FEDERAL RESERVE BANK OF SAN FRANCISCO

WORKING PAPER SERIES

Household Inflation Experiences in the U.S.: A Comprehensive Approach

Bart Hobijn
Federal Reserve Bank of San Francisco

Kristin Mayer
Princeton University

Carter Stennis and Giorgio Topa
Federal Reserve Bank of New York

September 2009

Working Paper 2009-19
http://www.frbsf.org/publications/economics/papers/2009/wp09-19bk.pdf

Household Inflation Experiences in the U.S.:

A Comprehensive Approach

BART HOBIJN
Federal Reserve Bank of San Francisco

KRISTIN MAYER
Princeton University

CARTIER STENNIS AND GIORGIO TOPA
Federal Reserve Bank of New York*

September 30, 2009

Abstract

We present new measures of household-specific inflation experiences based on comprehensive information from the Consumer Expenditure Survey (CEX). We match households in the Interview and the Diary Surveys from the CEX to produce both complete and detailed pictures of household expenditures. The resulting household inflation measures are based on a more accurate and detailed description of household expenditures than those previously available.

We find that our household-based inflation measures track aggregate measures such as the CPI-U quite well and that the addition of Diary Survey data induces small but significant differences in the measurement of household inflation. The distribution of inflation experiences across households exhibits a large amount of dispersion over the entire sample period. In addition, we uncover a significantly negative relationship between mean inflation and inflation inequality across households.

Keywords: Inflation, inequality, household inflation rates.

JEL-codes: C43, D12, D39

* Correspondence to Giorgio Topa, Research and Statistics Group, Federal Reserve Bank of New York, 33 Liberty street, 3rd floor. E-mail: Giorgio.Topa@ny.frb.org. All views expressed in this paper are those of the authors only and do not necessarily reflect those of the Federal Reserve Bank of New York, the Federal Reserve Bank of San Francisco, or those of the Federal Reserve System in general.

1. Introduction

The U.S. Consumer Price Index (CPI) is broadly used as a measure of the increase in the cost of living in the United States. It measures the percentage increase in the cost of a fixed set of expenditures over time. This set of expenditures is changed infrequently[1] and is based on the average expenditures of all U.S. households. However, expenditure patterns across households in the U.S. vary substantially. Consequently, the single CPI number might be masking a significant amount of variation in cost-of-living increases across U.S. households.

There is, by now, a substantial literature that has tried to quantify these inflation differentials for the U.S. Following Pollak (1980), most studies[2] have focused on constructing price indices that approximate the average trend in the cost of living for particular groups of households (e.g. poor households, elderly households, households with children). These are known as group-price indices. The Federal Reserve Bank of Chicago's IBEX project,[3] described in McGranahan and Paulson (2006), publishes group price indices for a broad set of groups as well as for a long sample period.

As discussed by Hobijn and Lagakos (2005), these group price indices can be interpreted as a particular summary statistic of the underlying distribution of household inflation rates. In this paper, as in Hobijn and Lagakos (2005), we construct new and improved estimates of the distribution of inflation rates across U.S. households for the period 1984-2004. The main improvement relative to the existing literature is that we use enhanced data on the expenditure patterns of individual households.

The CPI is constructed using aggregate U.S. consumer expenditure data taken from the Consumer Expenditure Survey (CEX), which is published by the Bureau of Labor Statistics (BLS).[4] This survey itself is composed of two different surveys: the Interview Survey and the Diary Survey. In the Interview Survey

[1] From December 2007 to now, for example, the set of expenditures used has been based on expenditures made in 2005 and 2006.

[2] See McGranahan and Paulson (2006), Hamilton (2001), Idson and Miller (1999), Garner, Johnson, and Kokoski (1996)), and Amble and Stewart (1994) for example.

[3] See http://www.chicagofed.org/community_development/chicagofed_ibex_consumer_price_index.cfm for more information and for the data.

[4] The CEX website can be found at http://www.bls.gov/cex/

individual households report their monthly expenditures on 'big ticket' items for four consecutive quarters. In the Diary Survey households keep a two-week diary on more frequently purchased, 'small ticket' items. The items covered by the Diary Survey are disproportionately food and personal care products. Ideally, one would combine information from these two surveys when analyzing household expenditures to get a complete and detailed picture of individual household expenditure patterns. The difficulty in this, however, is that the samples of households used in the two surveys are independent. This means that, although the CEX easily provides us with adequate estimates of aggregate household expenditures in the U.S., it does not allow us to easily estimate the complete expenditure patterns of individual households.

The most common practice, therefore, has been to only use the data from the Interview Survey for applied microeconomic analysis of the CEX for individual households. However, such an approach ignores the heterogeneity in expenditures that is only captured in the detailed expenditures reported in the Diary Survey. For an analysis of inflation differentials across households, this may lead to an understatement in inflation differentials in times of, for example, high food price inflation.

The enhancement of the expenditure data for individual households that we develop here is the matching of households in the Interview Survey with households in the Diary Survey. We match each household in the Interview Survey with households in the Diary Survey that have similar predicted changes in expenditures based on their demographic characteristics. The resulting matched data set contains imputed data on the total expenditures of 'hybrid' households. For each of these 'hybrid' households, the expenditures on 'big ticket' items are taken from one household from the Interview Survey and the expenditures on 'small ticket' items are taken from several households in the Diary Survey. Adding the information from the Diary Survey gives us a more comprehensive and accurate image of household level inflation experiences.

We find that our household-based inflation measures track aggregate measures such as the CPI-U quite well and that the addition of Diary Survey data induces small but significant differences in the measurement of household inflation. The distribution of inflation experiences across households exhibits a large amount of cross-sectional dispersion over the entire sample period, with the 90^{th}-10^{th} percentile interval ranging from

about 100 to more than 300 basis points. Further, we uncover a significantly negative relationship between mean inflation and inflation inequality across households; this is a novel finding made possible by our approach, and it opens up interesting macro-economic questions for future research.

We also use the household-based inflation rates to calculate group-level inflation measures. In these, we find that within-group inflation inequality is much higher than between-group differentials. The larger gaps occur between young and elderly households, with the latter consistently experiencing higher inflation than younger households. The cumulative gap over our sample period between these two groups is equal to about nine percentage points, or roughly 45 basis points per year, on average.

The rest of this paper is organized as follows. In the next section, we discuss the structure of the CEX and some of the main differences between the Interview and the Diary Surveys; we also discuss the advantages of adding Diary Survey data for this exercise. We outline our matching method in Section 3 and report some of the diagnostics we run to check the reasonability of the approach in Section 4. We discuss the details of the matching methodology in Appendix B. In Section 5 we present the results for the estimated household inflation distributions based on matched Interview and Diary Survey data. We do so in two parts. In the first part, we compare our time series and cross-sectional results with aggregate measures such as CPI inflation. In the second part, we present a set of demographic group price indices and inflation distributions based on the merged expenditure data. We conclude in Section 6.

2. Diary vs. Interview Survey

Because the main contribution of our analysis is to present household inflation estimates for individual households based on combined evidence from the two surveys that make up the CEX, we begin by discussing these two surveys and why it is preferable to merge the evidence from these surveys for our analysis.

The reason that the Consumer Expenditure Survey[5] consists of two surveys is, in large part, practical. It is simply not reasonable to expect respondents of the Interview Survey to keep a detailed diary of expenditures over the year that they are part of the sample. As a consequence, the BLS has set up the CEX such that data on frequent, but relatively inexpensive, purchases are collected in the Diary Survey rather than the Interview Survey.[6]

In the Diary Survey, participants record their expenditures on a daily basis for two weeks. Because of the frequency and time period, these tend to be frequent, inexpensive expenditures (e.g. personal care items or specific food expenditures). In any given month, approximately 600 people are completing diaries.[7]

In the Interview Survey, participants give quarterly reports of their expenditures over the past three months for four quarters. Expenditures reported in this survey tend to be infrequent, large expenditures (e.g. autos), although participants also report aggregate estimates for some smaller, more frequent purchases (e.g. food). For the Interview Survey, the BLS gets reports from approximately 7,600 households each quarter. Each quarter, 25 percent of those households rotate out of the sample and new ones rotate in.

The BLS makes available a correspondence table that indicates which expenditures it takes from the Diary Survey and which from the Interview Survey in constructing the expenditure baskets used to calculate the CPI. We used this table to classify 'items' into four particular categories: (*i*) items that are only covered in the Diary Survey, (*ii*) items that are reported in a more reliable fashion in the Diary Survey, (*iii*) items that are covered in more detail in the Diary Survey, (*iv*) items for which only Interview Survey data are used for the construction of the CPI.

The data that we use cover 247 items for which we have expenditure data with corresponding price data from the CPI. Table 1 summarizes how these items are classified. About half of the items come from the Interview Survey. Since these are generally big ticket items, however, they make up about three quarters of

[5] More information about both Surveys is available in BLS (2008, Chapter 16)

[6] Expenditure categories for which the BLS collects price quotes for the construction of the CPI are called item strata. The CEX equivalent of an item stratum is a Universal Classification Code (UCC). Throughout this paper we refer to both as 'items'.

[7] Note that these numbers refer to recent sample sizes; the samples have generally grown over time.

total expenditures. Any analysis solely based on evidence from the Interview Survey (i.e. using only items from the Interview Survey to calculate a household's total expenditures) would capture at most about 95 percent of expenditures, but that would use many items that are measured less precisely or in less detail than in the Diary Survey: in fact, slightly over one fifth of household expenditures are measured more accurately in the Diary Survey.

The other half of the 247 items we use come from the Diary Survey. Table 2 lists the fraction of items taken from the Diary Survey for the eight main CPI item categories. This shows that merging the Diary and Interview Surveys may especially affect the measured expenditure patterns on food, apparel and personal care items.[8]

If prices for all items always went up, or down, at the same rate, then differences in expenditure patterns would not lead to differences in cost-of-living increases. This is, of course, not the case. What this point emphasizes, however, is that if price changes of the Diary items are not much different than those of the Interview items they are replacing, then including these items will only have a minor impact on household inflation measures. We will discuss this in our results section.

For the price data we use 12-month (year-over-year) changes in prices for urban consumers for the 247 item strata in our sample for 1984-2004, taken from the underlying disaggregated CPI data.[9]

3. Matching methodology

We explain our matching method in three parts. We first discuss the main criterion that we use. Second, we focus on how we construct estimates to implement this matching criterion. Third, we discuss the actual steps taken to match households in the Interview Survey sample with households from the Diary Survey.

Criterion

To understand our choice of matching criterion, it is worthwhile to first consider the optimal case. In that case we would have the same respondents for the Interview and Diary Surveys, and we could combine their

[8] The Diary items in the "Other goods and services" category are mainly personal care items.

[9] A more detailed description of how we match price and expenditure data is contained in Appendix A.

two surveys to construct each household's overall expenditure basket. Since the two Surveys' samples are independent, this is not possible.

The next best approach would be to match households from the Interview and Diary Surveys that have similar expenditure patterns. However, we are not particularly concerned with expenditure differentials between matched households on items with very similar inflation rates; such differentials will likely not lead to very different household-specific changes in the cost-of-living. Therefore, we want to match households on the basis of items that exhibit similar price changes, or – in a more aggregate sense – on the basis of similar expenditure *changes*.

Let $e_{i,j,t-1}$ be the amount spent by household *j* on item *i* in year *t-1*, and let $\pi_{i,t}$ be the percentage change in the price of item *i* between year *t-1* and *t*, as given by the CPI data.[10] If there are *n* items, then the change in the total expenditures of household *j* between *t-1* and *t* such that the household is able to purchase the same set of items in period *t* as in period *t-1*is given by

$$\Delta e_{j,t} = \sum_{i=1}^{n} e_{i,j,t-1}\pi_{i,t} \ . \qquad (1)$$

We call this the *household expenditure change*. It is a summary measure of expenditures for each household that emphasizes the importance of the individual item strata based on their price changes.

These household expenditure changes form the basis of our matching criterion. That is, we match households in the Interview Survey with households in the Diary Survey that have similar household expenditure changes.

Construction of household expenditure changes

The problem with implementing this strategy, however, is that we do not have data on the same set of items for households from both the Interview and Diary surveys. We resolve this problem differently for each Survey.

[10] Similar to other studies of household inflation rates and group price indices, we assume that the price changes are not household specific. This may lead to underestimates of the dispersion in inflation experiences across households if, for instance, there are geographic differences in price changes. We do not have sufficient geographic detail in the CEX to pursue this.

As reported in Table 1, approximately 95 percent of the expenditures of the average U.S. household are covered in the Interview Survey; as we have seen, however, about 22 percent of household expenditures are measured more precisely in the Diary Survey, which is part of the motivation for incorporating the Diary Survey data. For each household *j* in the Interview Survey during year *t-1*, we calculate a *household expenditure change*, denoted by $\Delta e^*_{j,t}$, which is based on the individual expenditure items and related expenditures taken from the Interview Survey.

Table 1 also shows that only about half of the expenditure items of the average U.S. household are covered in the Diary Survey. As a result, household expenditure changes similar to $\Delta e^*_{j,t}$ cannot be constructed for Diary households. Instead, for these households, we construct *predicted* household expenditure changes based on their demographic characteristics. We denote the predicted household expenditure change for Diary household *k* at time *t*, by $\Delta \hat{e}^*_{k,t}$.

For this purpose, we select a set of demographic characteristics that are available for households in both the Interview and the Diary Survey.[11] Let this set of characteristics for household *j* be summarized by the vector $\boldsymbol{x}_j$.[12] Then the first step in our construction of the predicted household expenditure changes for the households in the Interview Survey is to estimate the equation

$$\Delta e^*_{j,t} = x'_j \hat{\beta}_t + u_{j,t} \tag{2}$$

for all households *j* that are respondents of the Interview Survey in year *t*. We then use this regression to calculate the predicted household expenditure change for households in the Diary Survey using

$$\Delta \hat{e}^*_{k,t} = x'_k \hat{\beta}_t \tag{3}$$

[11] The detailed set of characteristics is listed in Appendix B.

[12] This vector does not depend on time because we only consider characteristics that households report once while being part of the CEX sample.

and use this measure to implement our matching method. This would be the predicted household expenditure change of household *k* in year *t* if it were part of the Interview Survey sample. For comparability purposes, we also compute the predicted household expenditure change for households *j* in the Interview Survey.

We define the distance between household *j* from the Interview Survey and household *k* from the Diary Survey as the absolute difference in predicted expenditure changes:

$$\delta_t(j,k) = \left|\Delta\hat{e}^*_{j,t} - \Delta\hat{e}^*_{k,t}\right| \tag{4}$$

and use this as the basis for our matching method.

Matching method

For each household *j* in the Interview Survey in period *t*, our matching method consists of the following steps:

I. ***Split up the sample period*** We divide the four quarters that the household is in the Interview Survey sample into 12 months. For each of these periods, the household will be matched with a different Diary household.

II. ***Condition on income group*** Each of the households is only matched with other households in the same income category.[13] The predicted household expenditure change regressions are also run conditional on the income group.

III. ***Find a set of nearest neighbors*** For each monthly period find the 20 households, *k*, in the same income category from the Diary Survey sample with the smallest $\delta_t(j,k)$ from the "seed" Interview household *j*.

IV. ***Randomly match the household with one of the nearest neighbors*** Randomly choose one of the 20 closest neighbors of household *j* as a match for that monthly period.[14]

[13] We use four cells of CEX income data: income <$15,000, $15,000 <income<$40,000, $40,000<income, and incomplete response to income question.

[14] Each diary spans two weeks of expenditures; however, we only pick one diary match per month (and scale each to represent one full month of expenditures). Matching to more diaries would likely produce too many of the same diary matches across Interview households.

In any given month, the number of households participating in the Diary Survey is much smaller than in the Interview Survey, so our matching method results in a one-to-many match: Every Interview household is matched to only one Diary household in a monthly period, but each Diary household could be matched with more than one household in the Interview Survey. We use random matching among the 20 nearest neighbors to reduce the repetition of Diary matches without significantly increasing the distance between each Interview household and its Diary match.

The result of this matching method is that, for each household in the Interview Survey, we construct a 'hybrid' household, the expenditure data for which consists of the expenditures taken from the Interview Survey for the 121 items in the fourth row of Table 1 and of the expenditures taken by summing across the 12 matched Diary Survey households for the other 126 items. We construct such 'hybrid' households for all Interview Survey respondents that are part of the CEX files over the period 1984-2005.[15] This procedure yields, in total, data for 86,492 such 'hybrid' households.

For each of these 'hybrid' households, we then calculate the household-*j*-specific inflation rate in year *t*, denoted by $\pi^{h}_{j,t}$ using a fixed-weighted Laspeyres formula, such that

$$\pi^{h}_{j,t} = \sum_{i} s_{i,j,t-1}\pi_{i,t} \text{ , where } s_{i,j,t-1} = e_{i,j,t-1} \Big/ \sum_{m} e_{m,j,t-1} \tag{6}$$

4. Matching diagnostics

To illustrate the quality of the matches that we use to construct the hybrid households, we provide diagnostic statistics related to these matches. These diagnostics focus on two main issues. The first is the explanatory power of the demographic characteristics that we use for the predicted household expenditure changes. The second is a comparison of our matches with those obtained using other commonly used matching methods. In addition to these matching diagnostics, we also provide evidence on how frequently we matched households from the Interview Survey with Diary households with similar demographic characteristics, focusing on

[15] The resulting household-specific inflation measures cover the period 1984-2004.

those characteristics that are often used to construct group price indices. This allows us to compare our matches with those considered in such indices.

Goodness-of-fit of household expenditure regressions

Our aim is to match households with similar expenditure changes. Unfortunately, since the expenditure data for Diary Survey participants do not cover the whole range of expenditures, the best we can do is to impute the expenditure changes for households in the Diary Survey based on demographic characteristics. Hence, it is important to assess the quality of this imputation. We do so by presenting goodness-of-fit statistics for the regression in equation (2) for the different income groups over time.

Over the sample period, on average, 17 percent of the variation of household expenditure changes across Interview Survey respondents can be attributed to *between* income cell variations, while the remaining 83 percent is due to *within* income cell variations.

Table 3 contains the R^2's of equation (2) for each income group and all years in our sample. The first row lists the average R^2 for each group over the sample period. Within each income group, equation (2) explains on average between about one-fifth and one-third of the variation in household expenditure changes. Thus, including the demographic variables within each income cell substantially reduces the imputation error for the imputed household expenditure changes as compared to simply using the average household expenditure change within income cells.

In total, equation (2) explains about 40 percent of the variation in expenditure changes not explained by income across households in the Interview Survey.

Comparison with alternative matching methods

The above regression results thus suggest that matching Interview Survey households with the nearest neighbors from the Diary Survey within each income cell may substantially improve the quality of the matches as compared to randomly matching Interview households with Diary households within each income cell.

Table 4 lists the quality of the matches for our method and compares it with those for two alternative matching methods: (*i*) randomly match a given household with another household, without conditioning on income (or any other demographic characteristics); (*ii*) randomly match a given household with a household in the same income cell. We quantify the quality of the matches by the average match-distance under each method:

$$\frac{1}{J}\sum_{j=1}^{J}\bar{\delta}_t(j,m), \text{ where } \bar{\delta}_t(j,m) = E\left[\delta_t(j,m)\right] \text{ and } m \in \{\text{matched}, \text{random in cell}, \text{random}\} \qquad (7)$$

and *matched* indicates the matched household from the Diary Survey, *random in cell* means matched with a random household in the same income cell, and *random* means randomly matched with any household in the sample. In the case of our matching method, *E(.)* indicates the actual average match distance between an Interview household *j* and its 12 Diary matches; in the case of the alternative matching methods, *E(.)* denotes the expected value over the entire sample of possible random matches.

As can be seen from Table 4, our matching method (column 1) produces average match distances that are much lower than under the two alternative methods (columns 2 and 3). Unconditional random matching generates, on average across years, expected match distances equal to $590: this is the absolute difference in expenditure changes between two randomly matched households. Conditioning on income reduces expected match distances by about 14% on average relative to random matching. Our method – matching within income cells on predicted expenditure changes – reduces average match distances by about 88% vs. random within-cell match, down to $62. Together, these two steps reduce expected match distances by 90%, on average, relative to unconditional random matching.

Thus, our matching method seems to match households quite well with respect to the expenditure change metric we have chosen: in fact, the average expenditure change for an Interview household is about $740, on average across years, so our average match distance of $62 is very small (about eight percent) relative to that. Further, given our criterion, our matching method substantially improves upon two alternative matching

methods that do not use imputed household expenditure changes but rather just randomly match across different (sub-) samples.

Within-group matching

One interesting aspect of our matching strategy is that it does not condition on particular demographic characteristics, but instead uses regression analysis to quantify the effect of demographic characteristics on household expenditure changes. This approach is very different from the approach used to create group-price indices, like those presented in McGranahan and Poulter (2006), which are cost-of-living indices that are calculated conditional on households having particular demographic characteristics.

The particular groups for which group price indices are calculated are, most often, chosen based on particular policy or research questions rather than on statistical criteria. Because our matches, instead, are based on statistical evidence it is worthwhile to consider how often we match households with particular common demographic characteristics.

Table 5 reports the fraction of households that are actually matched with households with similar demographic characteristics for different categories as well as the fraction that would have been matched, on average, using random matching. Our matching protocol consistently improves the probability that two households share demographic characteristics relative to random matching, especially with respect to education categories. For instance, Interview households with college are matched on average with Diary households who also have completed college 42 percent of the time. The corresponding fraction under random matching would be 29 percent. This is a significant improvement considering that we do not explicitly match on demographics but rather on predicted expenditure changes.

Group price indices are a way to consider the mean inflation rate across households with different demographic characteristics. They do not, however, take into account within-group variation in household-specific inflation rates nor do they allow for a comparison of this variation with that in the overall economy. For such a comparison, one needs to consider the distribution of household-specific inflation rates,

unconditional and conditional on various household characteristics, which is what we do when we present our results in the next section.

5. Results

We present our results in two parts. First, we report some time series and cross-sectional evidence from our estimates of household-level inflation and compare it to historical patterns in various measures of aggregate inflation over time. Next, we construct estimates of the inflation distribution for households conditional on various demographic characteristics and compute cumulative inflation differentials across demographic groups over the sample period.

Comparison to overall CPI inflation

Figure 1 presents a time series comparison of our household-based inflation measures with two overall CPI measures: the CPI-U and the chain-weighted C-CPI (the latter is only available from 2000 onwards). In the plot, the shaded areas denote various percentiles of the distribution of inflation across households from our estimates; the red line represents the plutocratic index of overall inflation constructed from our estimates,[16] and the blue line denotes the median inflation across households.

The first thing to note is that our plutocratic index tracks CPI inflation quite well. The CPI-U index always falls within the [90^{th} – 10^{th}] percentile interval in the plot (i.e. within the bulk of our household inflation distribution). In the second half of our sample period (from 1995 onwards) our plutocratic index tends to be systematically lower than the CPI-U, although it again matches the CPI quite well in the last two years in our sample. Further, our index seems closer to the chain-weighted measure of the CPI (the C-CPI) than the official CPI-U measure – at least for the period over which the C-CPI is available. This is because our household inflation index is effectively an annually chain-weighted index and therefore resembles the C-CPI more than the CPI-U, which is fixed-weighted.

[16] The plutocratic index is a weighted average of inflation experiences across households, where the weights are each household's total expenditure level.

Secondly, our household inflation distribution exhibits a large amount of dispersion over the entire sample period under consideration. The width of the 90^{th}-10^{th} percentile interval ranges between about 100 and over 300 basis points over time: thus we observe a significant amount of inflation inequality across households, with large numbers of households experiencing significantly higher (or lower) inflation than the official CPI measure. For instance, in 2002 one in ten households was experiencing an annual inflation rate of about 0.4% or lower, whereas another 10 percent of households were experiencing annual inflation higher than 3.7%.

That inflation inequality is high is confirmed by the fact that the time series standard deviation of the CPI-U measure over this entire 21-year period is roughly comparable to the average cross-sectional standard deviation of household inflation from our estimates: the standard deviation of the CPI over time is 0.99 percentage points, while the average standard deviation of individual inflation across households is 0.94 percentage points. The economics profession typically focuses a lot of attention on variations in the CPI over time; here we show that the cross-sectional variations are comparably sized.

We compare our plutocratic inflation index with its democratic counterpart in Figure 2.[17] Just like Kokoski (2000), we find that the differences between the two indices are quite small, with a cumulative difference over the sample period of less than three percentage points. From 2000 onwards, our democratic index lies consistently above the plutocratic one: this is reflective of high gas prices in these years and of the larger weight that gasoline consumption has in poorer households' budgets relative to richer households. We will observe a similar pattern in our analysis by demographic groups.

Figure 3 contains a comparison of inflation estimates with and without Diary Survey information. The "without diary" measures are based solely on expenditure information from the Interview Survey. As mentioned earlier, the "with diary" measures do not simply add more expenditure items, but in some cases replace items in the "without diary" measure with ones from the Diary Survey that the BLS deems more detailed or more accurately measured (see Table 1).

[17] In a democratic index all households are equally weighted regardless of their expenditure levels.

The top panel of Figure 3 reports democratic means in the two cases. The without diary time series tends to be above the one with diary for most of the sample period. This is probably because the Interview-only measure is likely to over-weigh expenditure categories such as medical care and tuition, which tend to exhibit the highest inflation rates over this period. Thus, adding detailed expenditure information from the Diary Survey seems to be valuable in order to obtain a more accurate picture of time series trends, since it produces measures of household expenditures in which the weights of broad expenditure categories are closer to actual ones. The cumulative difference between the two mean measures is 5.4 percentage points over the sample period (on average, about 25 basis points per year); while small, this is an economically significant differential.

The bottom panel of Figure 3 reports the time series of the cross-sectional standard deviation of household inflation, with and without diary information. Interestingly, the with diary cross-sectional standard deviation is always smaller than that resulting from the Interview Survey alone, indicating that the addition of diary information results in less dispersion in inflation experiences across households. This is most likely because the variation in inflation rates across the food and personal care categories that are getting more weight when the Diary expenditures are added is less than that for the aggregate versions of those expenditures in the Interview Survey. Note that medical care and tuition have the highest average inflation rates and durable goods have the lowest inflation rates[18]; these are all categories that are taken exclusively from the Interview Survey.

Figure 4 reports our household inflation distributions for core and total inflation. Core inflation is defined over expenditure baskets that do not include food and energy. The dispersion of core inflation experiences across households is comparable to that of total inflation. Based on our estimates, a significant fraction of households (over 10 percent) actually experienced deflation in core expenditures over the period 2001-2003. This is consistent with the deflation scare of the early 2000s and is highly relevant in the current situation, following the 2008 financial crisis.

[18] See Hobijn and Lagakos (2005) for a detailed discussion of inflation differentials across item strata.

A novel contribution of our methodology, as we have demonstrated, is that it enables us to study the dispersion in inflation experiences across individual households, not just between broad demographic groups. We highlight this feature in Figure 5, where we report kernel density estimates of the household inflation distribution for two "typical" years: one, 1991, in which average inflation (and the CPI) was fairly high; the other, 2002, in which average inflation was rather low. The two distributions look very different. Inflation inequality across households is much higher in the "low inflation" year (2002) than in the "high inflation" example (1991).

This pattern is not unusual: in Figure 6 we report scatter-plots (with associated regression lines) of mean inflation vs. inflation inequality (measured as cross-sectional standard deviation in household inflation). There is a clear negative relationship between mean inflation and inflation inequality across households, for both total and core inflation, over our sample period. This runs contrary to the macroeconomic hypothesis, first posed by Okun (1971), according to which high average inflation leads to high inflation volatility (over time). The correlation between mean and variance of inflation in our cross-sectional analysis indicates the opposite is true. This kind of analysis is uniquely made possible by our approach that aims at constructing household level measures of inflation experiences.

Group analysis

We now turn to our analysis of household inflation experiences conditional on broad demographic groupings. The top panel of Figure 7 reports the time series of mean inflation (democratic indices) for "young" vs. "elderly" households.[19] Mean inflation is almost always higher for older households: this reflects elderly households' higher medical care costs. The cumulative inflation gap between the two groups is about nine percentage points over the entire sample period (or roughly 45 basis points per year, on average). While relatively small, this is still a significant differential between these two demographic groups.

[19] A young household refers to a household in which the reference person is younger than 61 years old. An elderly household refers to a household in which the reference person is older than 61 years old.

Interestingly, the inflation differential across groups is quite small relative to within-group inflation dispersion. The bottom panel of Figure 7 reports kernel density estimates for the distribution of young and elderly household inflation, in a representative year. The bulk of the two group distributions span roughly 400 basis points, which is much larger than the difference between the means of the two distributions. This suggests that group price indices might emphasize a dimension of heterogeneity that turns out to be relatively unimportant compared to the overall degree of inflation heterogeneity across households.

Other demographic groupings (reported in Figures 8, 9 and 10) indicate that young and elderly households exhibit the largest group differentials. The cumulative gap between poor and non-poor households over the entire sample period is only three percentage points; that between households with and without children is 4.3 percent; and that between low- and high-education households is 6.7 percent. As mentioned earlier, the gap between poor and non-poor households (Figure 8) exhibits a similar pattern to that of our plutocratic vs. democratic comparison in Figure 2. Poorer households experience higher inflation in the latter part of our sample period because of high gas and food prices over those years, which is accentuated by the fact that these items represent a larger share of poorer households' expenditures. Again, it is important to note that within-group inflation dispersion far outweighs between-group differentials, as evidenced in Figure 10 for households with different levels of education.

Limitations

Throughout, we have taken our results at face value. However, it is important to bear in mind some of the limitations in quantifying household-specific inflation rates. The limitations fall into three categories: (*i*) approximation bias in the price-index formula, (*ii*) possible mis-measurement of price increases the households face, and (*iii*) possible mis-measurement of the expenditures by households.

The Laspeyres price-index formula, (6), that we use has some well-documented limitations. Most notably it tends to underestimate changes in the cost of living because it does not take into account the welfare

benefits of the household's opportunity to change its consumption pattern in response to changes in relative prices.[20]

The main assumption that we make about price increases for households is that the percentage change in prices for each of the 247 items in our data is the same for all households. One reason households might face different price changes is because they reside in different parts of the country. In terms of price levels, Aten (2005) documents that prices in the NY suburbs in 2003 were the highest of any U.S. metropolitan area and were 48% higher than those in St. Louis, which had the lowest prices. Such level differences would not be an issue for our analysis if percentage changes in prices were the same. However, annual inflation rates also vary substantially across regions. For example, in 2003, the year for which Aten (2005) documents inter-area price differentials, prices in the Boston metropolitan area increased at a rate of 3.8 percent while those in Denver only increased by 1.1 percent, according to the BLS.

Ideally, we would take these inter-area inflation differentials into account in our calculation of household-specific inflation rates. This is infeasible, however, because the CEX does not report enough detail on the geographic location of individual households.

Another source of cross-household variation in price increases might be point of purchase and within-item-strata consumption behavior. In principle, such sources of variation can be picked up by scanner data on prices and sales at particular retail stores. However, no such data sets cover the whole spectrum of household spending. Recent evidence from one such scanner dataset that covers a limited set of purchases (see Broda and Romalis (2008)) suggests that, when one compensates for the benefit of new goods, the prices that low-income households pay for their non-durable consumption goods have risen less than those paid by their higher income counterparts.

There are two important drawbacks in the way we quantify household expenditures. The first is the way we deal with the purchases of durables. We do not correct for durable purchases and therefore introduce

[20] Boskin et. al. (1996) contains a detailed discussion of substitution bias in the CPI.

cross-household variation in inflation rates due to some households devoting a large share of their expenditures to buying a consumer durable, like a car or a couch, which they only buy infrequently.[21]

The second drawback is that, because the sample size of the Interview Survey is much bigger than the Diary Survey, the matching of these surveys might come at the cost of not capturing some of the cross-household variation in spending patterns because some Diary households get matched up with more than one Interview household. While this is mitigated by our specific matching protocol (which randomizes within the set of 20 nearest Diary potential matches), the results in Figure 3 suggest that this may indeed be a valid concern.

Finally, one important caveat is that our method of merging the information from the Interview and the Diary surveys in the CEX necessarily introduces some noise in the measurement of individual inflation experiences by households. It is not clear whether this additional noise, by itself, may increase or reduce the observed variation in inflation experiences.

6. Conclusion

In this paper we present new measures of household-level inflation experiences based on comprehensive information from the Consumer Expenditure Survey. Our approach matches households in the Interview and the Diary surveys of the CEX, using a two-step method that first conditions on income cells and then matches households within income cells based on predicted expenditure changes as a function of observed socio-economic attributes. We argue that this matching protocol yields good results in terms of our expenditure change metric and that the resulting household inflation measures are based on a more accurate and more complete description of household expenditures than previously available.

We find that our household-based inflation measures track aggregate measures such as the CPI-U quite well and that the addition of diary survey data introduces small but significant differences in the measurement of household inflation. The distribution of inflation experiences across households exhibits a

[21] This is a limitation of all commonly used price index formulas. Reis (2009) proposes an alternative measure of inflation that aims to solve this problem.

large amount of dispersion over the entire sample period, with the width of the 90^{th}-10^{th} percentile interval ranging from about 100 to over 300 basis points. Further, we uncover a significantly negative relationship between mean inflation and inflation inequality across households; this is a novel finding made possible by our approach, and it opens up interesting macro-economic questions for future research.

With regard to our group-level measures, we find that within-group inflation inequality is much higher than between-groups differentials. The larger gaps occur between young and elderly households, with the latter consistently experiencing higher inflation than younger households. The cumulative gap over our sample period between these two groups is equal to about 9 percentage points, or roughly 45 basis points per year on average.

Finally, it is important to note that, given the various potential sources of variation in price increases experienced by different households (be it for geographic or other demographic differences), our exercise is likely to under-estimate the degree of heterogeneity in household-specific inflation rates, since we impute the same price changes to all households. Thus the large amount of inflation inequality across households that we find in our analysis seems even more salient and warrants further investigation.

References

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

Aten, Bettina H. (2005), "Report on Interarea Price Levels," *Bureau of Economic Analysis WP2005-11*.

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.

Broda, Christian, and John Romalis (2008), "Inequality and Prices: Does China Benefit the Poor in America?," *mimeo*, University of Chicago.

Bureau of Labor Statistics (BLS) (2008), *Handbook of Methods*, BLS Publishing: Washington, DC.

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

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

Hobijn, Bart, and David Lagakos (2005) "Inflation-Inequality in the United States," *Review of Income and Wealth*, 51, 581-606.

Idson, Todd, and Cynthia Miller (1999), "Calculating a Price Index for Families with Children: Implications for Measuring Trends in Child Poverty Rates", *Review of Income and Wealth*, 45, 217-233.

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

McGranahan, Leslie, and Anna Paulson (2006), "Constructing the Chicago Fed Income Based Economic Index – Consumer Price Index: Inflation Experiences by Demographic Group: 1983-2005," *Federal Reserve Bank of Chicago WP 2005-20*.

Okun, Arthur M. (1971), "The Mirage of Steady Inflation," *Brookings Papers on Economic Activity*, 2:1971, 486- 98.

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

Reis, Ricardo (2009), “A Dynamic Measure of Inflation,” *mimeo*, Columbia University.

Appendix: Data Details

A. CPI-CEX matching

To construct the inflation rates, we match the expenditures collected in the CEX to the price series in the CPI. The following discusses the price data, the expenditure data, and the methodology we used to do this matching.

In this paper, we use seasonally unadjusted data from the CPI for all urban consumers (CPI-U) from 1984 to 2005. The CPI collects price data about a set of items – referred to as "item strata" – that are organized in a tree structure, with price data for detailed items aggregating up to produce price data for the broader items. All of the items the CPI collects data for ultimately aggregate up to a broad "All Items" category.[22] Over the period of time in our sample, the CPI has occasionally added new items to the tree structure and discontinued collecting data for some existing items. Thus, data is not available for all items in all of the years in our sample.

The CEX collects expenditure data about a set of expenditure categories – identified by "Universal Classification Codes", or UCCs – through the Interview Survey and the Diary Survey. In the Interview Survey, data is collected on broad expenditure categories that cover about 95% of an average household's expenditures. In the Diary Survey, data is collected on detailed expenditure categories. For example, aggregate expenditures on food are collected in the Interview Survey and expenditures on a variety of detailed food items are collected in the Diary Survey. By design, then, there is a significant amount of overlap between the data collected in the two surveys. Even within a single survey, there is some degree of overlap across the purchases covered in each category. Thus, when summing across a set of expenditure categories to calculate measures of total expenditures, we want to be careful to only include those UCCs that represent a unique set of purchases. For calculations using only data from the Interview Survey, we use the list of UCCs the BLS uses to calculate each Interview household's "Total Expenditures". For calculations that combine data from the Interview and Diary surveys, we use the BLS's "Integrated List" as a guide for

[22] This is the item from which total CPI inflation is calculated.

which UCCs to use from which survey. The BLS uses this list for their own calculations of total expenditures that combine the two surveys.[23]

To match the CEX expenditures to the CPI price data, we first associated every expenditure series collected in the CEX (i.e. each UCC) with an inflation rate for every year data was collected for that UCC. To do this, we manually assigned each UCC an item stratum from the CPI by finding the item stratum description that best matched the UCC's item description.[24] Note that this method allows multiple UCCs to be mapped to the same price level (if, for example, some of the UCC categories were more detailed than the CPI categories) but does not allow multiple price levels to be mapped to a single UCC.

We then assigned the UCCs the inflation rates from their associated item strata. For years in which the BLS did not collect data for a UCC's item stratum match, we used data for the next item stratum up in the tree that had data available in that year.

B. Diary-Interview Survey matching

We obtain data on household expenditures and demographics from the Consumer Expenditure Survey (CE). The CE has two components: (1) an interview survey, in which the BLS interviews households about broad expenditure categories once per quarter for four quarters and (2) a diary survey, in which the BLS asks households to keep a diary of detailed expenditures for two weeks.[25] Thus, the interview survey captures the bulk of household expenditures, but lacks the detail behind those expenditures (for example, participants report expenditures on food at home, but not on specific food items), while the diary survey captures the detail, but is likely to ignore large, infrequent expenditures.

Before we discuss our objective and methodology in combining the interview and diary surveys, it's helpful to note a few technical details about the surveys. First, both of the surveys are designed as rolling panels. In the interview survey, this means that the first of the four quarterly interviews can start on any day

[23] In particular, they use it to calculate the aggregate expenditure baskets used as weights in CPI calculations.

[24] Some UCCs were not assigned to a CPI item stratum because they did not have a close enough match in among the set of CPI item strata. Rather than assign them to the broad "All Items" series, we chose to leave them unmatched. Those UCC would, therefore, be automatically excluded from any inflation rate calculations.

[25] From here, participants in the two surveys will be referred to as "interviews" and "diaries", respectively.

of the year (the following three interviews then occur every three months after the first). In the diary survey, this means participants can start their two-week diary on any day of the year. Given this feature, to allocate the interviews into a single year and the diaries into a single month, we assign all interviews (diaries) whose expenditures were first reported in a given year (month) into that year (month). Second, the two surveys are independently constructed samples from the US population. In other words, the same people do not necessarily participate in both the interview and diary surveys. The following explanation of our matching methodology explains how we account for this feature.

Our objective is to obtain as complete and detailed a picture of household expenditures as possible; to do this, we combine the information about large, infrequent purchases from the interview survey with the information about detailed, frequent purchases from the diary survey. As the base of our sample, we use all of the interviews that completed all four quarters of the survey in order to examine yearly expenditures. We then match each interview to twelve different diaries, one from each month the interview reported its expenditures. To obtain yearly expenditures from those diary matches, we scale up each diary to represent a full month (each diary spans two weeks), then sum expenditures across the twelve diaries.[26] We then follow the BLS's protocol for aggregating expenditures across the interview and diary surveys to decide how to combine expenditures from each interview and its diary match. For example, we do not use food from the interview survey; instead, we use each of the detailed food expenditures from the diary survey.

The purpose of the matching is to use diaries to impute each interview's detailed expenditures, so we want to match each interview with diaries whose detailed expenditures best reflect what the interview would have reported. To do this, we use a two-step method based on demographic data collected in the surveys. First, we divide the interviews and diaries into four "cells" based on income brackets. Within each income

[26] This maintains most of the seasonality of expenditures, which would be progressively lost if diaries were scaled up to represent larger periods of time. At the same time, it maintains a relatively large number of diaries for the interviews to be matched with; as you force interviews to be matched to more diaries (e.g. if each interview were matched to 24 diaries, one for each two-week period of expenditures reported), the set of diaries to choose from for each match decreases.

cell, we regress yearly expenditure change[27] on a set of demographics for the interviews in that cell.[28] We then use the coefficients from the regression in a given income cell to predict yearly expenditure changes for both the interviews and the diaries in that cell. Using this two-step method (division into cells and predicting expenditure change), we create a single measure for all survey participants that incorporates all of the information contained across a wide set of demographics to the degree that each demographic attribute explains variations in yearly expenditure changes.

Once we have these predicted expenditures changes, we create a distance matrix within each income cell where the interviews are the rows, the diaries are the columns, and the values are the distance between each interview and each diary's predicted expenditure change. Then, for each interview and for each month that interview's expenditures span, we find the twenty "closest" diaries and randomly match the interview to one of those twenty diaries. As noted above, each interview will receive twelve diary matches using this method, one for each month the interview's expenditures spans. In some cases, multiple interviews may be matched to the same diary for one or more months, both because there is a random element in the match process and because there are usually not enough diaries in a given cell to assign each interview a unique diary in each month.

An obvious alternative to this method would be to divide the interviews and diaries into many demographic cells (e.g. a cell for low-income white households in the Midwest, a cell for high-income Asian households in the Northwest), then match interviews and diaries within those more specific cells. The difficulty with this method is that once the cells are defined by more than one demographic characteristic, the number of interviews and diaries within some of the cells becomes very small or even zero. Thus, this method would leave many interviews without any possible diary matches.

[27] Defined as the level change in yearly expenditures from one year to the next, using price changes for individual items from the CPI and the set of expenditures collected in the interview survey that the BLS considers to represent "total expenditures".

[28] Census region, urban/rural, housing tenure, number of persons in household younger than 18, number of persons in household older than 64, family type, number of earners, number of vehicles, age of the reference person, sex of the reference person, race of the reference person, marital status, family size, residence in a metro area.

Table 1. Classification of items across Diary and Interview Surveys

	Number of items	Percentage of items	Percentage of expenditures
Taken from Diary Survey	126	50	22
Only in Diary Survey	*14*	*5*	*4*
More reliable in Diary Survey	*25*	*10*	*6*
More detailed in Diary Survey	*87*	*35*	*12*
Taken from Interview Survey	121	50	78
Total	**247**	**100**	**100**

Note: Percentage of items based on the number of item strata used to construct the CPI.
Percentage of expenditures column based on 2004-2005 expenditure weights.

Table 2. Fraction of total items taken from Diary Survey for main categories

Category	Percentage of items	Percentage of expenditures
Food and beverages	95.0	80.4
Housing	15.2	4.3
Apparel	56.9	60.6
Transportation	6.8	1.8
Medical Care	14.7	1.8
Recreation	24.0	7.9
Education and communication	14.8	3.8
Other goods and services	38.7	9.0

Table 3. R^2 of equation (2) by year and income cell

Year	< $15,000	Between $15,000 and $40,000	> $40,000	Incomplete Response
Average	0.32	0.19	0.18	0.33
1984	0.42	0.20	0.19	0.43
1985	0.33	0.27	0.35	0.70
1986	0.42	0.20	0.19	0.46
1987	0.34	0.24	0.19	0.45
1988	0.41	0.23	0.21	0.35
1989	0.41	0.24	0.23	0.51
1990	0.36	0.14	0.21	0.21
1991	0.30	0.18	0.21	0.16
1992	0.32	0.16	0.28	0.29
1993	0.35	0.14	0.17	0.37
1994	0.34	0.15	0.11	0.38
1995	0.53	0.37	0.20	0.48
1996	0.24	0.14	0.14	0.28
1997	0.20	0.17	0.15	0.18
1998	0.20	0.11	0.12	0.20
1999	0.39	0.24	0.15	0.40
2000	0.24	0.22	0.18	0.19
2001	0.18	0.12	0.14	0.11
2002	0.30	0.13	0.13	0.20
2003	0.22	0.13	0.13	0.26

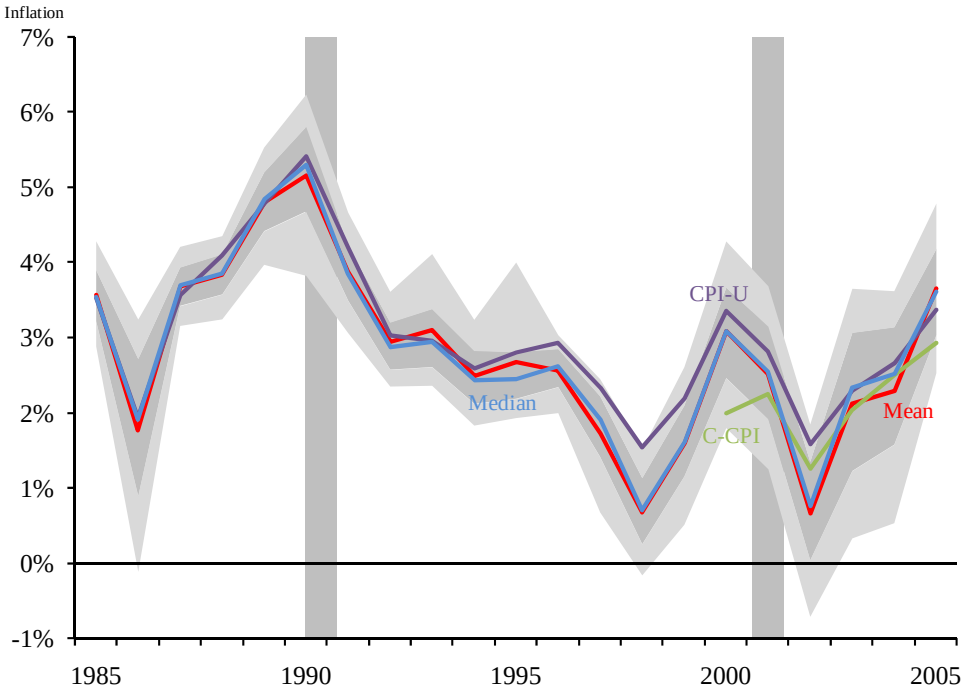
Table 4. Quality of matches for different matching methods

Year	Our matching method	Random matching within income cells	Random matching
Average	62	510	590
1984	56	375	464
1985	65	431	465
1986	54	433	549
1987	63	461	585
1988	83	561	714
1989	108	645	814
1990	79	550	653
1991	67	409	533
1992	43	530	542
1993	62	469	552
1994	70	527	606
1995	89	457	557
1996	65	441	482
1997	58	584	603
1998	37	397	424
1999	64	598	682
2000	47	463	502
2001	34	472	489
2002	44	571	615
2003	45	603	676
2004	58	723	882

Table 5. Fraction of households matched by demographic characteristics

Group	Subgroup	Fraction matched	Population share
Age of head of household	Young (61-)	79%	73%
	Old (61+)	33%	27%
Presence of kids	Yes	44%	37%
	No	67%	63%
Education level	Less than high school	36%	9%
	High school or some college	67%	61%
	College	42%	29%

Figure 1. Comparison of plutocratic index with CPI-U and C-CPI



Note: Vertical bars indicate NBER recession dates, shaded area denotes 10th,25th,75th, and 90th percentile of plutocratic distribution.

Figure 2. Comparison of democratic distribution and index with plutocratic index

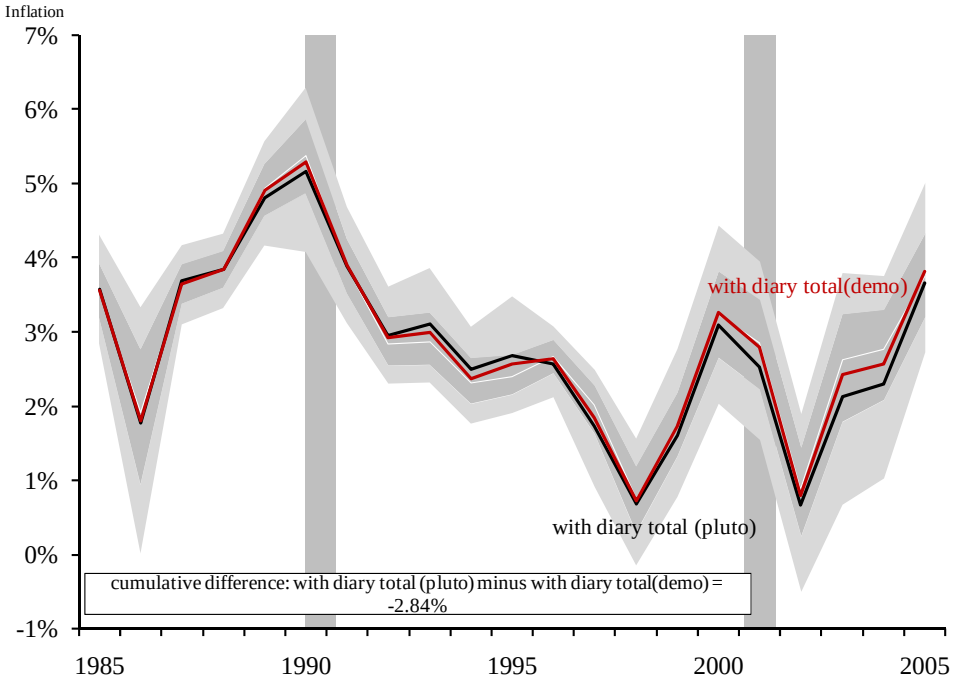


Figure 3. Mean and standard deviation of cross-household inflation distributions, with and without diary survey.

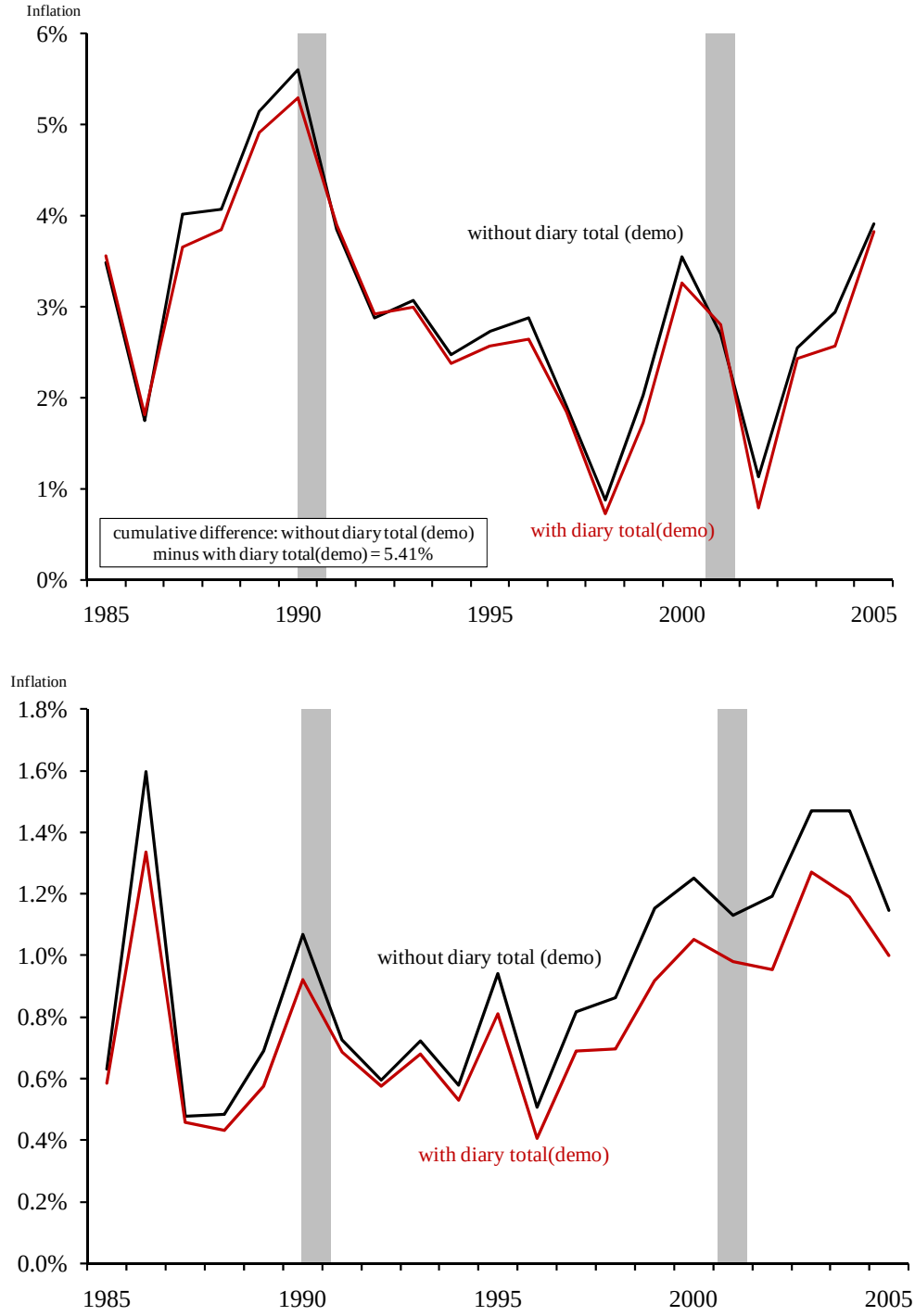
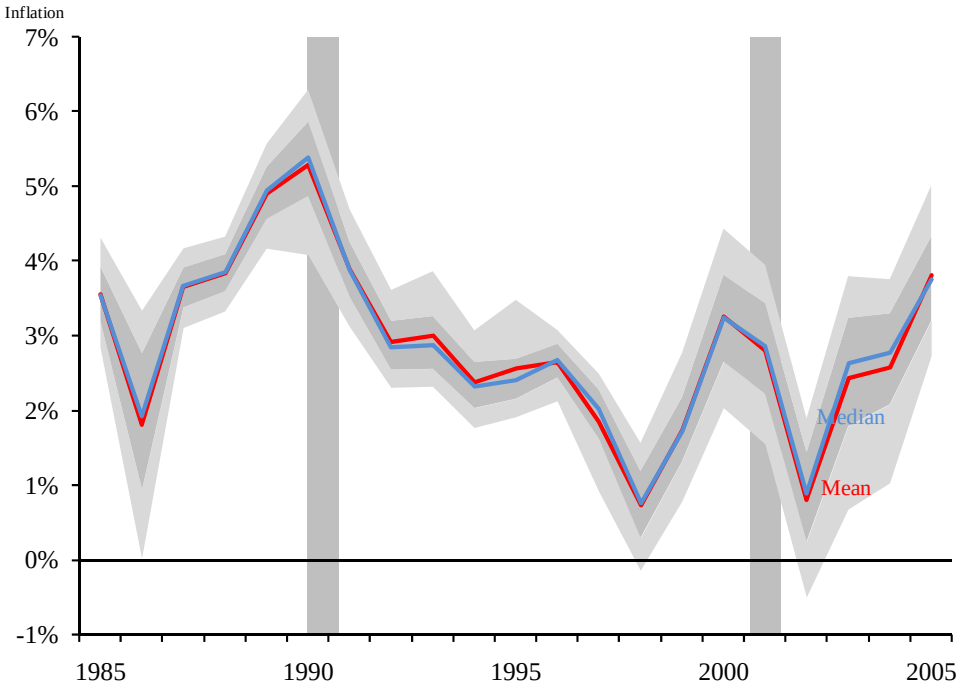
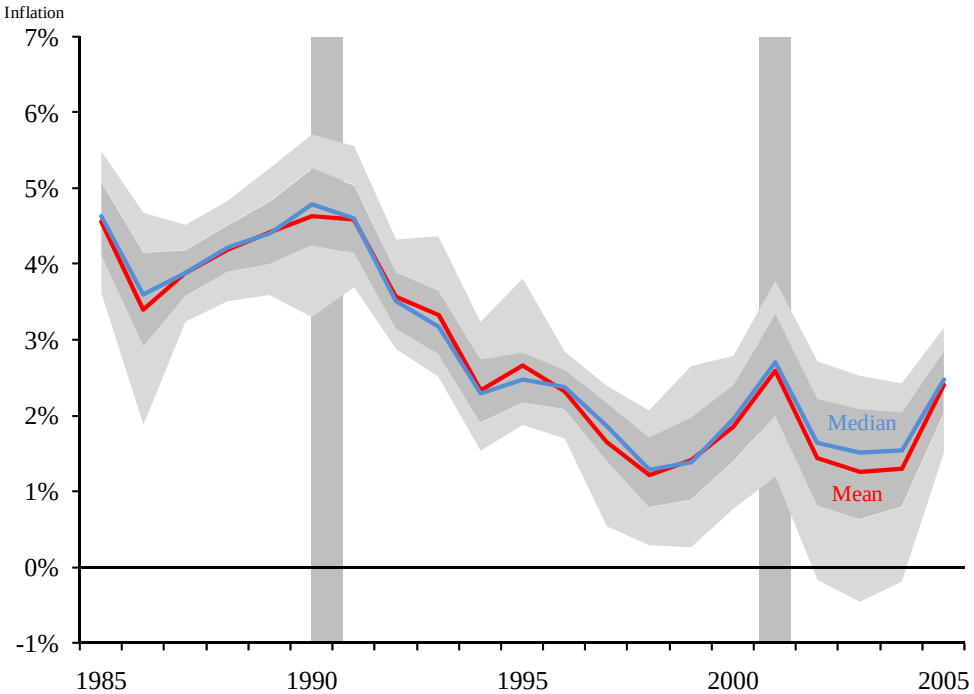


Figure 4. Distributions of core and non-core inflation rates



Total inflation, democratic distribution



Core inflation, democratic distribution

Figure 5. Estimated household inflation distributions, 1991 and 2002.

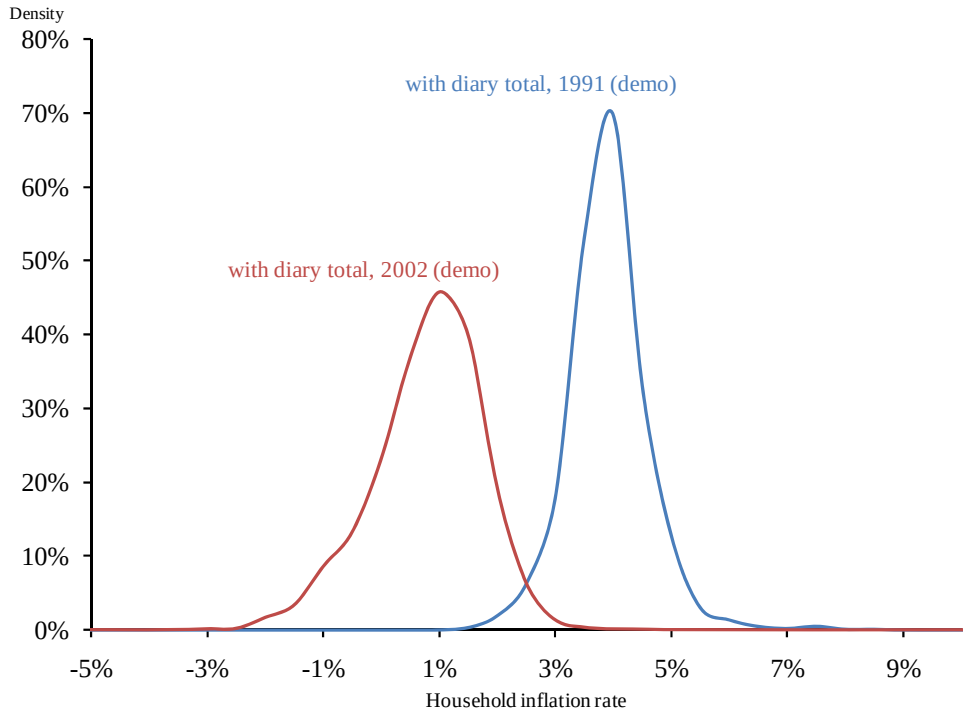


Figure 6. Mean inflation versus inflation inequality

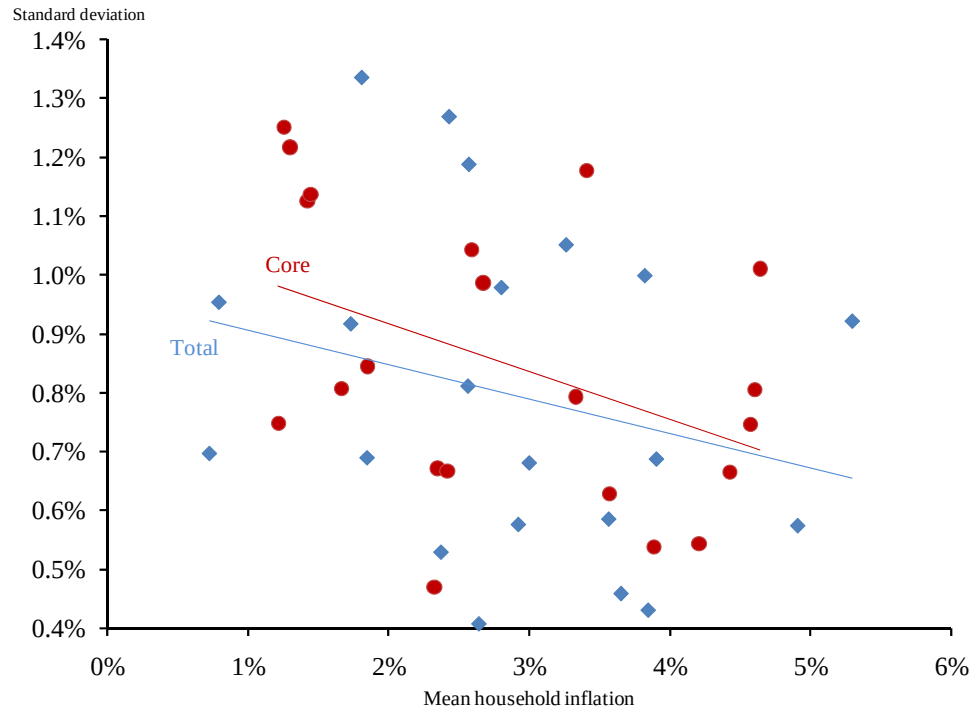
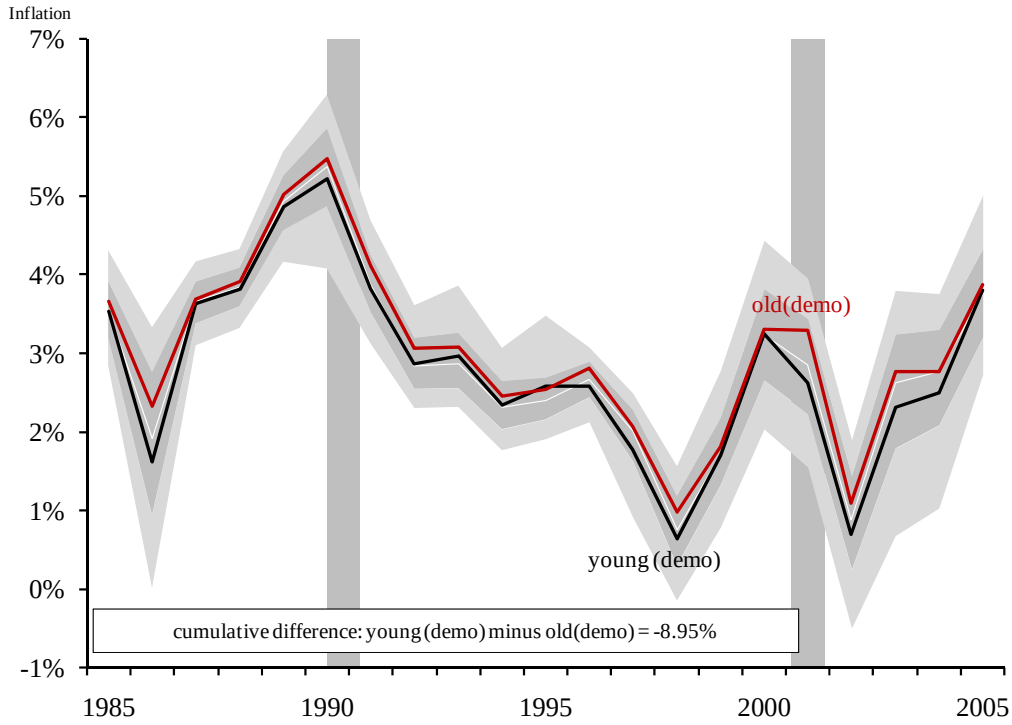


Figure 7. Households with old and young reference persons



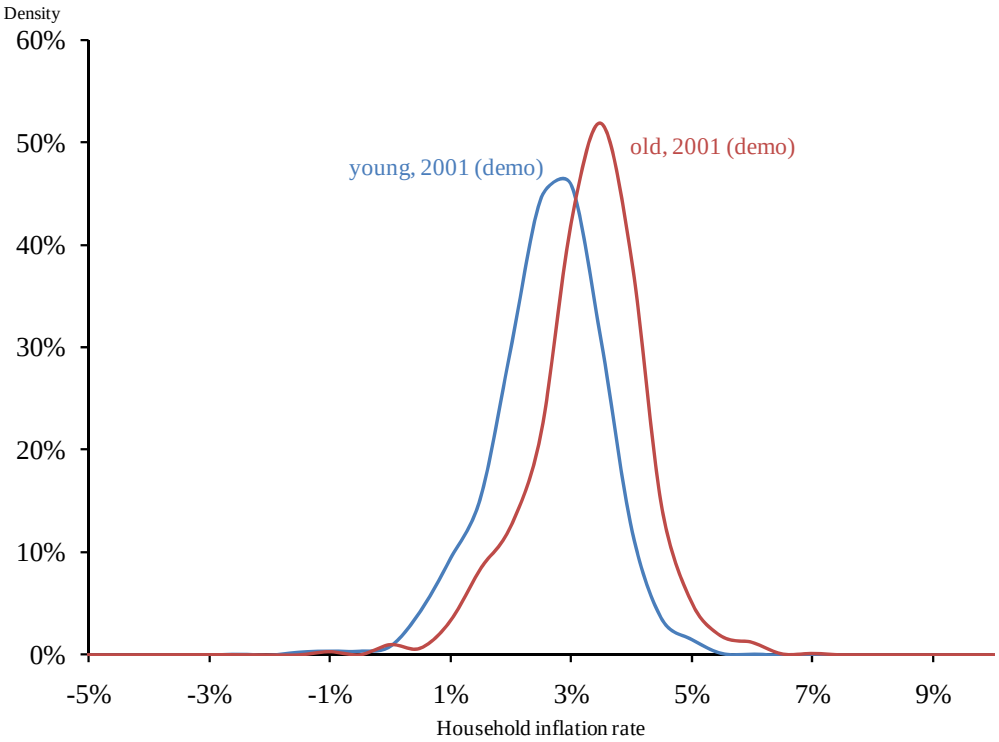


Figure 8. Poor vs. Non-Poor

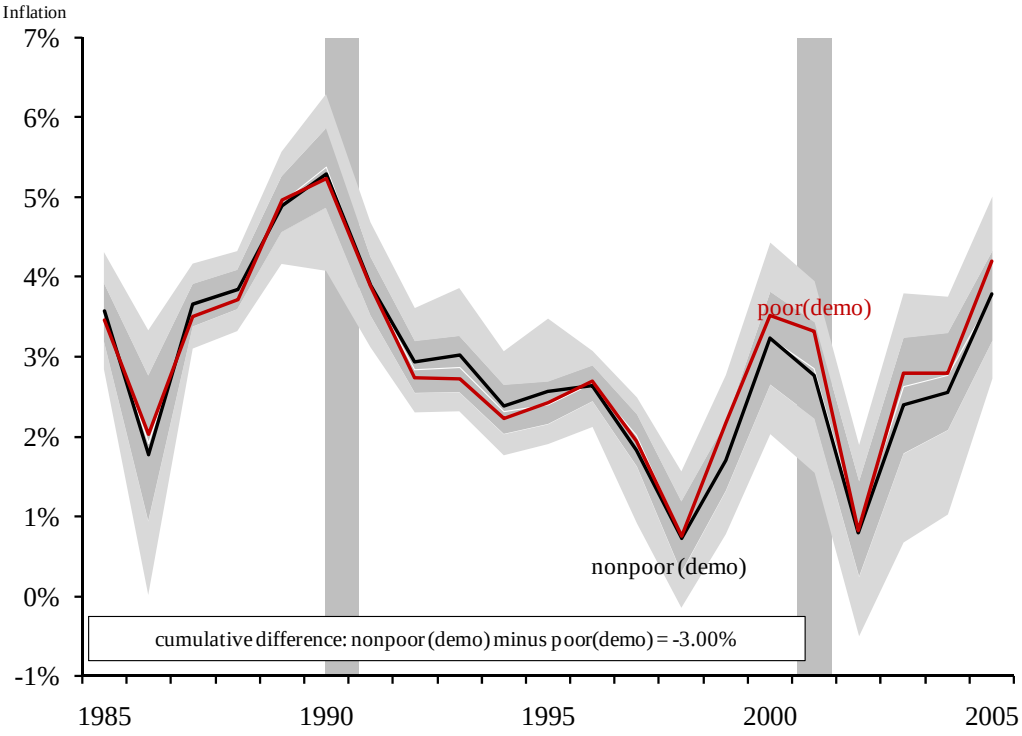


Figure 9. With-children and without children

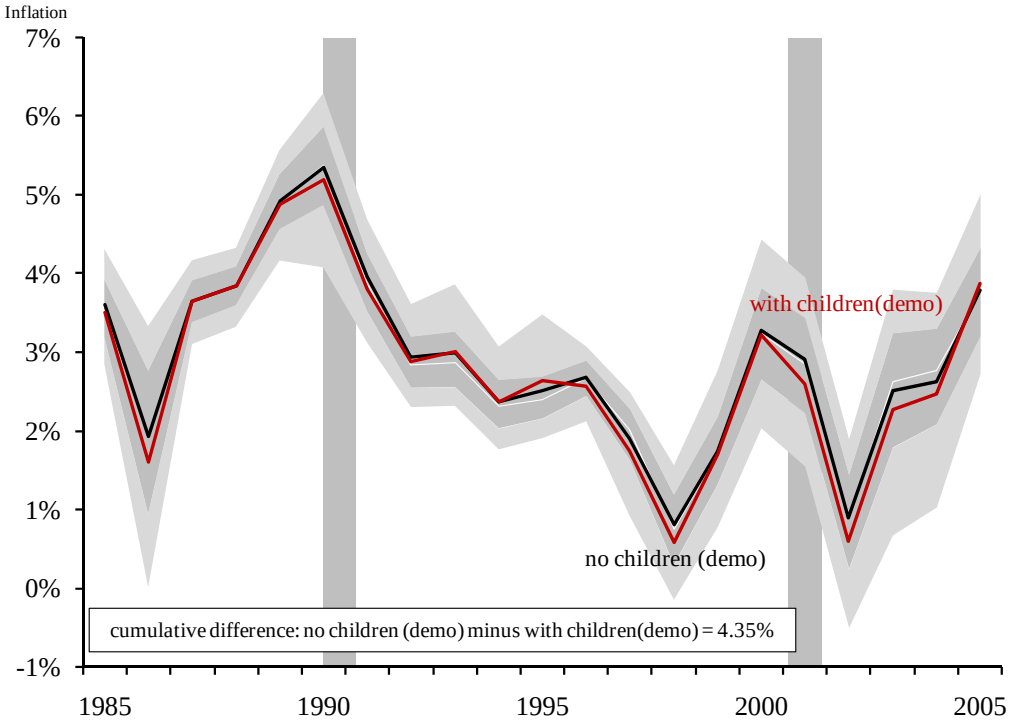


Figure 10. Less than High School vs. at least College households