BLS WORKING PAPERS

U.S. Department of Labor
U.S. Bureau of Labor Statistics
Office of Prices and Living Conditions

The Effects of Rounding on the Consumer Price Index

Elliot Williams, U.S. Bureau of Labor Statistics

Working Paper 397
August 2006

The Effects of Rounding on the Consumer Price Index

Elliot Williams*
Division of Price and Index Number Research
Bureau of Labor Statistics†

Abstract

The Bureau of Labor Statistics rounds the Consumer Price Index (CPI) to a single decimal place before releasing it, and the published CPI inflation series is calculated from those rounded index values. While rounding has only a relatively small effect on the *level* of the CPI series at present, it can have a significant effect on CPI inflation, the monthly *percent changes* in the CPI.

This paper estimates the impact of rounding error on the published CPI inflation for both contemporaneous and historical data. Using an unrounded CPI series from January 1986 to July 2005 as a benchmark, I find that published CPI inflation differs from its full-precision counterpart approximately 25% of the time, and that reporting the CPI levels to three decimal places would reduce these discrepancies to under 0.5%. Further, the variance introduced by rounding error is large when compared to the sampling variation in CPI inflation. I find that the BLS could reduce total CPI inflation error variance by 42% by simply reporting more digits in the CPI index, resulting in a significantly more accurate reflection of monthly inflation.

In order to extend these results to the CPI historical series, I derive the distribution of the rounding error component of inflation. From this analysis, it is possible to estimate the probability of large rounding errors for a given CPI level and rounding precision. Three regimes emerge. Before the 1970's inflation, discrepancies due to rounding were both frequent and frequently large relative to the underlying inflation rate. During the inflationary period of the mid-1970's to mid-1980's, both the probability and relative magnitude of discrepancies decrease dramatically. Finally, the last twenty years are characterized by a slowly falling probability of any rounding-induced error, but a roughly constant probability of an error of a given size.

*The author would like to thank David Johnson, John Greenlees, Robert McClelland, and DPINR for helpful comments and feedback on earlier drafts of this paper.

†The views expressed in this paper are the author's and do not represent those of the BLS.

1 Introduction

The U.S. Bureau of Labor Statistics rounds the Consumer Price Index (CPI) to a single decimal place before it is publicly released. In 1984, the current index was rebased to 100.0 and it stands near 200 today. Because the index value is so large, one might think that the difference between a CPI of 189.7 and 189.72 would be negligible. However, this is not the case for the percent change between two CPI values, CPI inflation. Because the actual changes in the CPI have been small recently (the rate of inflation has been relatively low) the small differences in rounding up or down can create a significantly misleading picture of monthly price inflation.

And this difference matters in the economy. For instance, this Reuters news article, "Consumer Prices Jump, Spur Inflation Woes," is representative of the impact that the release of the February 2005 index value had, signaling a surprisingly large increase in the rate of inflation to financial markets:

> The core CPI, which strips out volatile food and energy costs, rose 0.3 percent. It was the biggest rise in the core rate since September and broke a string of four straight 0.2 percent gains.
>
> Wall Street economists had braced for a milder 0.3 percent rise in overall consumer prices and had expected another 0.2 percent gain outside food and energy.
>
> The report added to financial market inflation jitters and increased speculation the Federal Reserve, which raised credit costs on Tuesday, might step up the pace of its rate rise to keep inflation under wraps.[1]

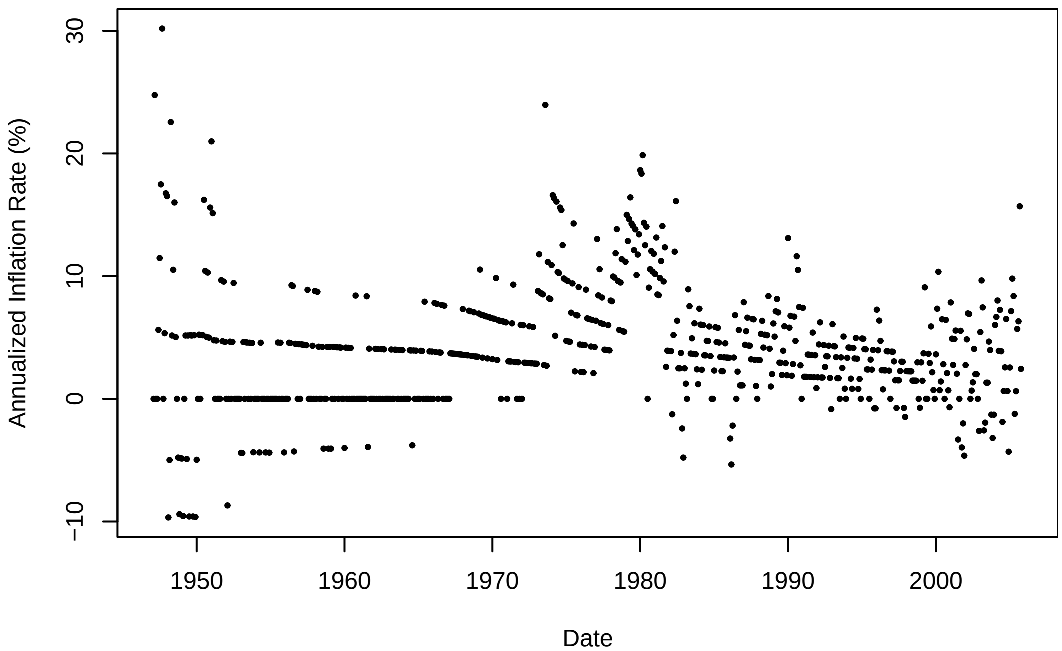
Both the stock and bond markets moved on the news that the index for all items less food and energy increased from inflating at a steady 0.2% rate to 0.3%, a relatively large growth in the rate of inflation. But in this case, the apparent increase is an artifact of using already-rounded index values to calculate the inflation rate. Calculating the "core" inflation rate using an unrounded CPI index series gives 0.2% instead of 0.3%, and would have constituted essentially no news for inflation projections or bond prices. This paper demonstrates how this can happen and investigates how frequently there is a discrepancy between inflation rates calculated from unrounded and rounded indexes under different possible rounding policies.

While the rounding error in recent month's CPI inflation can cause a passing stir in the financial markets, some effects of rounding are still more marked in the historical CPI series. Plotting the percent changes in the published CPI all-items series with points in addition to the usual lines makes the rounding apparent to the naked eye, as shown in Figure 1.

There is nothing fancy about this plot; each month a percent change is calculated and plotted. The fanning horizontal lines that the eye picks up are evidence of the fact that rounding the original series constrains changes in the level of the CPI to integer multiples

[1] Source: "Consumer Prices Jump, Spur Inflation Woes", Reuters, March 24, 2005

Figure 1: CPI Inflation



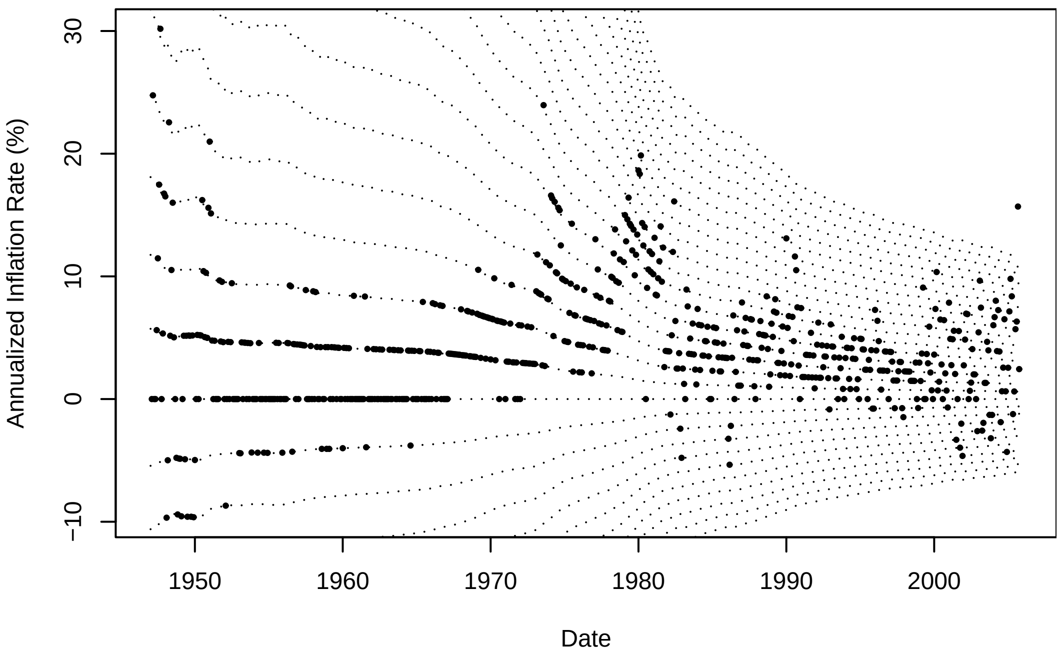
of 0.1. The post-rounding percentage changes are thus integer multiples of 0.1 divided by the level of the CPI in the beginning period. The horizontal lines which appear in the series correspond exactly to $\{\ldots, -0.2, -0.1,\ 0.0,\ 0.1,\ 0.2,\ \ldots\}/CPI_t$.

Figure 2 demonstrates this fact by overlaying the plot of points with lines corresponding to the different allowed rounded CPI inflation values in the historical 1984-base-year series. One can see that all of the percent changes in the reported series line up nicely, as they must.

Without doing any math, some features of the percent difference series are immediately apparent. Inflation takes on discrete values, which widen apart as the level of the CPI decreases. In the period between 1955 and 1970, inflation took on one of four values, and only two of them with any regularity. The most you can say about this period from the rounded data is that monthly inflation was at an annualized rate somewhere between 0% and 5%. Gradual inflation, especially in the earlier part of the series, is replaced with months of zero inflation followed by months with too-large inflation. This effect is visible on inspection; too many of the earlier months in the series register zero inflation. Over all the postwar data, 19.5% of the monthly changes are exactly zero. And each month that is rounded down to zero is offset by other months which are rounded upwards by the same amount. Thus rounding tends to inflate the time-series variance of inflation, making it appear that monthly inflation was swinging wildly during the period, when in fact it was relatively calmer.

In recent times, rounding error significantly increases the variability of CPI inflation.

Figure 2: CPI Inflation: Over-plotted with Allowed Levels



Because the BLS now collects a very large number of prices for goods and services every month, the sampling variation of CPI inflation is very small, on the order of 0.0036% monthly. Unfortunately, as shown below, rounding error adds an additional 0.0026% error variance to the reported figures – an additional 72%. In the case of monthly CPI all-items inflation, the current BLS rounding policy is obscuring a reliably estimated figure.

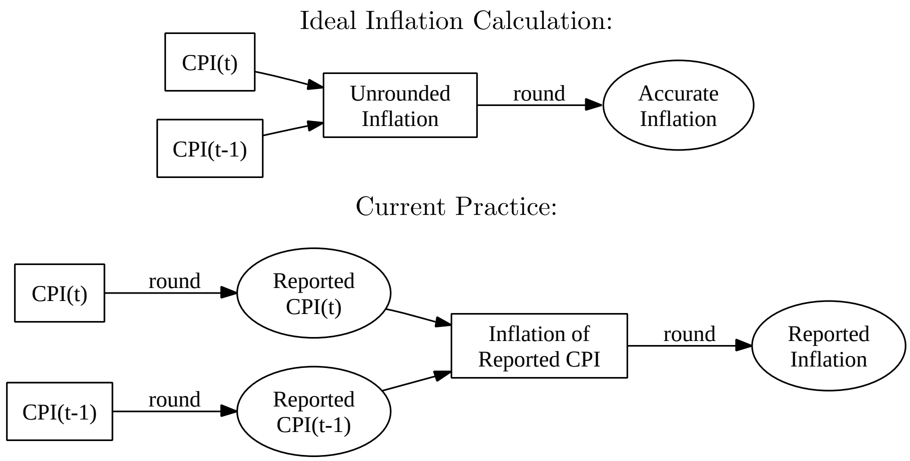
The picture of rounding in CPI inflation is not entirely bleak, however. In the long run, the rounding errors *do* average out, so rounding is not a source of long-term bias in the CPI. Rounding is also a less important source of error in the annual inflation series than the monthly series, because the average of twelve rounding errors is closer to zero than a single error, and the magnitude of a year's inflation (for most years) is larger than the magnitude of a single rounding error. Moreover, the BLS still makes a series in the 1967 base-year available, which is less subject to rounding error due simply to the fact that the index values are larger, and thus rounding to the tenths place is a smaller relative error. And finally, one could in principle calculate a more accurate monthly inflation series by going back to the original publications (when the values were higher because the series had not yet been rebased) and converting them to current values, retaining the extra precision.

Nevertheless, for short-run inflation using the current CPI series, the choice of rounding to the first decimal place has profound effects on the accuracy of contemporaneous and historical data. Section 2 details how the CPI series are rounded, and demonstrates by example how discrepancies can arise. Section 3 examines the effects of rounding error on recent inflation data for which an unrounded counterpart is available. Section 4 mathematically analyzes the effect of rounding error to extend these results to the entire historical CPI data series. Section 5 concludes.

2 Rounding the CPI

The BLS long ago standardized on one decimal place as the precision level to report all of its CPI series. Both the level of the CPI and the percent change in the CPI are rounded to the tenths place before being released to the public as official statistics. However, because the BLS desires to have the released inflation series match the released index series, CPI inflation is calculated from the *rounded* CPI index values. Figure 3 illustrates the way CPI inflation is calculated.

Figure 3: CPI Rounding Procedures



Notice that the final inflation figure has been rounded twice, once before taking a percent difference to calculate inflation and once afterward. The two stages of rounding the CPI inflation series have qualitatively different effects.

The final stage of rounding merely shortens the figure and provides a signal of how much confidence the BLS has in the inflation estimate. Indeed, it is possible to motivate the choice of rounding the inflation rate to the nearest 0.1% by appealing to the BLS's estimates of the sampling variation in the CPI. Approximate 95% confidence intervals can be constructed around the reported inflation figure by adding or subtracting 0.12%[2]. Thus when the BLS reports a value of 0.2%, one can be 95% confident that the true value lies between approximately 0.1% and 0.3%. Releasing the final inflation figure with less precision would obscure detail that the BLS measures well, while releasing more precision would give an appearance of more confidence in the inflation estimate than is warranted.

[2]See `http://www.bls.gov/cpi/cpivar2005.pdf`, for instance.

The first stage rounding – rounding the CPI index level – is the cause of the troubles documented here. A numerical example can help demonstrate how rounding the CPI index before calculating the inflation rate can result in discrepancies between rounded and unrounded figures; see table 1.

Table 1: A Demonstration of Rounding Error

	Unrounded		Rounded	
	CPI_t	CPI_{t-1}	CPI_t	CPI_{t-1}
CPI_t	192.345	192.770	192.3	192.8
ΔCPI_t	0.425		0.500	
$\%\Delta CPI_t$	0.425 / 192.770		0.5 / 192.8	
	= 0.221%		= 0.260%	
Rounded $\%\Delta CPI_t$	0.2%		0.3%	

The first column in Table 1 corresponds to the ideal inflation calculation method presented above. ΔCPI, the change in the CPI, is calculated by subtracting last period's unrounded CPI from this period's unrounded value. This difference is then divided by last period's CPI, and the final result rounded. The second column corresponds to BLS practice – the same procedure is carried out, but starting instead with the rounded CPI values.

Comparing the rounded and unrounded CPI levels, it is evident that they differ only in precision. However the change in the CPI, ΔCPI, calculated from the rounded data is different from ΔCPI calculated from the unrounded data. This discrepancy then carries over into the final percent change.

Note here that the difference between the two changes is small relative to the size of the index, and thus the difference between the rounded and unrounded percent changes is relatively small, 0.1%, though not insignificant. For the historical series, the relative size of the errors increases because, while the difference between the rounded and unrounded numbers remains constant as one moves back in time, the level of the index drops. This makes the error in the percent change series larger and larger as the level of the index falls.

For instance, if this example were based on 1960's data when the index level was around 30, one can imagine the unrounded CPI values being 30.345 and 30.770 respectively. The error in the change in levels remains the same at 0.075, while the error in the percent change would be $0.075/30 = 0.25\%$, which is twice as large as the average monthly inflation rate in 1960. A rounding error of this size would obscure the actual monthly changes in the inflation rate for a large part of the historical series.

Given that the BLS desires to use rounded indexes to calculate monthly CPI inflation, and desires to report an accurate statistic to 0.1% precision, how many digits should it retain in the CPI levels series? How often do the reported BLS inflation numbers differ from a measure of inflation calculated with the unrounded CPI figures? When the figures do differ, by how much? What effect does rounding the CPI levels have on the error variance of inflation?

To answer these questions, this paper takes two paths. Where rounded CPI data are available, I compare the published CPI inflation values to those calculated before rounding and ask how often they match at the reported level of precision. To address the frequency and magnitude of these differences in pre-1986 data, the paper relies on some simple statistical analysis.

3 Investigating Different Rounding Policies with Real Data

To construct a measure of inflation which is free from rounding error, this section uses the CPI's Research Database (RDB) index data files. This database includes all of the major indexes from January 1986 to July 2005 at the full level of precision used internally at the BLS. For this paper, the CPI all-items index and its top-level components are considered. Additionally, the information technology and personal computers indexes are included because they have seen rapid declines in price, and are probably the worst case scenario for rounding error in the post-1986 period.

A monthly benchmark inflation series is calculated from the unrounded data and then rounded to the one-tenth of a percent level, as in the ideal method presented above. Copying current and possible BLS procedures, the RDB data is also initially rounded to the one, two, and three decimal places, and inflation rates calculated. The resulting inflation series is then rounded to the tenths place in percentage terms. The only difference between the benchmark and rounded series is in the precision of the first stage of rounding.

Table 2 reports the percentage of the sample for which the inflation rates in the rounded data differ from the benchmark series at 0.1 precision. Results are presented for both the non-seasonally-adjusted series and the seasonally adjusted series (SA). Since the two series are similar but the rounding errors should be independent between the two series, the differences in the percentage estimates give an indication the variability of the estimated percentage.

Table 2 shows that, following the current practice of rounding the CPI index to the tenths place, the derived monthly inflation is materially different from the benchmark inflation rate roughly 25% of the time. This result is basically consistent across the various series, with a few exceptions. The relatively low percentage of differences in the medical and other goods and services indexes are due to the fact that those sectors have seen high inflation over the period 1986-2005, and have large index values for most of the period. Conversely, information technology and personal computers have decreased in price dramatically over the period, and so have very small index values, making the first-stage rounding error large enough to change the monthly inflation rate 75% of the time.

Turn attention to the columns corresponding to retaining two and three decimals in the CPI. By reporting the CPI rounded instead to three decimal places, the frequency of discrepancies between the inflation series can be reduced to nearly zero for most series, save the problematic personal computers series.

If the inflation series created from CPI data rounded to the tenths place differs from the

Table 2: Percentage Inaccurate Monthly CPI Inflation Values by Precision

	Not Seasonally Adjusted			Seasonally Adjusted		
Index	One	Two	Three	One	Two	Three
All Items	26%	1.7%	0.4%	24%	0.9%	0.0%
Food	23%	1.7%	0.0%	17%	0.9%	0.0%
Energy	32%	4.3%	0.4%	30%	2.1%	0.4%
All Items Less Food and Energy	25%	1.3%	0.4%	16%	3.4%	0.9%
Apparel	27%	1.7%	0.9%	26%	1.7%	0.0%
Education and Communication	35%	3.3%	0.0%	37%	2.0%	0.7%
Food and Beverages	21%	1.7%	0.0%	16%	3.0%	0.0%
Other Goods and Services	13%	1.3%	0.0%	13%	2.1%	0.4%
Housing	26%	3.0%	0.4%	19%	0.9%	0.0%
Medical	13%	0.9%	0.0%	12%	2.1%	0.0%
Recreation	23%	3.3%	0.7%	24%	3.3%	1.3%
Transportation	26%	2.1%	0.4%	24%	2.1%	0.0%
Information Tech	57%	10.3%	0.0%	-	-	-
Personal Computers	75%	18.7%	2.2%	-	-	-

benchmark series roughly 25% of the time, by how much is it off? Fortunately, the rounded data is precise enough that the difference is always limited to +/- 0.1% from 1986 to the present. In recent times, however, monthly inflation rates have been around 0.2%, which makes the rounding error as a percentage of the actual monthly change quite large indeed. Table 3 summarizes the distribution of the magnitude of the rounding errors relative to the unrounded inflation rate for the all-items index.

Reading the first column of Table 3, one sees that of the 234 total observations of the rounded CPI all-items inflation, 19 of them (8.1%) are in error by between 25% and 50% of the magnitude of the unrounded monthly change. Summing down the columns, one can see that 62 observations (25.5%) differ by more than 5% of the benchmark inflation rate. 19% of the time, the reported CPI inflation value is differs from the benchmark inflation rate by 25% or more. More than 6% of the time, inflation derived from the CPI rounded to one decimal place is off by 100% or more.

Reading across Table 3, one can see that raising the initial level of rounding to the hundredths place eliminates all of the gigantic relative errors. Reporting the index rounded to the thousandths place would reduce the frequency of discrepancies to under 1%, and the magnitude of the error would be greatly reduced.

An alternative measure of the importance of rounding error for CPI inflation is to compare rounding error variance with the intrinsic sampling error variance. Sampling error arises because the BLS is unable to collect all prices on all goods in the market and instead takes a sample of these prices. The BLS takes great pains to assure that the sample of prices

Table 3: Density of Relative Errors in CPI All Items Inflation

Relative Errors	Error Count (Percentage): One Digit		Two Digits		Three Digits	
0% - 5%	172	(73.5%)	230	(98.3%)	233	(99.6%)
5% - 15%	1	(0.4%)	0	(0.0%)	0	(0.0%)
15% - 25%	11	(4.7%)	3	(1.3%)	1	(0.4%)
25% - 50%	19	(8.1%)	1	(0.4%)	0	(0.0%)
50% - 100%	16	(6.8%)	0	(0.0%)	0	(0.0%)
100% - 200%	7	(3.0%)	0	(0.0%)	0	(0.0%)
≥200%	8	(3.4%)	0	(0.0%)	0	(0.0%)

Sample size: 234 observations.

collected reflects the universe of possible prices, and reports an estimate of error variance due to this sampling procedure. Currently, the monthly sampling error variance of all-items CPI inflation is around 0.0036%.

The sampling error variance estimates were created using unrounded figures, so adding rounding error to the CPI increases the variance of the reported inflation series relative to an unrounded series. Table 4 demonstrates that error variance due to rounding accounts for approximately 72% of the reported inflation series' total error variance. Reducing the rounding error variance would reduce the total error variance by 42%.

Table 4: Decomposition of Error Variance of CPI Inflation

	Decimals Reported: One Digit	Two Digits	Three Digits
Sampling Error Variance	0.0036%	0.0036%	0.0036%
Rounding Error Variance	0.0026%	0.0002%	0.0000%
Total Error Variance	0.0062%	0.0038%	0.0036%

4 Mathematical Analysis

To get a better feel for how rounding error affects the historical inflation record further back into the past, one would like to undertake the same experiments as above, comparing the percent changes in unrounded figures to their rounded counterparts. Unfortunately, the BLS doesn't produce a full-precision historical series. Instead, I turn to a mathematical analysis, which is in some sense approximate, but provides additional insight that the data alone could not supply. The analysis which follows parallels the steps the BLS takes in producing the CPI inflation figures as summarized above.

Given a rounded CPI index value, one knows a range in which the true (unrounded) CPI value must lie, but doesn't know the value of the digits which have been rounded away. For instance, if the reported CPI level is 145.2, the true value can lie anywhere between 145.15 and 145.25 with equal likelihood. The true and rounded levels can differ by one half of the precision in either way,

$$CPI_t = CPI_t^* + \epsilon_t,$$
$$\epsilon_t \sim \mathrm{U}\left(-\delta/2,\ \delta/2\right),$$

where CPI_t is the rounded CPI level in month t, CPI_t^* is the unrounded value, and the ϵ's are independent, uniformly-distributed random variables which take values between plus or minus one-half the first-stage rounding precision, δ.

After rounding the CPI levels, the difference between two adjacent month's values are calculated:

$$\begin{aligned}\Delta CPI_t &\equiv CPI_t - CPI_{t-1}\\ &= CPI_t^* + \epsilon_t - CPI_{t-1}^* - \epsilon_{t-1}\\ &\equiv \Delta CPI_t^* + \Delta\epsilon_t,\end{aligned}$$

where ΔCPI_t^* is defined as the difference between the two unrounded values, and $\Delta\epsilon_t$ is defined as the difference between the two errors.

Next, the percent change is calculated by dividing through last period's CPI value,

$$\frac{\Delta CPI_t}{CPI_{t-1}} = \frac{\Delta CPI_t^*}{CPI_{t-1}} + \frac{\Delta\epsilon_t}{CPI_{t-1}},$$

and then the resulting percent change is rounded again, this time at the final precision level, α, which is 0.1% in our example:

$$\frac{\Delta CPI_t}{CPI_{t-1}} = \frac{\Delta CPI_t^*}{CPI_{t-1}} + \frac{\Delta\epsilon_t}{CPI_{t-1}} + \nu_t, \tag{1}$$
$$\nu_t \sim \mathrm{U}\left(-\alpha/2,\ \alpha/2\right).$$

Equation 1 shows that the reported CPI inflation figure is the sum of three terms: the true CPI inflation figure,[3] plus the first-stage rounding error scaled by the CPI, plus the second-stage rounding error. The two error terms are qualitatively different. As the level of the CPI increases, the first-stage rounding error matters less and less, $\Delta\epsilon_t/CPI_{t-1}$ gets smaller and smaller, while ν_t stays of the same magnitude. Conversely, as is made obvious in the plots of inflation above, the first-stage rounding term increases in size as the value of the CPI shrinks moving backward in time. Alternatively, if the BLS increased rounding precision in the reported CPI, $\Delta\epsilon_t$ would be smaller and smaller, leaving only the final rounding error difference between the true and rounded inflation values.

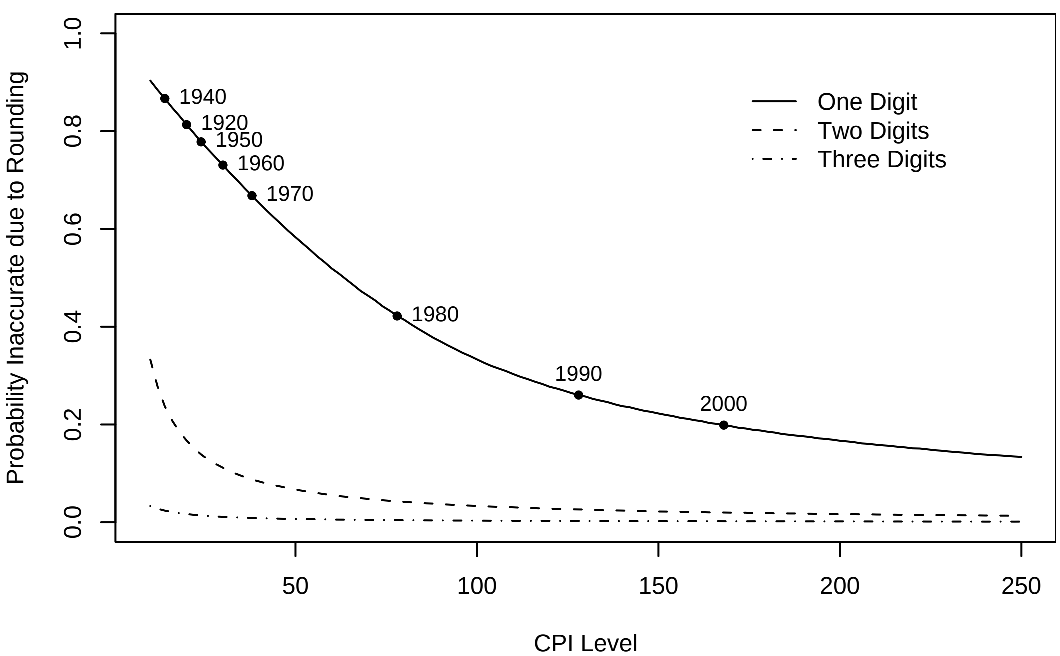
[3]This term is not really the true inflation rate, $\frac{\Delta CPI_t^*}{CPI_{t-1}^*}$, but the difference between this term and the true inflation rate is negligible compared to the two rounding errors. The analysis here takes $\frac{\Delta CPI_t^*}{CPI_{t-1}} = \frac{\Delta CPI_t^*}{CPI_{t-1}^*+\epsilon_t}$ as the true inflation for simplicity.

Now we are prepared to answer the question analytically which was previously posed to the data, how frequently does the total difference between the reported percent change and the true percent change cause the reported inflation rate to differ? That is, we wish to know

$$\text{Prob}\left(\left|\frac{\Delta\epsilon_t}{CPI_{t-1}} + \nu_t\right| > \frac{\alpha}{2}\right). \tag{2}$$

For a given first-stage precision, δ, and a given desired final precision, α, and at a given CPI level, equation 2 can be evaluated by computer simulation as follows. Draw two uniform random numbers from the interval $(-\delta/2,\ \delta/2)$, subtract one from the other, and divide the difference by the CPI value. Add to this difference a third uniform number drawn from the interval $(-\alpha/2,\ \alpha/2)$, and compare the result with $\alpha/2$. If this procedure is repeated many times, the average number of times that the absolute value of this sum exceeds $\alpha/2$ will be equal to the probability of a first-stage rounding error resulting in an wrong inflation report. The results of repeating this simulation, with 1,000,000 repetitions per CPI level, are presented in Figure 4.

Figure 4: Probability of an Inaccurate Percent Change due to Rounding



Note: Plotted points mark the CPI values in the 1984 base-year series for January of selected years.

One can see from the graph that extending the number of digits precision at which the CPI is reported will go a long way toward reducing the probability of rounding errors affecting the final result. For example, for an index value of 100, there is a 33% chance of an different figure when the series is rounded to one decimal place. The probability drops to 3.3% for two digits, and 0.36% for three digits. For an index value of 50, the chances are 58%, 6.7%, and 0.64% respectively. For reference the January CPI values for select years are plotted on the graph. One can see that the reported series should differ from the benchmark series more than 60% of the time prior to 1970.

Note from equation 2, however, that the only way to guarantee a CPI inflation series which is entirely free of rounding error is to not round the CPI levels at all before taking the percent change. Not rounding at all corresponds to $\Delta\epsilon = 0$. Since the second-stage rounding error, ν_t, is between $-\alpha/2$ and $\alpha/2$, the probability that its absolute value exceeds $\alpha/2$ is exactly zero. Conversely, *any* first stage rounding makes this probability greater than zero. Intuitively, when the second-stage rounding error is very close to being as large as it can be, $\alpha/2$, even a tiny first-stage error can push it over the edge, and the inflation figure will round the wrong way.

As with the real-data experiment above, we can also ask how big the rounding errors are relative to monthly inflation rates for the historical series. That is, we can estimate the distribution of

$$r_t^* = \frac{\left|\frac{\Delta\epsilon_t}{CPI_{t-1}} + \nu_t\right|}{\frac{\Delta CPI_t^*}{CPI_{t-1}}},$$

the magnitude of the rounding error relative to the benchmark inflation rate. The numerator can be calculated as in the simulation above, while the denominator is the benchmark inflation rate for the month in question. We can then ask what the relative percent error is as a fraction of the benchmark, or can look at how many errors exceed a given size relative to the true inflation rate.

Unfortunately we don't know the unrounded inflation rate for the denominator in r_t^*, and the reported inflation rate is unsuitable, which is indeed part of the motivation for this paper in the first place. Because the rounded series contains a (misleadingly) large fraction of months with no change, using the reported inflation series in the denominator results in dividing by zero in many months, exaggerating the relative size of the errors for those months.

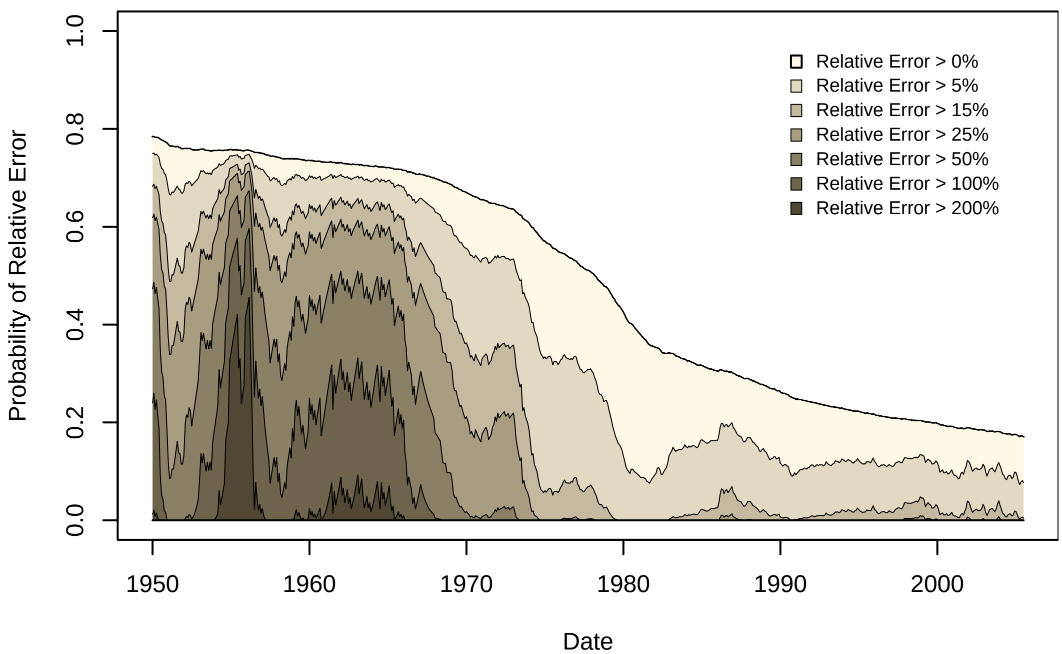
For the purpose of showing overall trends in rounding error, I use a smoothed version of the inflation series in place of the "true" inflation rate in the denominator of r_t^*.[4] For each month, one million samples are taken from the rounding error distribution corresponding to that month's CPI level as in the simulation above. Each of these million simulated errors is divided by that month's smoothed inflation value. The resulting simulated sample of a million values of r_t should approximate the true distribution of the relative errors very closely.

Figure 5 shows the frequencies of different values of the relative error on one plot. The

[4]Specifically, I use an exponentially-weighted monthly moving average of the reported CPI inflation rate with $\lambda = 0.05$ which behaves similarly to a 3-year moving average of inflation. Using the annual inflation rate in the denominator still shows significant rounding effects in the early periods, but gives otherwise similar results to the exponentially-weighted average.

topmost line represents the probability that $r_t > 0$, or the probability of any discrepancy at all. Consequently, the 0% line corresponds to the simulated results presented in figure 4. For comparison, at the 1990 CPI value of 130, Figure 4 shows that using the rounded CPI to calculate inflation differs from using an unrounded CPI 25% of the time. In Figure 5, for 1990, the chance of a relative error greater than zero percent in magnitude is approximately 25%.

Figure 5: Probability of Relative Errors by Size



The other lines in the figure the probabilities of errors of various sizes relative to the underlying smoothed inflation rate. For instance, in 1990, we see that errors larger than 5% of the true inflation rate occur roughly 15% of the time, and errors larger than 15% of the inflation rate occur very seldom – just around 2% of the time. In the past, particularly prior to 1975, the errors are not only much more frequent, but the magnitude of the errors relative to the inflation rate are much larger. For instance, through the early 1960's, the reported inflation rate differs from the benchmark roughly 70% of the time, with errors as large as 50% of the inflation rate occurring around 45% of the time. Startlingly, errors larger than the actual inflation rate occur around 20% of the time in the period from 1950 to 1968.

Two opposing tensions underlie the probability distribution of the relative size of the discrepancies due to rounding. On one hand, in modern times the value of the CPI is relatively large, so rounding error should be small as a fraction of the CPI level, and consequently differences should be infrequent. At the same time, the rate of inflation decreased through the 1990's so rounding errors, when they do occur, would be expected to be larger relative to an underlying low rate of inflation. Figure 5 shows how the two forces interact. In 1983, the probability of any error was around 35%, while the probability of errors larger than 5% of the inflation rate was around 10%. Because the CPI level increased over the next two decades, the chance of any error decreased to around 20%. But because the CPI level was increasing slowly, the chance of errors greater than 5% of the inflation rate remained virtually constant.

Going backwards in time, one can see that the two forces act in concert. The smaller CPI values pre-1970 lead to a larger probability of discrepancy, at 75%, and the periods of relatively low inflation in the 1950's through 1960's lead to errors which are large relative to the the actual inflation rate. Most errors are larger than 5% of the inflation rate at this time, with errors larger than the inflation rate itself occurring roughly 20% of the time.

The remaining time period, between 1970 and the early 1980's, was characterized by two periods of high inflation. Consequently, we see the level of the CPI rising, and the probability of a rounding error decreasing. At the same time, high monthly inflation rates make the size of errors smaller as a fraction of the inflation rate. One can see the mid-1970's inflation driving the probability of a 25% or larger relative error down from around 20% to nearly 0% over the course of two years. Although the period was horrible for the value of a dollar, it was fantastic for the accuracy of reported inflation statistics.

5 Conclusions and Remarks

This paper demonstrates how rounding a series before calculating its percent changes introduces an additional source of statistical error. Using an unrounded CPI dataset, we find that a representative inflation series differs from an unrounded benchmark approximately a quarter of the time, and that the differences can be large relative to the true underlying inflation rate.

Mathematical analysis and some simulation results demonstrate in more detail how the rounding-induced errors behave with both the the level of the CPI and with the inflation rate over time. Three regimes emerge. Before 1970, both the frequency of errors and their magnitude are large. The mid 1970's and early 1980's inflations cut the probability of a rounding error in half, and lead to the relative errors moderation. In the present period, a high CPI value makes the inflation rate accurate around 75% of the time, but the low underlying inflation rate has kept the probability of errors of a given relative size roughly constant and comparatively moderate. This certainly has implications for inflation research over these earlier periods.

Finally, the take-home message from the real-data experiments in section 3 show that increasing the precision of reported CPI levels would go a long way toward making the errors which arise from rounding error negligible. I show that publishing the CPI to three decimal places, for instance, will reduce the error variance of the CPI inflation series by 42% and will reduce the chance of disagreement between an unrounded index and the reported index from its current 25% to under 0.5%.