DISAGGREGATE EVIDENCE ON THE PERSISTENCE OF CONSUMER PRICE INFLATION

Todd E. Clark

DECEMBER 2003

RWP 03-11

Research Division
Federal Reserve Bank of Kansas City

Todd Clark is a vice president and economist at the Federal Reserve Bank of Kansas City. The views expressed herein are those of the author and do not necessarily reflect the views of the Federal Reserve Bank of Kansas City or the Federal Reserve System. The author gratefully acknowledges research assistance from Taisuke Nakata and the helpful comments of: Ted Juhl, Max Marcellino, Serena Ng, Mototsugu Shintani; seminar participants at the Federal Reserve Bank of Kansas City and University of Michigan; and participants of the 2003 MEG meetings.

Clark email: todd.e.clark@kc.frb.org

Abstract

This paper uses disaggregate inflation data spanning all of consumption to examine: (i) the persistence of disaggregate inflation relative to aggregate inflation; (ii) the distribution of persistence across consumption sectors; and (iii) whether persistence has changed. Assuming mean inflation to be unchanged within samples, the average persistence of disaggregate inflation is consistently below aggregate persistence. Taking into account an early 1990s shift in mean inflation identified by break tests—including tests applied to systems of disaggregate equations—yields much lower estimates of both aggregate and disaggregate persistence for 1984-02. But with the mean break taken into account, average disaggregate persistence is actually as great as aggregate inflation persistence. A factor model provides a natural framework for interpreting the relationship between aggregate and disaggregate persistence.

JEL classification: E31, E52, C22

Keywords: Inflation dynamics; structural breaks; relative prices; factor models

1. Introduction

The persistence of inflation has recently become the focal point of considerable research. At issue is not only the level of persistence but also whether persistence has changed over time. For many years, inflation was widely viewed as having near-unit root persistence in data starting in the late 1950s or early 1960s.[1] The most recent debate has been stirred in part by Cogley and Sargent (2001) and Stock (2001). Cogley and Sargent present evidence that inflation persistence had varied widely over time and recently fallen considerably.[2] Cogley and Sargent (2003) provide corroborating evidence, allowing for the time–varying volatilities that Sims (2001) and Stock (2001) suggest could account for Cogley and Sargent's earlier results.[3] Using an alternative econometric framework, Stock shows that persistence has remained consistent with a unit root in inflation. More recently, Pivetta and Reis (2003) have corroborated Stock's (2001) findings for a range of inflation measures and persistence estimates. Two recent multi–country studies, Benati and Kapetanios (2002) and Levin and Piger (2003), find that inflation persistence is generally well below unity once discrete shifts in mean inflation identified by standard break tests are taken into account. This persistence literature continues to grow quickly, as indicated by the surveys provided in such studies as Benati and Kapetanios (2002) and Levin and Piger (2003).

The rising interest in inflation persistence has been accompanied by growing interest in the behavior of inflation rates at more disaggregate levels, reflecting the potential for disaggregate analysis to shed light on models of price setting and the business cycle. Using a new data set on frequencies of price change in detailed consumer prices, Bils and Klenow (2002) examine the persistence of disaggregate consumer price inflation rates in relation to the measured frequencies of price change. They show that a simple dynamic model for inflation implied by the time–dependent

---

[1] Barsky (1987) documents a sharp rise in inflation persistence starting in 1960. Levin and Piger (2003) provide a more complete survey of subsequent studies documenting high persistence.

[2] Taylor (2000) and Brainard and Perry (2000), among others, also argue persistence has declined.

[3] Benati's (2002) results for TVP models estimated with frequentist methods also corroborate Cogley and Sargent's (2001) results based on Bayesian methods.

pricing model of Calvo (1983) — one that has become the workhorse of aggregate sticky price models of the business cycle — dramatically fails to fit the disaggregate data. With marginal cost characterized as a random walk, the Calvo model implies inflation in each sector should be a linear function of one lag of inflation and current marginal cost, with coefficients determined by the frequency of price adjustment. Bils and Klenow find the data to be sharply at odds with the Calvo–implied model. Among other problems, persistence appears to be far lower in the data than the measured frequencies of price adjustment imply, given the behavior of marginal cost.

In two other important examples of growing interest in the behavior of disaggregate prices, Erceg and Levin (2002) and Barsky, House, and Kimball (2003) examine the flexibility of the prices of durables and nondurables. Erceg and Levin focus on the monetary policy implications of the much greater interest rate sensitivity of the durables sector, using a two–sector DSGE model. Barsky, House, and Kimball argue that the more rapid response of durables prices to monetary policy shocks observed in the data poses a problem for sticky–price DSGE models. In such models, the greater flexibility of durables prices causes durables and nondurables consumption to move in offsetting directions, sharply reducing the effects of policy on aggregate output.

Motivated by these literatures and by the potential power gains of using disaggregate data to test for aggregate shifts (gains highlighted by Bai, Lumsdaine, and Stock (1998)), this paper examines the persistence of disaggregate consumer price inflation and the disaggregate evidence of any change in inflation persistence. More specifically, this paper uses disaggregate inflation series spanning all of consumption as measured in the NIPA accounts over the time period 1959-2002, to assess: (1) the persistence of disaggregate inflation relative to aggregate inflation; (2) the distribution of persistence across consumption sectors; and (3) whether inflation persistence has changed.[4] In assessing persistence, this paper departs from Bils and Klenow (2002) by using one

[4] This paper's contributions on the first and third issues build on Yoon's (2001) analysis of breaks and persistence in a set of 23 "randomly selected" CPI inflation series representing various levels of aggregation (but not spanning all of CPI spending). Yoon tests 1967-2000 data for evidence of multiple mean shifts and then finds

of the standard measures of persistence rather than the simple AR(1) measure Bils and Klenow use in light of their interest in the Calvo model, and by using Hansen's (1999) grid bootstrap to estimate accurate confidence intervals for persistence. In assessing the evidence of changes in mean inflation, break tests are applied to not only univariate specifications but also systems of disaggregate inflation equations, to further the existing evidence on aggregate breaks.

Assuming model stability within samples, this analysis shows that average persistence in disaggregate inflation rates is consistently below aggregate persistence, although the size of the gap hinges on how "average" is measured. Virtually all disaggregate series are lower in persistence than aggregate inflation, though a sizable proportion of disaggregate series are highly persistent. Series with higher persistence tend to account for larger shares of consumer spending. Contrary to what some might expect, there are no material differences in the persistence of durable goods, nondurable goods, and services: at the aggregate level, all three sectors display high inflation persistence. This paper's baseline results on disaggregate persistence relative to aggregate persistence differ from those of Bils and Klenow (2002) in certain respects because the AR(1) measure used by Bils and Klenow is systematically lower than persistence as measured in this paper by the sum of AR coefficients, especially at the monthly frequency used by Bils and Klenow. That said, any differences between this paper's baseline results and Bils and Klenow's do nothing to alter their fundamental findings that frequency of price change has little correlation with inflation persistence and that a simple inflation model implied by Calvo price staggering fails in characterizing the behavior of disaggregate inflation rates at the monthly frequency.

Rolling estimates of persistence from 1977 to 2002 indicate that persistence has drifted down only slightly. Despite the downward drift, aggregate persistence remains high enough that a unit root cannot be ruled out. While many disaggregate components display sizable reductions in persistence, the larger declines have tended to occur in components receiving relatively smaller

that allowing for the identified breaks lowers persistence as measured with simple autocorrelation functions.

weights. Weighted by expenditure shares, the decline in disaggregate persistence is comparable to that for aggregate data.

Taking into account the shifts in mean inflation evident in 1984-2002 data lowers the estimated persistence of both aggregate and disaggregate inflation. Break tests applied to aggregate and disaggregate inflation data yield consistent evidence of a mean shift in the early 1990s, but only weak evidence of shifts in the AR coefficients. In particular, tests for a common shift in the intercepts of AR equations applied to systems of equations yield strong evidence of a break in the early 1990s. In aggregate inflation rates, allowing an intercept shift in 1993:Q1 turns out to greatly reduce estimated persistence, but widens the estimated confidence bands enough that roots of unity often cannot be ruled out. Taking the 1993:Q1 mean shift into account also lowers the estimated persistence of disaggregate inflation rates, but not as sharply as in the aggregate data. Thus, once the significant 1993:Q1 break in mean inflation is taken into account, average disaggregate persistence is actually as great as aggregate inflation persistence.

To help interpret the results on disaggregate persistence relative to aggregate persistence, the paper concludes by examining the persistence of common and idiosyncratic components of inflation estimated with the factor model methods of Stock and Watson (2002a). Under the factor model interpretation, persistence in disaggregate inflation rates could be quite low while aggregate persistence is high if the common factor component is persistent but the idiosyncratic components, which typically account for most of the variation in inflation, are not. Estimates of factor models that assume mean inflation to be unchanged over the 1984-02 sample yield a common component with near–unity persistence and idiosyncratic components with much lower persistence. Taking a 1993:Q1 shift in mean inflation into account yields an aggregate or common component with estimated persistence of between .6 and .8, while having little effect on average disaggregate persistence relative to the stable–mean case.

The paper proceeds as follows. Section 2 describes the data. Section 3 presents the baseline persistence estimates, assuming no changes in mean inflation. Section 4 examines rolling estimates of persistence from 1977 through 2002 for evidence of changes in aggregate and disaggregate inflation persistence. Section 5 presents the results of break tests applied to AR models of aggregate and disaggregate inflation estimated with 1984-2002 data. Section 6 then reexamines the persistence of inflation taking into account the identified breaks. Section 7 uses a factor model to interpret the results. Section 8 concludes.

2. Data

The source data on price indexes and nominal expenditures for all components of consumption as measured in the NIPA accounts (hereafter referred to as PCE, for *personal consumption expenditure*) are taken from a spreadsheet available on the BEA's website. These data permit breakdowns at various levels of aggregation, ranging from a simple breakdown of PCE into three series on durables, nondurables, and services to a breakdown of PCE into more than 150 series such as new domestic autos, new foreign autos, televisions, etc. In most cases, the data are continuously available from 1959 through 2002. The results below focus on so–called *core* inflation — inflation excluding food and energy. Food and energy prices are widely known to be more flexible and volatile than other prices; monetary policymakers routinely focus on ex food and energy measures of inflation.

This analysis reports results for data broken into several levels of disaggregation, each spanning all of core PCE. At what the discussion below refers to as the "aggregate" level, results are reported for total PCE, durables, nondurables, and services inflation (all excluding food and energy).[5] What the discussion below refers to as "disaggregate" results are reported for four levels of aggregation, ranging from 11 to 156 components. The first, 11–component level breaks durables, nondurables,

[5] Because the BEA reports only total nondurables and services series, core series are constructed by removing the relevant food and energy components, using the simple Tornqvist approximation to an ideal price index recommended by Whelan (2002). As noted by Whelan, the Tornqvist approximation is very accurate. For example, a Tornqvist–based measure of core PCE inflation is only trivially different from the BEA's reported chain price index for core PCE.

and services inflation into their most aggregate components. The second, 46–component level drills down another layer of detail, breaking the 11 components into their most aggregate components. The third and fourth levels of aggregation drill further into the PCE detail. Reflecting the varying levels of detail available across sectors, the third and fourth levels of aggregation include some higher aggregation level components. For example, the third level of aggregation includes the second level series for (1) jewelry and watches, (2) shoes, and (3) tobacco products because no further detail on these categories is available in the PCE data.[6] Appendix Tables 1-3 list the series used in aggregation levels 1-3. In the interest of brevity, in presenting disaggregate results the paper focuses largely on the third level of aggregation, which consists of 109 series, although some basic results are presented for all levels. As the limited results presented for aggregation level 4 indicate, using more detailed data would yield findings very similar to those reported for level 3.

The results below use monthly and quarterly data from 1959 through the third quarter of 2002. The analysis focuses on quarterly data for consistency with most other research on persistence and to speed the computation of the bootstrap–based confidence intervals. Select monthly results are presented for comparison with those of Bils and Klenow (2002). For consistency across time periods, the analysis uses only those series with complete data for 1959-2002. At the third and fourth levels of aggregation, this requirement eliminates small or modest numbers of components. In particular, aggregation level 3 omits three PCE series for which complete data are not available (the count of 109 series given above does not include the three excluded series). These series represent just 3.3% of ex food and energy PCE expenditures in 2001. Aggregation level 4 omits 25 series, representing 9.8% of core PCE expenditures in 2001. The larger problem with missing data in aggregation level 4 partly motivates the focus on aggregation level 3.

Calculations that require shares of nominal expenditures accounted for by a given component or set of components use shares that are fixed over time. In particular, the shares used are based

[6] In total, the 109 series included in aggregation level 3 include 12 level 2 series.

on annual expenditures in 2001 (the share of each component as a percentage of core PCE is reported in Appendix Tables 1-3). Fixed rather than time–varying weights are used in part to simplify calculations and in part because it is unclear what time variation should be allowed in some of the calculations. In some key instances, such as the rolling persistence estimates based on disaggregate data, allowing time–varying weights produces (unreported) results virtually identical to the reported fixed–weight results.

3. Baseline Persistence Estimates

After first explaining the paper's approach to measuring persistence, this section presents aggregate and disaggregate persistence estimates for 1959-02 and 1984-02, assuming no changes in mean inflation occur within those sample periods. The section then compares the reported estimates to those of Bils and Klenow (2002).

3.1 Estimating persistence and confidence bands

Following others such as Andrews and Chen (1994) and Levin and Piger (2003), this paper measures persistence as the sum of autoregressive coefficients (SARC). The sum of AR coefficients is directly related to the cumulative long–run response of inflation to a shock and the spectral density at frequency zero, two intuitive indicators of persistence. Pivetta and Reis (2003) summarize the pros and cons of the SARC measure and two alternatives, the largest autoregressive root and the half–life of responses to shocks.[7] In order to facilitate comparison with the results of Bils and Klenow (2002), in this initial evaluation of aggregate and disaggregate inflation, persistence is also estimated with the simple AR(1) coefficient.

In constructing the SARC persistence estimates, the lag order in each autoregression is determined using the AIC, allowing a minimum order of 1 lag and maximum of 6 lags for quarterly data and 18 lags for monthly data. Accordingly, persistence estimates based on 1959-02 data always use a regression sample with a start point of 1959:2 plus the maximum number of lags allowed,

[7] Cochrane (1988) proposes a non–parametric approach to estimating persistence.

while persistence estimates based on 1984-02 data use a regression start point of 1984:1 plus the maximum number of lags allowed.

Confidence intervals for the SARC persistence estimate and median unbiased estimates of persistence are calculated using Hansen's (1999) grid bootstrap. Even if true persistence is not unity, OLS estimates of the AR coefficients are biased downward, more so when persistence is high, and confidence intervals based on the asymptotic normal distribution have poor coverage properties. Of course, when persistence is unity, the coverage problems of the asymptotic normal–based interval stem in part from the fact that the asymptotic distribution of the persistence estimate is non–normal. Monte Carlo results in Hansen and Monte Carlo experiments conducted with some of the inflation variables used in this paper show that, in all cases — including those in which persistence is quite low — the grid bootstrap yields accurate confidence intervals and accurate unbiased estimates. In all cases, the bootstrap calculations use 999 draws and 101 grid points over a range given by the sample persistence estimate plus or minus four (OLS) standard errors.

3.2 Aggregate inflation persistence estimates

The SARC persistence estimates in Table 1 indicate that aggregate inflation is highly persistent in both 1959-02 and 1984-02 data. In most cases, the point estimates of persistence are about .9, with 90% confidence intervals that include unity, and median unbiased estimates that are often .95 or higher. The persistence of core PCE inflation, for example, is estimated to be .930 in 1959-02 and .907 in 1984-02. Although the point estimates based on monthly data for 1984-02 suggest slightly lower persistence, the median unbiased estimates show that the lower sample estimates relative to quarterly data for the same period or relative to the longer sample are due to greater bias in the monthly 1984-02 estimates.

The results also show that, contrary to what some might expect, there are no material differences in the persistence of durable goods, nondurable goods, and services, or in the persistence of non–housing inflation relative to overall inflation.[8] In quarterly data for 1984-02, for example, the persistence estimates for durables, nondurables, and services are, respectively, .921, .878, and .855; the median unbiased estimates are all essentially at or above 1, and the confidence intervals are qualitatively comparable. Similarly, for 1984-02 core PCE inflation, the overall and ex housing sample persistence estimates are .907 and .904, respectively.

These findings run contrary to the expectation that services inflation should be more persistent than goods inflation, because services production is relatively labor intensive, the markets for some key services such as medical care are quite different from the markets for goods, etc. These results also seem to run counter to the findings in Erceg and Levin (2002) and Barsky, House, and Kimball (2003) that prices of durable goods are more flexible than prices of nondurable goods, in that durables prices respond faster to monetary policy shocks. The contrast in results stems largely from differences in definitions of flexibility and whether housing purchases are included in durables.[9] As Barsky, House, and Kimball show, housing prices respond sharply to changes in monetary policy. Consequently, including housing purchases in durables, as Erceg and Levin do, makes durables prices appear relatively flexible in responding to policy shocks. In this paper's data, housing costs are measured as the implicit flow value of housing services and included in services prices rather than durables. Nonetheless, even in comparing non–housing durables and nondurables, Barsky, House, and Kimball find that durables prices respond more sharply to changes in monetary policy.[10]

3.3 Disaggregate vs. aggregate persistence estimates

Tables 2 and 3 report various summary statistics of persistence in disaggregate inflation

[8] Durables, nondurables, and services account for 14.8, 15.5, and 69.9 percent, respectively, of core PCE spending in 2001. Housing accounts for 17.9 percent of core PCE spending in 2001.

[9] While focusing on responsiveness to changes in monetary policy, Barsky, House, and Kimball (2003) also consider autocorrelations of (non–housing) durables and nondurables inflation. Consistent with the results in this paper, the autocorrelations suggest overall durables inflation to be more persistent than nondurables.

[10] As is often the case, the identification of monetary policy shocks can affect such results. Barsky, House, and Kimball's (2003) findings are based on simple Romer dates. Including the core durables, nondurables, and services price measures used in this paper in a structural (recursively identified) VAR like those in Bernanke and Blinder (1992), Christiano, Eichenbaum, and Evans (1996), and Clark (1999) fails to produce any significant sectoral differences in price responses.

rates. For varying levels of aggregation and quarterly data for 1959-02 and 1984-02, Table 2 reports simple means, medians, and 75th and 90th percentiles of the cross–section distribution of estimates. Table 2 also reports weighted means, medians, and 75th and 90th percentiles, based on weights corresponding to nominal expenditure shares of each component in 2001.[11] [12] Finally, Table 2 reports, for each level of aggregation, the percentage of components for which the sample persistence estimate is less than the estimate for core PCE inflation, and the percentage of components for which the upper band of the estimated 90% confidence band is less than 1. Table 3 provides some additional detail for the level 3 estimates, reporting summary statistics for monthly and quarterly persistence estimates and AR(1) estimates. For estimates of medians, 75th percentiles, and 90th percentiles, Table 3 also reports the associated median unbiased estimate and 90% confidence bands. For example, for the disaggregate series representing the weighted 75th percentile of the distribution of persistence estimates, the table reports its median unbiased estimate and confidence interval.

The SARC persistence estimates reported in Tables 2 and 3 indicate that average persistence in disaggregate inflation rates is consistently below aggregate persistence, although the size of the gap and its stability over time hinge on how "average" is measured.[13] For both data frequencies and time periods, the unweighted means and medians and weighted means and medians are considerably below the aggregate persistence estimates. Consider, for example, quarterly data for 1959-02, for which the aggregate PCE persistence estimate is .930 (Table 1). At aggregation level 3, the unweighted mean and median of the disaggregate persistence estimates are .644 and .712,

---

[11] In population, taking the disaggregate models as the DGPs, aggregate persistence will be a function of the persistence of each disaggregate component, the weights each component receives in the aggregate series, and the residual variances and covariances. The simple share–weighted averages reported in this paper are not intended to measure the persistence of aggregate inflation implied by the disaggregate models.

[12] The weighted 75th percentile, for example, is calculated by ordering the persistence estimates from smallest to largest and cumulating the weights of the ranked components. The weighted 75th percentile is then just the first persistence estimate for which the cumulative weight is greater than or equal to 75%.

[13] Consistent with the aggregate results that show persistence to be similar for goods and services, there don't appear to be any striking differences between goods and services in the disaggregate data.

respectively; the weighted mean and median are .739 and .799. In quarterly data for 1984-02, the unweighted mean and median estimates fall to .357 and .408, respectively, while the weighted mean and median estimates decline to .525 and .708. As these comparisons make clear, unweighted means are consistently a bit below medians, indicating the cross–section distribution of persistence estimates to be slightly skewed to the left. Moreover, the weighted means and medians consistently exceed the unweighted means and medians, highlighting the tendency for components with relatively large weights to be more persistent. This tendency is confirmed by Figure 1 and the modest positive correlations — of roughly .22 for level 3 data — shown near the bottom of Table 2.

Accordingly, the gap between disaggregate and aggregate persistence looks smaller when (i) "average" disaggregate persistence is measured by medians than by means and (ii) the estimates are weighted than unweighted. The choice of a measure of "average" also affects whether the gap is different across the 1959-02 and 1984-02 periods. The means and unweighted median measures of "average" suggest the gap is larger in the 1984-02 period, but the weighted median indicates the gap is little changed. A simple median, for example, shows disaggregate persistence declining from .712 over 1959-02 to .408 over 1984-02. Yet the weighted median of disaggregate persistence estimates is .799 for 1959-02 and .708 for 1984-02. In comparison, aggregate persistence is roughly .9 in both periods.

Regardless of how "average" is measured, average disaggregate persistence is below aggregate persistence because virtually all disaggregate series are lower in persistence than aggregate inflation. As shown in Table 2, in quarterly data for 1959-02, 97.2% of the level 3 series have persistence below that of aggregate inflation; the corresponding figure for 1984-02 is 95.4%. While not shown in the table, making the same calculations with median unbiased estimates in place of the sample estimates produces very similar results. Moreover, the persistence estimates are significantly below unity for about 60% of the level 3 series (accordingly, the upper bounds of the confidence bands around the unweighted medians reported in Table 2 are always less than 1).[14] While not reported in the interest of brevity, standard tests confirm that a unit root is rejected for many of the disaggregate series. With lag selected by the AIC, the DF-GLS test recommended by Elliott, Rothenberg, and Stock (1996) rejects the null of a unit root for 67.0% of the level 3 series over 1959-02 and 56.4% over 1984-02.[15]

Of course, the high persistence of aggregate inflation is reflected in the high persistence of a sizable set of disaggregate inflation series. Overall, with roughly 60% of the series having persistence estimates significantly less than unity (see the preceding paragraph), the confidence bands for the other 40 percent of the series are large enough to preclude ruling out a unit root. Table 3 shows that, in quarterly data for aggregation level 3, the unweighted 75th percentile of the cross–section distribution of disaggregate persistence estimates is .814 for 1959-02 and .733 for 1984-02, and in both cases the estimated confidence interval includes unity. The weighted 75th percentiles are even higher, .892 for 1959-02 and .823 for 1984-02, again with confidence bands that include unity.

To shed further light on the cross–section distribution, Table 4 reports the percentages of persistence estimates that fall within certain bands. In quarterly data for 1959-02, about 36% of the level 3 series have sample persistence estimates less than or equal to .6, while 35% have estimates between .6 and .8, and 29% have estimates greater than .8. But reflecting the positive correlation between weights and persistence estimates, the series with persistence less than or equal to .6 account for only 24% of core PCE spending, while components with estimates between .6 and .8 account for 26% and series with estimates greater than .8 account for 50%. Over 1984-02, the percentage of series with "low" persistence is even higher: for example, about 65% of the level 3

[14] In the 1959-02 estimates, the components for which the upper band is less than unity account for 42.5% of PCE spending; in the 1984-2002 estimates, the below–unity components account for 35.4% of spending.

[15] Perhaps not surprisingly, these results are somewhat sensitive to the test and lag selection approach. Selecting lag length with Ng and Perron's (2001) MAIC instead of the AIC modestly lowers the rejection rates. Using Perron and Ng's (1996) MZ_α test in lieu of the DF-GLS test also lowers the rejection rates. As indicated by Ng and Perron (2001), the DF-GLS test is generally more powerful than the MZ_α test, but the MZ_α test tends to have better size properties. The MAIC will be preferable to the AIC if the data have large negative MA roots. For aggregate core PCE inflation, though, none of the tests reject the unit root null.

series have sample persistence estimates less than or equal to .6. Yet again, the higher persistence series account for the majority of PCE spending: in this case, the series with persistence greater than .6 account for 56% of core PCE spending.

3.4 Comparison with previous results on disaggregate persistence

Despite some differences in data coverage and sample periods, this paper's results for monthly data and an AR(1) measure of persistence are very to the results of Bils and Klenow (2002). Bils and Klenow report that, in monthly data, the persistence of an aggregate inflation series constructed from PCE series corresponding to non–housing CPI components is .63 over 1959-2000 and .20 over 1995-2000. The weighted mean of their disaggregate persistence estimates series is .26 over 1959-2000 and -.05 over 1995-2000. These estimates suggest a large gap in aggregate and disaggregate persistence for 1959-2000 and a modest gap over the more recent period. For the data and sample periods used in this paper — the data differ from Bils and Klenow's in that all of non–food and non–energy PCE is used — monthly AR(1) estimates are similar to Bils and Klenow's.[16] As shown in Table 3, in 1959-02 data, AR(1) persistence is .673 in aggregate data and .249 (weighted mean) in the disaggregate data. Consistent with the Bils and Klenow finding that persistence is lower and the aggregate–disaggregate gap smaller in more recent data, AR(1) estimates for monthly 1984-02 data yield aggregate persistence of .126 and average (weighted mean) disaggregate persistence of .138 (Table 3).

Overall persistence appears lower in the Bils and Klenow (2002) results than in the results discussed in the preceding section because the AR(1) estimates are systematically lower than the SARC estimates, more so for monthly than quarterly data and more so for 1984-02 than 1959-02. The difference between AR(1) and SARC estimates also appears to explain why the aggregate–disaggregate persistence gap falls over time in Bils and Klenow's results but rises or remains little

[16] Excluding non-CPI components from the disaggregate data (using the CPI-PCE mapping of Fixler and Jaditz (1997)) produces results very similar to those reported for all of core PCE.

changed in the results discussed above. For example, Tables 1 and 3 show that, in quarterly 1984-02 data, the AR(1) estimates yield aggregate and (weighted mean) disaggregate persistence of .788 and .391, respectively, compared to SARC estimates of .907 and .525, respectively. In monthly 1984-02 data, the AR(1) estimates indicate virtually no persistence in aggregate or disaggregate inflation, while the SARC estimates show aggregate persistence to be near unity (the point estimate is .834 and the median unbiased estimate is .987) and average disaggregate persistence to be .328. The dramatic falloff in the monthly aggregate AR(1) estimate between the 1959-02 and 1984-02 periods — from .673 to .126 — causes the aggregate–disaggregate persistence gap to fall markedly, contrary to what the SARC estimates show.

Nonetheless, any differences between this paper's results and Bils and Klenow's (2002) do nothing to alter the fundamental findings of Bils and Klenow. Essentially, their point is that, given the behavior of marginal cost and the measured frequencies of price change, the simple inflation model implied by Calvo price staggering dramatically fails in characterizing the behavior of disaggregate inflation rates at the monthly frequency. For example, contrary to the implications of the Calvo model, the data indicate the measured frequency of price change has little relationship (cross–sectionally) to persistence. There is nothing in this paper to contradict these essential points of Bils and Klenow.

4. Rolling Persistence Estimates

To evaluate whether persistence has changed over time, authors such as Stock (2001) and Pivetta and Reis (2003) rely on rolling estimates of persistence, using a rolling window of data to estimate persistence at each point in time. Following this precedent, this section compares rolling persistence estimates for aggregate and disaggregate PCE inflation. For each inflation series considered, time series of persistence (along with median unbiased estimates and confidence intervals) for 1977:Q3 through 2002:Q3 are estimated using a rolling window of 17 years of data. For each series, the AR model's lag length is fixed over time at the order determined with the AIC and the full sample of 1959-02 data.[17]

Consistent with the results of Stock (2001) and Pivetta and Reis (2003), the rolling estimates in Figure 2 indicate that aggregate persistence has drifted down a bit but remained high — high enough that a unit root still cannot be ruled out. For core PCE inflation, the sample persistence estimate has fallen from .929 in 1977:Q3 to .783 in 2002:Q3, and the median unbiased estimate has declined from 1.006 to .826. But the upper bound of the 90% confidence band has consistently exceeded 1, except for a handful of quarters since 1997. For durables, nondurables, and services inflation, the rolling estimates (sample and median unbiased) indicate persistence is higher now than it was in the late 1970s in the case of durables and nondurables and slightly lower now than in the late 1970s in the case of services. In 2002:Q3, the sample and median unbiased estimates of persistence are higher for goods than services — a finding again contrary to the view that services prices are more rigid than goods prices. In all cases, though, the confidence bands include unity.

In the disaggregate data, rolling persistence estimates indicate many components have experienced a more substantial decline.[18] But because the larger declines have tended to occur in components receiving relatively smaller expenditure weights, on a share–weighted basis disaggregate persistence has declined only slightly. In fact, on a weighted basis, movements in disaggregate persistence have paralleled those in aggregate persistence. As shown in the upper panel of Figure 3, the unweighted mean and median sample estimates of persistence for aggregation level 3 have declined sharply, relative to aggregate persistence, with most of the decline occurring in the late 1990s. For example, the median of the cross–section distribution declined from a high of about .7 in 1977:Q3 and 1995:Q3 to .410 in 2002:Q3. The rolling confidence band estimates for the unweighted median persistence estimate shown in the upper left panel of Figure 4 suggest that the decline is

[17] As noted by Levin and Piger (2003), using shorter windows tends to produce a bigger falloff in persistence in recent years.

[18] In total, persistence declined from 1977:Q3 to 2002:Q3 for 78% of the level 3 series.

statistically significant. After remaining above 1 for much of the period through the mid–1990s, the upper band now stands at about .6. But as shown in the lower panel of Figure 3, on an expenditure share–weighted basis, the decline in persistence has been much more modest, and probably not statistically significant. For example, the weighted median of the cross–section distribution has drifted down from .731 in 1977:Q3 to .544 in 2002:Q3. The upper right panel of Figure 4 shows that there has been no downward drift in the upper band of the 90% confidence interval around the persistence estimate representing the weighted median of the cross–section distribution.

5. Evidence on Structural Breaks

The baseline persistence estimates for 1959-02 and 1984-02 and the rolling persistence estimates for 1977-2002 presented above assume that, within a given regression sample over which persistence is estimated, mean inflation is constant (although the rolling estimates allow the mean to change from sample to sample). Yet, as many authors have argued, forces such as changes in the behavior of monetary policy may result in significant changes in mean inflation over time.[19] If ignored, variation in mean inflation could cause persistence estimates to overstate true persistence.[20] Accordingly, following Levin and Piger (2003), this section reports the results of tests for discrete breaks in the parameters of AR models fit to aggregate and disaggregate inflation rates.

This analysis uses data for just 1984-02, for two broad reasons. First, as Levin and Piger note, researchers seem to agree that persistence was very high in data from 1959 through the early 1980s; the debate is over persistence in data spanning the last 20 or so years. Second, the shorter sample period seems to reasonably limit the potential number of monetary policy–related breaks to at most 1, which permits the use of the Andrews (1993) test for a single break and simple simulation methods for inference, which is important in light of evidence presented below that

[19] Kozicki and Tinsley (2001), for example, present evidence of significant movement in long–term inflation expectations over time, attributable to markets learning about the changing long–run inflation goal of monetary policy.

[20] Although the inflation persistence literature focuses exclusively on the potential for unmodeled mean shifts to exaggerate persistence, Nunes, Kuan, and Newbold (1995, 1996) show that unit root persistence can create spurious findings of mean shifts.

using standard asymptotic critical values would yield significantly oversized tests. Over a longer period, there could be multiple breaks, so the tests of Bai and Perron (1998, 2003) would have to be used to try to identify the breakpoints. Testing for multiple breaks in the large number of series under consideration would be somewhat unwieldy. Testing for multiple breaks in the systems of disaggregate inflation equations would also be a complicated endeavor. Finally, with multiple breaks it would be much more difficult to use simulation methods to generate accurate critical values.

While this paper focuses on discrete shifts in reduced form inflation models, others such as Cogley and Sargent (2001, 2003), Erceg and Levin (2003), and Kozicki and Tinsley (2003) have treated any variation in the model coefficients as stochastic. For example, the gradual movement in long–term inflation expectations as measured by the Survey of Professional Forecasters could be seen as evidence that variation in mean inflation is better represented by a stochastic, time–varying mean. In the end, the data used in this study seem to favor the discrete representation, although the data are such that probably neither representation can be ruled out. As detailed below, this paper finds evidence of discrete breaks in mean inflation, while unreported Nyblom (1989) tests applied to the aggregate rates (total, durables, etc.) systematically fail to yield any evidence of TVP in the intercept or AR coefficients.[21] That said, Monte Carlo experiments in Cogley and Sargent (2003) indicate that Nyblom tests (and tests for discrete breaks) have poor power in their DGP that actually features TVP.

After explaining this paper's break test approach, this section presents the results of break tests applied to AR models for aggregate inflation rates and disaggregate inflation rates and, at tractable levels of aggregation, break tests applied to systems of AR equations for disaggregate inflation rates.

[21] As implemented here, the tests allow heteroskedasticity as in Hansen (1992) and use Hansen's asymptotic critical values.

5.1 Break test approach

In quarterly data for 1984-02, this paper uses Andrews' (1993) sup Wald statistic to test for breaks in the coefficients and residual variances of AR models for inflation. Following Stock and Watson (2002b) and Levin and Piger (2003), among others, heteroskedasticity–robust variances are used in forming the Wald statistics for shifts in the regression coefficients. The tests for shifts in the residual variance use simple least squares variances in forming the test statistics, as appropriate under the null. For each equation, break tests are computed for the intercept, the set of AR coefficients, and the full set of coefficients, as well as the residual variance. In addition, small systems of AR models for inflation — one three variable system for durables, nondurables, and services and an 11–variable system of aggregation level 1 models — are tested for a single shift in the set of intercepts, using the basic sup Wald methodology described by Bai, Lumsdaine, and Stock (1998). These system tests are constructed two ways: (i) estimating the model with SUR techniques, imposing homoskedasticity, as in Bai, Lumsdaine, and Stock; and (ii) estimating the model equations by OLS, allowing heteroskedasticity in computing the system variance matrix. Finally, small systems of squared residuals from the AR models are tested for a single break in variance, using SUR estimates to compute sup Wald tests for the systems. In all of these tests, the sample trim is set to $\pi_0 = .21$, to ensure that the pre-break and post-break periods include at least 15 observations, as recommended by Bai and Perron (2003).

As shown by Diebold and Chen (1996), comparing the Andrews (1993) break test against Andrews' asymptotic critical values can lead to significantly oversized tests. One asymptotic source of such size problems, noted by Hansen (2000), is that inference based on Andrews' critical values can be misleading if, contrary to assumptions underlying Andrews' asymptotic results, the marginal distributions of the regressors are not constant. In the types of models and breaks under consideration here, of course, the marginal distributions of the regressors may not be constant. Another, finite–sample, problem is that break tests applied to the intercepts of AR models in which true persistence is high may be oversized because of the difficulty of distinguishing between high persistence and a mean shift.

Monte Carlo experiments using DGPs drawn from 1984-02 estimates of AR models for select inflation variables confirm the Diebold and Chen (1996) finding that a simple bootstrap yields the most accurately sized tests. In particular, size distortions are smaller when the Andrews tests for shifts in the intercept and AR coefficients are compared against critical values generated with a parametric bootstrap than against Andrews' (1993) asymptotic critical values or against critical values generated with Hansen's (2000) fixed regressor bootstrap. Perhaps not surprisingly, though, the asymptotics–based test for a shift in residual variance is much closer to being correctly sized. In these size experiments, DGPs were estimated by fitting AR models to inflation data for total PCE, durables, nondurables, services, and recreation. The recreation series, a component in the level 1 aggregation, is included as an example because its persistence is much lower than that of the other series — roughly .5 rather than .9. The lag lengths were determined by the AIC using a regression sample of 1985:Q2 through 2003:Q3 (70 observations). For each inflation series, the estimated, stable AR model was used as the DGP in a Monte Carlo experiment, with 1000 simulations. The fixed regressor and simple bootstraps use 999 draws.[22]

As shown in Table 5, using Andrews' (1993) asymptotic critical values leads to dramatic size distortions in tests for shifts in the regression coefficients when applied in the small samples considered in this paper. In the DGP based on core PCE inflation, for example, the test for a break in the intercept has an empirical size of 56.2% (the nominal size is 10%), while the test for a break in the set of AR coefficients has an empirical size of 54.9%. Results are similar for the durables, nondurables, and services experiments, presumably reflecting the similar persistence of the series. The size distortions stemming from the use of asymptotic critical values are more modest

[22] The simple parametric bootstrap generates data using sample estimates of the AR model imposing the null of stability and resampled residuals of the model. The initial observations — those preceding the sample of data used to estimate the models — necessitated by the lag structure of the estimated model are selected by sampling from the actual data, following Stine (1987).

in the case of the recreation DGP, in which persistence is much lower: the sizes of the intercept and AR coefficients tests are 24.4% and 31.5%, respectively. Unreported results confirm that not only the high degree of persistence but also the small sample dimension contribute to the large size distortions. Simulations of test performance in artificial data for 1959-02 yield significantly smaller, though still substantial, size distortions.[23] For example, with the longer sample of data, the test for a single break in the intercept of the core PCE model has an empirical size of 32.8%.

Table 5 also shows that using critical values generated with Hansen's fixed regressor bootstrap reduces the size distortions somewhat, but the distortions remain very large in tests for a break in the intercept. In the core PCE DGP, for example, the fixed regressor bootstrap yields empirical sizes of 44.8% for the intercept test and 21.8% for the AR coefficients test. The empirical size of the intercept test is similar in the durables, nondurables, and services DGPs, but the size of the AR coefficients test is considerably better in the durables and nondurables cases. Again, the lower persistence of the recreation DGP reduces but does not eliminate the size distortion of the intercept test, pulling size down to 20.4%. The version of the fixed regressor bootstrap that allows heteroskedasticity, the so–called heteroskedastic bootstrap, yields qualitatively similar results.

Finally, the results in Table 5 show clearly that using critical values based on a simple parametric bootstrap yields the most accurate tests. The empirical size of the test for a break in the intercept is 13.8% in the PCE experiment, 15.0% for durables, 16.5% for nondurables, 13.7% for services, and 11.9% for recreation (again, the nominal size is 10%). The empirical size of the test for a break in the set of AR coefficients ranges from 10.0 to 11.7% across the five experiments. In light of these Monte Carlo findings, the break test results presented below use p–values based on a simple parametric bootstrap.[24]

---

[23] These longer–sample experiments used DGPs based on model estimates for 1959-02 rather than 1984-02. As shown above, persistence is similar across the samples, but the DGPs do differ in other aspects.

[24] The residual variance results also use bootstrap p–values since the results in Table 5 indicate the bootstrap is slightly more accurate than the asymptotic critical values.

5.2 Break test results

Collectively, break tests applied to aggregate and disaggregate inflation data for 1984-2002 yield consistent evidence of a mean shift in the early 1990s (while not reported, in virtually all cases the shift is a decline in mean inflation), but only weak evidence of shifts in the AR coefficients.[25] Table 6 reports simple bootstrap p–values for break tests applied to models for the four aggregate inflation variables and the 11 level 1 disaggregate series, along with least–squares estimates of the break date in those cases in which the break is significant at the 10% level. As shown in the upper panel of the table, the null of stability in the intercept is rejected for all of the aggregate inflation series except nondurables, with estimated break dates of 1992:Q1 for core PCE, 1995:Q1 for durables, and 1992:Q1 for services. In contrast, for these same aggregate inflation measures, none of the tests for a shift in the set of AR coefficients have significant p–values. The level 1 disaggregate series yield a similar pattern of results. According to the intercept tests, all three of the durables series and three of the six services series have a significant break (at 10% confidence), while neither of the two nondurables series have a break. For only three of the 11 series is stability in the AR coefficients rejected.

The system break test p–values in Table 7 yield strong evidence of a common shift in the intercepts of the AR equations. The sup Wald test for a common (in date) shift in the intercepts of the three equations for durables, nondurables, and services has a bootstrap p–value of .009 when the system is estimated by SUR and .049 when the system is estimated by OLS and heteroskedasticity is allowed. In both cases, the least squares date estimate is early 1992.[26] The sup Wald test for a common shift in the intercepts of the 11–equation system for the *level 1* disaggregate series is also highly significant, with a break date of 1992:Q2. In light of the analysis of Bai, Lumsdaine,

[25] Testing for AR coefficient breaks in models that impose an intercept break in 1993:Q1 yields qualitatively very similar results on the stability of the AR coefficients.

[26] Following Bai, Lumsdaine, and Stock (1998), the SUR estimate is simply the date at which the system likelihood function is maximized. The OLS estimate is simply the date at which the equally–weighted sum of squared residuals across the system of equations is minimized. If the residuals were uncorrelated across equations, this estimate would of course be the same as the SUR estimate.

and Stock (1998), these system results should be viewed as powerful evidence of an aggregate shift. Combining information across variables provides a way of substantially boosting the power of break tests relative to the alternative of single–equation tests. That said, it can't be claimed that the system tests used here are more powerful than a univariate test applied to core PCE inflation, because a system test based on disaggregate variables isn't necessarily bring more information to bear on the hypothesis of interest.[27] Rather, system tests are simply an alternative to aggregating the data into a single series and then performing a univariate test on the aggregate series. A seemingly reasonable interpretation — one supported by some simple Monte Carlo exercises — is that the system tests should have power comparable to those based on single equation, aggregate tests.

Table 8's simple summary of the break test results for more disaggregate data provides further evidence of an intercept shift in the early 1990s, with again comparatively weaker evidence of a shift in AR coefficients. For aggregation levels 2 and 3, the table simply provides a count of the series for which bootstrap p–values indicate a significant break (at 10% confidence) in the intercept and set of AR coefficients, with the count broken out by durables, nondurables, and services. At both aggregation levels, a significant intercept break occurs in roughly half of the series. At aggregation level 3, for example, 52 of the 109 series experience a significant break in the intercept. For those series with significant intercept breaks, the average of the least–squares date estimates, at both levels of aggregation, is 1993:Q3. Significant breaks in AR coefficients are much less common. At level 2, AR breaks occur in just 8 of the 46 series, while at level 3, AR breaks occur in 30 of the 109 series.

As to the secondary issue of the stability of the volatility of inflation, the results in Tables 6-8 provide modest evidence of a significant shift (a decline) in the residual variances of the AR

[27] In Bai, Lumsdaine, and Stock (1998), using a system of variables increases power in the sense that the system test is more powerful than univariate tests applied to each variable. The issue here is whether the system test is more powerful than a test based on an aggregate of the variables under consideration.

equations for aggregate and disaggregate inflation. As shown in Table 6, the null of stability in the residual variance is rejected for only one of the three aggregate series (nondurables) and only four of the 11 level 1 series (two nondurables components and two services components). More disaggregate data yield similar results, with rejections of stability occurring for about 40% of the level 2 and level 3 series (Table 8). The system tests nearly reject stability in the set of aggregate residual variance equations and clearly indicate instability in the system of level 1 disaggregate series.

6. Persistence Estimates Allowing Break in Intercept

In light of the above evidence of a break in mean inflation due to a shift in the intercept of AR models in the early 1990s, this section presents estimates of inflation persistence that condition on the identified break. In particular, this section reports estimates of persistence for 1984-02 data that allow a single shift in the intercept of each series' AR representation, occurring in 1993:Q1. This date of 1993:Q1 is roughly an average of the estimates of 1992:Q1 or 1992:Q2 produced by the system break tests (Table 7) and the estimate of 1993:Q3 obtained by averaging the univariate break date estimates from the levels 2 and 3 disaggregate data (Table 8). Because the identified break represents an *aggregate* shift, the 1993:Q1 break is imposed for all inflation series, regardless of whether the break test for that variable indicates a significant break. Imposing the 1993:Q1 break on only those series which have a statistically significant break yields very similar results, not reported in the interest of brevity.[28] In the estimates below, the AR models and resulting persistence estimates use the same lags as the estimates in Tables 1-3: the lag length is selected using the AIC assuming a constant mean. This approach tends to yield longer lag lengths than does the alternative of taking the break into account when determining the optimal lag order. In a large sample, using too many lags shouldn't affect persistence estimates, but in the data used here,

[28] For example, this alternative approach yields weighted mean and median persistence estimates of .405 and .522, respectively, compared to the reported figures of .353 and .519.

some aspects of the results (noted below) are sensitive to the lag selection approach.

As shown in Table 9, allowing the shift in average inflation greatly reduces estimated persistence, but widens the confidence bands enough that roots of unity often cannot be ruled out. For aggregate (core) PCE inflation, allowing the intercept break lowers estimated persistence from .907 (Table 1) to .402, with a 90% confidence interval of (.206,1.052). Allowing the break also significantly lowers the estimated persistence of inflation in nondurables and services, to .367 and .137, respectively (compared to Table 1's estimates of .878 and .855), but has little effect on the estimated persistence of durables inflation. In the case of aggregate PCE inflation, the estimated degree of persistence is highly sensitive to the lag length. If the intercept break is taken into account in lag selection, the AIC–optimizing lag order is just 1 (rather than 3), and the persistence estimate drops to .191, with a confidence interval of (.022, .449). Ultimately, the sample evidence based on aggregate data indicates persistence is well below unity, but statistically, the degree of uncertainty is so large as to preclude ruling out a unit root.

Taking the 1993:Q1 mean shift into account also lowers the estimated persistence of disaggregate inflation, but not as sharply as in the aggregate case. As shown in Table 9, imposing the break lowers the weighted mean and median of the disaggregate persistence estimates to .353 and .519, respectively (compared to .525 and .708 in Table 2's constant–mean estimates). The percentage of components for which persistence of unity can be rejected rises from the constant–mean estimate of 59.6% (Table 2) to 71.6%. These estimates, though, indicate that even with the break, persistence of unity cannot be rejected for nearly 30 percent of the disaggregate inflation rates. Unreported results show that using the alternative approach of taking the mean break into account when selecting the lag length lowers the average persistence estimates much as it does in some of the aggregate estimates. For example, this alternative lag selection approach yields weighted mean and median persistence estimates of .271 and .333, respectively.

Overall, the estimates suggest that, once the significant 1993:Q1 break in mean inflation is taken into account, average disaggregate persistence is actually as great as aggregate persistence. Persistence in disaggregate data is .353 by the weighted mean and .519 by the weighted median, compared to persistence of .402 for aggregate PCE inflation. This finding could be interpreted as indicating that the idiosyncratic components of disaggregate inflation rates driven by differences across sectors in product demand, technology, etc., are at least as persistent as the aggregate component driven by monetary policy and, in the short run, other aggregate forces.[29]

7. A Factor Model Interpretation

To provide a more formal interpretation of the relationship between aggregate and disaggregate persistence, this section examines the persistence of common and idiosyncratic components of inflation estimated with the factor model methods of Stock and Watson (2002a).[30] Under the factor model specification, inflation in each sector i is a function of a common aggregate component driven by such underlying forces as monetary policy and an idiosyncratic component driven by differences across sectors in product demand, technology, etc. The persistence of inflation in sector i reflects the persistence of both the aggregate component and the idiosyncratic component. Persistence in disaggregate inflation rates could be quite low while aggregate persistence is high if the common component is persistent but the idiosyncratic components, which in most sectors account for most of the variance in inflation, are not. The same basic factor model framework has been used in such studies as Bryan and Cecchetti (1993) and Dow (1994) to extract a measure of underlying, or core, inflation.

More specifically, this section reports SARC estimates of the persistence of a common factor and idiosyncratic components extracted from disaggregate inflation data. For 1984-2002 data, a

[29] Econometrically, it is not possible to simply decompose the persistence of a given sector's inflation rate into aggregate and idiosyncratic components. Measured persistence is a complicated function of the persistence of each component, depending on variance shares, etc.

[30] In light of the evidence in Kapetanios and Marcellino (2003) that state space–based estimation methods are sometimes superior to the principal component–based approach of Stock and Watson (2002a), the EM algorithm of Quah and Sargent (1993) was also used to estimate dynamic factor models for disaggregate inflation data. In general, this EM–based approach yielded qualitatively similar results, although the specific results were somewhat sensitive to the lag orders allowed for the common and idiosyncratic components.

common factor is estimated as the first principal component from standardized inflation rates. The idiosyncratic component for each sector i is estimated as the residual from a regression of standardized inflation in sector i on the common factor. In light of the evidence of an early 1990s shift in mean inflation, factor model estimates are constructed both treating mean inflation as constant and allowing a mean shift in 1993:Q1 (through the standardization of the raw data). For the resulting estimates of the common factor and idiosyncratic components, *persistence* is measured by the sum of AR coefficients, with the lag length for each model selected using the AIC. In the case of the common factor, 90% confidence intervals for the persistence estimates are computed with Hansen's (1999) grid bootstrap. Shintani (2003) establishes the validity of such asymptotic procedures applied to estimates of the common factor.

Given Boivin and Ng's (2003) findings on the potential for the size of the cross–section to affect the results, factor estimates are generated from two different aggregation levels — levels 2 and 3. Admittedly, both data sets have some of the features that Boivin and Ng show to be potentially problematic for accurate factor estimation: dispersion across sectors in the explanatory power of the common factor for inflation (as measured by the $\bar{R}^2$ of a regression of inflation in sector i on the common factor); and some cross–sector correlation of idiosyncratic components. The problems are likely greater for aggregation level 3 than 2, as cross–sector correlation is higher for aggregation level 3. Nonetheless, estimation of the common factor using Boivin and Ng's SWb version of weighted principal components yields results very similar to those reported (based on unweighted principal components).

As shown in Table 10, estimates of factor models that assume mean inflation to be unchanged over the 1984-02 sample yield a common component with near–unity persistence and idiosyncratic components with much lower persistence. Using either aggregation level 2 or 3, the estimated persistence of the common factor is about .96. For aggregation level 2, the weighted mean and median of the persistence estimates for the idiosyncratic components of disaggregate inflation are

.294 and .372, respectively. Using aggregation level 3 data yields a weighted mean and median of .224 and .298 for the idiosyncratic components.

Taking a 1993:Q1 shift in mean inflation into account causes the persistence of the aggregate or common component to fall, but has little effect on average disaggregate persistence. As reported in Table 10, with a mean shift allowed in the data, the estimated persistence of the common factor is .597 in the level 2 data and .808 in the level 3 data. The average of persistence in the idiosyncratic components of aggregation level 2 data is .267 according to the weighted mean and .320 according to the weighted median. The corresponding estimates for aggregation level 3 are .160 (weighted mean) and .286 (weighted median).

In light of the aforementioned potential concerns about the accuracy of the factor model estimates, it bears noting that a simple alternative based on relative inflation rates yields qualitatively similar results on the persistence of the idiosyncratic components of inflation. In this approach, the idiosyncratic component of inflation for each sector i is simply the inflation rate for good i less the aggregate rate, measured as core PCE inflation. For each such relative inflation rate, persistence is estimated just as in the raw data, by fitting an AR model using an AIC-determined lag length, Hansen's grid bootstrap to estimate the confidence band, etc. In 1984-02 data for aggregation level 3, the weighted mean and median of these relative inflation persistence estimates are .363 and .465, respectively (compared to .525 and .708 in the raw data results presented in Table 2).[31]

8. Conclusions

In light of the strong interest in inflation persistence and growing interest in the behavior of disaggregate prices shown in recent research, this paper studies the persistence of disaggregate consumer price inflation and the disaggregate evidence of any change in inflation persistence. More specifically, this paper uses disaggregate inflation series spanning all of PCE to examine (i) the

[31] In estimates based on 1959-02 data, the weighted mean and median of the relative inflation persistence estimates are .498 and .502, respectively, compared to .739 and .799 in the raw inflation rates (Table 2). For this sample, persistence is significantly below unity for 90.8% of the relative rates, up from 60.6% in the raw inflation rates.

persistence of disaggregate inflation relative to aggregate inflation; (ii) the distribution of persistence across consumption sectors; and (iii) whether inflation persistence has changed.

Assuming mean inflation to be unchanged within samples, the baseline results indicate that the average persistence of disaggregate inflation is consistently below aggregate persistence, although the size of the gap hinges on the definition of "average." Nearly all disaggregate series are lower in persistence than aggregate inflation, but a sizable proportion of disaggregate series are highly persistent, and the more persistent series tend to represent larger shares of consumer spending. The results also indicate that, at the aggregate level, inflation persistence is very similar across the durable goods, nondurable goods, and services sectors. Rolling estimates from 1977 to 2002 show that persistence has drifted down slightly, but aggregate persistence remains high enough that a unit root cannot be ruled out.

The estimated persistence of both aggregate and disaggregate inflation over the 1984-02 period is significantly lower when an early 1990s shift in mean inflation is taken into account. Break tests applied to aggregate and disaggregate data yield consistent evidence of a mean shift in the early 1990s. In particular, tests for a common shift in the intercepts of AR equations applied to systems of equations yield strong evidence of a break in the early 1990s. In aggregate data, allowing an intercept shift in 1993:Q1 greatly reduces estimated persistence, but widens the confidence bands enough that roots of unity cannot be ruled out. Taking the intercept break into account also lowers the estimated persistence of disaggregate inflation, but not as sharply as in the aggregate case. With the mean break accounted for, average disaggregate persistence is actually as great as aggregate inflation persistence. Allowing a 1993:Q1 shift in mean inflation, factor model estimates yield a common aggregate component with persistence between .6 and .8 and idiosyncratic components with considerably less persistence.

These results, in conjunction with those of other studies such as Bils and Klenow (2002), Erceg and Levin (2002), and Barsky, House, and Kimball (2003), suggest some important challenges for future research. One, perhaps insurmountable, is to resolve the debate about the overall persistence of inflation and whether persistence has changed. A second challenge is to better understand why measured inflation persistence does not seem to be correlated (across sectors) with the measured frequency of price change and why a simple inflation model implied by Calvo staggering does not seem to fit the disaggregate data. Yet another is to develop general equilibrium business cycle models that accurately capture the empirical behavior of the prices and quantities of durables, nondurables, and services.

REFERENCES

Andrews, Donald W.K., 1993, "Tests for Parameter Instability and Structural Change with Unknown Change Point," *Econometrica* 61 (July), pp. 821-56.

Andrews, Donald W.K., and W.K. Chen, 1994, "Approximately Median-Unbiased Estimation of Autoregressive Models," *Journal of Business and Economic Statistics* 12 (April), pp. 187-204.

Bai, Jushan, Robin L. Lumsdaine, and James H. Stock, 1998, "Testing for and Dating Common Breaks in Multivariate Time Series," *Review of Economic Studies* 65 (July), pp. 395-432.

Bai, Jushan, and Pierre Perron, 1998, "Estimating and Testing Linear Models with Multiple Structural Changes," *Econometrica* 66 (Jan.), pp. 47-78.

Bai, Jushan, and Pierre Perron, 2003, "Computation and Analysis of Multiple Structural Change Models," *Journal of Applied Econometrics* 18 (Jan.-Feb.), pp. 1-22.

Barsky, Robert B., 1987, "The Fisher Hypothesis and the Forecastability and Persistence of Inflation," *Journal of Monetary Economics* 19 (Jan.), pp. 3-24.

Barsky, Robert, Christopher L. House, and Miles Kimball, 2003, "Do Flexible Goods Prices Undermine Sticky Price Models?" NBER Working Paper #9832, July.

Benati, Luca, 2002, "Investigating Inflation Persistence Across Monetary Regimes," manuscript, Bank of England.

Benati, Luca, and George Kapetanios, 2002, "Structural Breaks in Inflation Dynamics," manuscript, Bank of England, December.

Bernanke, Ben S., and Alan S. Blinder, 1992, "The Federal Funds Rate and the Channels of Monetary Transmission," *American Economic Review* 82 (Sept.), pp. 901-21.

Bils, Mark, and Peter J. Klenow, 2002, "Some Evidence on the Importance of Sticky Prices," NBER Working Paper #9069, July.

Boivin, Jean, and Serena Ng, 2003, "Are More Data Always Better for Factor Analysis?" *Journal of Econometrics*, forthcoming.

Brainard, William, and G. Perry, 2000, "Making Policy in a Changing World," in G. Perry and J. Tobin, eds., *Economic Events, Ideas, and Policies: The 1960s and After*, Washington, D.C.: Brookings Institution, pp. 43-69.

Bryan, Michael F., and Stephen G. Cecchetti, 1993, "The Consumer Price Index as a Measure of Inflation," Federal Reserve Bank of Cleveland, *Economic Review*, Fourth Quarter.

Calvo, Guillermo A., 1983, "Staggered Pricing in a Utility Maximizing Framework," *Journal of Monetary Economics* 12 (Sept.), pp. 383-98.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans, 1996, "The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds," *Review of Economics and Statistics* 78 (Feb.), pp. 16-34.

Clark, Todd E., 1999, "The Responses of Prices at Different Stages of Production to Monetary Policy Shocks," *Review of Economics and Statistics* 81 (Aug.), pp. 420-33.

Cochrane, John H., 1988, "How Big Is the Random Walk in GNP?" *Journal of Political Economy* 96 (Oct.), pp. 893-920.

Cogley, Timothy, and Thomas J. Sargent, 2001, "Evolving Post-World War II U.S. Inflation Dynamics," *NBER Macroeconomics Annual* 16, pp. 331-73.

Cogley, Timothy, and Thomas J. Sargent, 2003, "Drifts and Volatilities: Monetary Policies and Outcomes in the Post-WWII U.S.," manuscript, Arizona State University, April.

Diebold, Francis X., and Celia Chen, 1996, "Testing Structural Stability with Endogenous Breakpoint: A Size Comparison of Analytic and Bootstrap Procedures," *Journal of Econometrics* 70 (Jan.), pp. 221-41.

Dow, James P., Jr., 1994, "Measuring Inflation Using Multiple Price Indexes," manuscript, University of California-Riverside.

Elliott, Graham, Thomas J. Rothenberg, and James H. Stock, 1996, "Efficient Tests for an Autoregressive Unit Root," *Econometrica* 64 (July), pp. 813-36.

Erceg, Christopher, and Andrew T. Levin, 2002, "Optimal Monetary Policy with Durable and Non-Durable Goods," International Finance Discussion Paper #748, Board of Governors of the Federal Reserve System, July.

Erceg, Christopher, and Andrew T. Levin, 2003, "Imperfect Credibility and Inflation Persistence," *Journal of Monetary Economics* 50 (May), pp. 889-914.

Fixler, Dennis, and Ted Jaditz, 1997, "An Examination of the Difference Between the CPI and the PCE Deflator," manuscript, Bureau of Labor Statistics, January.

Hansen, Bruce E., 1992, "Testing for Parameter Instability in Linear Models," *Journal of Policy Modeling* 14 (Aug.), pp. 517-33.

Hansen, Bruce E., 1997, "Approximate Asymptotic P Values for Structural-Change Models," *Journal of Business and Economic Statistics* 15 (Jan.), pp. 60-67.

Hansen, Bruce E., 1999, "The Grid Bootstrap and the Autoregressive Model," *Review of Economics and Statistics* 81 (Nov.), pp. 594-607.

Hansen, Bruce E., 2000, "Testing for Structural Change in Conditional Models," *Journal of Econometrics* 97 (July), pp. 93-116.

Kapetanios, George, and Massimiliano Marcellino, 2003, "A Comparison of Estimation Methods for Dynamic Factor Models of Large Dimensions," Working Paper No. 489, Queen Mary University of London, April.

Kozicki, Sharon, and P.A. Tinsley, 2001, "Shifting Endpoints in the Term Structure of Interest Rates," *Journal of Monetary Economics* 15 (June), pp. 613-52.

Kozicki, Sharon, and Peter A. Tinsley, 2003, "Alternative Sources of the Lag Dynamics of Inflation," in *Proceedings of a Conference on Price Adjustment and Monetary Policy*, Bank of Canada, forthcoming.

Levin, Andrew T., and Jeremy Piger, 2003, "Is Inflation Persistence Intrinsic in Industrial Economies?" Working Paper 2002-023B, Federal Reserve Bank of St. Louis, February.

Ng, Serena, and Pierre Perron, 2001, "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power," *Econometrica* 69 (Nov.), pp. 1519-54.

Nunes, Luis C., Chung–Ming Kuan, and Paul Newbold, 1995, "Spurious Break," *Econometric Theory* 11 (Oct.), pp. 736-49.

Nunes, Luis C., Chung–Ming Kuan, and Paul Newbold, 1996, "Spurious Number of Breaks," *Economics Letters* 50 (Feb.), pp. 175-78.

Nyblom, Jukka, 1989, "Testing for the Constancy of Parameters Over Time," *Journal of the American Statistical Association* 84 (March), pp. 223-30.

Perron, Pierre, and Serena Ng, 1996, "Useful Modifications to Unit Root Tests with Dependent Errors and their Local Asymptotic Properties," *Review of Economic Studies* 63 (July), pp. 435-65.

Pivetta, Frederic, and Ricardo Reis, 2003, "The Persistence of Inflation in the United States," manuscript, Harvard University, April.

Quah, Danny, and Thomas J. Sargent, 1993, "A Dynamic Index Model for Large Cross Sections," in J. Stock and M. Watson, eds., *Business Cycles, Indicators, and Forecasting*, Chicago: University of Chicago Press, pp. 285-310.

Shintani, Mototsugu, 2003, "Finite Sample Performance of Two–Step Estimators for Dynamic Factor Models," manuscript, Vanderbilt University.

Sims, Christopher A., 2001, "Comment on 'Evolving Post-World War II U.S. Inflation Dynamics," *NBER Macroeconomics Annual* 16, pp. 373-79.

Stine, Robert A., 1987, "Estimating Properties of Autoregressive Forecasts," *Journal of the American Statistical Association* 82 (Dec.), pp. 1072-78.

Stock, James H., 2001 "Comment on 'Evolving Post-World War II U.S. Inflation Dynamics," *NBER Macroeconomics Annual* 16, pp. 379-87.

Stock, James H., and Mark W. Watson, 2002(a), "Macroeconomic Forecasting Using Diffusion Indexes," *Journal of Business and Economic Statistics* 20 (Apr.), pp. 147-62.

Stock, James H., and Mark W. Watson, 2002(b), "Has the Business Cycle Changed and Why?" *NBER Macroeconomics Annual* 17, pp. 159-218.

Taylor, John, 2000, "Low Inflation, Pass-Through, and the Pricing Power of Firms," *European Economic Review* 44 (June), pp. 1389-1408.

Whelan, Karl, 2002, "A Guide to the Use of Chain Aggregated NIPA Data," *Review of Income and Wealth* 48 (June), pp. 217-33.

Yoon, Gawon, 2001, "Inflation Persistence with Structural Breaks," manuscript, Pusan National University.

Figure 1: Weights vs. Persistence Estimates

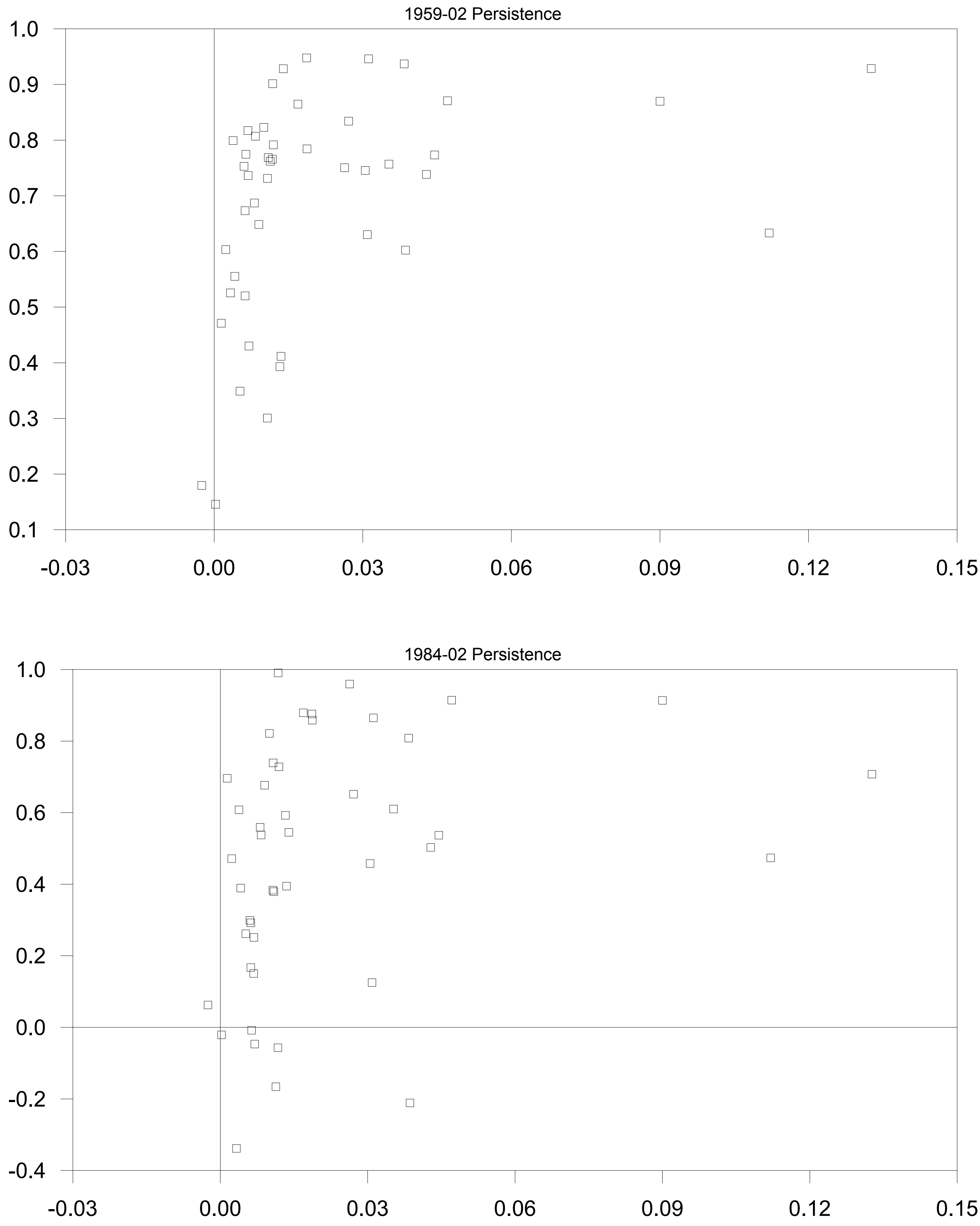
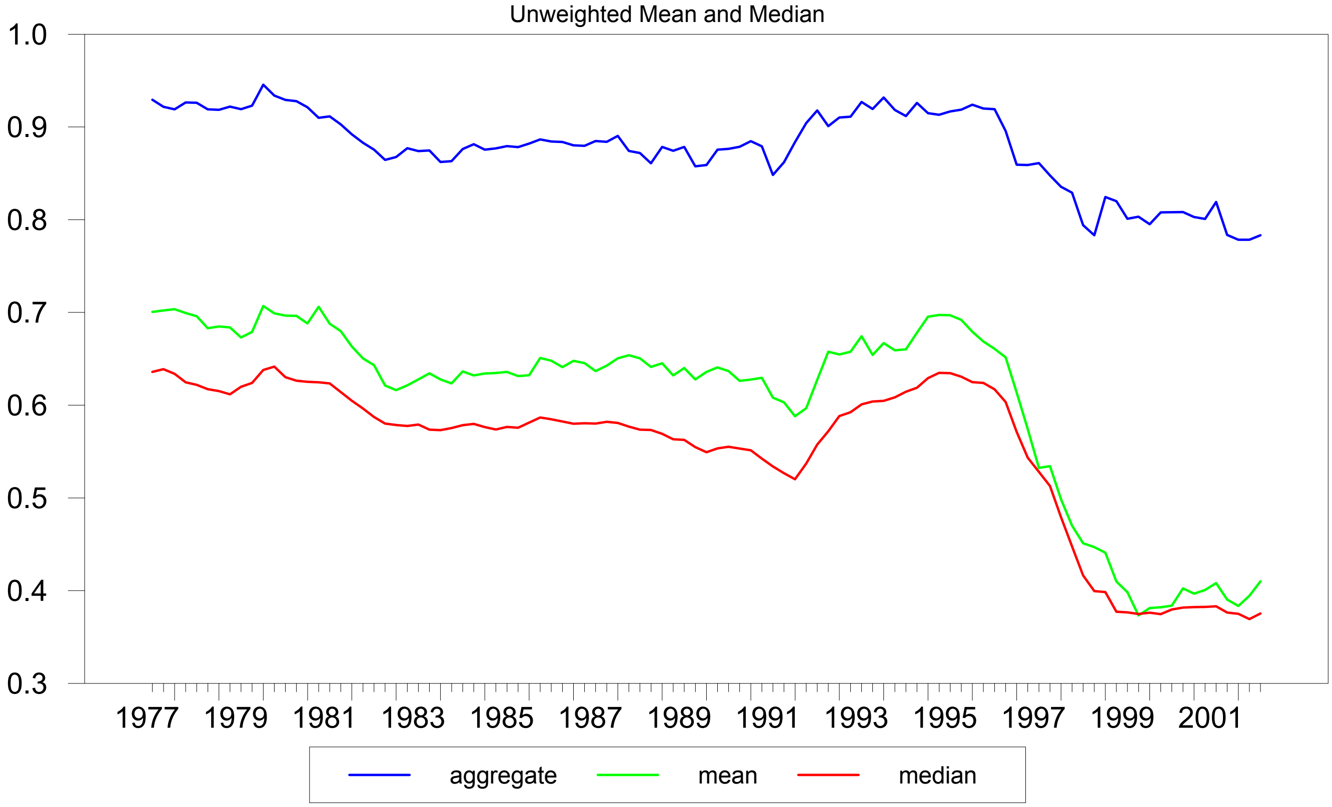


Figure 2: Rolling Persistence Estimates for Aggregate Inflation

Figure 3: Rolling Disaggregate Persistence Estimates, Mean and Median



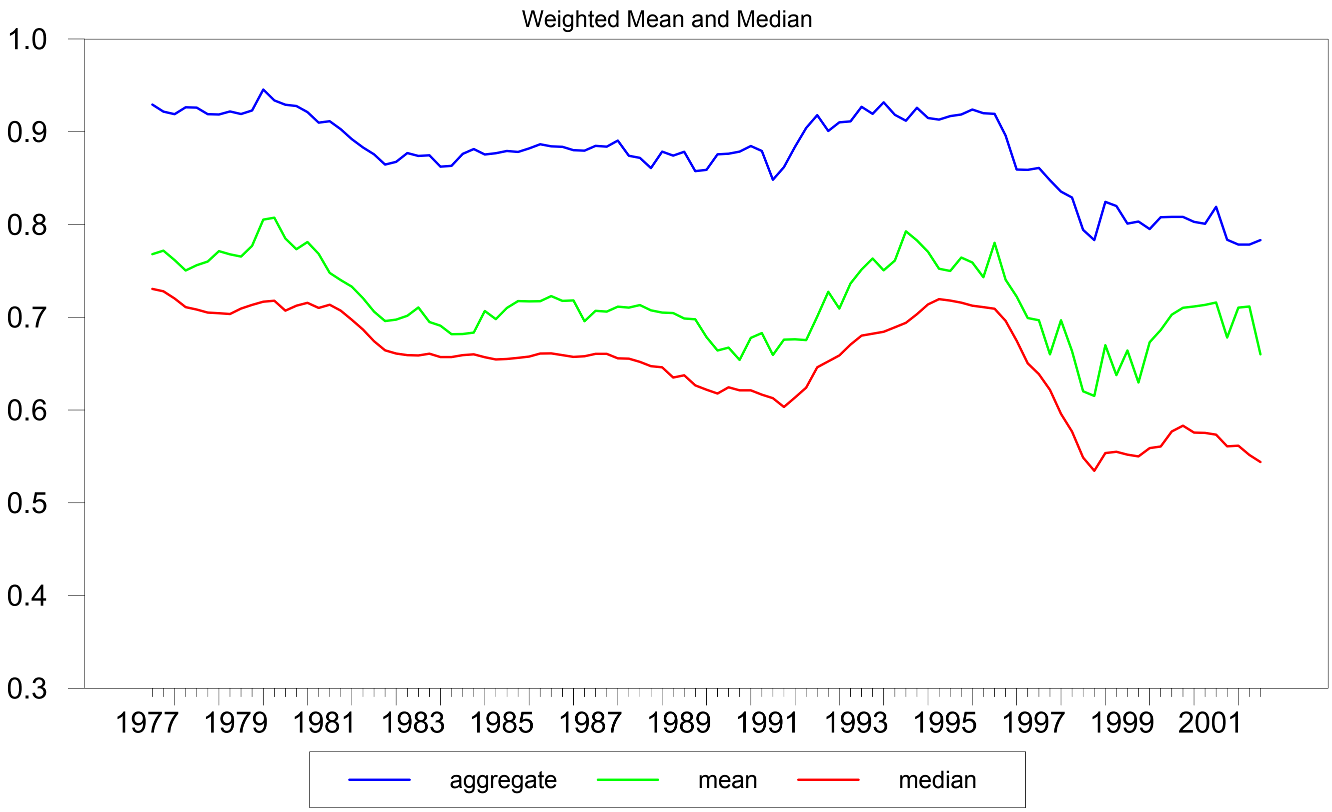


Figure 4: Rolling Persistence Estimates, Percentiles and Confidence Intervals

Table 1: Persistence Estimates for Aggregate Inflation						
	persistence estimate (90% band)	*median unbiased estimate*	*AR(1) esti-mate*	*persistence estimate (90% band)*	*median unbiased estimate*	*AR(1) esti-mate*
	Monthly data, 1959-02			**Quarterly data, 1959-02**		
PCE	.928 (.893,1.018)	.950	.673	.930 (.898,1.012)	.950	.931
PCE ex housing	.900 (.860,1.007)	.920	.604	.903 (.866, .999)	.917	.904
Durables	.897 (.843,1.028)	.936	.531	.901 (.849,1.029)	.938	.844
Nondurables	.822 (.745,1.005)	.857	.321	.878 (.813,1.029)	.917	.793
Services	.934 (.900,1.035)	.980	.520	.949 (.922,1.031)	.994	.906
	Monthly data, 1984-02			**Quarterly data, 1984-02**		
PCE	.834 (.741,1.114)	.987	.126	.907 (.856,1.096)	1.021	.788
PCE ex housing	.826 (.725,1.120)	1.000	.081	.904 (.848,1.103)	1.028	.751
Durables	.858 (.779,1.090)	.996	.335	.921 (.868,1.127)	1.047	.766
Nondurables	.802 (.668,1.232)	1.070	.004	.878 (.799,1.159)	1.054	.568
Services	.698 (.560,1.090)	.825	-.100	.855 (.766,1.108)	.996	.678

Notes:
1. All data exclude food and energy.
2. *Persistence* is measured by the sum of AR coefficients. The AR lag length for each model is selected using the AIC. The *AR(1)* entries are simple coefficient estimates from an AR(1) model.
3. The 90% confidence bands and median unbiased estimates are estimated with Hansen's (1999) grid bootstrap.

Table 2: Persistence Estimates by Level of Aggregation *(quarterly data)*								
	1959-02				**1984-02**			
	aggregation level				*aggregation level*			
	1	2	3	4	1	2	3	4
mean	.822	.688	.644	.619	.576	.463	.357	.338
median	.838	.750	.712	.657	.681	.503	.408	.305
75%ile	.901	.817	.814	.791	.821	.728	.733	.689
90%ile	.942	.901	.891	.876	.876	.876	.859	.815
weighted mean	.842	.774	.739	.729	.612	.588	.525	.520
weighted median	.851	.773	.799	.787	.768	.652	.708	.708
weighted 75%ile	.901	.871	.892	.900	.821	.859	.823	.806
weighted 90%ile	.942	.929	.929	.929	.876	.914	.911	.914
% pers. < agg. pers.	.818	.935	.972	.974	1.000	.913	.954	.942
% with upper band < 1	.545	.652	.606	.660	.364	.500	.596	.679
corr(weight,pers.)	.380	.347	.226	.225	.185	.293	.229	.220
# series	11	46	109	156	11	46	109	156

Notes:
1. All data exclude food and energy.
2. *Persistence* is measured by the sum of AR coefficients. The AR lag length for each model is selected using the AIC.
3. The mean, median, 75%ile, and 90%ile and their weighted counterparts represent summary statistics from the cross–section of persistence estimates at each level of aggregation.
4. The row "% pers. < agg. pers." reports the percentage of series at each level of aggregation for which estimated persistence is less than the estimated persistence of core PCE inflation (for each sample period).
4. The row "% with upper band < 1" reports, for each sample period, the percentage of series at each level of aggregation for which the upper band of the estimated confidence interval around the persistence estimate, computed with Hansen's (1999) grid bootstrap, is less than 1.

Table 3: Persistence Estimates for Disaggregate Inflation, Level 3						
	persistence estimate (90% band)	*median unbiased estimate*	*AR(1) estimate*	*persistence estimate (90% band)*	*median unbiased estimate*	*AR(1) estimate*
	Monthly data, 1959-02			**Quarterly data, 1959-02**		
mean	.537		.119	.644		.470
median	.623 (.494, .833)	.663	.119	.712 (.602, .933)	.757	.495
75%ile	.790 (.693,1.040)	.852	.278	.814 (.726,1.030)	.861	.658
90%ile	.868 (.799,1.049)	.932	.491	.891 (.838,1.014)	.918	.744
weighted mean	.669		.249	.739		.577
weighted median	.736 (.622, .962)	.778	.274	.799 (.713,1.021)	.842	.628
weighted 75%ile	.881 (.829,1.011)	.909	.424	.892 (.835,1.041)	.940	.837
weighted 90%ile	.910 (.862,1.034)	.953	.580	.929 (.891,1.036)	.973	.860
	Monthly data, 1984-02			**Quarterly data, 1984-02**		
mean	.108		.072	.357		.268
median	.121 (.004, .245)	.121	.019	.408 (.200, .772)	.473	.239
75%ile	.670 (.537, .944)	.732	.182	.733 (.577,1.232)	1.029	.468
90%ile	.802 (.664,1.142)	1.015	.463	.859 (.774,1.115)	1.014	.682
weighted mean	.328		.138	.525		.391
weighted median	.670 (.537, .944)	.732	.144	.708 (.568,1.032)	.769	.344
weighted 75%ile	.750 (.638,1.278)	1.070	.305	.823 (.747,1.049)	.886	.618
weighted 90%ile	.896 (.828,1.129)	1.046	.465	.911 (.862,1.074)	1.010	.811

Notes:
1. See the notes to Tables 1 and 2.
2. For those sample statistics that represent an observation in the cross–section distribution of persistence estimates — the reported median, for example, is the persistence estimate for a particular PCE sector — the table reports the associated median unbiased estimate and 90% confidence interval, estimated with Hansen's (1999) grid bootstrap.

Table 4: Distribution of Disaggregate Persistence Estimates, Level 3 *(quarterly data)*		
Unweighted: Percentage Distribution		
	1959-02	*1984-02*
persistence <= .2	0.046	0.358
.2 < persistence <= .4	0.101	0.138
.4 < persistence <= .6	0.211	0.156
.6 < persistence <= .8	0.349	0.147
.8 < persistence <= 1.	0.294	0.202
Weighted: Expenditure Share-Based Distribution		
	1959-02	*1984-02*
persistence <= .2	0.013	0.205
.2 < persistence <= .4	0.071	0.086
.4 < persistence <= .6	0.159	0.148
.6 < persistence <= .8	0.259	0.253
.8 < persistence <= 1.	0.499	0.309

Notes:
1. *Persistence* is measured by the sum of AR coefficients. The AR lag length for each model is selected using the AIC.
2. The upper panel reports the percentage of (109) components for which the sample persistence estimate falls within the indicated range.
3. The lower panel reports the 2001 (nominal) expenditure shares of those components for which the sample persistence estimate falls within the indicated range.

Table 5: Monte Carlo Results on the Size of Break Tests *(nominal size = 10%)*				
	Empirical size based on p-values from:			
	asymptotic distribution	*fixed bootstrap*	*heteroskedastic bootstrap*	*simple bootstrap*
PCE DGP				
intercept	.562	.448	.529	.138
set of AR coefs.	.549	.218	.412	.113
all coefs.	.750	.276	.509	.125
resid. var.	.070			.104
Durables DGP				
intercept	.571	.460	.536	.150
set of AR coefs.	.545	.130	.311	.103
all coefs.	.819	.240	.467	.141
resid. var.	.072			.098
Nondurables DGP				
intercept	.576	.456	.525	.165
set of AR coefs.	.626	.097	.274	.100
all coefs.	.848	.188	.409	.134
resid. var.	.068			.093
Services DGP				
intercept	.511	.409	.485	.137
set of AR coefs.	.595	.170	.364	.119
all coefs.	.796	.209	.436	.134
resid. var.	.070			.097
Recreation DGP				
intercept	.244	.204	.245	.119
set of AR coefs.	.315	.150	.261	.117
all coefs.	.469	.168	.325	.112
resid. var.	.077			.105

Notes:
1. For the inflation rates of aggregate (core) PCE, durables, nondurables, services, and recreation, DGPs are specified from estimates of AR models fit over the regression sample of 1985:Q2 to 2002:Q3, with lag lengths determined using the AIC, assuming no change in the intercept.
2. Each DGP is used to generate 1000 artificial data sets of length equal to 70 (the 1985:Q2 to 2002:Q3 sample consists of 70 observations) plus the number of lags, using draws from the normal distribution. Each artificial data series is used to fit an AR model and test it for a break in the intercept, set of AR coefficients, all coefficients, and the residual variance. In line with the empirical analysis, the Andrews break tests use a sample trim of $\pi_0 = .21$.
3. In each draw, the fixed and heteroskedastic bootstraps developed by Hansen (2000) and a simple parametric bootstrap are used to generate p–values, based on 999 bootstrap replications (see footnote 22 in the text for details of the simple bootstrap).
4. The table reports the percentage of Monte Carlo trials in which asymptotic, fixed regressor bootstrap, heteroskedastic bootstrap, and simple bootstrap p–values indicate the null of stability is rejected at 10% confidence. The asymptotic p–values are computed using Hansen's (1997) approximation.

Table 6: Univariate Break Test Results, Aggregate Inflation and Level 1 Disaggregate Inflation

(quarterly data, 1984-2002)

Aggregate Inflation

		bootstrap p-values:			
	lag	*intercept*	*AR coefs.*	*all coefs.*	*resid. var.*
PCE	3	.029 (92:01)	.130	.016 (92:01)	.457
Durables	4	.019 (95:01)	.575	.162	.250
Nondurables	5	.281	.281	.617	.021 (90:02)
Services	4	.039 (92:01)	.142	.169	.442

Disaggregate Inflation, Level 1

		bootstrap p-values:			
	lag	*intercept*	*AR coefs.*	*all coefs.*	*resid. var.*
Durables					
Motor vehicles and parts	3	.016 (96:01)	.600	.224	.368
Furniture and household equipment	2	.001 (94:03)	.007 (94:03)	.040 (94:03)	.606
Other durables	3	.035 (91:04)	.323	.256	.474
Nondurables					
Clothing and shoes	5	.119	.885	.423	.027 (90:02)
Other nondurables	4	.220	.126	.152	.034 (98:04)
Services					
Housing	3	.245	.246	.526	.007 (91:04)
Household operation	1	.852	.957	.871	.067 (97:04)
Transportation	5	.070 (93:02)	.050 (92:04)	.089 (93:01)	.219
Medical care	3	.373	.379	.562	.404
Recreation	2	.036 (91:02)	.110	.301	.259
Other services	1	.018 (92:01)	.005 (92:01)	.017 (92:01)	.489

Note:
1. The table reports simple parametric bootstrap p-values for Andrews break tests applied to the parameters of AR models estimated for each inflation series, using a lag length determined by the AIC (indicated in the table). The regression sample is 1984:Q1+lag order to 2002:Q3. The tests use a sample trim of $\pi_0 = .21$.
2. The number of bootstrap replications is 999.

Table 7: System Break Test Results, Aggregate Inflation and Level 1 Disaggregate Inflation *(quarterly data, 1984-2002)*	
Aggregate Inflation (n = 3)	
	bootstrap p-value
intercept, SUR	.009 (92:01)
intercept, OLS	.049 (92:02)
resid var, SUR	.105 (90:02)
Disaggregate Inflation, Level 1 (n = 11)	
	bootstrap p-value
intercept, SUR	.001 (92:02)
intercept, OLS	.036 (92:02)
resid var, SUR	.016 (91:02)

Notes:
1. The table reports simple parametric bootstrap p-values for sup Wald break tests applied to systems of inflation equations. The top panel provides results for three–variable systems, consisting of equations for durables, nondurables, and services. The lower panel provides results for 11–variable systems, consisting of equations for the components listed in Appendix Table 1.
2. The top two rows of each panel present results for tests of the stability of the set of intercepts in a set of AR equations for the indicated variables. The lag lengths of the equations are set separately for each one based on the AIC; they are listed in Table 7. The system of AR equations is then estimated over the sample of 1984:Q1+(longest lag) to 2002:Q3. In one case, the system is estimated using SUR methods; in the other, using equation–by-equation OLS. In the SUR case, the Wald tests impose homoskedasticity in each equation's residuals. In the OLS case, the Wald tests are computed with heteroskedasticity–robust variances. The SUR break date estimate corresponds to the date at which the system likelihood function is maximized. The OLS break date estimate corresponds to the date at which the sum of squared residuals for the system is minimized.
3. The third row of each panel presents the result of a test of the stability of the residual variances of the set of AR equations for the indicated variables. The system of equations for the residual variances is estimated with SUR methods.

Table 8: Summary of Univariate Break Test Results Disaggregate Inflation, Levels 2 and 3 *(quarterly data, 1984-2002)*			
Level 2 (n = **46**)			
	number of series with rejections:		
	intercept	*AR coefs.*	*resid. var.*
Durables ($n = 13$)	8	1	3
Nondurables ($n = 13$)	8	1	5
Services ($n = 20$)	8	6	10
total ($n = 46$)	24	8	18
avg. date, signif. breaks	93:3	92:2	92:3
Level 3 (n = **109**)			
	number of series with rejections:		
	intercept	*AR coefs.*	*resid. var.*
Durables ($n = 23$)	7	3	6
Nondurables ($n = 27$)	20	6	7
Services ($n = 59$)	25	21	26
total ($n = 109$)	52	30	39
avg. date, signif. breaks	93:3	92:3	91:3

Notes:
1. See the notes to Table 6.
2. The table simply provides a count of the number of series for which bootstrap p-values reject the null of stability, with the count broken down by durables, nondurables, and services. The counts are based on a 10% significance level.
3. The bottom row of each panel reports the average of the least squares date estimates for those breaks that are statistically significant (at 10%).

Table 9: Persistence Estimates for Aggregate and Disaggregate Inflation (Level 3), Allowing Mean Shift in 1993:1

(quarterly data, 1984-2002)

Aggregate Estimates		
	persistence estimate (90% band)	*median unbiased estimate*
PCE	.402 (.206,1.052)	.538
Durables	.800 (.753,1.124)	1.029
Nondurables	.367 (.175,1.039)	.559
Services	.137 (-.103, .633)	.249
Disaggregate Estimates, Level 3		
	persistence estimate (90% band)	*median unbiased estimate*
mean	.192	
median	.226 (.018, .786)	.370
75%ile	.478 (.350,1.134)	.750
90%ile	.728 (.659,1.182)	1.044
weighted mean	.353	
weighted median	.519 (.371,1.081)	.693
weighted 75%ile	.661 (.536,1.078)	.807
weighted 90%ile	.728 (.650,1.167)	1.025
% with upper band < 1	.716	

Notes:
1. All data exclude food and energy.
2. *Persistence* is measured by the sum of AR coefficients, in a model that allows the intercept to shift in 1993:Q1. The AR lag length for each model is selected using the AIC, assuming a stable intercept.
3. The mean, median, 75%ile, and 90%ile and their weighted counterparts represent summary statistics from the cross–section of disaggregate persistence estimates.
4. For those sample statistics that represent an observation in the cross–section distribution of persistence estimates — the reported median, for example, is the persistence estimate for a particular PCE sector — the table reports the associated median unbiased estimate and 90% confidence interval, estimated with Hansen's (1999) grid bootstrap.
5. The row "% with upper band < 1" reports, for each sample period, the percentage of series at each level of aggregation for which the upper band of the estimated confidence interval around the persistence estimate, computed with Hansen's (1999) grid bootstrap (treating the break date as fixed and known), is less than 1.

Table 10: Factor Model–Based Persistence Estimates *(quarterly data, 1984-2002)*				
Models Assuming Constant Mean Inflation				
	common factor persistence		idiosyncratic persistence	
	sample est. (90% band)	*median unbiased est.*	*weighted mean*	*weighted median*
agg. level 2	.957 (.918,1.091)	1.035	.294	.372
agg. level 3	.966 (.934,1.086)	1.036	.224	.298
Models Allowing 1993:Q1 Shift in Mean Inflation				
	common factor persistence		idiosyncratic persistence	
	sample est. (90% band)	*median unbiased est.*	*weighted mean*	*weighted median*
agg. level 2	.597 (.434, .903)	.658	.267	.320
agg. level 3	.808 (.720,1.039)	.858	.160	.286

Notes:
1. The results are based on common and idiosyncratic components estimated with the factor model approach of Stock and Watson (2002). In particular, for a given aggregation level, a common factor is estimated as the first principal component from standardized inflation rates, either assuming a constant mean over the 1984-2002 sample or allowing a mean shift in 1993:Q1. The idiosyncratic component for each sector i is estimated as the residual from a regression of standardized inflation in sector i on the common factor.
2. For the estimated common factor and idiosyncratic components, *persistence* is measured by the sum of AR coefficients. The AR lag length for each model is selected using the AIC.
3. For the aggregate factor, the table reports the associated median unbiased estimate of persistence and the 90% confidence interval, estimated with Hansen's (1999) grid bootstrap.

Appendix Table 1
Components in Aggregation Level 1

	Line #	Component	Weight
1.	44	Motor vehicles and parts	6.213
2.	58	Furniture and household equipment	5.264
3.	85	Other durables	2.896
4.	140	Clothing and shoes	5.420
5.	164	Other nondurables	9.543
6.	197	Housing	17.443
7.	213	Household operation	6.986
8.	241	Transportation	4.666
9.	262	Medical care	18.436
10.	282	Recreation	4.675
11.	310	Other services	18.458

Notes:
1. "Line #" refers to the "line number" the BEA assigns each component in the detailed PCE data file.
2. The weights given in the table represent the listed components' shares of core PCE (nominal) spending in 2001, in percentage terms. The weights sum to 100.

Appendix Table 2

Components in Aggregation Level 2

	Line #	Component	Weight
1.	45	New autos	1.871
2.	48	Net purchases of used autos	1.071
3.	52	Other motor vehicles	2.632
4.	55	Tires, tubes, accessories, and other parts	.809
5.	59	Furniture, including mattresses and bedsprings	1.130
6.	60	Kitchen and other household appliances	.638
7.	63	China, glassware, tableware, and utensils	.602
8.	64	Video and audio goods, including musical instruments, etc.	1.865
9.	76	Other durable house furnishings	1.172
10.	86	Ophthalmic products and orthopedic appliances	.381
11.	87	Wheel goods, sports and photographic equipment, boats, etc.	1.073
12.	97	Jewelry and watches	.900
13.	98	Books and maps	.621
14.	141	Shoes	.830
15.	142	Women's and children's clothing and accessories except	3.049
16.	148	Men's and boys' clothing and accessories except shoes	1.689
17.	165	Tobacco products	1.348
18.	166	Toilet articles and preparations	.999
19.	170	Semidurable house furnishings	.683
20.	171	Cleaning and polishing preparations, and miscellaneous	1.091
21.	175	Drug preparations and sundries	3.116
22.	180	Nondurable toys and sport supplies	1.178
23.	184	Stationery and writing supplies	.416
24.	187	Net foreign remittances	.025
25.	192	Magazines, newspapers, and sheet music	.621
26.	195	Flowers, seeds, and potted plants	.327
27.	198	Owner-occupied nonfarm dwellings–space rent	13.264
28.	201	Tenant-occupied nonfarm dwellings–rent	3.835
29.	205	Rental value of farm dwellings	.141
30.	206	Other housing	.679
31.	217	Other household operation	4.448
32.	242	User-operated transportation	3.860
33.	254	Purchased local transportation	.233
34.	257	Purchased intercity transportation	.700
35.	263	Physicians	4.710
36.	264	Dentists	1.193
37.	265	Other professional services	2.711
38.	270	Hospitals and nursing homes	9.001
39.	278	Health insurance	1.324
40.	283	Admissions to specified spectator amusements	.519
41.	287	Other recreation	4.284
42.	311	Personal care	1.395
43.	323	Personal business	11.204
44.	351	Education and research	3.089
45.	361	Religious and welfare activities	3.526
46.	369	Net foreign travel	-.253

Note: See the notes to Appendix Table 1.

Appendix Table 3
Components in Aggregation Level 3

	Line #	Component	Weight
1.	46	New domestic autos	1.118
2.	47	New foreign autos	.817
3.	49	Net transactions in used autos	.650
4.	50	Used auto margin	.488
5.	51	Employee reimbursement	-.032
6.	53	Trucks, new and net used	2.472
7.	54	Recreational vehicles	.250
8.	56	Tires and tubes	.369
9.	57	Accessories and parts	.467
10.	59	Furniture, including mattresses and bedsprings	1.169
11.	61	Major household appliances	.577
12.	62	Small electric appliances	.083
13.	63	China, glassware, tableware, and utensils	.622
14.	65	Video and audio goods, including musical instruments, etc.	1.328
15.	77	Floor coverings	.322
16.	78	Durable housefurnishings n.e.c.	.625
17.	81	Writing equipment	.073
18.	82	Hand tools	.191
19.	86	Ophthalmic products and orthopedic appliances	.394
20.	88	Sport and photo equipment, cycles	.845
21.	94	Pleasure boats and aircraft	.265
22.	97	Jewelry and watches	.931
23.	98	Books and maps	.642
24.	141	Shoes	.858
25.	143	Clothing and sewing for females	3.082
26.	147	Luggage for females	.071
27.	149	Men's and boys' clothing and luggage	1.741
28.	154	Standard clothing issued to military personnel	.006
29.	165	Tobacco products	1.394
30.	167	Soap	.101
31.	168	Cosmetics and perfumes	.339
32.	169	Other person hygiene goods	.593
33.	170	Semidurable house furnishings	.706
34.	172	Cleaning preparations	.555
35.	173	Lighting supplies	.143
36.	174	Paper products	.430
37.	176	Prescription drugs	2.384
38.	177	Nonprescription drugs	.663
39.	178	Medical supplies	.103
40.	179	Gynecological goods	.072
41.	181	Toys, dolls, and games	.877
42.	182	Sport supplies, including ammo	.237
43.	183	Film and photo supplies	.105
44.	185	Stationery and school supplies	.171
45.	186	Greeting cards	.259
46.	188	Expenditures abroad by U.S. residents	.066
47.	191	(Less) Personal remittances in kind to nonresidents	-.040
48.	193	Magazines and sheet music	.342
49.	194	Newspapers	.301
50.	195	Flowers, seeds, and potted plants	.339
51.	199	Owner-occupied mobile home	.537
52.	200	Owner-occupied station homes	13.179
53.	202	Tenant-occupied mobile homes	.111
54.	203	Tenant-occupied station homes	3.744
55.	204	Tenant-landlord durables	.110
56.	205	Rental value of farm dwellings	.145
57.	207	Hotels and motels	.499
58.	208	Clubs and fraternity housing	.012
59.	209	Higher education housing	.170
60.	210	Elementary and secondary education housing	.003

Appendix Table 3, Continued

Components in Aggregation Level 3

	Line #	Component	Weight
61.	211	Tenant group room and board	.017
62.	212	Tenant group employee lodging	.001
63.	218	Water and other sanitary services	.926
64.	221	Telephone and telegraph	2.494
65.	228	Domestic service	.267
66.	231	Other other household operation	.913
67.	243	Repair, greasing, washing, parking, storage, rental, etc.	3.317
68.	251	Other user-operated transportation	.675
69.	255	Mass transit systems	.173
70.	256	Taxicab	.068
71.	258	Railway	.017
72.	259	Bus	.027
73.	260	Airline	.593
74.	261	Other intercity purchased transportation	.088
75.	263	Physicians	4.870
76.	264	Dentists	1.234
77.	271	Hospitals	7.779
78.	275	Nursing homes	1.528
79.	279	Medical care and hospitalization	1.219
80.	280	Income loss	.034
81.	281	Workers' compensation	.116
82.	284	Motion picture theaters	.160
83.	285	Legitimate theaters and opera, and entertainments	.193
84.	286	Spectator sports	.184
85.	288	Radio and television repair	.077
86.	289	Clubs and fraternal organizations	.315
87.	290	Commercial participant amusements	1.339
88.	296	Pari-mutuel net receipts	.088
89.	297	Other other recreation	2.610
90.	312	Cleaning, storage, and repair of clothing and shoes	.285
91.	317	Barbershops, beauty parlors, and health clubs	.595
92.	320	Other personal care	.563
93.	324	Brokerage charges and investment counseling	1.356
94.	332	Bank service charges, trust services, and safe deposit	1.342
95.	337	Services furnished without payment by financial intermediaries	4.740
96.	340	Expense of handling life insurance and pension plans	1.898
97.	341	Legal services	1.285
98.	342	Funeral and burial expenses	.320
99.	343	Other personal business	.644
100.	352	Higher education	1.601
101.	355	Nursery, elementary, and secondary schools	.649
102.	358	Other education and research	.945
103.	362	Political organizations	.013
104.	363	Museums and libraries	.134
105.	364	Foundations to religion and welfare	.076
106.	365	Social welfare	2.502
107.	368	Religion	.921
108.	370	Foreign travel by U.S. residents	1.394
109.	374	Expenditures in the U.S. by nonresidents	-1.655

Note: See the notes to Appendix Table 1.