Unequal Inflation Rates Across American Households

BART HOBIJN AND DAVID LAGAKOS
Federal Reserve Bank of New York*

Version 1, June 2002
Preliminary and Incomplete
Please do not quote

Abstract

In this paper we consider whether households face very different inflation experiences because of their varying expenditure patterns and, if so, whether these differences in inflation can be partly explained by different household characteristics. We consider the evolution of the cross-household inflation distribution over the period 1987.02-2000.12 and revisit and update some of the main results on this question in the literature. We then proceed by estimating a consumer expenditure model, which allows us to calculate counterfactual inflation rates.

We have two main results. First of all, in all months in our sample we observe a substantial variation in inflation rates across households. Secondly, both conventional methods as well as our counterfactual inflation rates suggest that there are no particular household characteristics that have a persistent and large effect on household specific inflation rates.

Therefore, we conclude that measured inflation is not a structurally biased measure of inflation for specific groups of households.

Keywords: Consumption price inflation, expenditure models, household inflation rates.

Jel-codes: D1, E3, D6

* Correspondence to Bart Hobijn, Domestic Research Function, Federal Reserve Bank of New York, 33 Liberty street, 3rd floor. E-mail: bart.hobijn@ny.frb.org. All views expressed in this paper are those of the authors only and neither necessarily reflect those of the Federal Reserve Bank of New York nor those of the Federal Reserve System in general.
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. Because the basket actually purchased by each household differs (or at least potentially differs) from the CPI’s basket, the inflation rate faced by any given household might be very different from the CPI-inflation rate. Given this fact, it is worth considering the question of whether or not certain subsets of the population, like say, the poor, get hit consistently harder by inflation than the average American.

In this paper we examine the distribution of inflation rates across American households for the period from 1986 through 1999, and try to shed some light on the question just posed. We estimate an expenditure share model that explains how expenditure patterns are a function of household characteristics, and use it to calculate the contribution to inflation over our time period of a number of different demographic characteristics. This will help us better understand exactly why certain households face different inflation rates. Understanding whether and why certain households face persistently higher (or lower) inflation than the average American is important for several reasons.

First of all, the CPI is used for the indexation of several federal programs. Social Security and Supplemental Security Income (SSI) are two of them. If the beneficiaries of these programs persistently face higher than average inflation, then indexing these programs to the CPI will lead to a decrease in purchasing power for these households over time. This concern has spurred members of the House of Representatives (2001) to propose to construct a specific CPI for the elderly for the indexation of Social Security.

Secondly, we can gain insight in the magnitude of unintended re-distributive effects of monetary policy. The concern here is that a monetary expansion might adversely affect people who tend to buy goods and services whose prices react relatively quickly to increases in the monetary supply, while benefiting those who buy goods with rigid prices. See Piachaud (1978) for a more complete description of this phenomenon.

We are not the first to consider the topic of heterogeneity in inflation rates. Theoretical contributions were made by Prais (1959) and Pollak (1980).

One of the initial empirical contributions, by Tipping (1970), shows that those at the lower end of the income distribution faced higher inflation in Britain during the 1950s. For the U.S., however, Garner et al. (1996) find that there is little difference between the inflation rates for the poor and non-poor in the period 1984 through 1994. Besides income, other studies have focused on characteristics like race and age. For race, Hamilton (2001) finds that cumulative inflation for blacks was between 10 and 15 percentage points lower than for whites. His analysis only distinguishes between food and other expenses. For age, Amble and Stewart (1994) construct an experimental price index for the elderly and find that between 1987 and 1993, the price index for the elderly rose about 0.4% per year more than the CPI. Michael (1979) and Hagemann (1982) calculate price indexes for individual households for the period 1967 to 1974. They estimate the relationship between differences over time in these indexes and certain household demographic characteristics. They each conclude that, although for certain years certain household characteristics were
associated with higher inflation, these relationships did not hold up over time. In other words, no one group suffered disproportionately from inflation.

Our analysis in this paper differs from the previous literature in several important dimensions. First of all, we create a general framework in which the different methods applied in the previous literature can be interpreted. This framework can be used to show how all previously applied methods can essentially be interpreted as summary statistics of the joint distribution of household characteristics and household specific inflation rates. Secondly, we measure the previously proposed summary statistics as well as an estimate of the overall underlying relevant distribution. Thirdly, we provide estimates of counterfactual inflation rates based on an empirical consumption expenditure model. Finally, we use a more recent time period than previous studies, namely 1986-1999. Because our sample period is more recent than in most of the existing literature, we spend a considerable portion of this paper recalculating previous results for our sample period.

The structure of the rest of this paper is as follows. In Section 2 we introduce the general framework, based on the joint distribution of household characteristics and household specific inflation rates, which sets the stage for all of our subsequent analyses. In Section 3 we apply this general framework to address the question of whether households face very different individual inflation rates. We then calculate group specific price indexes to consider whether there are particular groups, like the elderly or the poor, which face persistently higher inflation rates. The empirical part of this section is largely an update and review of the existing literature on group specific inflation rates. In Section 4 we introduce our demand system based approach to estimating differences in inflation rates. The empirical results obtained using this method are presented in Section 5. In Section 6 we discuss the limitations and caveats of our empirical results and give some suggestions for future research. Finally, in Section 7 we conclude.

2. General framework

There have been several approaches to measuring differences in inflation rates. Our aim in this section is to provide a general framework that encompasses these methods and gives an insight into which are the underlying dimensions that are of interest.

At the heart of our analysis is the concept of a ‘household specific inflation rate’. We begin this section by explaining our definition of this term, starting from the definition of inflation as commonly calculated in the CPI. Next we develop a mathematical framework for explaining and understanding heterogeneity in inflation rates, as well as group price indexes. We show how many previous results concerning the inflation rates faced by specific groups, like the elderly, can be interpreted as summary statistics of the joint distribution of household inflation rates and household characteristics. Interpreting these previous results as summary statistics allows us to consider their specific focus, assumptions, and limitations.

Household specific inflation rates

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. For this
paper, we will assume that the relevant change in the cost of living is calculated by combining the price changes of \( m \) goods categories\(^1\).

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 constructed slightly differently from the overall CPI. Before we present how we calculate household specific inflation rates, let us first review how the BLS calculates the overall CPI.

The overall CPI measures inflation in period \( t \), which we will denote by \( \pi_{\pi} \), as the ratio of weighted averages of the percentage price increases 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

\[
\pi_{\pi} = \sum_{j=1}^{m} \frac{w_{j,b} \cdot \frac{p_{j,t}}{p_{j,b}}}{\sum_{j=1}^{m} w_{j,b} \cdot \frac{p_{j,b}}{p_{j,b}} - 1}
\]

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. 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 ways.

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

By updating the base period in every period for the overall CPI we would obtain an alternative measure of inflation of the form

\[
\pi_{\pi} = \sum_{j=1}^{m} w_{j,t-1} \cdot \frac{p_{j,t}}{p_{j,t-1}} - 1 = \sum_{j=1}^{m} w_{j,t-1} \cdot \left( \frac{p_{j,t}}{p_{j,t-1}} - 1 \right) = \sum_{j=1}^{m} w_{j,t-1} \cdot \pi_{j,t}
\]

where \( \pi_{j,t} \) is the inflation measured for item stratum \( j \).

When (2) is applied 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

---

\(^1\) A goods category is also often referred to as an ‘item stratum’ (plural is ‘item strata’). We will both terms interchangeably.
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.

If \( t \) indexes time in months, making this change to (2) gives us an inflation measure of the form

\[
\pi_t^{(H)} = \sum_{j=1}^{m} w_{j,t-12} \left( \frac{p_{j,t}}{p_{j,t-12}} - 1 \right) = \sum_{j=1}^{m} w_{j,t-12} \pi_{j,t}
\]

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

We will focus on household specific inflation rates, but (3) does not contain any household specificity. In principle, we would like to measure household specific price changes and household specific expenditure weights and apply them to (3). For each household, which we will index by \( i \), we 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 changes, \( \pi_{j,t} \), 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 indexes, 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 households 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 specific inflation rate. We will denote this inflation rate as \( \tilde{\pi}_{i,t} \) for household \( i \) in month \( t \). That is:

\[
\tilde{\pi}_{i,t} = \sum_{j=1}^{m} w_{i,j,t-12} \pi_{j,t}
\]

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.

**Distributional framework**

Various studies have proposed several different ways of looking at group specific cost of living indexes. Garner, Thesia, Johnson and Kokoski (1996) compute a group price index for the poor, and find relatively little difference between their index and the index for the overall population. Hamilton (2001) computes group price indexes, based on an Almost Ideal Demand System, for blacks and whites, and finds that whites face higher inflation than do blacks.
It turns out to be useful to consider these different proposed methods in a more general framework that nests all of them. As a basis for this general framework 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 \( y_{i,t} \), and its other characteristics, which we will denote by the vector \( \mathbf{x}_{i,t}^* \). We will denote the joint density of the household specific inflation rate and these characteristics by \( g(\mathbf{π}_{i,t}, y_{i,t}, \mathbf{x}_{i,t}^*) \).

General aggregate price indexes do not consider household characteristics, besides total expenditures, in their calculation. Hence, in order to interpret some of the aggregate price indexes 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 \( g_{\mathbf{x}^*, y}(\mathbf{π}_{i,t}, y_{i,t}) \). This joint density is obtained by integrating out the household specific characteristics other than total expenditures. That is,

\[
\int g(\mathbf{π}_{i,t}, y_{i,t}, \mathbf{x}_{i,t}^*) d\mathbf{x}_{i,t}^* = g_{\mathbf{x}^*, y}(\mathbf{π}_{i,t}, y_{i,t})
\]

where the densities in the integral are the relevant conditional and marginal densities that result from the application of Bayes’ rule.

Based on this density, one can calculate two aggregate inflation rates. The first, known as the plutocratic price index, is what is generally calculated as an aggregate price index. It measures the aggregate inflation rate as a weighted average of the inflation rates over different goods categories. The CPI, for example, is a plutocratic price index. The weights used are the weights of these goods categories in aggregate expenditures. In practice these weights are easy to calculate, since they involve measuring only total expenditures on each goods category for the sample as a whole.

In a plutocratic index, each household’s contribution to these aggregate expenditure weights is proportional to the household’s total expenditure level. Mathematically, the measured inflation corresponding to the plutocratic price index \( \bar{\pi}_i^P \), is the sample equivalent of the population moment

\[
\bar{\pi}_i^P = \int \left[ \bar{\pi}_{i,t} g_{\mathbf{x}^*, y}(\mathbf{π}_{i,t}, y_{i,t}) \right] d\mathbf{x}_{i,t}^* \left( \frac{y_{i,t-12} g_y(y_{i,t-12})}{\int y g_y(y) dy} \right) dy_{i,t-12}
\]

Here, the plutocratic weight

\[
\left( \frac{y_{i,t-12} g_y(y_{i,t-12})}{\int y g_y(y) dy} \right)
\]

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 \( y_{i,t} \).

As an alternative one could weigh all households equally. When one does so, the resulting inflation rate is known as derived from a democratic price index, which we will denote by \( \bar{\pi}_i^D \). It corresponds to the sample equivalent of the population moment.
That is, the democratic price index measures inflation as the mean of the marginal distribution of household specific inflation rates. Prais (1959) already distinguished between the plutocratic and democratic price indexes and noted that the latter was the one that gave all households equal weight.

The final alternative would be to consider a specific percentile of the distribution of inflation rates across households, which is defined as \( \bar{\pi}^\alpha \), where

\[
\alpha = \frac{1}{\int_{ WM} \left( \bar{\pi}^\alpha \right) d\bar{\pi}_{ij}}
\]

One obvious alternative to the democratic and plutocratic price indexes, which are (weighted) means, is to consider the median household inflation rate and choose \( \alpha=0.5 \). Note that this is very different from the ‘median CPI’ concept applied by Cechetti (1997). Our definition of median CPI is the median inflation rate across households. Cechetti’s (1997) definition is the weighted median inflation rate across item strata.

The bottom-line of the above derivations is that many of the aggregate inflation measures proposed and studied before can be derived as straightforward sample equivalents of specific population moments that are derived from the general joint distribution of household characteristics and household specific inflation rates. Thus, a study of the heterogeneity underlying aggregate inflation should focus on this distribution.

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

\[
\bar{\pi}_{ij} = \sum_{j=1}^{m} W_{i,j,t-12} \pi_{j,t}
\]

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

\[
\bar{\pi}^D = \sum_{j=1}^{m} \int_{ WM} W_{i,j,t-12} \pi_{j,t} d\bar{\pi}_{ij} = \sum_{j=1}^{m} \mu W_{i,j,t-12} \pi_{j,t}
\]

where \( \mu_{w,j,t-12} \) is the average expenditure share of item stratum \( j \) across households. This representation allows us to write the deviation of a specific household’s inflation rate from the mean as
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.

Secondly, households must have different-than-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 (12) 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.

### Conditioning on specific household characteristics

Instead of looking at how individual households fare compared to the average household, one could compare how entire groups fare in terms of inflation. Pollak (1980) was one of the first to propose the use of group price indexes. Amble et al. (1994) and Garner et al. (1996) calculate such group indexes for elderly and the poor respectively. In order to be comparable with the CPI, which is a plutocratic index, these group indexes 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 \( x_{i,t-1} \in G \). The inflation rate based on their plutocratic group price index is given by

\[
\tilde{\pi}_i^G = \int_{x_{i,t-1} \in G} \int \left[ \tilde{\pi}_{i,t} x_{i,t-1} \pi_{i,t-1} \left( \frac{y_{i,t-12}}{g_{x_{i,t-1}}} \right) dy \right] d\tilde{\pi}_{i,t}\]

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

\[
E[\tilde{\pi}_{i,t} | y_{i,t-12}, x_{i,t-12}] = \int \left[ \tilde{\pi}_{i,t} x_{i,t-1} \pi_{i,t-1} \left( \frac{y_{i,t-12}}{g_{x_{i,t-1}}} \right) dy \right] d\tilde{\pi}_{i,t}\]

(14)
is the expected inflation rate faced by a household with expenditures \( y_{i,t-12} \) and characteristics \( x_{i,t-12} \). The second part

\[
E[\pi_{i,t}] y_{i,t-12} x_{i,t-12} \left( \int_{y} g_{x,i} y_{i,t-12} d\pi \right) dy_{i,t-12}
\]  

(15)

is the inflation rate given by a plutocratic index for households with characteristics \( x_{i,t-12} \). Finally (13) is obtained by integrating out the household characteristics over the set of characteristics, \( G \), that defines the group under consideration.

We have shown how most existing methods used to study distributional effects of inflation can be interpreted in a general framework that is based on the joint distribution of household inflation rates and household characteristics. In the next section we will apply these methods are applied to actual data.

3. Descriptive statistics

In this section we will apply the framework of the previous section to the data. This section basically consists of three parts. In the first part we discuss how combine data from the Consumer Expenditure Survey and the Consumer Price Index to calculate our household specific inflation rates. In the second part we consider the extent to which households inflation rates differ. We will do so in two steps. In the first step we will consider the cross strata variation in inflation rates. That is, we consider part (B) of the decomposition in equation (12). In the second step we look that the overall distribution of household inflation rates, i.e. \( g_{\pi}(\cdot) \), and how it evolves over time. In the final part of this section, we present some results on plutocratic, democratic, and group price indexes. It updates and reviews the results in Amble et. al. (1994), Garner et. al. (1996), Kokoski (2000), and Hamilton (2001).

**Measuring household specific inflation rates using the Consumer Expenditure Survey**

Before we can estimate any of the population moments considered in the previous section, we first have to make our concept of a household specific inflation rate operational. First of all, we must choose how many and exactly which item strata we would like to consider. We also need household level expenditure data to calculate the expenditure weights for each household for each of the \( m \) goods categories we choose. In addition, we need monthly price data in order to calculate \( \pi_{j,t} \), the inflation rate for each goods category.

Data on household expenditures and demographic characteristics are obtained from the Consumer Expenditure Survey (CES). The CES 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 month, one third of the quarter’s panel is interviewed. Each household reports on the expenditures made over the three months prior to the interview month.

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 CES and the CPI series, we ended up with \( m=13 \) categories. They are: Food, Alcohol, Shelter, Utilities, Household Furnishings and Operations, Apparel, Transportation, Medical, Entertainment, Personal Care, Reading, Education, and Tobacco. Appendix A contains a detailed description of these categories. They match up closely with those used in most other studies using the CES.

Using these data, we construct household specific inflation rates in the following way. We use equation (4) and measure the expenditure shares, \( w_{i,j,t-12} \), as the shares of the expenditures in three months before the interview month. The item strata inflation rates \( \pi_{i,j} \) are measured as the percentage changes in the average price index values over these three months and the same average for these three months a year later.

**Cross-strata variation in inflation**

As we showed in the decomposition in equation (12), large differences in household specific inflation rates can only occur when (A) households have substantially different expenditure patterns, and (B) there are large variations in cross-strata inflation rates. 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. Table 1 lists 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. 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. 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 of about 1.5% there are household operations and equipment as well as apparel, while on the high end we find health care (5.74%), education (6.83%), and tobacco (8.72%). Tobacco inflation rates, however, are mainly driven by 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 (12) 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-correlations 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

---

more unrelated persons living together who pool their income to make joint expenditure decisions. Students living in university-sponsored housing are also included in the sample as separate CU’s.
believes that common price shocks, due to monetary policy actions, are the most important cause of inflation. There are two categories that exhibit much lower correlations with other strata. These two are transportation, for which the bulk of fluctuations is caused by changes in gas-prices, and tobacco, for which most of the fluctuations are caused by changes in tax laws.

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 (B) of decomposition (12) has the biggest potential of driving cross-household heterogeneity in inflation rates.

Figure 1 plots the inflation rates for the 13 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 and shoots up in the 1998-2000 period in reaction to increases in the taxes on tobacco. For the other two 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%. This observation will become relevant when we consider the inflation rate of a group price index for the elderly later on in this section. 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 children might have been hit harder by inflation for our period.

There are many categories whose inflation rates follow the overall CPI-U fairly closely. Most importantly, this is true for shelter. Partly, this is due to the fact that, as noted in Table 1, shelter makes up about 30% of the overall CPI. However, other categories, like food, entertainment, personal care, reading, and transportation show, in general, relatively small (between −2% and 2%) deviations from the overall CPI as well. Furthermore, these deviations are not very persistent over time. Transportation is an interesting category, because a large part of its fluctuations are driven by changes in gas-prices.

The shelter, food, entertainment, personal care, reading, and transportation strata have a total 1994-1995 expenditure share of 72.43%. 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 72.43% of expenditures, part (B) of decomposition (12) will probably not be a significant source of the differences in household inflation rates.

Hence, the conclusion of our analysis of cross-strata differences in inflation rates is that for goods categories that make up 72.43% of the budget the inflation rates track the overall CPI-U relatively closely. For these categories expenditure differences between households are very unlikely to generate large differences in household specific inflation rates. For the other 27.57% of expenditures, inflation rates are very different from the overall CPI-U and are a likely source of cross-household differences in inflation rates. The most important item strata that are part of these 27.57% of expenditures are medical care, education, utilities, household operations and equipment, and apparel. These are the categories for which part (B) of decomposition (12) is a potential source of differences in inflation rates across households.
Cross-household variations in inflation

Now that we have considered which item strata are the most probable sources of cross-household variations in inflation, it is time to consider the magnitude of these cross-household variations. The aim of this subsection is to document the most important facts on ‘overall’ cross-household variations in inflation. The term ‘overall’ refers to the fact that this section focuses on the unconditional distribution of household specific inflation rates. That is, in terms of the notation of the previous section, this subsection documents the evolution of the marginal density of household specific inflation rates, \( g_\pi(\bar{x}_{ij}) \), over time.

Figure 2 depicts the evolution of some of the moments of \( g_\pi(\bar{x}_{ij}) \) over time. The top panel plots the democratic mean, as defined in (8), as well as the median, and the 5\textsuperscript{th} and 95\textsuperscript{th} percentile, as defined in (9). The bottom panel plots the standard deviation.

The top panel suggests that the mean and median are almost equal to each other for the whole sample period, suggesting that the inflation distribution is fairly symmetric across households. The standard deviation of the distribution, plotted in the bottom panel, fluctuates over time but is mostly between 0.5 and 0.8 percent. Two notable exceptions are the start of the sample period, in which decreasing transportation prices resulted in an increase in the variance of household specific inflation rates and the period in 1998-1999 during which increases in tobacco taxes raised tobacco inflation to unprecedented levels, yielding high inflation rates for smokers. The tax-increases in the latter period coincided with high fluctuations in gas prices, leading to large fluctuations in inflation rates for households that spend a big portion of their money on transportation. Finally, the range between the 5\textsuperscript{th} and 95\textsuperscript{th} percentiles is slightly bigger than 2% over most of the time period, with the exception of the two periods mentioned before. To get an idea of the overall shape of \( g_\pi(\bar{x}_{ij}) \) we included Figure 3, which contains kernel estimates of the density for the start, middle and end of the sample period.

One of the aims of this paper is to consider which household characteristics would be able to predict part of the deviation of a household specific inflation rate from the average inflation rate. The main intuition behind this is that if households with the same characteristics continuously face higher (or lower) inflation than the overall inflation rate, then the CPI-U might not accurately reflect the change in their cost of living. In order to for households to ‘continuously’ face higher (or lower) than average inflation, the deviation of a household’s inflation rate from overall inflation should be persistent. In order to assess the degree of persistence in household inflation rates in our data, we present the results of a non-parametric AR(1) regression.

Figure 4 plots the kernel estimates of the distribution of the deviation of a household’s inflation rate from average inflation three months hence, conditional on its current deviation from average inflation. The thin lines depict isoprobs of the estimated conditional density, while the thick line is the conditional expectation of the deviation of a household’s inflation rate from the average 3 months in the future, conditional on its current deviation. We present the results of a non-parametric regression, because a standard parametric AR(1) regression would assume that this conditional expectation is a linear function of the initial deviation. However, our non-parametric estimates reveal that there is a distinct non-linearity in this relationship.
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 three months from now. This implies that on the downward side we do not find much persistence in a household’s inflation rate at all. However, on the high side we do find some persistence. A rough linear approximation of our non-linear estimation results suggests that the expected deviation from average inflation 3 months from now for a household that currently faces above average inflation is about a third of its current deviation. In a linear AR(1) regression, this would imply a 0.33 estimate of the auto-correlation coefficient. Hence, on the upside we do find persistence, but it is not very high.

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. Two important caveats about the estimates presented here are worthwhile mentioning, however.

First of all, this analysis is sensitive to seasonal fluctuations in expenditure patterns. Consider the following example. Suppose we observe a household that consists of two parents and two kids in college. The tuition for the education of these children is paid in full in September. Let us assume that we observe this household in September, in which a far above average share of its expenditures are devoted to education, and then in December in which no tuition payments are made and the household consumes the average goods basket. In that case, the household faces an above average inflation rate in September because it a large part of its expenditures are on education, which has a high inflation rate. However, in December the household faces the average inflation rate. Consequently, seasonality in the household’s expenditure pattern results in a low estimate of persistence in quarter-to-quarter deviations in a household’s inflation rate from the average.

In principle, it would have been preferable to compare year-to-year deviations in a household’s inflation rate from the average, such that we compared expenditures in the same three months. This is not possible with the CES, however, since it follows households only for four consecutive quarters at maximum.

The second caveat in the persistence analysis is that the household specific expenditure shares include expenditures on consumer durables, many of which fall into strata with below average inflation. Consider the example of a household that always consumes the average goods basket. However, in one quarter it uses part of its savings to buy a new computer. In that specific quarter it will have an above average expenditure share of household operations and equipment, which has a below average inflation rate. As a result, the household will have a below average inflation rate.

The current setup does not allow us to take into account the seasonal effects and the effects of consumer durables. Seasonal effects we deal with in our model in the next section. Durable goods remain an area for further improvement, as we discuss in our section on limitations and extensions.

Aggregation method

Thus far, we have considered the unweighted mean across households and unweighted percentiles across households when we considered the cross-household density of inflation rates. That is, we considered the marginal density $g_s(\tilde{\pi}_{ij})$. Alternatively, we could also consider the ‘plutocratic’ marginal that does not
weigh all households equally, but instead weights them proportional to their expenditure levels. That is, we could consider the moments of the density

\[ g_\pi^p(\bar{\pi}_{ij}) = \int g_p(\bar{\pi}_{ij}) \frac{y_{ij-12}g_y(y_{ij-12})}{\int yg_y(y)dy} dy_{ij-12} \]  

(16)

The sample mean of this density would be the inflation rate calculated using a plutocratic price index, while the mean of \( g_\pi^p(\bar{\pi}_{ij}) \) is that based on a democratic index. Prais (1959) argued that a democratic index is more appropriate than a plutocratic one because it weighs all households equally. Because the overall CPI-U is a plutocratic index, there is a concern that a democratic index would find a very different inflation rate. Kokoski’s (2000) results suggest that there is no big difference between a plutocratic and democratic index.

Our data, based on a more recent sample, confirm Kokoski’s result. Figure 5 plots the evolution of the means of \( g_\pi^p(\bar{\pi}_{ij}) \) and \( g_\pi^d(\bar{\pi}_{ij}) \) over time. That is, it plots the plutocratic and democratic inflation rates. As can be seen from the figure, there seems to be very little difference between the two of them. Furthermore, we also compared the standard deviations and percentiles of \( g_\pi^p(\bar{\pi}_{ij}) \) and \( g_\pi^d(\bar{\pi}_{ij}) \) over time. For brevity purposes we decided not to present these results in detail. However, the gist of them was the same as for the means. Namely, they are very similar. This leads us to conclude that the densities \( g_\pi^p(\bar{\pi}_{ij}) \) and \( g_\pi^d(\bar{\pi}_{ij}) \) are very similar over time and that it thus does not matter which aggregation method, democratic or plutocratic, one uses to calculate an overall index.

**Group price indexes**

Our final set of results in this section focuses on group specific price indexes. We will focus on three divisions of our sample that have been focused on previously in the literature. These are poor vs. non-poor, white vs. other, and elderly vs. non-elderly.

Previous work on a price index for the poor, for example Garner et. al. (1996), found very little evidence of major differences in the cost of living changes of the poor\(^3\) and non-poor. Our results, which are found in the top-left panel of Figure 6, confirm theirs.

Using PSID data and only two goods categories, i.e. food and the rest, Hamilton (2001) found that blacks tend to face lower inflation than whites. However, using our extended dataset as well as a more recent sample, we do not find this at all. The top-right panel of Figure 6 depicts the plutocratic group price indexes for white and non-white households, which are households for which the reference person self-identifies as white or not. From this panel there does not appear to be significant differences in the inflation rates for these two groups.

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

\(^3\) We consider a household poor whenever its reported income level is below its applicable poverty level, as applied by the Census Bureau. We are aware, however, that this way of selecting the ‘income-poor’ using the CES is imperfect because of the low quality of the income data in the CES and the self-selection of households on whether or not to report their income levels.
overall CPI does not properly reflect the cost of living changes that the elderly face, then the CPI-W would not be the appropriate price index for Social Security indexation. In fact, a current proposal in the House of Representatives (H.R.2035,2001) would require the Bureau of Labor Statistics to produce a separate CPI for the elderly (CPI-E).

Would the inflation measured by such a CPI-E be very different from the overall CPI? 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 CPI-E and found that for the end of the 1980’s the inflation measured by the CPI-E was between 0.2% and 0.4% higher than the inflation measured by the overall CPI-W.

The bottom two 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 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 1987-1993 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. However, after 1993 the difference between health care price inflation and the overall CPI dropped from the 3%-6% range to the 0%-2% range. Consequently, the health care expenditures of the elderly were much less of a source that caused them to have face higher than average inflation. What seems to dominate the differences between the elderly and the non-elderly inflation rates after 1993 is actually gas prices. Elderly spend a smaller share on transportation and thus their inflation rates are less susceptible to fluctuations in gas prices. Hence, when relative transportation prices decline, as in the 1997-1999 period, then the elderly face a slightly higher than average inflation because the do not benefit as much from this relative price decline as other consumers. However, when relative transportation prices rise, as in the 1999-2001 period, the elderly are hit less hard by the price increase.

Consequently, for the 1993-2001 period we do not observe that the elderly face a persistently higher inflation rate than everyone else. In fact, in the 1999-2001 period they actually faced lower inflation.

4. Model of expenditure shares and household characteristics

The main aim of this paper is to get a better understanding of whether the inflation experiences of individual households are very different and whether these differences are partly due to the underlying household characteristics. In the previous section we presented results calculated using the commonly used methodology of group price indexes. Although group price indexes allow us to consider the dependence of the household specific inflation rate on a particular household characteristic, they do not allow us to systematically compare the joint effect of multiple household characteristics.
The aim of this paragraph is the development of a methodology that would allow us to do so. This methodology turns out to involve the estimation of a consumption expenditure system that depends on household characteristics. Such a system would allow us to explain part of the observed differences in the consumption patterns across households, i.e. the differences in the $w_{i,j,t}$’s for different $i$’s, from differences in household characteristics.

In order to see why the estimation of such a system would be useful, consider the following. We would like to relate a household’s specific inflation rate to its characteristics. Hence, we are essentially interested in whether and how the conditional expectation

$$E[	ilde{\pi}_{i,t}|x_{i,t-12}]$$

depends on the household characteristics, $x_{i,t-12}$. For this conditional expectation, we obtain that

$$E[	ilde{\pi}_{i,t}|x_{i,t-12}] = E\left[\sum_{j=1}^{N} w_{i,j,t-12} \pi_{j,t} | x_{i,t-12}\right] = \sum_{j=1}^{N} E[w_{i,j,t-12} \pi_{j,t} | x_{i,t-12}] = \sum_{j=1}^{N} \pi_{j,t} E[w_{i,j,t-12} | x_{i,t-12}]$$

where in the last step we have assumed that $\pi_{j,t}$ is orthogonal to the expectation error

$$w_{i,j,t-12} - E[w_{i,j,t-12} | x_{i,t-12}]$$

Thus, in order to understand the dependence of household specific inflation rates on household characteristics, we will have to consider how the expenditure shares, $w_{i,j,t}$, depend on the household characteristics, $x_{i,t}$. We will study this dependence by estimating a consumption expenditure model that relates a household’s expenditure pattern to its characteristics. Using such an expenditure model allows us to obtain a different perspective on inflation differences compared to the conventional group price indices that we considered in the previous section.

Instead of focusing on a single household characteristic, as group indexes do, estimating a consumption expenditure model allows for jointly taking into account a whole set of characteristics. Hence, contrary to group indexes, it allows us to consider the effect of a particular household characteristic on inflation, conditional on all other characteristics. The estimation of a consumption expenditure model will thus make it feasible to consider actual marginal inflation effects of particular household characteristics.

In the rest of this section we will introduce the consumption expenditure model and describe how we will use it to estimate average marginal inflation effects of various household characteristics.

**Model specification**

For each household we observe at least one period of its expenditures. We will model the expenditure pattern of the household using an expanded translog functional form. The translog functional form derives the share equations as resulting from a flexible functional form of the indirect utility function. For the standard translog model this flexible functional form, as introduced by Christensen, Jorgenson and Lau (1975), consists of a second order Taylor approximation of the indirect utility function in the prices and expenditures. However, we will include also a set of demographic variables in the model. Therefore we will derive our...
expenditure equations as resulting from a second order approximation of the indirect utility function not only in the prices and total expenditures, but also in the relevant set of demographic variables.

Let \( V_{i,t} \) denote the utility level of household \( i \) at time \( t \), then we will assume that

\[
V_{i,t} = -\sum_{j=1}^{m} \alpha_j \ln(p_{j,i} / y_{i,t}) - \frac{1}{2} \sum_{j=1}^{m} \sum_{k=1}^{m} \beta_{j,k} \ln(p_{j,i} / y_{i,t}) \ln(p_{k,i} / y_{i,t}) \\
- \sum_{k=1}^{s} \phi_k x_{i,k} - \frac{1}{2} \left\{ \sum_{j=1}^{m} \sum_{k=1}^{m} \theta_{j,k} x_{i,j} x_{i,k} + \sum_{j=1}^{m} \sum_{k=1}^{m} (\varphi_{j,k} + \varphi_{k,j}) \ln(p_{j,i} / y_{i,t}) x_{i,k} \right\}
\]

where the indirect utility function satisfies the adding up restrictions

\[
\sum_{j=1}^{m} \alpha_j = 1 \quad \text{and} \quad \sum_{j=1}^{m} \varphi_{j,k} = \sum_{j=1}^{m} \varphi_{k,j} = 0
\]

Since some of the parameters in this approximation can be interpreted as values of a Hessian of the indirect utility function, these parameters also have to satisfy the symmetry restrictions

\[
\beta_{i,j,k} = \beta_{i,k,j} \quad \text{for all} \quad j,k=1,\ldots,m, \quad (22)
\]

\[
\varphi_{j,k} = \varphi_{k,j} \quad \text{for all} \quad j=1,\ldots,m \quad \text{and} \quad k=1,\ldots,s, \quad (23)
\]

\[
\theta_{j,k} = \theta_{k,j} \quad \text{for all} \quad j,k=1,\ldots,s
\]

(24)

Applying Roy’s identity to this indirect utility function yields the following expenditure share equations

\[
\hat{\omega}_{i,t} = -\frac{\partial V_{i,t} / \partial \ln p_{j,i}}{\partial V_{i,t} / \partial \ln y_{i,t}}
\]

\[
= \frac{\alpha_j + \frac{1}{2} \sum_{k=1}^{s} (\varphi_{j,k} + \varphi_{k,j}) x_{i,k} + \frac{1}{2} \sum_{k=1}^{m} (\beta_{j,k} + \beta_{k,j}) \ln(p_{k,i} / y_{i,t})}{1 + \frac{1}{2} \sum_{j=1}^{m} \sum_{k=1}^{m} (\beta_{j,k} + \beta_{k,j}) \ln(p_{k,i} / y_{i,t})}
\]

\[
= \frac{\alpha_j + \sum_{k=1}^{s} \varphi_{j,k} x_{i,k} + \sum_{k=1}^{m} \beta_{j,k} (\ln(p_{k,i}) - \ln(y_{i,t}))}{1 + \sum_{j=1}^{m} \sum_{k=1}^{m} \beta_{j,k} (\ln(p_{k,i}) - \ln(y_{i,t}))}
\]

(25)

These are simply the expenditure share equations that result from the expanded translog indirect utility function above. We denote the expenditure shares implied by the model with a circumflex.

We will estimate this model using GMM. Our identifying assumptions for the GMM estimation are based on the difference between the expenditure shares predicted by the estimated model and the ones that are actually observed rescaled by the denominator of the share equation. That is, on
\[ e_{i,j,t} = (w_{i,j,t} - \hat{w}_{i,j,t}) \left( 1 + \sum_{j=1}^{m} \sum_{k=1}^{m} \beta_{j,k} (\ln(p_{j,t}) - \ln(y_{j,t})) \right) \]  

(26)

The reason that we re-scale the prediction error is that this will lead to a GMM moment condition that is linear in all the parameters, which given the amount of data that we have, is the only feasible form in which the equation is estimable. The linearity of (26) follows from rewriting it to obtain

\[ e_{i,j,t} = w_{i,j,t} + \sum_{j=1}^{m} \sum_{k=1}^{m} \beta_{j,k} (\ln(p_{j,t}) - \ln(y_{j,t})) \nu_{i,j,t} \]

\[ -\alpha_j - \sum_{s=1}^{s} \phi_{j,k} x_{i,k} - \sum_{k=1}^{m} \beta_{j,k} (\ln(p_{j,t}) - \ln(y_{j,t})) \]

(27)

**Moment conditions**

For our GMM estimation we will assume the following identifying moment conditions

\[ E[e_{i,j,t}] = 0 \text{ for all } i,j,t \]  

(28)

\[ E[e_{i,j,t} x_{i,s}] = 0 \text{ for all } i,j,s,t \]  

(29)

\[ E[e_{i,j,t} (\ln p_{j,t} - \ln y_{j,t})] = 0 \text{ for all } i,j,k,t \]  

(30)

\[ E[e_{i,j,t} \pi_{j,t+1}] = 0 \text{ for all } i,j,t \]  

(31)

\[ E\left[ \sum_{j=1}^{m} e_{i,j,t} \pi_{j,t+1} \right] = 0 \text{ for all } i,j,t \]  

(32)

Expenditure models suffer from the additivity restriction that the sum of the expenditure shares over all item strata equals one by definition. For that reason, we will apply the commonly used approach of estimating the model only for the first \( m-1 \) strata and then inferring the parameters for the \( m \)th stratum from the adding-up identity.

We will estimate the model without imposing the symmetry restrictions on the parameters. Given the amount of available data, consistency of the parameter estimates would imply that these restrictions will hold approximately anyway if they are valid. By not imposing the symmetry restrictions, the moment conditions (28) through (32) make the model exactly identified. Which simplifies our parameter estimation. The reason that we make the last two identifying assumptions is explained below.

**Seasonality and time specific fixed effects**

We would like the expenditure shares to depend on seasonal fluctuations. Furthermore, the translog approximation is a local approximation of the indirect utility function. However, our sample covers 24 years. It is hard to argue that for this long a period, such a local approximation is appropriate. In order to
compensate for these two factors, we will include time specific fixed affects in the $\alpha_{ij}$ coefficients. These time specific fixed effects can capture both seasonal fluctuations in the expenditure shares as well as a re-normalization of the expansion point of the translog indirect utility function approximation.

Let $T$ denote the total number of months in our sample, then our resulting model contains

$$(m-1) \cdot (T+s)+m^2$$

(33)

parameters to be estimated as well as moment conditions.

Calculating marginal inflation effects

We are ultimately not interested in the particular parameter estimates we obtain from our expenditure model but rather in using these parameter estimates to predict household specific inflation rates. The aim of this section is the description of how we go about doing so. We first consider under which assumptions our expenditure model estimates can be used to get an insight in the effects of various household characteristics on household specific inflation rates. Then, we describe how we will aggregate over the predicted household specific inflation rates to obtain estimates of ‘marginal inflation effects’. Finally, we describe how the results of our procedure compare to those obtained using standard group price indexes.

Identifying assumption

We are interested in using our expenditure model to predict household specific inflation rates. Let $\hat{W}_{i,j,t}$ denote household $i$’s predicted budget share in period $t$ for stratum $j$. We will construct a predicted inflation rate for household $i$ in period $t$, which we will denote by $\hat{\pi}_{i,t}$, by calculating

$$\hat{\pi}_{i,t} = \sum_{j} \hat{W}_{i,j,t} \pi_{j,t}$$

(34)

Using this prediction, the prediction error is the difference between the predicted and the actual household specific inflation rate. This difference equals (where we have assumed that the parameters have been estimated consistently)

$$\hat{\pi}_{i,t} - \hat{\pi}_{i,t} = \sum_{j} (W_{i,j,t} - \hat{W}_{i,j,t}) \pi_{j,t} = \sum_{j} e_{i,j,t} \left[ \frac{\pi_{j,t+12}}{1+\sum_{k=1}^{m} \beta_{j,k} \left( \ln(p_{k,j}) - \ln(y_{k,j}) \right)} \right]$$

(35)

For the expectation of this prediction error to be zero and our predicted household specific inflation rate to be an unbiased estimate of the actual household specific inflation rate, it must hold that
We will simply assume this to be true, which is why we imposed the last two moment conditions described above. However, we presented this because we think it is important to realize that when the expenditure model is applied to predict inflation rates, one makes an additional orthogonality assumption.

**Average marginal inflation effects**

The value added of estimating an expenditure model instead of doing group price indexes is that the expenditure model allows us to calculate counterfactuals conditioning not only one variable, but taking into account all household characteristics included in the dataset. This allows us to calculate ‘average marginal inflation’ effects of different characteristics.

To illustrate our concept of an ‘average marginal inflation’ effect of a characteristic, consider the following. Suppose we would like to know what is the effect of being elderly on inflation, conditional on all other characteristics. We could then ask ourselves the question: ‘What would the average inflation rate have been if all households in our sample would have been elderly?’ For the households that are elderly, this would not change anything. However, for households with non-elderly, this would require calculating a counterfactual household specific inflation rate that predicts, conditional on all its other characteristics, what inflation rate the household would have faced if it was elderly.

Calculating such a predicted counterfactual is fairly straightforward with our expenditure model. The model allows us to predict the counterfactual expenditure shares, \( \hat{w}^{\text{CF}}_{ij,t} \), and then use these to predict household specific inflation. This yields the counterfactual household specific inflation rate

\[
\hat{\pi}_{tij}^{\text{CF}} \approx \sum_{j=1}^{m} \hat{w}^{\text{CF}}_{ij,t} \pi_{j,t+12}^{\text{CF}}
\]

The difference between the counterfactual inflation rate of household \( i \) and its actual inflation rate can be decomposed in two parts as follows

\[
\hat{\pi}_{tij} - \hat{\pi}_{tij}^{\text{CF}} \approx \frac{1}{1 + \sum_{j=1}^{m} \sum_{k=1}^{m} \beta_{j,k} (\ln(p_{tij}) - \ln(y_{tij}))}
\]

\[
\sum_{j=1}^{m} \hat{w}^{\text{CF}}_{ij,t} \pi_{j,t+12}^{\text{CF}}
\]

We will use these to predict household specific inflation. This yields the counterfactual household specific inflation rate

\[
\hat{\pi}_{tij}^{\text{CF}} = \sum_{j=1}^{m} \hat{w}^{\text{CF}}_{ij,t} \pi_{j,t+12}^{\text{CF}}
\]

The difference between the counterfactual inflation rate of household \( i \) and its actual inflation rate can be decomposed in two parts as follows

\[
\hat{\pi}_{tij} - \hat{\pi}_{tij}^{\text{CF}} = \left( \hat{\pi}_{tij}^{\text{prediction error}} - \hat{\pi}_{tij}^{\text{CF}} \right) + \left( \hat{\pi}_{tij}^{\text{marginal inflation effect}} - \hat{\pi}_{tij}^{\text{CF}} \right)
\]

The first part is the prediction error, which is the difference between the actual household specific inflation rate and that predicted by the expenditure model. The second part is the marginal inflation effect that we are interested in. It is the difference between the predicted inflation rate for household \( i \) and its counterfactual.

Comparing this for all households is simply not feasible, so we will consider the average marginal effect of a characteristic on inflation. This is simply obtained by integrating out over the distribution of household characteristics at time \( t \). That is, \( \hat{\pi}_{tij} \) and \( \hat{\pi}_{tij}^{\text{CF}} \) are functions of the vector of household characteristics \( \mathbf{x}_{i,t-12} \). Hence, by integrating out (38) over the distribution of household characteristics, we obtain
\[ \hat{\pi}_t^D - \hat{\pi}_t^{D,CF} = \left( \hat{\pi}_t^D - \hat{\pi}_t^D \right) + \left( \hat{\pi}_t^D - \hat{\pi}_t^{D,CF} \right) \]

where \( \hat{\pi}_t^D \) is the inflation rate implied by a democratic price index and

\[ \hat{\pi}_t^D = \int \hat{\pi}_t^D (x_{j-12}) k(x_{j-12}) dx_{j-12} \]

is the democratic inflation rate predicted by the expenditure model, while

\[ \hat{\pi}_t^{D,CF} = \int \hat{\pi}_t^{D,CF} (x_{j-12}) k(x_{j-12}) dx_{j-12} \]

is the counterfactual democratic inflation rate predicted by the expenditure model.

The ‘average marginal effect’ of being elderly described above can be interpreted as the average increase in inflation that the households would face if they all would be elderly. Obviously, this is a partial equilibrium calculation, because we do not take into account the effect of everyone becoming elderly on the goods specific inflation rates.

Our framework allows us to consider a wide variety of similar counterfactual experiments. As a comparison with the elderly group price index, for example, we could calculate a counterfactual to answer the question ‘How high would the inflation rate for the group of elderly have been if they all would have been younger than 60?’ The answer to this question allows us to consider filter out the individual effect on inflation of just being elderly, while a group index would be affected by the covariation of being elderly with other characteristics, like not having kids.

5. Empirical results

6. Limitations

7. Conclusion
References


A. Data appendix

Matching Expenditure Categories with Price Data
Household expenditure, income and other characteristics are taken from the family file of the Consumer Expenditure Survey (CES). The primary income variable used is FINCATAx, or household after-tax income. Expenditures are aggregated to a quarterly level. Expenditure categories (eg apparel) are taken as given from the CES, and match, in most cases, with the categories which comprise the Consumer Price Index (CPI). In cases where CES and CPI categories do not match exactly, we calculate our own price index using an amalgamation of CPI series and their appropriate weights, (which can be found in the table ‘Relative Importance in the CPI’). Table A.1 details the series used in our analysis:

Households and Characteristics
We use the word ‘household’ to mean ‘consumption unit,’ a term defined in the CES. 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 CES variable NEWID, which is unique for each household, as a household identifier.

Household Inflation Rates
Each household has expenditure data for at most four consecutive quarters. For each quarter, a household’s inflation rate is the dot product of two vectors. The first is the expenditure shares for each of our 13 categories. The second is the year/year inflation rate for each of the 13 categories, calculated using the CPI.

The 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 1999. 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.
Table A.1. Description of goods categories and matching between CES and CPI

<table>
<thead>
<tr>
<th>CES Expenditure Category</th>
<th>CPI Series (and description if necessary)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1</strong> Food (food purchased at grocery or other food stores, or restaurants, excluding alcoholic beverages)</td>
<td>Food</td>
</tr>
<tr>
<td><strong>2</strong> Alcoholic Beverages (beer, wine, and other liquor purchased for consumption at or away from home)</td>
<td>Alcoholic Beverages</td>
</tr>
<tr>
<td><strong>3</strong> Shelter (includes interest on mortgages, rent, property taxes, insurance related to housing, refinancing charges, expenses for repairs and maintenance, and other housing expenses)</td>
<td>Shelter</td>
</tr>
<tr>
<td><strong>4</strong> Utilities (includes electricity, natural gas and other fuels, water, garbage collection, telephone charges)</td>
<td>Fuels and Utilities</td>
</tr>
<tr>
<td><strong>5</strong> Household Equipment (furniture, household decorations, personal computers, household appliances)</td>
<td>Household Equipment and Operations</td>
</tr>
<tr>
<td>Household Operations (domestic services, child care, etc)</td>
<td>Information Processing other than Telephone</td>
</tr>
<tr>
<td><strong>6</strong> Apparel (clothing purchases and upkeep)</td>
<td>Apparel</td>
</tr>
<tr>
<td><strong>7</strong> Transportation (includes vehicle purchases and rentals, finance charges, gasoline, insurance, public transportation)</td>
<td>Transportation</td>
</tr>
<tr>
<td><strong>8</strong> Health Care (health insurance, medical services, drugs, medical supplies)</td>
<td>Medical Care</td>
</tr>
<tr>
<td><strong>9</strong> Entertainment (includes fees and admissions, television, audio and video equipment, pets, toys, hobbies, other entertainment equipment)</td>
<td>Recreation (1993 and later)</td>
</tr>
<tr>
<td><strong>10</strong> Personal Care (includes hair products and services, cosmetic &amp; bath products, other personal goods and services)</td>
<td>Personal Care (hair, dental and shaving goods and services, funeral expenses, financial services, laundry services)</td>
</tr>
<tr>
<td><strong>11</strong> Reading (includes magazine and newspaper subscriptions, books)</td>
<td>Recreational Reading Materials</td>
</tr>
<tr>
<td><strong>12</strong> Education (includes tuition and fees for universities, primary, secondary, and nursery schools; textbooks, educational equipment)</td>
<td>Educational Books and Supplies</td>
</tr>
<tr>
<td><strong>13</strong> Tobacco</td>
<td>Tobacco</td>
</tr>
</tbody>
</table>
Table 1. Average, standard deviation, CPI weight and correlations of strata inflation rates

<table>
<thead>
<tr>
<th>Strata</th>
<th>Mean (%), Standard Dev (%)</th>
<th>CPI weight (%)</th>
<th>Food</th>
<th>Alcohol</th>
<th>Shelter</th>
<th>Utilities</th>
<th>House Op+Eq</th>
<th>Apparel</th>
<th>Transportation</th>
<th>Health</th>
<th>Entertainment</th>
<th>Personal Care</th>
<th>Reading</th>
<th>Education</th>
<th>Tobacco</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food</td>
<td>3.16% 1.42% 15.22%</td>
<td>1</td>
<td>0.22</td>
<td>0.61</td>
<td>0.23</td>
<td>0.42</td>
<td>0.48</td>
<td>0.26</td>
<td>0.51</td>
<td>0.64</td>
<td>0.60</td>
<td>0.28</td>
<td>0.29</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>Alcohol</td>
<td>3.44% 2.25% 0.98%</td>
<td>0.22</td>
<td>1</td>
<td>0.54</td>
<td>0.41</td>
<td>0.46</td>
<td>0.50</td>
<td>0.07</td>
<td>0.63</td>
<td>0.66</td>
<td>0.49</td>
<td>0.58</td>
<td>0.52</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>Shelter</td>
<td>3.83% 0.96% 30.25%</td>
<td>0.61</td>
<td>0.54</td>
<td>1</td>
<td>0.15</td>
<td>0.46</td>
<td>0.62</td>
<td>0.00</td>
<td>0.76</td>
<td>0.73</td>
<td>0.74</td>
<td>0.47</td>
<td>0.65</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>Utilities</td>
<td>1.60% 1.92% 7.44%</td>
<td>0.23</td>
<td>0.41</td>
<td>0.15</td>
<td>1</td>
<td>0.28</td>
<td>0.25</td>
<td>0.64</td>
<td>0.31</td>
<td>0.42</td>
<td>0.27</td>
<td>0.21</td>
<td>0.21</td>
<td>-0.18</td>
<td></td>
</tr>
<tr>
<td>House Op+Eq</td>
<td>1.46% 0.70% 4.77%</td>
<td>0.42</td>
<td>0.46</td>
<td>0.46</td>
<td>0.28</td>
<td>1</td>
<td>0.57</td>
<td>0.16</td>
<td>0.58</td>
<td>0.56</td>
<td>0.46</td>
<td>0.66</td>
<td>0.58</td>
<td>-0.21</td>
<td></td>
</tr>
<tr>
<td>Apparel</td>
<td>1.46% 2.28% 4.45%</td>
<td>0.48</td>
<td>0.50</td>
<td>0.62</td>
<td>0.25</td>
<td>0.57</td>
<td>1</td>
<td>0.20</td>
<td>0.72</td>
<td>0.85</td>
<td>0.57</td>
<td>0.49</td>
<td>0.72</td>
<td>-0.03</td>
<td></td>
</tr>
<tr>
<td>Transportation</td>
<td>2.52% 3.16% 17.57%</td>
<td>0.26</td>
<td>0.07</td>
<td>0.00</td>
<td>0.64</td>
<td>0.16</td>
<td>0.20</td>
<td>1</td>
<td>0.14</td>
<td>0.27</td>
<td>0.17</td>
<td>-0.03</td>
<td>-0.03</td>
<td>-0.01</td>
<td></td>
</tr>
<tr>
<td>Health</td>
<td>5.74% 2.07% 5.81%</td>
<td>0.51</td>
<td>0.63</td>
<td>0.76</td>
<td>0.31</td>
<td>0.58</td>
<td>0.72</td>
<td>0.14</td>
<td>1</td>
<td>0.82</td>
<td>0.68</td>
<td>0.64</td>
<td>0.89</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Enter</td>
<td>2.15% 2.26% 5.33%</td>
<td>0.64</td>
<td>0.66</td>
<td>0.73</td>
<td>0.42</td>
<td>0.56</td>
<td>0.85</td>
<td>0.27</td>
<td>0.82</td>
<td>1</td>
<td>0.75</td>
<td>0.45</td>
<td>0.73</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>Personal Care</td>
<td>2.88% 0.97% 3.46%</td>
<td>0.60</td>
<td>0.49</td>
<td>0.74</td>
<td>0.27</td>
<td>0.46</td>
<td>0.57</td>
<td>0.17</td>
<td>0.68</td>
<td>0.75</td>
<td>0.38</td>
<td>0.50</td>
<td>0.30</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td>Reading</td>
<td>3.63% 1.53% 0.58%</td>
<td>0.28</td>
<td>0.58</td>
<td>0.47</td>
<td>0.21</td>
<td>0.66</td>
<td>0.49</td>
<td>-0.03</td>
<td>0.64</td>
<td>0.45</td>
<td>0.38</td>
<td>1</td>
<td>0.65</td>
<td>-0.21</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>6.83% 1.53% 2.80%</td>
<td>0.29</td>
<td>0.52</td>
<td>0.65</td>
<td>0.21</td>
<td>0.58</td>
<td>0.72</td>
<td>-0.03</td>
<td>0.89</td>
<td>0.73</td>
<td>0.50</td>
<td>0.65</td>
<td>1</td>
<td>-0.12</td>
<td></td>
</tr>
<tr>
<td>Tobacco</td>
<td>8.72% 7.84% 1.31%</td>
<td>0.03</td>
<td>0.19</td>
<td>0.08</td>
<td>-0.18</td>
<td>-0.21</td>
<td>-0.03</td>
<td>-0.01</td>
<td>0.04</td>
<td>0.13</td>
<td>0.30</td>
<td>-0.21</td>
<td>-0.12</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Note: All results based on 1985.12-2000.12 data, except ‘CPI weights’. These are 1993-95 weights of strata in CPI-U
Figure 1. Goods specific year over year inflation rates in deviation from the CPI-U
Figure 2. Moments of household inflation distribution
Figure 3. Cross-household distribution of inflation at beginning, middle, and end of our sample period

Figure 4. Persistence of household’s inflation rate from average
Figure 5. Democratic and plutocratic inflation rates

Figure 6. Group price indexes for the poor, racial groups, and elderly

Note: The solid line is for the group of interest (poor, white, elderly) and the dashed line for its complement.