Charles W. Calomiris

Charles W. Calomiris is an associate professor at the University of Illinois, Urbana-Champaign, and is a co-director for the university's Office for Banking Research. He is also a faculty research fellow, National Bureau of Economic Research, and a visiting scholar at the Federal Reserve Bank of St. Louis.

Commentary

IN THEIR PAPER, Anderson and Kavajecz provide the rare public service of a careful examination of the construction of monetary data. The paper is important because the data on monetary aggregates are central to academic and policy research in macroeconomics. I expect that by the metric of the percentage of monetary economists who will have this paper in their file cabinets 10 years from now, this will be one of the most successful works in monetary economics. Data are forever. Like the Federal Reserve Board's *Banking and Monetary Statistics* and *All Bank Statistics*, Friedman and Schwartz's *Monetary Statistics of the United States*, Capie and Wood's *Monetary Statistics of the United Kingdom*, and Eisner's *How Real is the Federal Deficit?*, Anderson and Kavajecz pose and answer descriptive questions of lasting interest to macroeconomists. Judging from the paucity of this kind of work, its importance seems to be underestimated.

I found little to quibble with in the way the authors organized their description. In my discussion I will focus on the question of why researchers doing empirical monetary economics should care about the details of how monetary aggregates are measured and how those measurements have changed over time.

Five sets of issues seem central in motivating the potential usefulness of this exercise. First, and most obviously, any attempt to construct monetary aggregates for long stretches of time must do its best to ensure comparability of measures. This means coming to grips not only with financial innovations that affect the range of definitions of money, but also with changes in sampling procedures, seasonal adjustment, and other choices made by the data constructors.

Second, the Fed's procedure of revising data retrospectively to maintain consistent definitions and seasonal adjustment factors—which sometimes have produced large retrospective revisions of the aggregates—makes it difficult to compare empirical research of different vintages. For example, two studies of M1 money demand over the same period, performed at different dates, may differ not only because of specification, but because of the vintage of data used in each. It would be worthwhile to ask how much of the differences across studies of money demand can be attributed to retrospective revision of data, as opposed to the incorporation of additional periods of data, or differences in specification.

Third, there should be an objective outside evaluation of the Federal Reserve Board's choices of definitions of money and methods of seasonal and benchmark adjustment. Prior to this study, this was not feasible because relatively little was known about the Fed's procedures. Anderson and Kavajecz suggest that the Fed's decisions regarding what to include in the monetary aggregates are often influenced by whether adding a new component helps to stabilize the relationship between money and economic activity. While this procedure may make sense, in general, given the Fed's desire to use monetary aggregates as targets, it would be interesting to describe clearly how the Fed decides (and how it should decide) that the improvement in stability will persist (that is, it reflects a lasting behavioral change rather than a temporary statistical coincidence). How long should the Fed wait before incorporating new (apparently stabilizing) elements of money into its definitions of aggregates? Would an increased emphasis on Divisia indices be warranted in light of the difficulties posed by having to make an all-or-nothing decision about whether to include financial assets in one or more of the aggregates? Regarding seasonal adjustment, it would be interesting to consider how the Fed should determine when a change in seasonal factors has occurred—and how far back retrospective changes in seasonality should be made. What is the optimal choice of the period over which deterministic seasonals should be estimated? How much less relevant is distant information for estimating seasonal compared to recent information? Anderson and Kavajecz have provided researchers interested in these questions with a wealth of detail that will allow them to construct counterfactual rules for defining monetary aggregates, and to compare these with those adopted by the Fed.

Fourth, if the Fed attempts to keep monetary aggregates "on track" relative to economic activity (by altering definitions and adjustment factors), then this makes the reported aggregates unsuitable for performing hypothesis tests about the stability of money demand. Researchers interested in whether money demand is stable, therefore, should perform sensitivity analysis to examine whether reasonable counterfactual definitions of the aggregates lead to different conclusions about the stability of money demand.

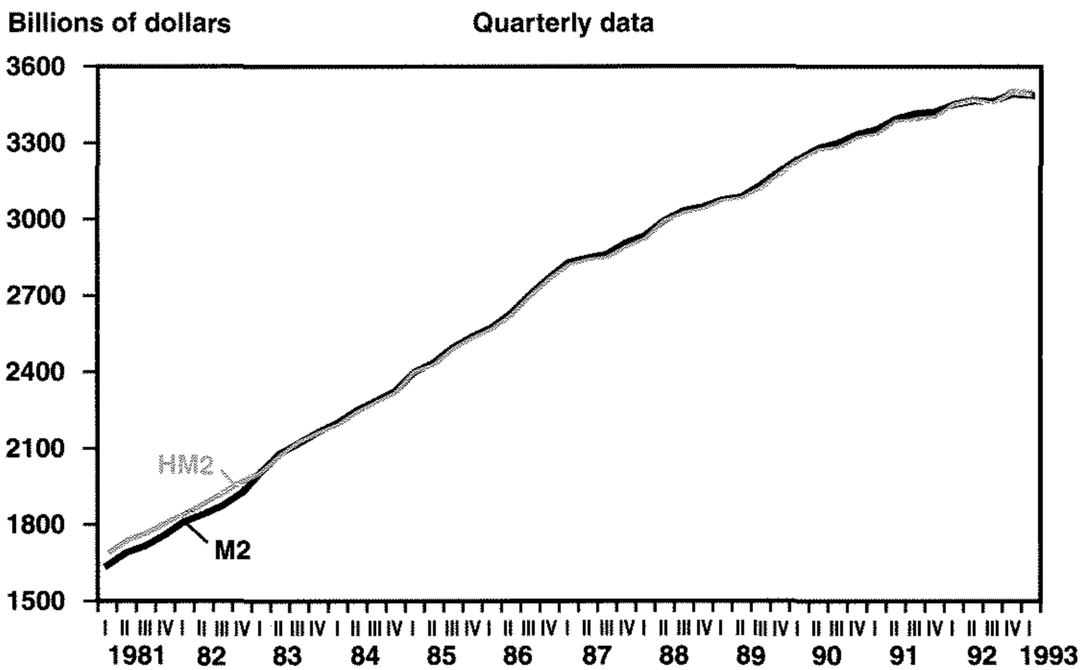
Finally, there is a problem I will label the "expectations error effect." The essence of this problem is that (unforecastable) errors in measurement, which affect expectations of agents in "real time," may weaken the apparent connection between money and output using ex post (corrected) monetary data. Assume for the sake of argument (unrealistically) that the current retrospective data on the monetary aggregates are "correct." That is, assume that all definitions, seasonal adjustments and revisions that have been made so far are perfectly accurate, and that no further revisions will be made in the series. Furthermore, assume that we can agree upon an econometric procedure for measuring the closeness of the relationship between money and output (using, for example, a "structural VAR" model of money, output and other variables).

Even under these ideal circumstances, the measured relationship between "true" money (measured ex post) and economic activity will be biased toward zero if money is initially measured with error. The reason is that "true" money, as well as errors in measuring aggregate money, will elicit responses that affect economic activity and money subsequently. Money and output are linked through "fundamental" structural links and through "expectational" effects. For example, an increase in an individual's holdings of money may lead him to rebalance his portfolio (putting pressure on interest rates to fall and output to rise in the standard IS-LM model). Second, estimates of monetary aggregates (which include initial measurement error) will also be taken into account by the public in economic activity if aggregates are used as economic indicators.

In the absence of measurement error, the individual agent can observe without error not only his own money, but also the aggregate. In the presence of measurement error, the aggregate is observed with error, and these errors will elicit real responses from agents. So, both announced and true money will be linked to output. Neither will be as strongly linked to output as true money in the absence of measurement error, and empirical analysis using ex post data (after removing errors) may underestimate the link between money and output.

Thus, temporary inaccuracies in monetary aggregate estimation (which elicit real responses) might explain weak correlations between money and output using ex post (accurate) data. How can one come to grips with this problem empirically to decide whether bias arising from "expectations error effects" is important for conclusions about the role of money in the economy? One simple first step is to compare various measures of monetary aggregates (and monetary growth rates) for a given period reported at different dates. If the differences among these measures are small, then the problem of potential bias is of little practical importance. If the differences

Figure 1
M2 Against HM2



are large, then one would have to take on the much harder job of measuring the extent of the bias by gauging the reaction of the public to measurement errors.

As Anderson and Kavajecz point out, there are several types of potential error, and each involves a different correction horizon. First-published numbers (which appear one or two weeks after the fact) are updated within a month or so because of the arrival of new data. They are changed (roughly) annually to adjust for changes in benchmarks and seasonals, and change with new definitions of the aggregates as well.

As an illustrative exercise, I chose the easiest case—M2 from January 1981 to January 1993. I chose this period because, as Anderson and Kavajecz show, there was no important change in the definition of M2 during this period, so that one can focus on the role of revisions from new data, benchmark changes, and changes in seasonal factors as sources of error. I constructed measures for this period using three different timings of measurement. I used the first date of publication of monthly M2 in the *H.6* statistical release as my definition of the initial measure of M2. This was released roughly two weeks after the end of the month. My second date of measurement is the M2 figure reported in the *Federal Reserve Bulletin*, which appears with a two-month lag. My third measure is the retrospective series as of January 1993.[1]

Using seasonally adjusted data from these sources for January, April, July and October, I constructed measures of the level of M2 and of

[1]Table 3 in Anderson and Kavajecz decomposes revisions in money into three different adjustments and expresses them in absolute terms. This is interesting for some purposes, but not for my purpose. I am interested in whether errors coming from all sources are potentially large relative to the actual number.

Figure 2
DM2 Against DHM2

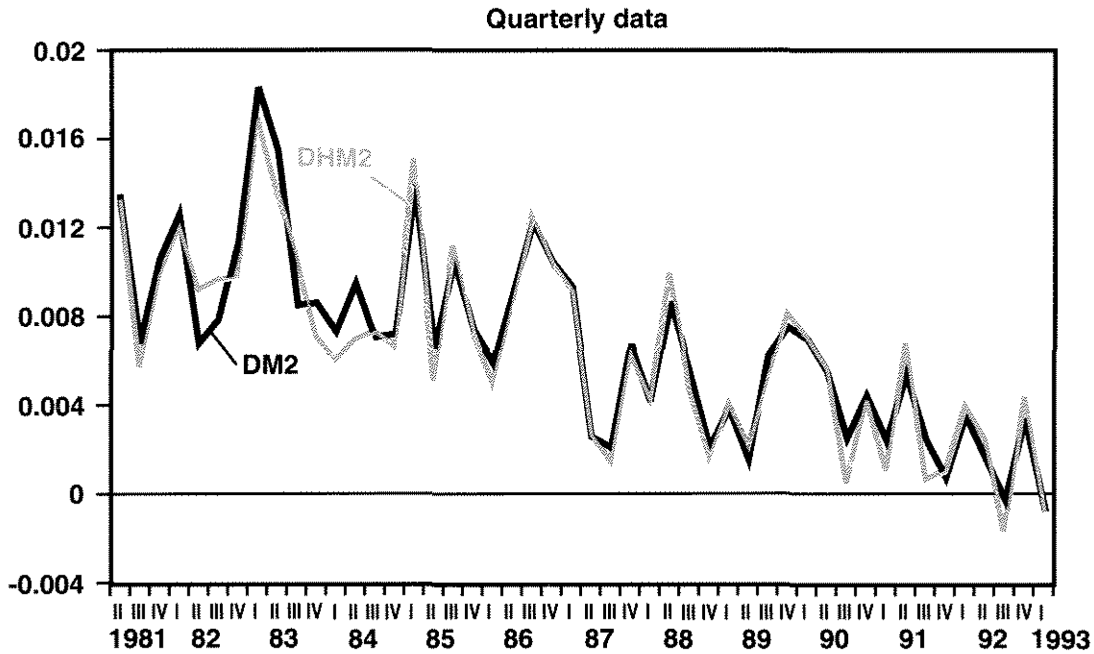
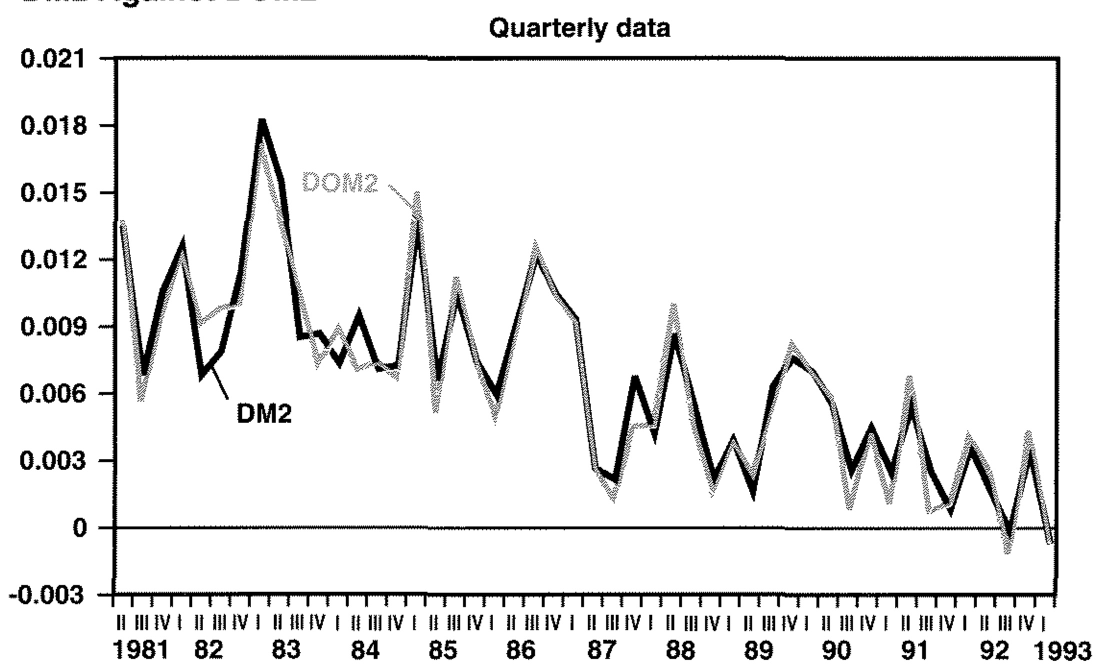


Figure 3
DM2 Against DOM2



that quarter's growth in M2 measured *at the time* the M2 number was reported. For example, M2 growth for the first quarter of 1982 according to the *H.6* release is the log difference between the first *H.6* number for M2 in April and the January 1982 number reported in that same release. Figures 1-3 compare these definitions of money and money growth using these three measurement horizons. The level and growth data from the 1993 series are labeled M2 and DM2; the data from *H.6* are labeled HM2 and DHM2; and the data from the *Bulletin* are labeled OM2 and DOM2.

These figures indicate that revision of M2 has been trivial in the 1980s, and so I conclude that for these series over this period, "expectations error effects" were not important. To the extent revisions did matter, long-term retrospective changes (the difference between M2 and O2, or DM2 and DOM2) are more important than those occurring within two months of initial publication.

One conclusion to draw from these findings is that, if there has been a breakdown in the relationship between M2 and output during the past decade, it cannot be attributed to temporary mismeasurement of money. Whether similar conclusions would be reached for M1 in the 1980s, or for these and other aggregates during other periods, remains an open question. As Anderson and Kavajecz note, Depository Institution Deregulation and Monetary Control Act (DIDMCA) improved the accuracy of monetary statistics in the 1980s, and M2 tends to be a smoother series than M1. Thus, my results may understate the importance of measurement error for other aggregates and earlier periods.

REFERENCES

Board of Governors of the Federal Reserve System. *All-Bank Statistics 1986–1955.* Board of Governors of the Federal Reserve System, 1959.

________. *Banking and Monetary Statistics* 1914-1941. Board of Governors of the Federal Reserve System, 1976.

________. *Banking and Monetary Statistics 1941-1970.* Board of Governors of the Federal Reserve System, 1976.

________. *Federal Reserve Bulletin.*

________. *Money Stock, Liquid Assets, and Debt Measures,* supplement to the Federal Reserve statistical release, *H.6.*

Capie, Forrest, and Alan Webber. *Monetary History of the United Kingdom, 1870–1982.* Allen & Unwin, 1985.

Eisner, Robert. *How Real is the Federal Deficit?* Free Press, 1986.

Friedman, Milton, and Anna J. Schwartz. *Monetary Statistics of the United States.* Columbia University Press, 1970.