mathbf{q}_\tau^h} \frac{p_{it}}{p_{i0}} = \frac{\sum_i p_{it}(p_{i\tau}/p_{i0}) q_{i\tau}^h}{\sum_i p_{i0}(p_{i\tau}/p_{i0}) q_{i\tau}^h}$$

$$= \frac{\sum_i p_{it} \eta_i^h}{\sum_i p_{i0} \eta_i^h} = \frac{\mathbf{p}_t \cdot \boldsymbol{\eta}^h}{\mathbf{p}_0 \cdot \boldsymbol{\eta}^h} = \ell(\mathbf{p}_t, \mathbf{p}_0; \boldsymbol{\eta}^h) \qquad (4)$$

where $w_{i\tau}^h = x_{i\tau}^h / x_\tau^h$ is the expenditure weight of good i, and $\boldsymbol{\eta}^h = (\eta_1^h, \ldots, \eta_N^h)$ with:

$$\eta_i^h = \frac{p_{i\tau}}{p_{i0}} q_{i\tau}^h \qquad (5)$$

Thus, the Spanish IPC for each h is a SPI where the reference vector $\boldsymbol{\eta}^h$ is the vector $\mathbf{q}^h$ multiplied by the price change between period τ and period 0.

At the aggregate level, let $\boldsymbol{\eta}$ be the average reference vector with generic element $\bar{\eta}_i = (1/H) \sum_h \eta_i^h$. Then, the aggregate IPC is given by:

$$IPC_t \equiv \ell(\mathbf{p}_t, \mathbf{p}_0; \boldsymbol{\eta}) = \frac{\mathbf{p}_t \cdot \boldsymbol{\eta}}{\mathbf{p}_0 \cdot \boldsymbol{\eta}} \qquad (6)$$

Define now the plutocratic weights:

$$\phi_\tau^h = \frac{1}{H} \frac{\mathbf{p}_0 \cdot \boldsymbol{\eta}^h}{\mathbf{p}_0 \cdot \boldsymbol{\eta}} = \frac{1}{H} \frac{\mathbf{p}_\tau \cdot \mathbf{q}_\tau^h}{\mathbf{p}_\tau \cdot \mathbf{q}_\tau} = \frac{x_\tau^h}{\sum_h x_\tau^h}$$

---

[3] For simplicity, in what follows geographical details will be ignored and the subindex j dropped. However, it should be kept in mind throughout that statistical agencies gather prices by geographical areas and, therefore, all formulas should reflect this – see, for example, Izquierdo *et al.* (2002).

Then, again it is seen that:

$$IPC_t = \sum_h \phi_\tau^h \, ipc_t^h \tag{7}$$

That is, on one hand, the IPC for the population as a whole is an aggregate SPI which takes $\boldsymbol{\eta}$ as the reference vector, which is the vector of mean quantities actually purchased in period τ multiplied by the price change between period τ and period 0 – see Equation 6. On the other hand, the general IPC is the plutocratic weighted mean of the household–specific IPCs – see Equation 7.

This construction poses, at least, three problems. In the first place, at the individual level the index ipc_t^h is unrelated to the modified COLI $\kappa(\mathbf{p}_t, \mathbf{p}_0; u_\tau^h)$. Consequently, the normative basis for the aggregate IPC_t is lost.

In the second place, suppose that it is wished to establish whether, for example, mean household expenditures have risen or not in real terms between period τ and period $t > 0$. Denote by $\mathbf{x}_{\tau,t} = (x_{\tau,t}^1, \dots, x_{\tau,t}^H)$ the distribution of household expenditures in period τ at prices of period t, where for each household $x_{\tau,t}^h = x_\tau^h \kappa(\mathbf{p}_t, \mathbf{p}_\tau; u_\tau^h)$. The change in mean household expenditures at prices p_t can be expressed as $\Delta(\mathbf{p}_t) = \mu(\mathbf{x}_t) - \mu(\mathbf{x}_{\tau,t})$, where $\mu(\mathbf{x})$ denotes the mean of distribution $\mathbf{x}$. Similarly, the change in mean household expenditures at prices of period τ can be expressed as $\Delta(\mathbf{p}_\tau) = \mu(\mathbf{x}_{t,\tau}) - \mu(\mathbf{x}_\tau)$. If there is a proper household-specific Laspeyres price index for each h, then $\hat{x}_{\tau,t}^h = x_\tau^h \ell(\mathbf{p}_t, \mathbf{p}_\tau; \mathbf{q}_\tau^h) = \mathbf{p}_t \cdot \mathbf{q}_\tau^h \geq x_{\tau,t}^h$ can be computed for each h. Thus, $\mu(\hat{\mathbf{x}}_{\tau,t}) \geq \mu(\mathbf{x}_{\tau,t})$ and a lower bound can be provided for $\Delta(\mathbf{p}_t)$. Similarly, $\hat{x}_{t,\tau}^h = x_t^h / \ell(\mathbf{p}_t, \mathbf{p}_\tau; \mathbf{q}_t^h) = \mathbf{p}_\tau \cdot \mathbf{q}_t^h \geq x_{t,\tau}^h$ can be computed for each h, and that distribution used to obtain an upper bound for $\Delta(\mathbf{p}_\tau)$. However, if there are only household-specific ipc_t^h's, then the only thing that can be done to express the household total expenditures in period τ at prices $\mathbf{p}_t$ is to multiply x_τ^h by the index defined in Equation 4:

$$x_\tau^h \, ipc_t^h = (\mathbf{p}_\tau \cdot \mathbf{q}_\tau^h) \frac{\mathbf{p}_t \cdot \boldsymbol{\eta}^h}{\mathbf{p}_0 \cdot \boldsymbol{\eta}^h} = \mathbf{p}_t \cdot \boldsymbol{\eta}^h \tag{8}$$

However, in this operation $\mathbf{p}_t \cdot \mathbf{q}_\tau^h$ cannot be recovered as desired. Similarly, with such indexes $\mathbf{p}_\tau \cdot \mathbf{q}_t^h$ cannot be recovered either. Hence, by this route neither a lower bound for $\Delta(\mathbf{p}_t)$ nor an upper bound for $\Delta(\mathbf{p}_\tau)$ can be provided.

Finally, let π_t be the inflation rate according to the IPC_t defined in Equation 7. Measuring inflation in this way, one incurs in some bias relative to the alternative of using a modified Laspeyres group price index CPI_t as defined in Equation 2 – let $\hat{\pi}_t$ be the corresponding inflation rate. Then the Laspeyres bias is defined by $(\pi_t - \hat{\pi}_t)$.

What can be expected about the sign of the Laspeyres bias in a given period? The answer depends on the behaviour of prices during that period and during the time interval which goes from the survey's collection period τ to the base period 0. It has been seen that all group indexes in this section are weighted averages of the individual indexes with weights proportional to household total expenditures. Alternatively, a democratic group index can be defined in which all household-specific indexes weigh equally (Prais, 1959). The plutocratic gap in the measurement of inflation is then defined as the difference between the inflation estimated according to a plutocratic and a democratic group index. When the inflation in a given period is greater (smaller) for the rich than for the poor households, then the plutocratic weighted mean would be greater (smaller) than the simple mean. Therefore, the plutocratic gap is expected to be positive (negative) when prices behave in an anti-rich (anti-poor) way.

Suppose now that between periods τ and 0 all prices have risen, so that, for each h, $\eta_i^h < q_{i\tau}^h$ for all i – see Equation 5. The greater the inflation experienced by a particular good i, the greater the difference between η_i^h and $q_{i\tau}^h$. Suppose that the prices of luxuries (goods with total expenditures elasticity greater than 1) have increased by more than the price of necessities (goods with total expenditures elasticity smaller than 1); or in other words, suppose that prices have been anti-rich so that the plutocratic gap has been positive. Then, for each h, the relative importance of luxuries in $\boldsymbol{\eta}^h$ is less than in $\mathbf{q}_\tau^h$, while the opposite will be the case for necessities. Suppose further that the same price pattern obtains between period 0 and period t, that is, suppose that prices are again anti-rich. Then, for each h, the SPI that takes as reference the vector $\boldsymbol{\eta}^h$ would tend to understate the inflation which has taken place according to a modified Laspeyres price index which takes the vector $\mathbf{q}_\tau^h$ as reference. In this case, the Laspeyres bias would have a negative sign. On the contrary, given that it is assumed that prices are anti-rich from period τ to period 0, if between period 0 and period t the plutocratic gap is negative, then the Laspeyres bias can be expected to be positive.[4]

Nonetheless, this relationship between the Laspeyres bias and the plutocratic gap – although it holds for the data analysed in this paper as seen below – is not theoretically guaranteed, and it is possible to construct counter-examples when $N \geq 3$.

III. EMPIRICAL RESULTS

Laspeyres price indexes are constructed for all households surveyed in the two latest household surveys conducted in

---

[4] Suppose instead that from period τ to period 0 prices are anti-poor. Then the relative importance of luxuries in $\boldsymbol{\eta}^h$ would be greater than would be the case in $\mathbf{q}_\tau^h$, while the opposite would be the case for necessities. Then the Laspeyres bias would tend to be negative (positive) according to whether prices from period 0 to period t behave in an anti-poor (anti-rich) way.

Table 1. *The Laspeyres bias (in percent per year)*

Subperiods	Inflation		Laspeyres bias	Plutocratic gap
	π_t	$\hat{\pi}_t$		
Base 1983				
1980–81/83				0.025
Aug 85 to Dec 85	6.896	6.852	0.033	−0.157
Dec 85 to Dec 86	8.217	8.230	−0.023	−0.038
Dec 86 to Dec 87	4.622	4.587	0.029	0.381
Dec 87 to Dec 88	5.895	5.866	0.021	−0.025
Dec 88 to Dec 89	6.908	6.923	−0.019	−0.032
Dec 89 to Dec 90	6.589	6.641	−0.049	0.228
Dec 90 to Dec 91	5.576	5.603	−0.037	0.312
Dec 91 to Dec 92	5.364	5.486	−0.128	0.593
Aug 85 to Dec 92	6.848	6.867	−0.026	0.186
Base 1992				
1990–91/92				0.088
Jan 93 to Jan 94	5.228	5.267	−0.039	0.105
Jan 94 to Jan 95	4.600	4.619	−0.018	−0.080
Jan 95 to Jan 96	4.047	4.078	−0.031	−0.050
Jan 96 to Jan 97	3.108	3.176	−0.068	0.090
Jan 97 to Jan 98	2.376	2.484	−0.108	0.125
1993 to Jan 98	4.120	4.181	−0.061	0.038

Spain from April 1990 to March 1991, and April 1980 to March 1981.[5] They are referred to as the 1990–1991 EPF, and the 1980–1981 EPF, respectively. The data sources are explained in the Appendix.

Let IPC_t be the group index defined in Equation 4 which compares the vector of prices in January of year t with the vector of prices in the base period – for example, $0 = 1983$, and 1992 in the two panels of Table 1. The inter-annual inflation rate in year t is denoted by $\pi_t = ((IPC_{t+1}/IPC_t) - 1) \times 100$. Let CPI_t be the modified Laspeyres group index defined in Equation 2 referring to the same base. Denote by $\hat{\pi}_t = ((CPI_t/CPI_{t-1}) - 1) \times 100$ the corresponding inter-annual inflation rate. The Laspeyres bias for year t is defined by $\pi_t - \hat{\pi}_t$. Table 1 presents the estimates for π_t, $\hat{\pi}_t$ and the corresponding Laspeyres bias, $t = 1985, \ldots, 1997$. For each of the two periods 1985–1992 and 1993–1998, the Laspeyres bias is equal to $\Pi - \hat{\Pi}$, where Π and $\hat{\Pi}$ are the average annual inflation rates using the IPC and the CPI, respectively – shown on the bottom row of each panel in Table 1.

Denote by $\tilde{\pi}_t$ the inter-annual inflation rate and $\tilde{\Pi}$ the average annual inflation rate when a democratic price index is used to measure inflation, $(1/H)\sum_h cpi_t^h$, instead of the plutocratic index defined in Equation 3. The pluto-cratic gap is then defined by $\pi_t - \tilde{\pi}_t$ for each year, and by

$\Pi - \tilde{\Pi}$ for the two periods 1985–1992 and 1993–1998. The estimates of the plutocratic gap are shown in the last col-umn of Table 1.

In the first place, it is observed that for the 1993–1998 subperiod as a whole, the Laspeyres bias is equal to −0.061% per year (−0.026 for the 1985–1992 subperiod). How can this negative sign be explained? Notice that from the 1990–1991 EPF's collection period to the base year 1992, prices behave in an anti-rich way: the plutocratic gap is equal to 0.088% per year (0.025 from the 1980–1981 survey period to the base year 1983). As seen in the last section, this means that the official IPC would tend to give less weight to luxuries and more weight to necessities than a modified Laspeyres group index. On the other hand, from 1992 to January 1998 price behaviour is again anti-rich: the plutocratic gap is equal to 0.038% per year (0.186 from August 1985 to December 1992). Consequently, the IPC would tend to register a smaller inflation, Π, than the modified Laspeyres alternative, $\hat{\Pi}$. This explains the nega-tive sign of the Laspeyres bias that is found on the bottom row of each panel in Table 1.

To appreciate the variability of the Laspeyres bias during the entire period considered in this paper, the top panel in Figure 1 shows a series of monthly observations on the evolution of the inter-annual Laspeyres biases and

---

[5] For the survey conducted during 1973–1974 the only information available on the inflation between the survey period and the base year is at a very aggregated level, five goods, which does not allow us to estimate the Laspeyres bias.

Inter-annual Laspeyres bias and plutocratic gap
(month by month)

Legend: Laspeyres Bias — Plutocratic Gap
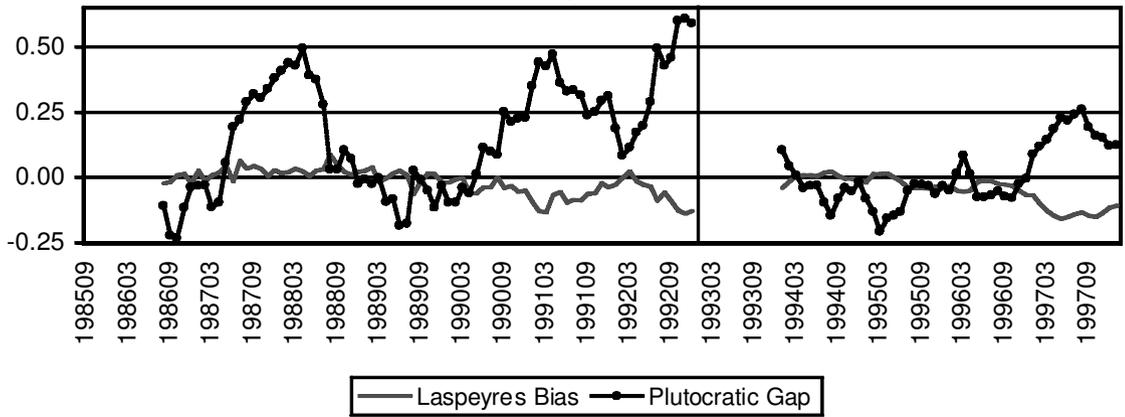
Monthly Laspeyres bias (annualized)
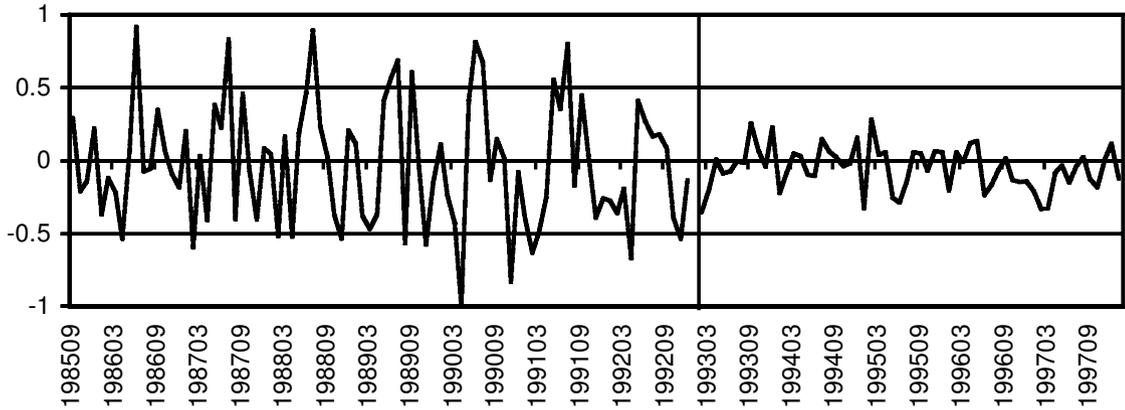
Monthly plutocratic gap (annualized)

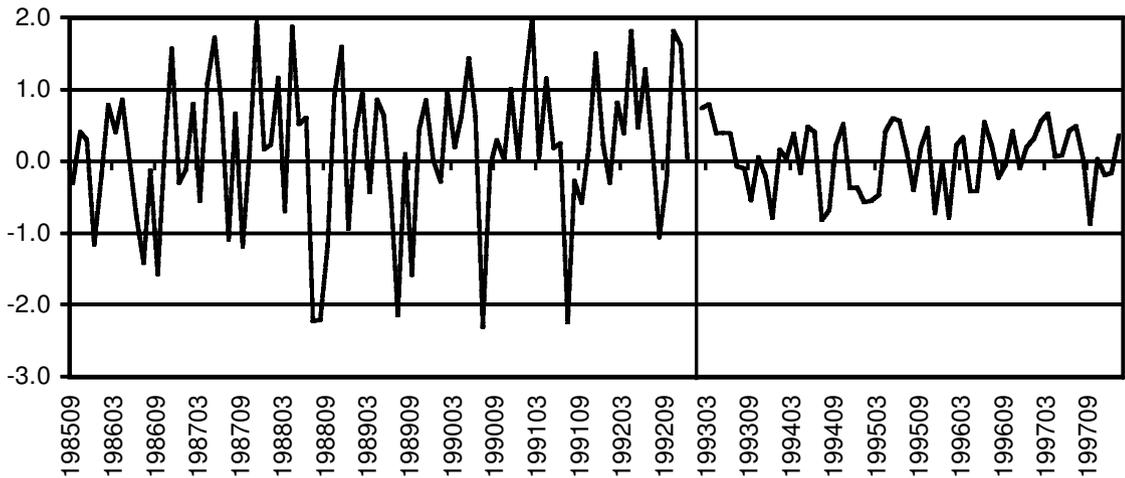
Fig. 1. *The Laspeyres bias and the plutocratic gap (in per cent per year)*

Table 2. *The Laspeyres bias vs the plutocratic gap*

	Base year	
	1992	1983
Intercept	3.5E-5	1.6E-5
S.E.	(1.2E-5)	(3.0E-5)
Plutocratic gap	−0.1766	−0.2485
S.E.	(0.046)	(0.035)
Durbin-Watson	2.12	1.62
$\bar{R}^2$	0.29	0.36
$F(1, T-2)$	24.1	49.25
T	59	87

plutocratic gaps for October 1986 to January 1998. Since each annual inflation is a moving average of the inflation of the 12 previous months, the information contained on each data point overlaps significantly with the adjacent observations (the series are integrated). The bottom panels show the first differences of these series – which are nothing but monthly inflation biases. The monthly series have been annualized in order to facilitate its interpretation. These annualized series display very large magnitudes, reaching close to $\pm 1\%$ per year in some instances in the 1980s. These large biases of different signs tend to cancel off over longer periods and inter-annual biases show smaller magnitudes.

Given the anti-rich price bias from the survey's period to the base year in the two cases considered, it is known that the official IPC takes as reference a vector of aggregate quantities where luxuries receive less weight and necessities receive more weight than they would in a modified Laspeyres construction. Therefore, if in a given period the plutocratic gap is positive (negative), reflecting an anti-rich (anti-poor) bias, then the corresponding Laspeyres bias is expected to move in the opposite direction. This is indeed what is observed in Figure 1.[6] Table 2 shows the results of the regression of the Laspeyres bias against the corresponding plutocratic gap, using inter-monthly inflation data. The negative relationship between the Laspeyres bias and the plutocratic gap is displayed in an estimated coefficient of about −0.2 with a standard error of about 0.04 for the 1992 and 1983 base systems.

IV. CONCLUSIONS

The CPI compares the cost of acquiring a reference quantity vector at current and base prices. Such reference vector

is the vector of mean quantities actually bought by a reference population, whose consumption patterns are investigated during a period τ prior to the index base period 0. This paper has shown that unless one takes into account the price change between these two dates, each component of the reference quantity vector will be multiplied by the ratio of the price of the good in period τ and in period 0. As a consequence, the CPI ceases to be a proper SPI of the Laspeyres type.

This has several negative consequences: (1) The link between the CPI and a group index based on the COLIs of the reference population breaks down; (2) the possibility of expressing the consumption expenditures in period τ at prices of other periods disappears, and, more importantly, (3) it produces a bias in the measurement of inflation which we have called the 'Laspeyres bias'. The relation of this bias with the plutocratic gap (Ley, 2002; Izquierdo *et al.*, 2002) during a particular period t depends on whether prices exhibit an anti-rich or an anti-poor behaviour from period τ to period 0, and from period 0 to the period t in question.

Inflation targets constitute a policy objective of paramount importance. For example, the Maastrich agreements in 1992 singled out an inflation objective as one of the three criteria for European Union members to become part of the European Monetary Union. Moreover, thanks to the ample publicity received by the report to the US Senate by a commission headed by Michael Boskin, it has been forcefully reminded about the dramatic economic consequences of a relatively small bias in the measurement of inflation – see Boskin *et al.* (1996). Consequently, statistical offices must ensure that the CPI preserves its alleged properties and that its measurement is as free as possible from any bias.

Of course, the urgency of the problem at hand depends on its quantitative importance. This paper has presented some evidence on the Laspeyres bias in Spain for the CPI systems based in 1983 and 1992. It has been shown that this bias has a predominantly negative sign for an extended period of time which expands from 1985–1998. Essentially, this is explained by the overall anti-rich bias exhibited by the evolution of prices in Spain during this period. This does not preclude that the Laspeyres bias takes a positive sign during specific subperiods characterized by an anti-poor price behaviour. Finally, the Laspeyres bias has displayed a considerable size during certain periods of time. For instance, from 1992–1998, the size of the Laspeyres bias was 0.061% per year, or about 6% of the overall bias from five sources estimated

[6] In Table 3.2 of Fry and Pashardes (1986, p. 26) the importance of the Laspeyres bias can be observed in the UK. Qualitatively, the difference with Spain is that, in the UK the bias in every year from 1977–1984 has a positive sign – except for the period 1975–1977 in which the bias is zero or slightly negative – reaching a maximum value of 0.8% in 1978. According to the discussion in Section II, the explanation is clear: as these authors and others have documented – see also Crawford (1994) and Muellbauer (1974 a,b) – during the 1970s price behaviour in the UK was anti-poor.

by the Boskin commission for the USA, which is equal to 1.1% per year.[7] The Laspeyres bias in shorter time periods has reached 0.122, and 0.108% per year in 1992, and 1997, respectively.

The practical message of the paper is clear: when the household budget survey's collection period τ differs from the CPI base year 0, it is necessary to gather information on the evolution of prices from period τ to period 0 in order to express the expenditures incurred in period τ at base period prices. Only in this case is it possible to construct (modified) Laspeyres price indexes which take as reference the mean commodity vector actually acquired by consumers during period τ.

The difficulty lies in the fact that the data collected in the household budget survey is essential for deciding on the characteristics of the new base in relation to the item space, product specifications, and the establishments where price quotes should be taken. How is it possible to record goods prices from period τ to period 0 according to the new methodology at the same time that such a methodology is being decided upon? Surely, some compromises should be adopted in order to find an answer to this practical question. There seems to be no doubt that the statistical agency responsible for the CPI is the best prepared to carry out this task.

Finally, it is worthwhile to emphasize that, once the comparison of the old and the new base is indirectly established through this process, the statistical agency is in a good position to provide the best possible reconstruction of past inflation according to the new methodology. This is potentially very important for those analysts in charge of predicting the short-run CPI behaviour immediately after a change of base.

REFERENCES

Catasús, V., Malo de Molina, J. L., Martínez, M. and Ortega, E. (1986) Cambio de base del Indice de Precios de Consumo, mimeo, Madrid, Banco de España.

Crawford, I. (1994) *UK Household Cost-of-Living Indexes*, Institute for Fiscal Studies, London.

Fry, V. and Pashardes, P. (1985) *The RPI and the Cost of Living*, Report Series No. 22, Institute for Fiscal Studies, London.

García España, E. and Serrano, J. M. (1980) *Indices de Precios de Consumo*, Instituto Nacional de Estadística, Madrid: Ministerio de Economía y Comercio, Madrid.

INE (1983) *Encuesta de Presupuestos Familiares 1980–81. Metodología y resultados*, Instituto Nacional de Estadística, Madrid.

INE (1985) *Indice de Precios de Consumo. Base 1983. Monografía Técnica*, Instituto Nacional de Estadística, Madrid.

INE (1992) *Encuesta de Presupuestos Familiares 1990–91. Metodología*, Instituto Nacional de Estadística, Madrid.

INE (1994) *Indice de Precios de Consumo. Base 1992. Metodología*, Instituto Nacional de Estadística, Madrid.

Izquierdo, M., Ley, E. and Ruiz-Castillo, J. (2002) The plutocratic gap in the CPI: evidence from Spain, *IMF Staff Papers*, forthcoming.

Konüs, A. A. (1924) The problem of the true index of the cost of living, English version in *Econometrica*, **7**, 10–29.

Ley, E. (2002) Whose inflation? A characterization the CPI plutocratic gap, mimeo, available online: http://econwpa. wustl.edu/

Lorenzo, F. (1998) *Modelización de la inflación con fines de predicción y diagnóstico*, PhD Dissertation, Universidad Carlos III de Madrid, Madrid.

Moulton, B. (1996) Constant elasticity cost-of-living index share relative form, mimeo, Bureau of Labour Statistics.

Muellbauer, J. (1974a) Prices and inequality: the United Kingdom experience, *Economic Journal*, **84**, 32–55.

Muellbauer, J. (1974b) The political economy of price indexes, Birbeck Discussion Paper 22.

Prais, S. (1958) Whose cost of living?, *The Review of Economic Studies*, **26**, 126–34.

Ruiz-Castillo, J., Higuera, C., Izquierdo, M. and Sastre, M. (1999a) Series de precios individuales para las EPF de 1973–74, 1980–81 y 1990–91 con base en 1976, 1983 y 1992, mimeo, available online: http://www.eco.uc3m.es/investigacion/epf.html

Ruiz-Castillo, J., Ley, E. and Izquierdo, M. (1999b) *La Medición de la Inflación en España*, la Caixa, Barcelona.

APPENDIX: THE CONSTRUCTION OF HOUSEHOLD-SPECIFIC LASPEYRES SPIs

In order to construct a series of household-specific Laspeyres price indexes for a given period, the following three pieces of information are needed: (1) The household budget survey which serves to estimate the aggregate weights of the official CPI; (2) a set of price subindexes for the period in question at a certain level of commodity and spatial disaggregation; and (3) a set of estimated price changes – that shall be called 'adjustment factors' – between the survey collection period τ and the official base period 0.

In the Spanish case, Laspeyres price indexes are constructed for all households surveyed in the two latest EPFs gathered in 1990–1991 and 1980–1981, respectively. These are large comparable samples consisting of 21 155,

---

[7] In Ruiz-Castillo *et al.* (1999b) it is estimated that this overall bias in Spain is of the order of 0.60% per year. Thus, the Laspeyres bias during the 1990s is about 10% of our best estimate of the overall bias in the Spanish economy.

and 23 972 household sample points, respectively. These samples represent a population of, approximately, 11 or 10 million households and 38 or 37 million persons, respectively, occupying residential housing in all of Spain. People living in collective housing, such as residences for the aged, hospitals, prisons, hotels, and the like, are excluded from the EPFs. The two surveys cover household expenditures on 893, and 614 commodities, respectively. In this and other respects, the later the survey period the more complete the survey is. However, they all share the same sample stratification design, and the same methodology to investigate household expenditures: all household members of 14 or more years of age are supposed to record all expenditures that take place during the sample week; then, in-depth interviews are conducted to register past expenditures over reference periods beyond a week and up to a year – for further details, see INE (1992), and INE (1983). From this information the statistical office estimates annual expenditures on all goods.

As indicated in the text, these EPFs have been used to estimate the corresponding aggregate weights of the Spanish IPC systems based in 1992 and 1983. The information on price subindexes and adjustment factors is best treated separately for each period.

The 1992 IPC: from January 1993 until the present

The INE collects elementary price indexes for a commodity basket consisting of 471 items in each of 52 provinces. For confidentiality reasons, the INE does not publish this information at the maximum spatial disaggregation level. Instead, from January 1993 it publishes on a monthly basis price subindexes for a commodity breakdown of 110 *subclases*, 57 *rúbricas*, 33 *subgrupos* and 8 *grupos* at the national level, the *rúbricas*, *subgrupos* and *grupos* at the 18 Autonomous Communities level, and the *subgrupos* and *grupos* at the 52 provincial level – for further details, see INE (1994).

For any commodity breakdown, it is possible to reconstruct the official IPC series using an appropriately defined vector of aggregate weights or budget shares. Similarly, defining a budget share vector for every household in the 1990–1991 sample, obtain a series of household specific IPCs can be obtained for any commodity breakdown. In principle, the only difference between alternative specifications of the commodity space, is that the dispersion of the set of individual IPCs should be greater the greater the disaggregation level of the price information used in their construction. Unfortunately, in spite of using the same informational basis as the INE – namely, the 1990–1991 EPF – some small discrepancies are found between our

estimates of the aggregate budget share vectors and those published by the INE. Thus, the CPI series which can be reconstructed varies slightly depending on the different commodity breakdowns characterizing the price information used – for an analysis of these discrepancies see Ruiz-Castillo *et al.* (1999a). In Ruiz-Castillo *et al.* (1999b), it is found that the specification consisting of the 21 food *rúbricas* at the Autonomous Community level, and the 32 non-food *subgrupos* at the provincial level outperforms the rest of the alternatives according to various statistical and economic criteria.

It should be emphasized that our series of household-specific price indexes defined over this 53 commodity space differs from the series underlying the official IPC. The reason is that there are a number of aspects in the official definition of total household expenditures for which what is believed to be superior alternatives are used: (1) the definition of housing expenditures for households occupying non-rental housing; (2) the inclusion of imputations for home production, wages in kind and subsidized meals, and (3) the estimation of annual food and drink expenditures using all the available information on bulk purchases in the 1990–1991 EPF.[8]

As pointed out in Section II, because the INE does not use any adjustment factors for taking into account the price change between the EPF's collection period and the base period, the official index $\ell(\mathbf{p}_t, \mathbf{p}_0; \boldsymbol{\eta})$ does not coincide with the modified Laspeyres index $\ell(\mathbf{p}_t, \mathbf{p}_0; \bar{\mathbf{q}}_\tau)$. Fortunately, the analysts devoted to short run forecasting of the economy need sufficiently long price series drawn with a common methodology in order to do their work. Thus, when there is a change of base in the system they estimate the price changes (p_{it}/p_{i0}), with $t < 0$, where the commodity space, as well as the item specifications correspond, as best as possible, to the methodology of the new CPI base. Taking into account the methodological changes adopted by the INE for the current IPC base of 1992, Lorenzo (1998) provides such information on a monthly basis for the 110 *subclases*, at the national level from January 1983 until 1992. For each of the quarters ($\tau =$ Spring, Summer, Autumn of 1990, and Winter of 1991), using Lorenzo (1998) data on the adjustment factors $p_{i\tau}/p_{i0}$ for each of the 110 *subclases* the price ratios $p_{it}/p_{i\tau} = (p_{it}/p_{i0})/(p_{i\tau}/p_{i0})$ are computed, where p_{it}/p_{i0} are the price subindexes provided by the INE on a monthly basis from January 1993 to January 1998. Given the Laspeyres indexes $\ell(\mathbf{p}_t, \mathbf{p}_\tau; \mathbf{q}_\tau^h)$, a series of modified Laspeyres price indexes is constructed from January 1993 until January 1998, based in that period $0 = 1992$, which takes as reference the commodity vector $\mathbf{q}_\tau^h$ actually acquired during the interview quarter τ.

---

[8] The joint impact of these modifications is important: according to Izquierdo *et al.* (2002), the official CPI understates the true Spanish inflation from 1992 to January of 1998 by 0.241% per year.

The INE collects elementary price indexes for a commodity basket consisting of 428 items in each of 52 provinces. It publishes on a monthly basis price subindexes for a commodity breakdown of 106 *subclases*, 57 *rúbricas*, 29 *subgrupos* and 8 *grupos* at the national level. Complete information at the Autonomous Communities level is only available for the eight *grupos* – for further details see INE (1985). The information is used for the 106 *subclases* at the national level.

This case does not depart from the official definition of household total expenditures it agrees with the way non-rental housing is treated. Nevertheless, there are some discrepancies between the official aggregate weights published by the INE and those presented here. In the first place, the information on all households interviewed in the 1980–1981 EPF is used, while the INE restricts itself to a reference population which excludes single-person households and those multi-person ones with total income below the 1980—1981 minimum wage or above a certain amount. These restrictions mean that the official IPC refers to 79% of all households, 86% of all persons, and 85% of all household expenditures. In the second place, even when this factor is taken into account, some minor discrepancies are found, as before, between our estimates of the aggregate weights and those published by the INE – for an analysis of these discrepancies see Ruiz-Castillo *et al.* (1999a).

As far as the adjustment factors, $p_{i\tau}/p_{i0}$, the monthly series for 60 goods is used at the national level provided by Catasús *et al.* (1986) from January 1978 to July of 1985. A set of 52 goods is worked with which constitute the minimum common denominator between the 57 official *rúbricas* and the 60 goods in Catasús (1986) – for the details of this construction see Ruiz-Castillo *et al.* (1999a).