Federal Reserve Bank of New York
Staff Reports

# Inflation Inequality in the United States

Bart Hobijn
David Lagakos

Staff Report no. 173
October 2003

**Title: Inflation Inequality in the United States**
Authors: Bart Hobijn and David Lagakos
*Federal Reserve Bank of New York Staff Reports*, no. 173
October 2003
JEL classification: C43, D12, D39

**Abstract**

Inflation is often assumed to affect all people in the same way. In practice, differences in spending patterns across households and differences in price increases across goods and services lead to unequal levels of inflation for different households. In this paper, we measure the degree of inequality in inflation across U.S. households for the period 1987-2001.

Our results suggest that the inflation experiences of U.S. households vary significantly. Most of the differences can be traced to changes in the relative prices of education, health care, and gasoline. We find that cost of living increases are generally higher for the elderly, in large part because of their health care expenditures, and that the cost of living for poor households is most sensitive to (the historically large) fluctuations in gasoline prices. To our surprise, we also find that those households that experience high inflation in one year do not generally face high inflation in the next year. That is, we do not find much household-specific persistence in inflation disparities.

Keywords: consumption price inflation, inequality, household inflation rates.

---

Correspondence to Bart Hobijn, Domestic Research Function, Federal Reserve Bank of New York, 33 Liberty Street, New York, NY 10045 (e-mail: bart.hobijn@ny.frb.org). The authors would like to thank Conchita D'Ambrosio, Rebecca Sela, and seminar participants at DIW, Berlin, and the Federal Reserve System Meetings on Microeconomics for their comments and suggestions. 

# *1. Introduction*

The Consumer Price Index (CPI) measures the continuously changing cost of the basket of goods and services purchased by the 'typical' American household. Price changes in this 'typical' goods basket have fluctuated dramatically over the past five decades. Inflation peaked at about 15% in the early eighties and was actually negative in the early fifties. Inflation is generally considered a macroeconomic variable and most of macroeconomic theory treats CPI inflation as being faced by all households.

In this paper we dive under the surface of the CPI as a measure of aggregate inflation and address the question of how representative the 'typical' household is on which the aggregate CPI measure is based. Because the basket actually purchased by each household potentially differs from the CPI's basket, the inflation rate faced by any given household might be very different from the CPI-inflation rate. The focus of this paper is on the evolution of the distribution of household specific inflation rates over time, namely over our sample period of 1987-2002.

Ever since the report by Boskin et al. (1996) there has been a renewed emphasis on the potential measurement biases in the CPI. The report was followed by, among others, Bils and Klenow (2001) and Lebow and Rudd (2003). However, the question whether the CPI actually reflects the inflation experience of the average American household has gotten less attention. We are not the first to address the issue of the effect of unequal inflation across American households. Earlier studies by Michael (1979) and Hagemann (1982) have also focused on differences in changes in the cost of living across U.S. households. Later studies, like Amble and Stewart (1994), Garner et al. (1996), Idson and Miller (1999), and Hamilton (2001) have focused on the inflation experience of particular groups.

Our analysis in this paper is in large part an update and refinement of the results presented in Michael (1979) and Hagemann (1982). Both of these studies use data from the 1970's. Our approach here is closely related to a recent study for the United Kingdom by Crawford and Smith (2002).

Throughout this paper we present a broad array of evidence on the degree of inequality of inflation across American households. In order to place this evidence in the right context, we illustrate how our results can be interpreted as summary statistics of the cross-household distribution of inflation rates. In our analysis we take a stepwise approach. In each step we address a slightly different question about inflation-inequality that we address, we review and introduce the relevant summary statistics of the cross-household inflation distribution, and then present the results for these statistics for our data for 1987-2001. Our data are taken from both the Consumer Expenditure Survey as well as the Consumer Price Index.

The results in this paper confirm some of the earlier results on U.S. inflation inequality from the 1970's. Namely, there are large differences in the inflation experience across households in the United States over the period 1987-2001. Major contributors to these disparities are, on the upside, increasing costs of education, and health care. Downside contributors are apparel prices and prices of household equipment. A large part of the fluctuations in these disparities turns out to be due to the most volatile price off all, which is the gasoline price index. Especially elderly households as well as households without kids are generally confronted with a higher than average inflation rate. Because they spend a relatively large part of their income on gas, poor households seem to be particularly hard hit in terms of inflation when gasoline price

skyrocket. To our surprise, we do not find a high degree in persistence of household specific inflation rates. That is, households that are confronted with higher than average inflation in one year are not extremely likely to be confronted with the same inflation disparity the year afterwards.

The structure of the rest of this paper is as follows. In Section 2 we introduce our definition of a household specific inflation rate and illustrates in which respects it differs from the CPI-U, as measured by the Bureau of Labor Statistics. In Section 3 we present alternative measures of aggregate inflation and discuss the reasons for their differences. The main reason is the underlying inequality in household inflation rates. In Section 4 we describe the main properties of the cross-household inflation density over our sample period. Section 5 focuses on the sources of inequality in household inflation rates and documents disparities in expenditure patterns as well as price changes across various goods categories. In Section 6 we consider whether there are certain groups of households that tend to face higher inflation than the rest of the population. Finally, in Section 7 we consider whether households that face higher than average changes in their cost of living in one year have the same tendency the next year. Finally, in Section 8 we conclude.

## *2. CPI and household specific inflation rates*

At the heart of our analysis in this paper is the concept of a 'household specific inflation rate'. In this section we explain our definition of this term and show how it is related to the definition of inflation as commonly calculated by the CPI. As it turns out, it is easiest to start off with the definition of inflation as it is measured by the CPI and then illustrate the respects in which our definition of household inflation deviates from the CPI measure.

In principle we would like to measure the proper changes in the cost of living for each household. It is well known from price index theory that calculating an exact index of the cost of living is not feasible, however. See Diewert (2001) for an extensive survey of the Consumer Price Index and index number theory.

To allow for comparison with the published CPI constructed by the Bureau of Labor Statistics (BLS), we will approximate a household's change in its cost of living by a Laspeyres price index. The inflation rates that we consider for each household are slightly different from those constructed for the CPI by the BLS.

Throughout, we will assume that the relevant change in the cost of living is calculated by combining the price changes of $m$ goods categories[1]. The overall CPI measures inflation in period $t$, which we will denote by $\pi_t^{CPI}$, as the ratio of weighted averages of the percentage price changes of each of the item strata between period $t$ and a base period in the numerator and between period $t-1$ and a base period in the denominator. Let $p_{j,t}$ be the price index for item stratum $j$ at time $t$, and let $t=b$ denote the base period. Furthermore, let $w_{j,b}$ be the aggregate expenditure share of goods category $j$ in the base period. Using this notation, CPI inflation is measured as

[1] A goods category is also often referred to as an 'item stratum' (plural is 'item strata'). We will both terms interchangeably.

$$\pi_t^{CPI} = \frac{\sum_{j=1}^{m} w_{j,b} \frac{p_{j,t}}{p_{j,b}}}{\sum_{j=1}^{m} w_{j,b} \frac{p_{j,t-1}}{p_{j,b}}} - 1 \tag{1}$$

The above equation implies that CPI inflation measures the percentage change in the price of the base period goodsbasket between periods *t-1* and *t*.

The Bureau of Labor Statistics updates the base period *b* relatively infrequently. Greenlees and Mason (1996) list the changes in the expenditure base period that have occurred for the CPI since 1940. From 1940 through 2000, the expenditure base period was changed 5 times. This infrequent updating of the base period is widely thought to be a major source of substitution bias in the CPI, because it implies that the CPI does not properly account for people substituting goods that become relatively cheaper for more expensive goods. See Lebow and Rudd (2001) for a recent discussion of this bias.

Our approach to calculating household specific inflation rates will differ from the approach chosen for the overall CPI in three important ways

One difference between our household specific inflation rates and the CPI inflation rate is that we update the expenditure weights in every time period, rather than use weights from some base period. Besides limiting the dependence of our results on the particular choice of base period, it also reduces the substitution bias in our calculations.

By updating the base period in every period, i.e. by setting *b=t-1* in (1), for the overall CPI we obtain an alternative measure of inflation of the form

$$\pi_t^{(I)} = \sum_{j=1}^{m} w_{j,t-1} \frac{p_{j,t}}{p_{j,t-1}} - 1 = \sum_{j=1}^{m} w_{j,t-1} \left( \frac{p_{j,t}}{p_{j,t-1}} - 1 \right) = \sum_{j=1}^{m} w_{j,t-1} \pi_{j,t} \tag{2}$$

where $\pi_{j,t}$ is the inflation measured for item stratum *j*. The result is that we are actually calculating a chain-weighted index.

When one applies (2) to monthly price data that are not seasonally adjusted, as we will do in section 3, there are large seasonal fluctuations in inflation rates. However, these seasonal fluctuations are not what we are interested in. There are, in principle, many ways to get rid off the seasonality in the calculated inflation rates. The approach that we choose in this paper is to consider annual inflation rates. That is, we do not compare current prices with those a month earlier, but rather twelve months earlier. This is a second major difference between CPI inflation and our household inflation rates. If *t* indexes time in months, making this change to the (2) gives us an inflation measure of the form

$$\pi_t^{(II)} = \sum_{j=1}^{m} w_{j,t-12} \left( \frac{p_{j.t}}{p_{j,t-12}} - 1 \right) = \sum_{j=1}^{m} w_{j,t-12} \pi_{j,t} \tag{3}$$

where the item stratum specific inflation, $\pi_{j,t}$, is now measures as a year/year inflation rate.

By using annual inflation rates, we get rid off the seasonal fluctuations in the stratum specific inflation rates, $\pi_{j,t}$. However, (3) still yields seasonal fluctuations in $\pi_t^{(II)}$ because of seasonal fluctuations in the

spending patterns reflected in the budget shares $w_{j,t}$. In practice, the remaining seasonal fluctuations turn out to be minimal.

Essentially, (3) represents the way we will define our household inflation rates. However, we will focus on household specific inflation rates and (3) does not contain any household specificity. In principle, we would like to measure household specific expenditure weights as well as household specific price changes. For each household, which we will index by $i$, we can and do observe its specific expenditure shares, $w_{i,j,t}$ for each of the $m$ goods categories. However, we are not able to observe the specific prices that households pay for the item strata. Therefore, we must assume that all households face the same price increases, $\pi_{j,t}$, for each item stratum. This is not to say that each item stratum has the same price increase in a given period, but that each household faces the same price increase as all other households *for any particular goods category* at each point in time. This is an assumption that is commonly made when constructing group price indices, as in Amble and Stewart (1994) and Garner et. al. (1996). There is very little empirical work that addresses the question whether different groups of household face different price changes for specific goods categories. The only evidence that we are aware of is the study by Berndt et. al. (1997) which suggests that the elderly do not face very different price changes in prescription drugs than other people, even though they tend to use very different drugs.

When we apply the assumption above, namely that households face the same price increases but that they choose different expenditure patterns in response to these prices, to (3) we arrive at our definition of a household inflation rate, which we donate as $\tilde{\pi}_{i,t}$ for household $i$ in month $t$. That is,

$$\tilde{\pi}_{i,t} = \sum_{j=1}^{m} w_{i,j,t-12} \pi_{j,t} \tag{4}$$

Here $w_{i,j,t-12}$ is household $i$'s expenditure share on good category $j$ twelve months before month $t$, while $\pi_{j,t}$ is the inflation in goods category $j$ over the year preceding month $t$. The summation runs over $j=1,\ldots,m$ because we have assumed that there are $m$ goods categories.

In sum, the household inflation rate that we measure represents the change in the price, over the past year, of the goods basket that a household bought a year earlier. We assumed that the increases in the prices of the various goods categories that the household faces are equal to the national average.

## *3. Aggregate inflation and the underlying distribution*

Now, on the one hand, we have the aggregate inflation rate measured by the CPI and, on the other hand, we have a set of household specific inflation rates. The latter are a sample from the distribution of household specific inflation rates that is the focus of this paper. It turns out that the CPI as well as other calculations presented in this paper can all be interpreted as summary statistics of this distribution of household specific inflation rates.

As a basis for the analysis in this paper we consider the joint density of household specific inflation rates and household characteristics. It is convenient to split the vector with household characteristics into the household's total expenditures, which we will denote by $y_{i,t-12}$, and its other characteristics, which we will denote by the vector $\mathbf{x}^*_{i,t-12}$. We will denote the joint density of the household specific inflation rate and these

characteristics by $g(\tilde{\pi}_{i,t}, y_{i,t-12}, \mathbf{x}^*_{i,t-12})$. Note that we consider the household characteristics at the beginning of the year over which the inflation rate is calculated.

General aggregate price indices do not consider household characteristics, besides total expenditures, in their calculation. Hence, in order to interpret some of the aggregate price indices that are studied and published it suffices to consider the joint distribution of household specific inflation rates and total expenditures. We will denote the density associated with this distribution by, $g_{\pi,y}(\tilde{\pi}_{i,t}, y_{i,t-12})$.

Based on this density, one can calculate two aggregate inflation rates. The first is known as the plutocratic price index while the second is known as the democratic price index.

The plutocratic index is a weighted average of all the household specific inflation rates, where each household's contribution is proportional to the household's total expenditure level. Mathematically, the measured inflation corresponding to the plutocratic price index $\tilde{\pi}^P_t$, is the sample equivalent of the population moment

$$\tilde{\pi}^P_t = \int \left[ \int \tilde{\pi}_{i,t} g_{\pi|y}\left(\tilde{\pi}_{i,t} \middle| y_{i,t-12}\right) d\tilde{\pi}_{i,t} \right] \left( \frac{y_{i,t-12} g_y\left(y_{i,t-12}\right)}{\int y g_y(y) dy} \right) dy_{i,t-12} \tag{5}$$

Here, the plutocratic weight

$$\left( \frac{y_{i,t-12} g_y\left(y_{i,t-12}\right)}{\int y g_y(y) dy} \right) \tag{6}$$

represents the share of households with total expenditure level $y_{i,t-12}$ in aggregate expenditures. The expression in square brackets represents the average inflation rate for households with total expenditure level $y_{i,t}$.

In practice, the calculation of the sample equivalence of (5) does not require the calculation of all the household specific inflation rates. Instead, it amounts to calculating a weighted sum of the item strata specific inflation rates. The weight of each stratum is equal to its share in aggregate expenditures. This is also essentially the way that the CPI is calculated, except that, as can be seen in (1), the CPI is not chained but instead is normalized to a baseyear. Because the CPI is also a plutocratic index, it weighs each household by a weight that its proportional to its expenditure level.

Prais (1959) already argued that there is, in principle, no reason to treat households asymmetrically in the measure of aggregate inflation. As an alternative to the plutocratic index, Prais (1959) proposed an index that measures aggregate inflation as an unweighted average of the household inflation rates. This aggregate is known as a democratic price index, which we will denote by $\tilde{\pi}^D_t$. It corresponds to the sample equivalent of the population moment

$$\tilde{\pi}^D_t = \int \left[ \int \tilde{\pi}_{i,t} g_{\pi|y}\left(\tilde{\pi}_{i,t} \middle| y_{i,t-12}\right) d\tilde{\pi}_{i,t} \right] g_y\left(y_{i,t-12}\right) dy_{i,t-12} = \int \tilde{\pi}_{i,t} g_\pi\left(\tilde{\pi}_{i,t}\right) d\tilde{\pi}_{i,t} \tag{7}$$

Thus, the democratic price index measures inflation as the mean of the marginal distribution of household specific inflation rates.

So, does it really matter whether we use the CPI, i.e. (1), our plutocratic index, i.e. (5), or the democratic index, i.e. (7), as measure of aggregate inflation? There is little or no theory to guide us to answer this question. It is simply a matter of what the data say. The data that we will use throughout this paper are obtained from two sources.

Data on household expenditures and demographic characteristics are obtained from the Consumer Expenditure Survey (CEX). The CEX is a quarterly rolling panel of about 5,000 households. Each household in the panel reports expenditure data for 4 consecutive quarters (if they respond on all interviews). In addition to the four interviews, the households also participate in an initial interview, in which they report demographic characteristics of the household and its members. In each quarter, one third of the panel is interviewed in each month. Each household reports on the expenditures made over the previous three months.

Price data are obtained from the CPI series for all urban consumers, for the specific goods categories that we choose. Matching the expenditure categories reported in the CEX and the CPI series, we ended up with $m=19$ categories. They are: Food at Home, Food Away from Home, Alcohol, Owned Dwellings, Rented Dwellings, Other Lodging, Utilities, Household Furnishings and Operations, Apparel, Vehicles, Gasoline, Other Vehicle Expenses, Public Transportation, Medical, Entertainment, Personal Care, Reading, Education, and Tobacco. Appendix A contains a detailed description of these categories, as well as details of the mapping between the CPI and CEX used to construct these categories. They match up closely with those used in most other studies using the CEX.

In principle, the results of the calculations presented in this paper depend on the level of aggregation of the strata. That is, our method takes into account between strata differences in expenditure patterns across households but not within strata differences. We have tried to use an as detailed as possible level of aggregation and the $m=19$ categories we came up with is the best match between CPI and CEX we were able to find. The resulting measured inflation rates cover the period of January 1987 through December 2001.

Figure 1 depicts the evolution of the CPI-U, as well as the plutocratic and democratic inflation rates over our sample period. Three main observations stand out from this figure.

First of all, the differences between these three measures of aggregate inflation are small compared to the fluctuations in these measures over time. Secondly, the differences between the plutocratic and democratic indexes seem negligible. The democratic index is higher in the period 1990-1992 and 1999-2001. These were both periods in which gasoline price inflation reached double digit levels. Hence, when gas prices go up at a double-digit rate, households that spend a relatively larger share of their expenditures on gas, read the households with lower expenditure levels, tend to be hit harder in terms of their cost of living. As we will see in much more detail later on, gas prices play an important role as a source of inflation inequality. Our results for the U.S. difference between the plutocratic and democratic index are similar to those obtained for Spain for 1992-1997 by Izquierdo et. al. (2002). They find that the average difference between the demo- and plutocratic indexes for Spain was less than 0.055%. Over the same period, we find that the average difference in the U.S. is 0.065%. However, over the whole sample period the difference is on average 0.275%, because the two periods with high gas price inflation are not included in the 1992-1997 period. Kokoski (2000) also concludes that there is not a substantial difference between inflation as measured by the

two indexes. Her results suggest an average difference of –0.05%, with yearly differences ranging from around –0.6% to 0.3%.

Finally, the difference between the plutocratic index and the CPI-U is on average 0.377%. It reached its maximum during the sample period at 1.108% in August 1991. The average difference between the CPI-U and our plutocratic index can be interpreted as an estimate of the average substitution bias in the CPI. Our estimate is very much in the range of the estimates of substitution bias found in the literature. See Lebow and Rudd (2003) for a comprehensive review of these estimates. One thing is worthwhile noting. That is that in some periods inflation as measured by the plutocratic index outpaces CPI-U inflation. This is true in the last quarter of 2000 for example. According to standard classical price index theory, introduced by Konüs (1939) and Frisch (1936), this can not be possible. However, in practice demand for some goods, gasoline in particular, might be so inelastic, or rather inflexible, that increases in the prices of these goods possibly result in a negative substitution bias. This is the case at the end of the 1990's when the aggregate expenditure share of gasoline went up over the decade and consumers were hit more by gas price increases than reflected in the CPI.

## *4. Shape and moments of cross-household inflation distribution*

In the previous section we have presented several alternative measures of aggregate inflation. We have shown how each of these can be interpreted as a summary statistic of the underlying distribution of household inflation rates, $g(\tilde{\pi}_{i,t}, y_{i,t-12}, \mathbf{x}^*_{i,t-12})$. In this section we dive below the surface of this distribution and consider what its main properties are and whether and how it has changed over time.

To start off we present the democratic mean that coincides with the inflation rate measured by a democratic price index together with some of the percentiles of $g_\pi(\tilde{\pi}_{i,t})$. The evolution of the mean, median, 5th and 95th percentiles of $g_\pi(\tilde{\pi}_{i,t})$ over time are plotted in Figure 2. The first thing that jumps out from this figure is that the mean and median inflation rates[2] are virtually identical over time, suggesting that the cross-household distribution of inflation rates is rather symmetric. The second observation from this figure is that there is a substantial variation in inflation rates across households. The width of the range between the 5th and 95th percentiles of the distribution varies between about 1% in 1996 to 6% in 1991.

In order to put this variation in perspective, we consider the behavior of the standard deviation of $g_\pi(\tilde{\pi}_{i,t})$ over time. This is depicted in Figure 3. The standard deviation varies between about 0.2% in early 1996, on the lower end, and 1.8% in the beginning of 1991 on the other hand. As it turns out here, as well as for many other results in this paper, the driving force behind the behavior of this standard deviation seems to be in large part gasoline prices. We will discuss the path and effect of gasoline prices over our sample period in more detail later.

The summary statistics above give us a perspective on some of the properties of the inflation distribution across households. However, they do not provide us with an insight in the particular shape of the distribution.

[2] The median inflation rate considered here is very different from the 'median CPI' introduced by Cechetti (1997). Our definition of median inflation considers the median across households, while the 'median CPI' measures the (weighted) median across item strata.

In order to consider the shape of the distribution, we present kernel density estimates of the inflation distribution for the beginning, middle, and end of our sample. That is, we present estimates of the distribution for February 1987, June 1994, and December 2001. These estimates are plotted in Figure 4. As the results for the mean and median already suggested, the distribution is relatively symmetric in all three months. Its tails are relatively thin, while in the middle it is almost triangular. The thinness of the tails is not surprising since by definition

$$\min_{j=1,\ldots,m} \pi_{j,t} \leq \tilde{\pi}_{i,t} \leq \max_{j=1,\ldots,m} \pi_{j,t} \tag{8}$$

Although our results cover a different period, it is worthwhile to compare them with the results presented by Michael (1979). Michael finds that the cross household inflation distribution for 1973 has many of the same properties that we found for the same distribution over our sample period. He also finds that that mean and median are almost the same, 0.1% apart, and finds a standard deviation of 2.33%. This is slightly higher than the standard deviation that we found and might be partly due to the oil crisis and the associated gas price inflation in 1973.

The general picture that emerges from the results in this section is that inflation rates did not only vary a lot over time but also across households. It is thus important to consider what causes these variations across households and whether there are particular types of households that tend to face higher or lower than average inflation rates. We will consider the sources of cross-household heterogeneity in inflation rates in the next section.

## *5. Sources of heterogeneity*

If households face different inflation rates, a natural question is whether we can pinpoint the source of this heterogeneity. In order to be able to do so, it turns out to be illustrative to reconsider (4). This equation implies that we measure a specific household's inflation rate as

$$\tilde{\pi}_{i,t} = \sum_{j=1}^{m} w_{i,j,t-12} \pi_{j,t} \tag{9}$$

such that the average inflation rate, i.e. the democratic mean, is

$$\begin{aligned} \pi_t^D &= \int \tilde{\pi}_{i,t} g_\pi \left( \tilde{\pi}_{i,t} \right) d\tilde{\pi}_{i,t} \\ &= \sum_{j=1}^{m} \left[ \int w_{i,j,t-12} g_{w_j} \left( w_{i,j,t-12} \right) dw_{i,j,t-12} \right] \pi_{j,t} \\ &= \sum_{j=1}^{m} \mu_{w_{j,t-12}} \pi_{j,t} \end{aligned} \tag{10}$$

where $\mu_{w_{j,t-12}}$ is the average expenditure share of item stratum *j*. This representation allows us to write the deviation of a specific household's inflation rate from the mean as

$$\begin{aligned}\left(\tilde{\pi}_{i,t}-\tilde{\pi}_t^D\right) &= \sum_{j=1}^{m}\left(w_{i,j,t-12}-\mu_{w_{j,t-12}}\right)\pi_{j,t} \\ &= \sum_{j=1}^{m}\underbrace{\left(w_{i,j,t-12}-\mu_{w_{j,t-12}}\right)}_{(A)}\underbrace{\left(\pi_{j,t}-\tilde{\pi}_t^D\right)}_{(B)}\end{aligned} \tag{11}$$

This decomposition illustrates that there are two things necessary for heterogeneity in household specific inflation rates.

First of all, there must be differences in inflation rates across item strata, as represented by part (B) of this decomposition. Since household specific inflation rates are a weighted average of the inflation rates of the item strata, if there is no difference in the cross-strata inflation rates then this weighted average does not depend on what weights are applied. Deviations for item strata specific inflation rates from the average inflation rate essentially represent changes in the relative prices of the various good categories. Hence, differences in household specific inflation rates occur as a result of differences in household expenditure patterns for goods for which their relative prices change.

Secondly, households must have different-from-average expenditure patterns, otherwise each household's inflation rate is based on the same expenditure weights and is thus the same. This is represented by part (A) in the decomposition, which is the deviation of the household's expenditure share from the average expenditure share.

The decomposition (11) also illustrates that what matters is the interaction between the deviations from average expenditures and inflation rate deviations. That is, if a household differs in its expenditure pattern in strata for which the inflation rate does not differ very much from the average inflation rate, then this will not cause its specific inflation rate to deviate far from the average. However, if a household spends a large share of its income on a goods category that exhibits far above average inflation, then this could possibly lead to a household specific inflation rate that is much higher than the average.

Hence, in order to get an idea what underlies the differences in inflation rates across households two things are of particular interest. First of all, we would like to know how big the variation is in expenditure shares for various strata across households. Basically, we would like to know which strata could be potentially important based on part (A) of the decomposition in equation (11). Secondly, we are interested in the relative price change for the different item strata over our sample period. That is, we would like to get some insight in the magnitude of part (B) of equation (11) for the goods categories in our sample. The results that pertain to these two issues our in Table 1 and Figure 5. The top line of Table 1 lists CPI-U weights of the various item strata. These are the 1993-1995 base period weights, i.e. the $w_{j,b}$'s from (1), which were used to calculate CPI inflation from 1998 through 2000.

### *Variation in expenditure patterns*

For our analysis of part (A) of equation (11) we performed an ANOVA of the expenditure shares of the households. That is, we decomposed the expenditure share variance according to

$$
\begin{aligned}
s^2 &= \frac{1}{\sum_{t=1}^{T} n_t} \sum_{t=1}^{T} \sum_{i=1}^{n_t} \left( w_{ijt} - \overline{w}_j \right)^2 \\
&= \underbrace{\frac{1}{\sum_{t=1}^{T} n_t} \sum_{t=1}^{T} \sum_{i=1}^{n_t} \left( w_{ijt} - \overline{w}_{jt} \right)^2}_{\text{within period variance}} + \underbrace{\frac{1}{\sum_{t=1}^{T} n_t} \sum_{t=1}^{T} n_t \left( \overline{w}_{jt} - \overline{w}_j \right)^2}_{\text{between period variance}}
\end{aligned} \tag{12}
$$

where $n_t$ is the number of households in the sample at time $t$,

$$
\overline{w}_{jt} = \frac{1}{n_t} \sum_{i=1}^{n_t} w_{ijt} \tag{13}
$$

is the average expenditure share of category $j$ at time $t$, while

$$
\overline{w}_j = \frac{1}{\sum_{t=1}^{T} n_t} \sum_{t=1}^{T} n_t \overline{w}_{jt} \tag{14}
$$

is the average expenditure share over the whole sample period.

The ANOVA gives us an insight into the relative importance of fluctuations of average budget shares of different strata over time, i.e. the between period variance, versus the variation in budget shares across households, i.e. the within period variance. Equation (11) suggests that it is the latter that is necessary for cross-household heterogeneity of inflation rates.

The second row of Table 1 lists the average shares $\overline{w}_j$, while the third and fourth row list the within and between period variance of the expenditure shares respectively.

It turns out that the CPI weights for some categories deviate quite a lot from the average expenditure shares in our sample. The most remarkable differences are the expenditures on food at home as well as on owned dwellings. There are a few aspects that contribute to these differences. First of all, the average share that is measure for the ANOVA analysis is not the same as the expenditure weighted CPI-weight. Essentially, the CPI-weights are plutocratic average shares while the average shares that we report are democratic average shares. Secondly, for owned dwellings the expenditure measures in the CEX do not coincide with the imputed rents used for the CPI's owners equivalent rent measure. Another important difference between the CPI weights and the average shares that we report is that for gasoline. For gasoline the average share measured is much higher than the CPI-weight in large part because the average share assigns a relatively higher weight to people with lower expenditure levels. These are the people who tend to spend a bigger chunk of their money on gas.

For the variances we find that the cross-household variation of expenditure shares by far outweighs the variation in average expenditure shares over time. That is, the within period variance for most categories is more than 100 times higher than the between period variance. Consequently, part (A) of decomposition (11)

gives rise to potentially big differences in cross household rates. Several item strata stand out as having a particularly high within period variation of shares.

Food at home has a high within period variation. So do owned and rented dwellings. This is the case because a household is most likely either to own or to rent their home. Hence, a household's expenditure shares of these categories are generally positive for one and zero for the other. Finally, vehicle expenditures tend to vary a lot across households because households that buy a new or used car tend to spend practically their whole budget on the car in a month while households that are just driving their cars spend much less, if not zero.

### *Cross-strata variation in inflation*

As we showed in the decomposition in equation (11), large differences in household specific inflation rates can only occur when (A) households have substantially different expenditure patterns, and (B) there are large relative price changes across item strata. In this subsection we will focus on part (B) and present some descriptive statistics on the inflation time series for the various goods categories.

This evidence is presented in two forms. The fifth and sixth rows of Table 1 list the average inflation rate and standard deviation of the inflation rate for each of these series for the period December 1985 through December 2000. The importance of these series in the CPI is reflected by the listed CPI weights. Finally, Table 1 also contains the correlations between the item strata inflation rates over the same time period.

It is evident from Table 1 that there are substantial differences between the average inflation rates across strata. On the lower end, with an average inflation rate below 2%, there are household operations and equipment, apparel and vehicles. On the high end we find health care (5.57%), education (6.61%), other lodging (5.26%) and tobacco (8.75%) [3]. Tobacco inflation rates, however, are mainly driven by excise tax increases, as we will see later on in this subsection. These major differences in the average inflation rates of the various goods categories suggest that part (B) of the decomposition in (11) at least gives rise to large potential inflation differences.

Only considering average inflation rates, however, does not tell the whole story. The standard deviations reported in Table 1 suggest that there are large fluctuations over time in the inflation rates of the various goods categories. Furthermore, these fluctuations are highly correlated for the various categories. Many strata inflation rates have cross-correlation coefficients of 0.6 or higher. Suggesting that the inflation rates for these categories have more than 60% of their fluctuations in common. This is not completely surprising if one believes that common price shocks, due to monetary policy actions, are the most important causes of inflation. There are a few categories that exhibit much lower correlation with other strata. These are the gas prices, and the price of public transportation. Oddly enough these are not highly correlated with each other either. Furthermore, tobacco price changes, for which most of the fluctuations are caused by changes in tax laws, hardly seem to be correlated with price changes for other strata.

Since there are such big fluctuations in cross-strata inflation rates, it seems worthwhile to consider inflation behavior over time for these categories. Doing so allows us to determine the periods in which part

[3] 'Other lodging' is comprised mainly of hotel and motel expenses, and living accommodations away from home for students.

(B) of decomposition (11) has the biggest potential of driving cross-household heterogeneity in inflation rates.

Figure 5 plots the inflation rates for the 19 goods categories in deviation from overall CPI-U inflation. It also plots CPI-U inflation for the sample period, as well as the sample standard deviation of the cross-strata inflation rates over time. It is worth noting that the cross-strata standard deviation is fairly constant at about 2.5% for most of the sample period. It shoots up in 1991 in response to gas price increases. It also increases during the 1998-2000 period in reaction to gas price increases again as well as taxes on tobacco. For two of the other item strata that had high average inflation besides tobacco, i.e. health care and education, we find the following. Medical care inflation was between 3% and 6% higher than CPI-U inflation for most of the 1987-1994 period. However, after 1994 the inflation rate for medical care exceeded overall inflation by only between about 1% and 2%. For education, we find that its inflation rate is persistently between 2% and 6% higher than overall inflation over the whole sample period, suggesting that families with (college-age) children might have been hit harder by inflation for our period.

There are many categories whose inflation rate fairly closely follows the overall CPI-U. Perhaps not surprisingly, this is true for owned dwellings, which constitutes about 21% of the overall CPI (see Table 1). Other such categories include food at and away from home, rented dwellings, entertainment, personal care and reading show, in general, relatively small (between –2% and 2%) deviations from the overall CPI.

These categories comprise a total of 53.13% of the CPI. Because these strata have inflation rates that do not differ that much from the overall CPI, differences in expenditure shares of households for these categories will cause very little differences in the specific inflation rates that these households face. That is, for these 52.13% of expenditures, part (B) of decomposition (11) will not be a significant source of the differences in household inflation rates.

Of the remaining categories, inflation rates seem to be substantially different from the overall CPI-U (at least at certain points) and are a likely source of cross-household differences in inflation rates. In other words, these are the categories for which part (B) of decomposition (11) is a potential source of different inflation rates across households. The most important item strata in this regard are those that are consistently higher or consistently lower than the CPI. From figure 1 it can be seen that medical care, education and tobacco (consistently higher) and apparel, household operations and equipment, new & used vehicles (consistently lower) are the categories that most fit this description. Gasoline prices are by far the most volatile.

## *6. Inflation differentials and household characteristics*

So far, we have used our results to illustrate that there is a very substantial inequality in the changes of the cost of living that individual households are confronted with. Is it possible to pinpoint particular groups of households that especially and consistently face a very different change in their cost of living than the representative household captured in the CPI-U? In order to answer this question we slice our sample according to various household characteristics and calculate group price indices.

Pollak (1980) was one of the first to propose the use of group price indices. Amble et. al. (1994) and Garner et. al. (1996) calculate such group indices for elderly and the poor respectively. In order to be

comparable with the CPI, which is a plutocratic index, these group indices are generally constructed as plutocratic indexes as well.

Let a specific group be defined as having household characteristics in a set $G$. That is, these are the households with $\mathbf{x}_{i,t-12} \in G$. The inflation rate measured by their plutocratic group price index is the sample equivalent of the population moment

$$\tilde{\pi}_t^G = \int_{\mathbf{x}_{i,t-12} \in G} \left[ \int \tilde{\pi}_{i,t} g_{\pi|y,\mathbf{x}} \left( \tilde{\pi}_{i,t} \middle| y_{i,t-12}, \mathbf{x}_{i,t-12} \right) d\tilde{\pi}_{i,t} \right] \left( \frac{y_{i,t-12} g_{y|\mathbf{x}} \left( y_{i,t-12} \middle| \mathbf{x}_{i,t-12} \right)}{\int y g_{y|\mathbf{x}} \left( y \middle| \mathbf{x}_{i,t-12} \right) dy} \right) dy_{i,t} \left( \frac{g_{\mathbf{x}} \left( \mathbf{x}_{i,t-12} \right)}{\int_{x \in G} g_{\mathbf{x}} (\mathbf{x}) d\mathbf{x}} \right) d\mathbf{x}_{i,t-12} \quad (15)$$

This equation consists of three parts. The first part, i.e.

$$E\left[ \tilde{\pi}_{i,t} \middle| y_{i,t-12}, \mathbf{x}_{i,t-12} \right] = \left[ \int \tilde{\pi}_{i,t} g_{\pi|y,\mathbf{x}} \left( \tilde{\pi}_{i,t} \middle| y_{i,t-12}, \mathbf{x}_{i,t-12} \right) d\tilde{\pi}_{i,t} \right] \quad (16)$$

is the expected inflation rate faced by a household with expenditures $y_{i,t-12}$ and characteristics $\mathbf{x}_{i,t-12}$. The second part is

$$\int E\left[ \tilde{\pi}_{i,t} \middle| y_{i,t-12}, \mathbf{x}_{i,t-12} \right] \left( \frac{y_{i,t-12} g_{y|\mathbf{x}} \left( y_{i,t-12} \middle| \mathbf{x}_{i,t-12} \right)}{\int y g_{y|\mathbf{x}} \left( y \middle| \mathbf{x}_{i,t-12} \right) dy} \right) dy_{i,t} \quad (17)$$

which is the inflation rate given by a plutocratic index for households with characteristics $\mathbf{x}_{i,t-12}$. Finally (15) is obtained by integrating out the household characteristics over the set of characteristics, $G$, that defines the group under consideration.

We will focus on five particular divisions of the sample of households: (1) households that live in an urban versus a rural area, (2) households whose reference person is white versus non-white, (3) the elderly versus the non-elderly (a.k.a. youngerly), (4) poor versus other households, and (5) households with kids less than 18 years old and other households. The results for the group price indexes are for these five groups are presented in Figure 6. This figure consists of two columns of panels and a column explaining the definition of each of the groups on which the index is conditioned. The first column of panels plots the path of the estimated inflation rates for the group in question and its complement. The second column of panels plots the path of the difference in the estimated inflation rates. This difference is the dashed line of the left-column panel minus the solid line. Because of the relatively small sample sizes that we are dealing with, these differences tend to be a bit noisy over time.

The top row of panels of Figure 6 contain our results for urban versus rural households. It turns out that neither of these groups structurally faces a higher inflation rate than the other[4]. When one compares the inflation differential plot with the deviation of gasoline price inflation from the overall CPI-U, i.e. panel 11 from Figure 5, it is obvious that the main determinant of the inflation differential between urban and rural

[4] It has to be noted that the stratum specific inflation rates that we use for this calculation are assumed to be equal for urban and rural areas and are assumed to be equal to the nationwide urban stratum specific inflation rate.

areas is gas prices. It roughly seems that for every 10% that gas price increases are higher than CPI-U inflation rural households face about a 0.15% higher increase in the cost of living than urban ones.

Using PSID data and only two goods categories, i.e. food and the rest, Hamilton (2001) found that blacks tended to face lower inflation than whites over the period 1974 through 1991. However, when we compare whites and non-whites and use more recent data, we do not arrive at the same finding for our sample period. The second row of panels in Figure 6 depicts the plutocratic group price indices for white and non-white households, which are households for which the reference person self-identifies as white or not. The results in these panels suggest that there is no particular persistent difference in the changes in the cost of living faced by whites and non-whites.

The question (addressed in the introduction) of whether the elderly face different inflation from other groups is particularly important because Social Security benefits are indexed to the CPI-W. Hence, if the CPI-W does not properly reflect the cost of living changes that the elderly face, then the CPI-W would not be the appropriate price index on which to base Social Security indexation. In fact, a current proposal in the House of Representatives (*H.R.2035,*2001) would require the Bureau of Labor Statistic to produce a separate CPI for the elderly (CPI-E).

Would the inflation measured by such a CPI-E be very different from the CPI-W? Given our analysis, we are able to address this question. We use CPI-U data rather than CPI-W, but for the rest we can construct our equivalent of a CPI-E. This follows up on earlier results by Amble and Stewart (1994), who also calculated a inflation rate for the elderly and found that for the end of the 1980's the inflation by this measure was between 0.2% and 0.4% higher than the inflation measured by the overall CPI-W[5].

The fourth row of panels of Figure 6 depict our results on a group price index for the elderly. Similar to Amble and Stewart (1994), we define 'the elderly' as households that have at least one reference person of age 60 or older. The lower-left panel of the figure is a similar to the ones we considered for poor vs. non-poor and white vs. other. It plots measured inflation for the elderly and the non-elderly. The lower-right panel plots the difference between these two inflation rates. Our measure of the difference between the two measured inflation rates is a bit noisy because of the relatively small size of our monthly samples.

Our results for the period 1987-1994 are very similar to those of Amble and Stewart (1994). The elderly faced an inflation rate that was generally between 0.2% and 0.4% higher than that of others. This can be seen from the lower-right panel of Figure 6. This difference appears to be mainly driven by higher medical care expenditure shares for the elderly. This share is about 10% for the elderly, which is almost twice as high as for the overall sample. Crawford and Smith (2002) find that, because of the presence of the National Health System in the U.K., elderly households in the U.K. do not face the same higher increases in their cost of living as in the U.S.. In fact, they face a slightly lower increase in their cost of living than other households.

The final group price index results that we present in the bottom row of Figure 6 are those for households with kids younger than 18 years old. These households seem to face lower inflation than other households in general. This is mainly due to these households generally having lower health care expenditures and not

[5] Hobijn and Lagakos (2003) also compare the CPI-E with the CPI-W and find that the average difference in inflation is about 0.38% over the period 1984-2001.

spending as big a share on education as households with college age kids. These households generally devote a big chunk of their expenditures on tuition. Hence, households with kids below 18 years old are less exposed to strata which exhibit the highest relative price increases and therefore generally tend to face lower household specific inflation rates. This confirms the evidence in Idson and Miller (1999) who also find that households with children face lower inflation over their 1969-1987 sample period..

Hence, we do find that some groups, most notably the elderly, generally face higher increases in their cost of living than other households. Instead of looking at particular groups we could also ask whether it is true that individual households that face higher than average inflation in one year tend to face higher than average inflation in the next.

## *7. Persistence of inflation differentials*

In this final section we present results about the persistence of household specific inflation rates. We address the issue of whether households that were hit by above average inflation in one year tend to suffer higher cost of living increases in the next year as well.

We are not the first to consider this issue. Michael (1979) found that for 1972-1973 there generally seemed to be a relatively high degree of persistence. The correlation between the household specific inflation rates in 1972 and 1973 was 0.65. He backs up this correlation with the estimation of an inflation transition matrix which considers the transition probabilities of households from one quintile of the inflation density in 1972 to another in 1973.

Quah (1993) performed a similar transition matrix estimation for the persistence of cross-country disparities in per capita income levels and then improved on it by doing a non-parametric transition density estimation in Quah (1997). We will follow Quah (1997) and present non-parametric kernel density estimates for the conditional distribution of the current deviation of a households inflation rate from average inflation conditional on last year's deviation of the households inflation rate from average inflation.

The problem of performing this analysis using the consumer expenditure survey is that households are only in the panel for at maximum a year. In order to compare household *i*'s inflation rate now with that a year from now using equation (4), we would need $w_{ijt-12}$ as well as $w_{ijt}$ for $j=1,\ldots,m$. However, because a household is only in the sample for a maximum of twelve months, we do not observe both of them. For our comparison of persistence we therefore compare fixed weighted measures of inflation by using $w_{ijt-24}$ for $j=1,\ldots,m$ to calculate $\tilde{\pi}_{i,t}$ as well as $\tilde{\pi}_{i,t+12}$. This means that our measure of $\tilde{\pi}_{i,t+12}$ (not $\tilde{\pi}_{i,t}$) for this analysis is more subject to substitution bias than that in previous parts of this paper. Because we do not fully account for substitution possibilities for $\tilde{\pi}_{i,t+12}$, we are likely to find more persistence of household specific inflation than is actually in the data.

Figure 7 plots the isoprobs of the conditional distribution (thin lines) as well as the conditional expectation (thick line). It also contains two dashed lines. The lines represent two polar cases. The first line is the 45-degree line which represents full persistence. The second line is the zero deviation from average inflation a year from now line and represents no persistence.

Full persistence would imply that the conditional expectation would equal the current deviation from the mean. In our figure this would mean that the conditional expectation would be on the 45-degree line plotted in the figure. What we observe is far from full persistence. In fact, for households that currently face below average inflation we find that the expected deviation from average inflation is virtually zero a year from now. This implies that on the downward side we do not find much persistence in a household's inflation rate at all. What is nice about the non-parametric estimation method that we apply is that it indicates that on the downside the conditional distribution of household inflation in deviation from the average a year from now conditional on the current deviation from average inflation is actually bimodal. One mode is close to zero and drive the result that there is little persistence when the conditional expectation is considered. The second mode is close to the 45-degree line and suggests that there is a fraction of households for which inflation is highly persistent. On the high side we do find some persistence, but very little.

Taken at face-value this result suggests that one does not need to have a major concern that particular households face consistently higher inflation than the overall CPI-U. We definitely find much less persistence of inflation rates across households than Michael (1979) did for 1972 and 1973.

It is important to bear in mind that this result is for the entire population. It does not preclude the possibility that certain segments of the population *do* in fact face persistence in above (or below) average inflation, but due to their small size as a group, this does not show up in our graph.

## *8. Conclusion*

Does the CPI-U capture the inflation experience of the average American household? The results in this paper seem to suggest that the answer to this question is affirmative. Apart from the substitution bias induced by the infrequent updating of the expenditure weights, the CPI methodology yields inflation estimates that closely follow the mean, median, and mode of the cross-household distribution of inflation rates.

However, household specific inflation rates tend to vary substantially around this mean. The disparities in household specific inflation rates are in large part due to relative price changes of three goods categories. These are education and health care which both exhibit persistently higher than average inflation and gasoline prices for which inflation tend to fluctuate wildly.

We find that the cost of living increases are generally higher for the elderly, in large part because of their health care expenditures, and that the cost of living of poor households is most sensitive to (the historically large) fluctuations in gasoline prices.

To our surprise we find that individual households that are confronted with high inflation in one year do not generally face high inflation in the subsequent year as well. That is, we do not find much household specific persistence in inflation disparities.

This combination of results leads us to believe that the CPI-U is a reasonable measure of aggregate inflation, but that one has to be careful when assuming that CPI-U inflation also accurately represents cost of living changes for particular subgroups, like the elderly. Furthermore, one might want to take into account that changes in the relative prices of certain goods, especially of gasoline, have very different effects on the cost of living of different households in the U.S..

## *References*


Amble, Nathan and Kenneth Stewart (1994), "Experimental Price Index for Elderly Consumers", *Monthly Labor Review*, May 1994, 11-16.

Berndt, Ernst, Iain Cockburn, Douglas Cocks, Arnold Epstein, and Zvi Grilliches (1997), "Is Price Inflation Different For The Elderly? An Empirical Analysis of Prescription Drugs", *NBER working paper 6182*.

Bils, Mark and Peter Klenow (2001), "Quantifying Quality Change", *American Economic Review*, 91, 1006-1030.

Boskin, Michael, Ellen Dulberger, Robert Gordon, Zvi Grilliches, and Dale Jorgenson (1996), "Toward a More Accurate Measure of the Cost of Living", *Final Report to the Senate Finance Committee*, December 4, 1996.

Bureau of Labor Statistics (1997), *Handbook of Methods*, Chapter 16: Consumer Expenditures and Income, http://stats.bls.gov/hom

Crawford, Ian, and Zoë Smith (2002), "Distributional Aspects of Inflation", *Institute for Fiscal Studies, Commentary 90*.

Cechetti, Stephen (1997), "Measuring Short-Run Inflation for Central Bankers", *Federal Reserve Bank of St. Louis Review*, May/June 1997, 143-155.

Diewert, Erwin (2001), "The Consumer Price Index and Index Number Theory: A Survey", *Department of Economics UBC discussion paper 01-02*.

Garner, Thesia, David Johnson, and Mary Kokoski (1996), "An Experimental Consumer Price Index for the Poor", *Monthly Labor Review*, September 1996, 32-42.

Greenlees, John, and Charles Mason (1996), "Overview of the 1998 Revision of the Consumer Price Index", *Monthly Labor Review*, December 1996, 3-9.

Frisch, Ragnar (1936), "Annual Survey of General Economic Theory: The Problem of Index Numbers", *Econometrica*, 4, 1-38.

Hagemann, Robert P. (1982), "The Variability of Inflation Rates across Household Types", *Journal of Money Credit and Banking*, 14, 494-510.

Hamilton, Bruce (2001), "Black-White Difference in Inflation: 1974-1991", *Journal of Urban Economics*, 50, 77-96.

Hobijn, Bart, and David Lagakos, (2003), "Social Security and the Consumer Price Index for the Elderly", *Current Issues in Economics and Finance*, May 2003, volume 9, number 5.

House of Representatives (2001), "Consumer Price Index for Elderly Consumers Act", *H.R. 2035.* (107th Congress)

Idson, Todd, and Cynthia Miller (1999), "Calculating a Price Index for Families with Children: Implications for Measuring Trends in Child Poverty Rates", *working paper*.

Izquierdo, Mario, Eduardo Ley and Javier Ruiz-Castillo (2002), "The Plutocratic Gap in the CPI: Evidence from Spain", *Working Paper*

Kokoski, Mary (2000), "Alternative CPI Aggregations: Two Approaches", *Monthly Labor Review*, November 2000, 31-39.

Konüs, A.A. (1939), "The Problem of the True Index of the Cost of Living", *Econometrica*, 7, 10-29.

Lebow, David and Jeremy Rudd (2003), "Measurement Error in the Consumer Price Index: Where Do We Stand?", *Journal of Economic Literature*, 41, 159-201.

Michael, Robert (1979), "Variation across Households in the Rate of Inflation", *Journal of Money, Credit, and Banking*, 11, 32-46.

Piachaud, David (1978), "Inflation and Income Distribution", in *The Political Economy of Inflation*, Fred Hirsch and John Goldthorpe (eds.), London: Martin Robertson & Co..

Pollak, Robert (1980), "Group Cost of Living Indexes", *American Economic Review: Papers and Proceedings*, 70, 273-278.

Pollak, Robert and Terence Wales (1992), *Demand System Specification & Estimation*, New York: Oxford University Press.

Prais, S.J. (1959), "Whose Cost of Living?", *Review of Economic Studies*, 26, 126-134.

Quah, Danny (1993), "Empirical Cross-Section Dynamics in Economic Growth", *European Economic Review*, 37, 426-434.

Quah, Danny (1997), "Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs", *Journal of Economic Growth*, 2, 27–59.

# *A. Data appendix*

## *Households*

We use the word 'household' to mean 'consumption unit,' the term used in the CEX. Any individual that makes his/her purchasing decisions alone, or any such group of people, be they related or unrelated, comprises a consumption unit. A married couple is one example. Two roommates that pool money and make purchasing decisions together is another. Each record in our data contains personal and expenditure data for one household. We use the CEX variable NEWID, which is unique for each household, as a household identifier.

## *Consumer Expenditure Survey*

The Consumer Expenditure survey is a comprehensive survey on the buying patterns of American Consumers. The survey consists of two components, a quarterly Interview Survey and a weekly Diary Survey. For this paper, we use only the quarterly Interview survey, for the years 1986 through 2000. In each quarter, approximately 5,000 households are interviewed. Each household in the survey is interviewed for five quarters consecutively (every three months). In the initial interview, information is collected on demographic and family characteristics, among other things. In the subsequent four interviews, expenditure data is collected for the three months prior to the month of the interview.

## *Price Indexes*

As much as possible we have matched up expenditure categories with CPI price indexes. In cases where expenditure categories do not match exactly with available CPI price indexes, we create our own indexes by combing the CPI series that most appropriately match our categories. The weights used to combine the CPI series are the base 'relative importance weights' in a base year, and those base year weights adjusted for changes in relative prices in all other years. In most cases base weights can be found in the BLS table 'Relative Importance in the CPI.' In all cases the CPI series we use are the non-seasonally adjusted 'US City Average' series for all urban consumers.

**Table A1.** Description of goods categories and matching between CEX and CPI

| | CEX Expenditure Category(s) | CPI Series |
|---|---|---|
| 1 | **Food at home** (food purchased at grocery or other food stores, excluding alcoholic beverages) | **Food at home** |
| 2 | **Food away from home** (excluding alcoholic beverages) | **Food away from home** |
| 3 | **Alcoholic Beverages** (purchased for consumption at or away from home) | **Alcoholic Beverages** |
| 4 | **Owned Dwellings** (includes interest on mortgages, property taxes, insurance related to housing, refinancing charges, expenses for repairs and maintenance, and other housing expenses) | **Owners Equivalent Rent of Primary Residence** |
| 5 | **Rented Dwellings** (includes rent, renter's insurance, expenses for repairs and maintenance, and other housing expenses) | **Rent of Primary Residence** |
| 6 | **Other Lodging** (lodging away from primary residence) | **Lodging away from home** (since 1997)<br>**Lodging while out of town** (before 1997) |
| 7 | **Utilities** (includes electricity, natural gas and other fuels, water, garbage collection, telephone charges) | **Fuels and Utilities**<br>**Telephone Service** |
| 8 | **Household Equipment** (furniture, household decorations, personal computers, household appliances)<br>**Household Operations** (domestic services, child care, etc) | **Household Equipment and Operations**<br>**Information Processing other than Telephone** (personal computers and peripherals, computer hardware & software) |
| 9 | **Apparel** (clothing purchases and upkeep) | **Apparel** |
| 10 | **Vehicles** (new and used cars, trucks and other vehicles) | **New and Used Motor Vehicles** (since 1993)<br>**New Motor Vehicles** (before 1993)<br>**Used Motor Vehicles** (before 1993) |
| 11 | **Gasoline** (and motor oil) | **Motor Fuel** |
| 12 | **Other Vehicle Expenses** | **Vehicle Parts and Equipment**<br>**Vehicle Maintenance and Repair**<br>**Motor Vehicle Insurance**<br>**Motor Vehicle Fees** |
| 13 | **Public Transportation** | **Public Transportation** |
| 14 | **Health Care** (health insurance, medical services, drugs, medical supplies) | **Medical Care** |
| 15 | **Entertainment** (includes fees and admissions, television, audio and video equipment, pets, toys, hobbies, other entertainment equipment) | **Recreation** |

**Table A1 (continued).** Description of goods categories and matching between CEX and CPI

| | CEX Expenditure Category(s) | CPI Series |
|---|---|---|
| 16 | **Personal Care** (includes hair products and services, cosmetic & bath products, other personal goods and services) | **Personal Care** (hair, dental and shaving goods and services, funeral expenses, financial services, laundry services) |
| 17 | **Reading** (includes magazine and newspaper subscriptions, books) | **Recreational Reading Materials** |
| 18 | **Education** (includes tuition and fees for universities, primary, secondary, and nursery schools; textbooks, educational equipment) | **Educational Books and Supplies**<br>**Tuition, Fees, and Child Care** |
| 19 | **Tobacco** | **Tobacco** |

**Table 1.** Average, standard deviation, CPI weight and correlations of strata inflation rates

| *j* | 1 Food at home | 2 Food away from home | 3 Alcoholic beverages | 4 Owned dwellings | 5 Rented dwellings | 6 Other lodging | 7 Utilities | 8 Household equip. and oper. | 9 Apparel | 10 Vehicles | 11 Gasoline | 12 Other vehicle expenditures | 13 Public transportation | 14 Health care | 15 Entertainment | 16 Personal care | 17 Reading | 18 Education | 19 Tobacco |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *CPI weight and ANOVA of expenditure shares* | | | | | | | | | | | | | | | | | | | |
| CPI weight (%) | 9.61 | 5.74 | 0.99 | 21.06 | 7.30 | 2.06 | 7.08 | 4.68 | 4.24 | 7.33 | 2.58 | 5.38 | 1.36 | 6.00 | 5.33 | 3.51 | 0.58 | 2.91 | 1.40 |
| avg. share (%) | 16.49 | 4.86 | 1.17 | 11.61 | 10.24 | 1.16 | 10.83 | 5.13 | 5.04 | 3.63 | 4.77 | 7.03 | 1.11 | 7.14 | 5.12 | 1.22 | 0.69 | 1.32 | 1.46 |
| $s^2_{within\ periods}$ * | 11.24 | 3.11 | 0.69 | 20.81 | 26.90 | 1.52 | 5.80 | 6.55 | 3.14 | 20.35 | 1.94 | 5.54 | 1.25 | 8.57 | 3.79 | 0.26 | 0.12 | 3.30 | 0.94 |
| $s^2_{between\ periods}$ ** | 5.93 | 1.15 | 0.33 | 17.38 | 2.05 | 0.90 | 5.55 | 2.64 | 8.07 | 3.47 | 2.20 | 1.37 | 0.16 | 2.02 | 2.40 | 0.03 | 0.10 | 3.04 | 0.22 |
| *mean and standard deviation of strata specific inflation rates over sample period* | | | | | | | | | | | | | | | | | | | |
| mean | 3.31 | 2.88 | 3.30 | 3.71 | 3.27 | 5.26 | 2.49 | 0.87 | 1.27 | 1.83 | 3.82 | 3.80 | 4.09 | 5.57 | 2.79 | 2.82 | 3.45 | 6.61 | 8.75 |
| std.dev. | 1.78 | 1.01 | 2.20 | 0.88 | 0.75 | 3.68 | 2.67 | 1.06 | 2.42 | 1.81 | 13.06 | 1.62 | 4.10 | 2.04 | 1.36 | 0.95 | 1.58 | 1.48 | 7.84 |
| *Correlations of strata specific inflation rates over sample period* | | | | | | | | | | | | | | | | | | | |
| 1 | 1 | 0.71 | 0.15 | 0.59 | 0.51 | 0.20 | 0.13 | 0.42 | 0.34 | 0.06 | 0.22 | 0.44 | 0.31 | 0.42 | 0.61 | 0.56 | 0.24 | 0.24 | -0.05 |
| 2 | 0.71 | 1 | 0.46 | 0.80 | 0.75 | 0.39 | 0.04 | 0.41 | 0.60 | -0.12 | 0.09 | 0.57 | 0.15 | 0.56 | 0.72 | 0.72 | 0.27 | 0.38 | 0.24 |
| 3 | 0.15 | 0.46 | 1 | 0.41 | 0.36 | 0.73 | 0.20 | 0.47 | 0.51 | -0.03 | 0.07 | 0.48 | 0.10 | 0.63 | 0.62 | 0.46 | 0.56 | 0.51 | 0.21 |
| 4 | 0.59 | 0.80 | 0.41 | 1 | 0.71 | 0.39 | 0.00 | 0.62 | 0.70 | 0.11 | 0.07 | 0.72 | 0.16 | 0.74 | 0.77 | 0.75 | 0.44 | 0.63 | 0.11 |
| 5 | 0.51 | 0.75 | 0.36 | 0.71 | 1 | 0.23 | 0.31 | 0.28 | 0.26 | -0.12 | 0.11 | 0.44 | 0.05 | 0.44 | 0.40 | 0.59 | 0.05 | 0.21 | 0.22 |
| 6 | 0.20 | 0.39 | 0.73 | 0.39 | 0.23 | 1 | 0.04 | 0.48 | 0.57 | 0.05 | 0.23 | 0.44 | 0.32 | 0.60 | 0.52 | 0.31 | 0.49 | 0.44 | 0.09 |
| 7 | 0.13 | 0.04 | 0.20 | 0.00 | 0.31 | 0.04 | 1 | 0.04 | -0.13 | -0.18 | 0.43 | 0.00 | 0.16 | 0.07 | 0.05 | 0.15 | -0.19 | -0.10 | -0.07 |
| 8 | 0.42 | 0.41 | 0.47 | 0.62 | 0.28 | 0.48 | 0.04 | 1 | 0.72 | 0.50 | 0.01 | 0.80 | 0.13 | 0.72 | 0.76 | 0.48 | 0.68 | 0.76 | -0.30 |
| 9 | 0.34 | 0.60 | 0.51 | 0.70 | 0.26 | 0.57 | -0.13 | 0.72 | 1 | 0.24 | 0.05 | 0.74 | 0.21 | 0.74 | 0.79 | 0.57 | 0.55 | 0.77 | -0.02 |
| 10 | 0.06 | -0.12 | -0.03 | 0.11 | -0.12 | 0.05 | -0.18 | 0.50 | 0.24 | 1 | -0.10 | 0.42 | 0.01 | 0.36 | 0.23 | -0.02 | 0.41 | 0.48 | -0.39 |
| 11 | 0.22 | 0.09 | 0.07 | 0.07 | 0.11 | 0.23 | 0.43 | 0.01 | 0.05 | -0.10 | 1 | -0.12 | 0.35 | 0.07 | 0.04 | 0.16 | -0.22 | -0.11 | 0.15 |
| 12 | 0.44 | 0.57 | 0.48 | 0.72 | 0.44 | 0.44 | 0.00 | 0.80 | 0.74 | 0.42 | -0.12 | 1 | 0.13 | 0.85 | 0.79 | 0.48 | 0.61 | 0.87 | -0.22 |
| 13 | 0.31 | 0.15 | 0.10 | 0.16 | 0.05 | 0.32 | 0.16 | 0.13 | 0.21 | 0.01 | 0.35 | 0.13 | 1 | 0.30 | 0.25 | 0.24 | 0.05 | 0.16 | -0.04 |
| 14 | 0.42 | 0.56 | 0.63 | 0.74 | 0.44 | 0.60 | 0.07 | 0.72 | 0.74 | 0.36 | 0.07 | 0.85 | 0.30 | 1 | 0.81 | 0.68 | 0.64 | 0.90 | 0.06 |
| 15 | 0.61 | 0.72 | 0.62 | 0.77 | 0.40 | 0.52 | 0.05 | 0.76 | 0.79 | 0.23 | 0.04 | 0.79 | 0.25 | 0.81 | 1 | 0.68 | 0.70 | 0.76 | -0.08 |
| 16 | 0.56 | 0.72 | 0.46 | 0.75 | 0.59 | 0.31 | 0.15 | 0.48 | 0.57 | -0.02 | 0.16 | 0.48 | 0.24 | 0.68 | 0.68 | 1 | 0.35 | 0.49 | 0.32 |
| 17 | 0.24 | 0.27 | 0.56 | 0.44 | 0.05 | 0.49 | -0.19 | 0.68 | 0.55 | 0.41 | -0.22 | 0.61 | 0.05 | 0.64 | 0.70 | 0.35 | 1 | 0.66 | -0.18 |
| 18 | 0.24 | 0.38 | 0.51 | 0.63 | 0.21 | 0.44 | -0.10 | 0.76 | 0.77 | 0.48 | -0.11 | 0.87 | 0.16 | 0.90 | 0.76 | 0.49 | 0.66 | 1 | -0.09 |
| 19 | -0.05 | 0.24 | 0.21 | 0.11 | 0.22 | 0.09 | -0.07 | -0.30 | -0.02 | -0.39 | 0.15 | -0.22 | -0.04 | 0.06 | -0.08 | 0.32 | -0.18 | -0.09 | 1 |

*Note:* * *multiplied by* $10^3$, ** *multiplied by* $10^5$

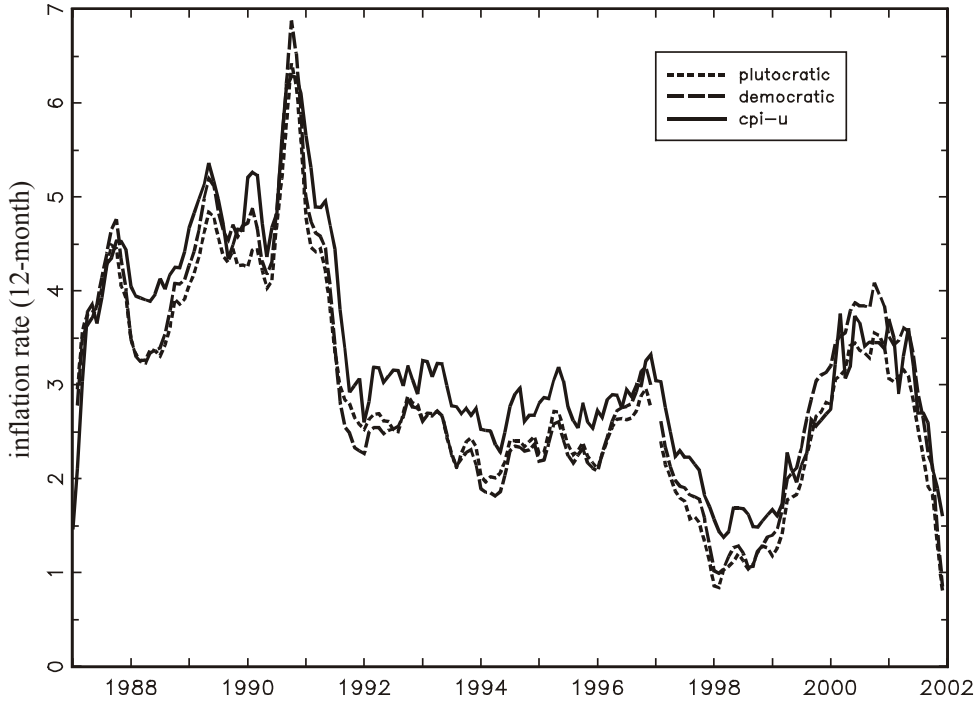


**Figure 1.** CPI-U and plutocratic and democratic inflation measures

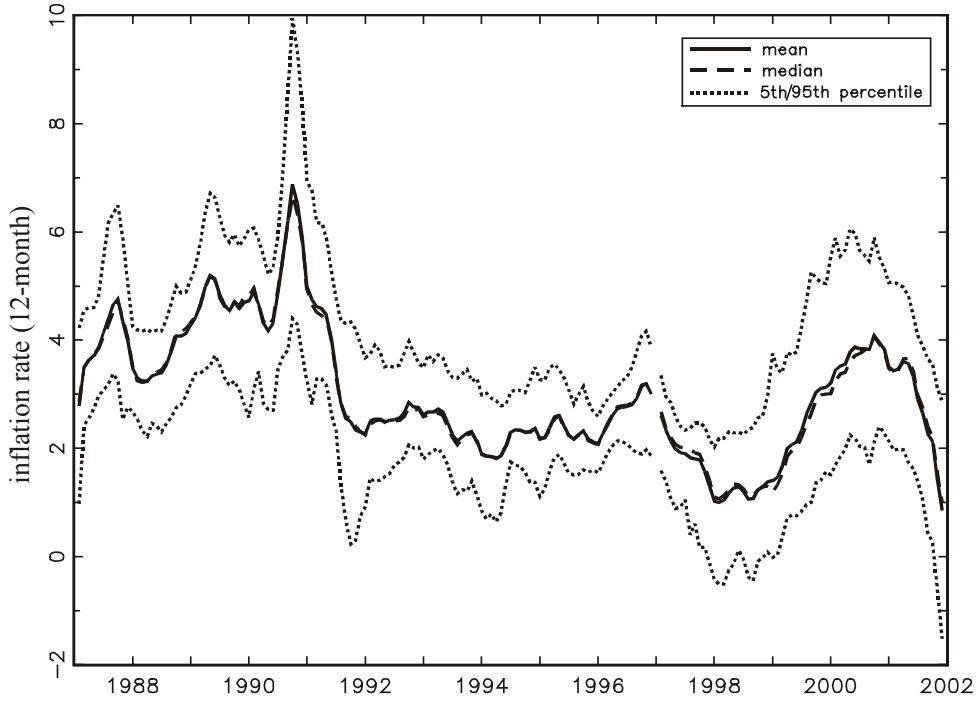


**Figure 2.** Moments of household inflation distribution

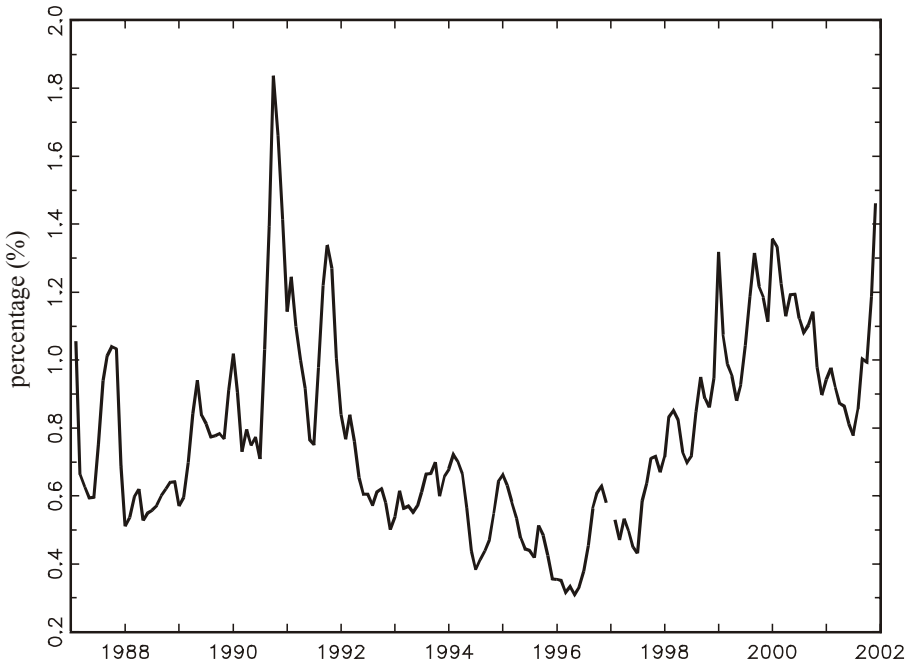


**Figure 3.** Standard deviation of the cross-household distribution of inflation

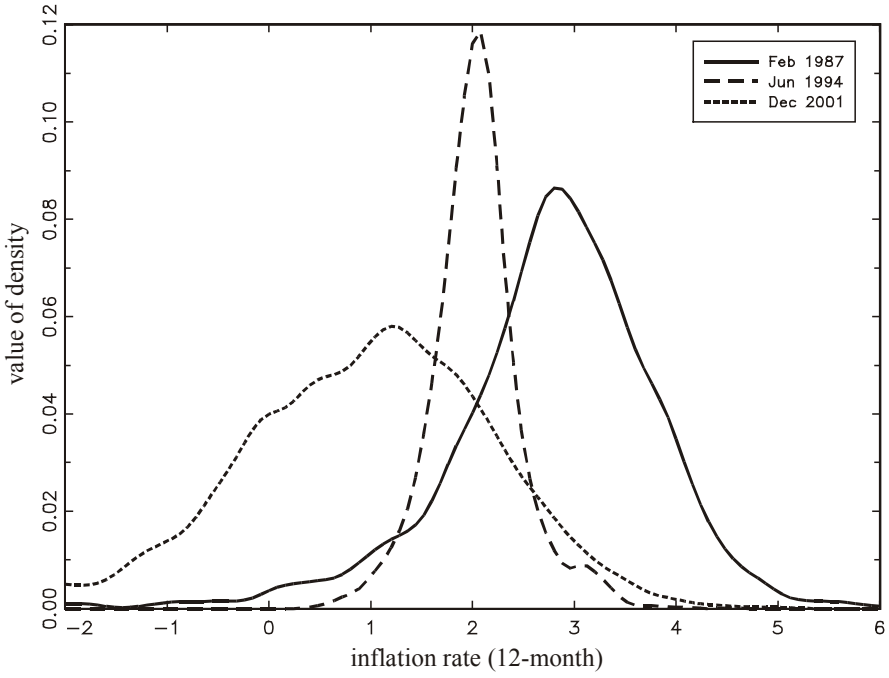


**Figure 4.** Cross-household distribution of inflation at beginning, middle, and end of our sample period

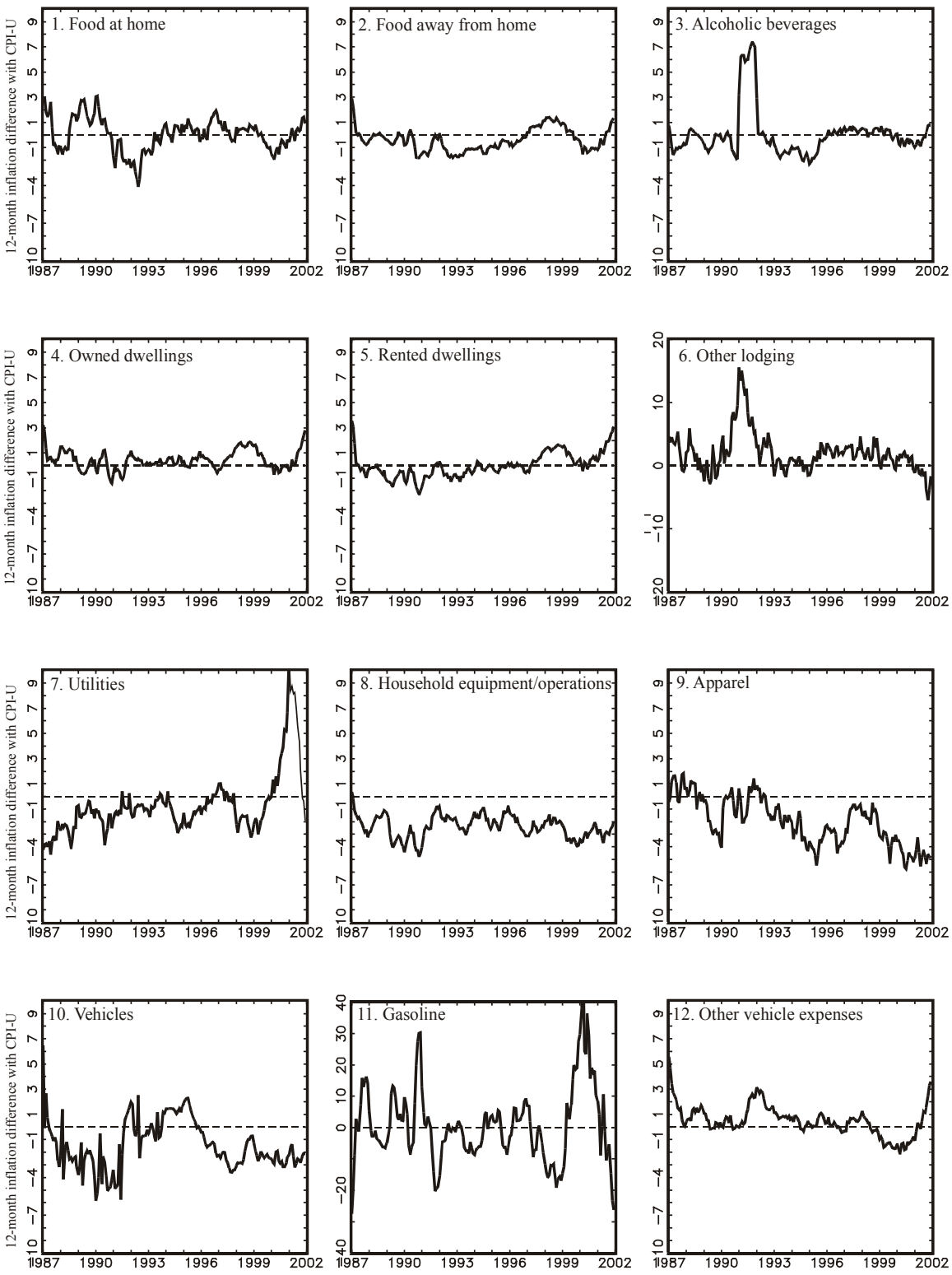


**Figure 5.** Cross-strata deviation of inflation from overall CPI-U

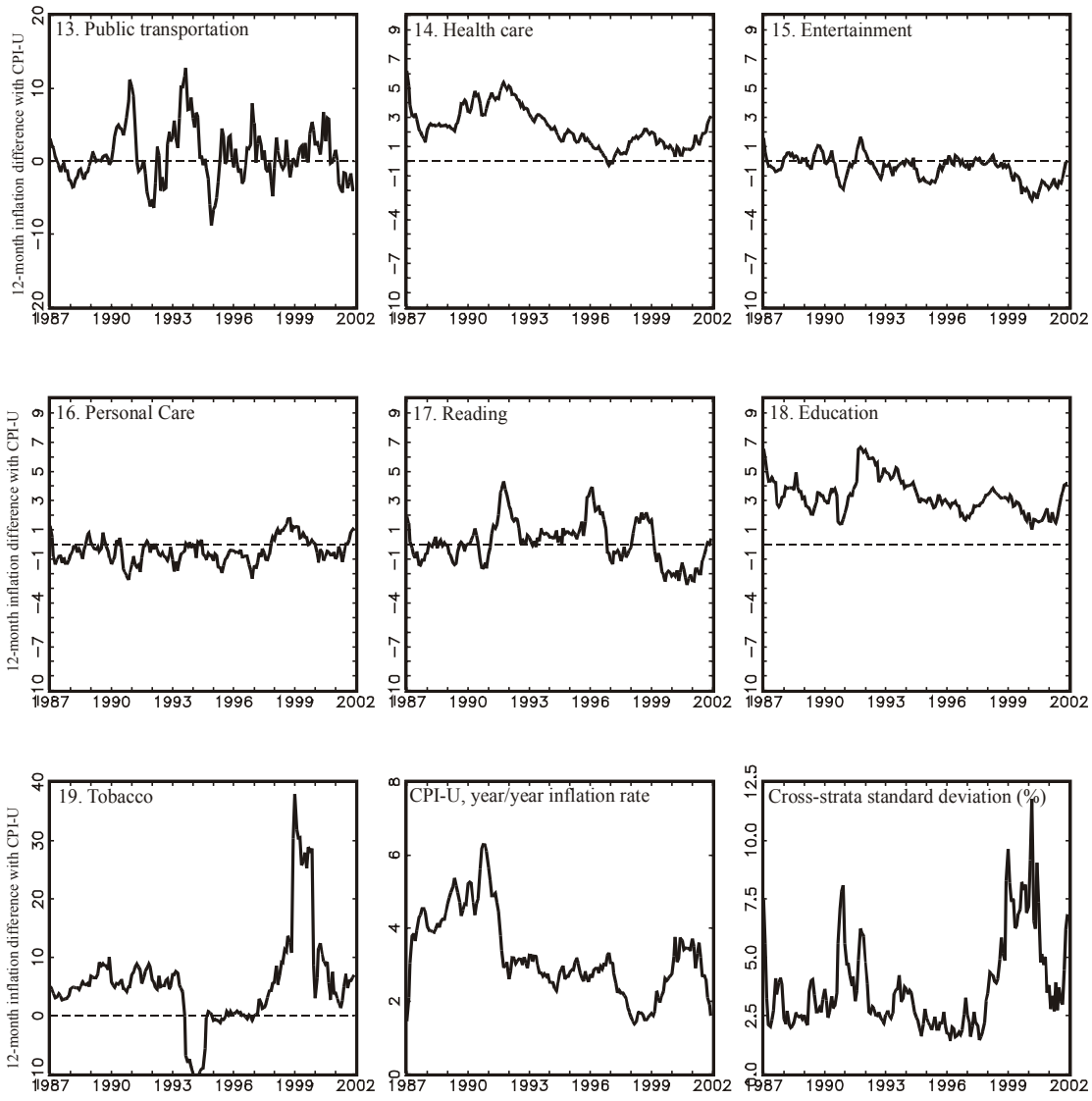


**Figure 5 continued.** Cross-strata deviation of inflation from overall CPI-U

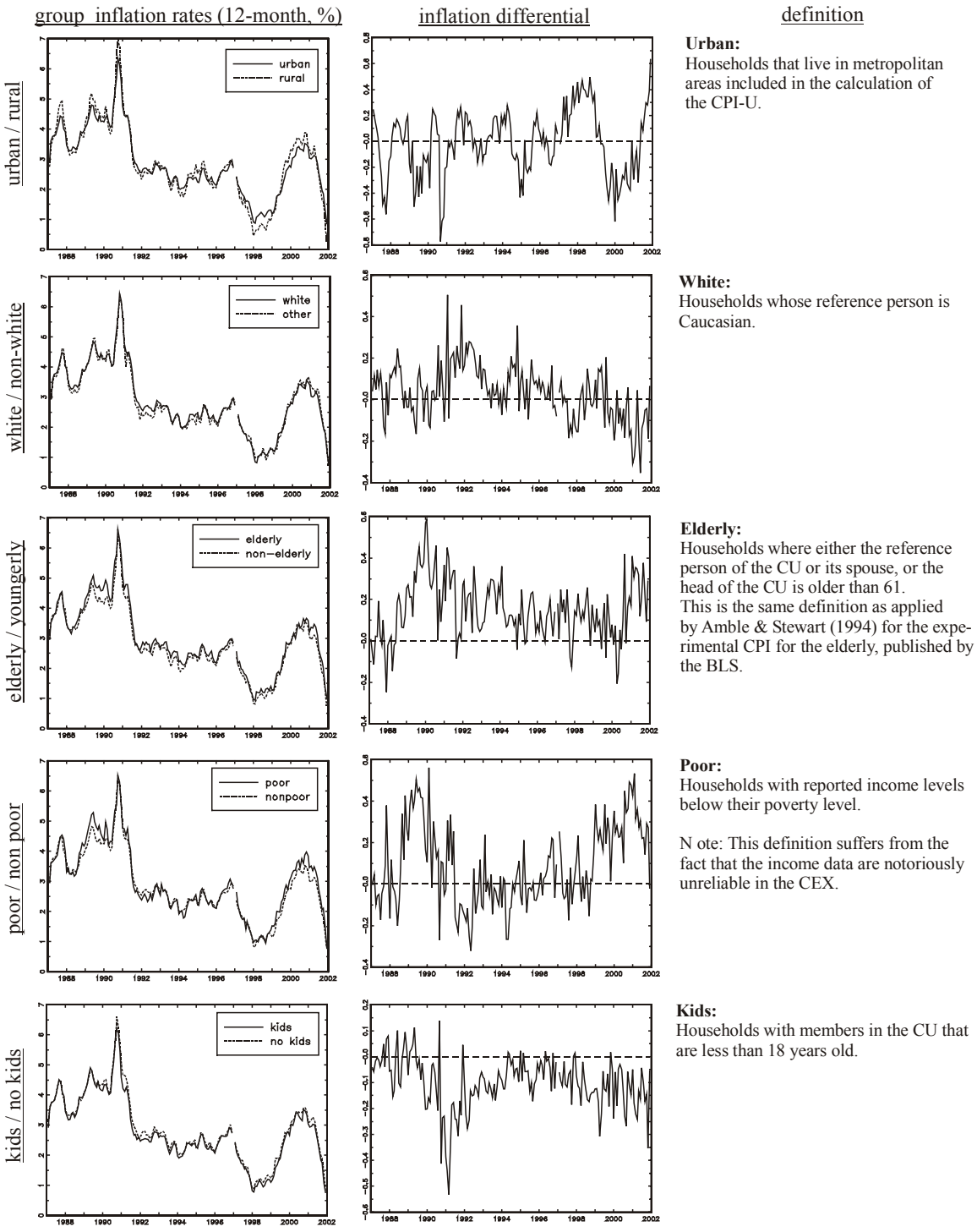


**Figure 6.** Group price indices and inflation differentials

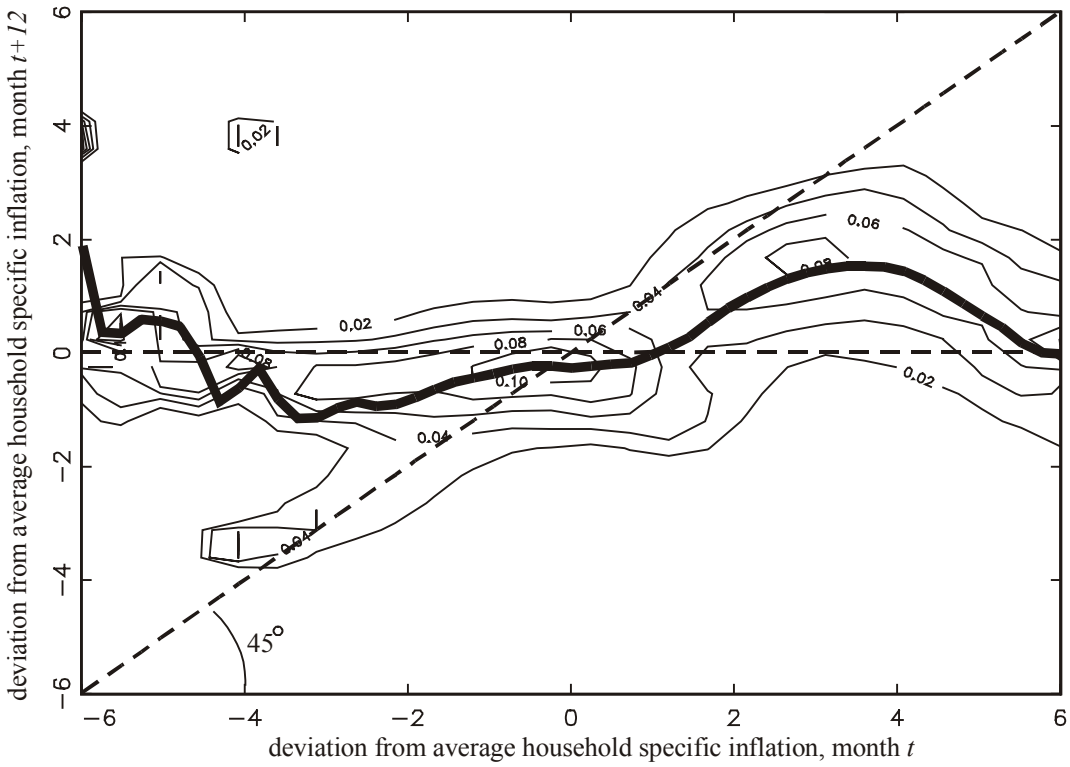


**Figure 7.** Persistence of inflation differentials across households