Cowles Foundation Discussion Papers   Cowles Foundation

2-15-2019

Variable Mismeasurement in a Class of DSGE Models: Comment

Ray C. Fair
Yale University

Recommended Citation

Fair, Ray C., "Variable Mismeasurement in a Class of DSGE Models: Comment" (2019). *Cowles Foundation Discussion Papers*. 210.
https://elischolar.library.yale.edu/cowles-discussion-paper-series/210

VARIABLE MISMEASUREMENT IN A CLASS OF DSGE MODELS: COMMENT

By

Ray C. Fair

February 2019

Revised July 2020

Variable Mismeasurement in a Class of DSGE Models: Comment

Ray C. Fair[*]

July 2020

Abstract

This comment points out mismeasurement of variables in the DSGE model in Smets and Wouters (2007) and in models that follow the Smets-Wouters measurement procedures. The mismeasurement errors appear to be large.

1   Introduction

The DSGE model in Smets and Wouters (2007) (SW) has had an important influence on macroeconomic research. In particular, the variable measurements introduced in SW have been used in much of the research that followed. The following nine studies are in this category: Edge and Gürkaynak (2010), Kolasa, Rubaszek, and Skrzypczyński (2012), Wolters (2013), Del Negro, Giannoni, and Schorfheide (2015), Kaplan, Moll, and Violante (2018), Wolters (2018), Anzoategui, Comin, Gertler, and Martinez (2019), Beraja, Hurst, and Ospina (2019), and Auclert, Rognlie, and Straub (2020).

---

[*]Cowles Foundation, Department of Economics, Yale University, New Haven, CT 06520-8281. Phone: 203-980-0646; e-mail: ray.fair@yale.edu; website: *fairmodel.econ.yale.edu*.

There are seven observable variables in the SW model: consumption, investment, output, hours, inflation, real wage, and interest rate. This comment points out that three of these variables—consumption, investment, and hours—are mismeasured.

2 Mismeasurement

Real consumption is measured in SW as nominal consumption divided by the GDP deflator, and real investment is measured as nominal investment divided by the GDP deflator.[1] However, nominal consumption should be divided by the consumption deflator, and nominal investment should be divided by the investment deflator.

Let C denote nominal consumption divided by the consumption deflator and CZ denote nominal consumption divided by the GDP deflator. Figure 1 plots the ratio of CZ to C for the 1966:1–2018:3 period.[2] As can be seen, there is considerable variation in this ratio. It ranges from 0.947 in 1972:4 to 1.006 in 2008:3. The variation is particularly large in the 2007–2009 period. The mean of the ratio is 0.983 with standard deviation 0.016.

Let I denote nominal investment divided by the investment deflator and IZ denote nominal investment divided by the GDP deflator. Figure 2 plots the ratio of IZ to I for the 1966:1–2018:3 period. There is also considerable variation in

---

[1]These are quarterly variables from the national income and product accounts. Consumption is total personal consumption expenditures, and investment is total fixed private investment. The latter includes both nonresidential and residential investment, but not inventory investment.

[2]The data for Figures 1 and 2 were taken from the Bureau of Economic Analysis website on February 6, 2019. They are from Tables 1.1.5 and 1.1.9. They were collected from 1948:1 through 2018:3, although Figures 1 and 2 begin in 1966:1, which is the first quarter of the SW model's estimation period. The estimation period for the SW model ends before 2018:3, but data through 2018:3 have been presented in the figures for completeness. None of the conclusions in this comment depend on the last few observations.

this ratio, with a large downward trend. It ranges from 0.979 in 2017:4 to 1.425 in 1980:1. The mean is 1.208 with standard deviation 0.146.

The hours variable in the SW model is measured as average weekly hours of all persons in the nonfarm business sector (H) times total civilian employment (E). The first is from the Bureau of Labor Statistics (BLS) establishment survey and the second from the household survey.[3] The total number of hours in the economy (LZ) is then taken to be the product of the two: $LZ = H \times E$. There are two problems with this measurement. The first is that weekly hours excludes farm workers and government workers (both federal and state and local). The implicit assumption is that average weekly hours for farm workers and government workers is the same as for private nonfarm workers, which is not true. The second is that civilian employment measures the number of people employed, not the number of jobs. Some people have two jobs and so are counted twice in the establishment survey but only once in the household survey. I will call the difference between the total number of jobs and the total number of people employed the number of "moonlighters," although there are a few other differences between the two surveys.

One can get from the BLS quarterly data on the number of hours in the total economy and various subsectors.[4] Ideally the hours should include all workers in the economy, including military workers, since the services of military workers are in GDP. However, LZ above does not include military workers, and so for comparability I have subtracted military hours, which are available on the BLS

---

[3]In the BLS notation H is CES0500000007 and E is LNS12000000. The data are monthly seasonally adjusted. Quarterly variables are constructed by summing the relevant three months and dividing by 3.

[4]The BLS site is *https://www.bls.gov/lpc/tables.htm*. Click the XLSX spreadsheet for "Total U.S. Economy - all workers." The hours are seasonally adjusted in billions.

website, from total hours. Let L denote the total number of hours in the economy less military hours.

The unit of LZ is thousands of hours per week, and the unit of L is billions of hours per year. To make the units comparable, let LZ now denote the old LZ multiplied by 52 and divided by 1,000,000. Figure 3 plots the ratio of this new LZ to L for the 1966:1–2018:3 period. There is again considerable variation. The ratio ranges from 0.951 in 1999:3 to 1.030 in 1976:3. The mean is 0.992 with standard deviation 0.017. One of the reasons for the fluctuations is that the number of moonlighters fluctuates. In my macroeconometric model—Fair (2018)—I have an equation explaining the number of moonlighters, where the number depends in part on the state of the economy—there are more moonlighters in tight labor markets. The ratio in Figure 3 is low in the late 1990s, which in part reflects the fact that the number of moonlighters was large because of the booming economy (so L is large relative to LZ).

One last issue concerns population, which is used in the SW model to put the variables in per capita terms. Monthly population data are available from the BLS, which are converted to quarterly data by summing the three relevant monthly values and dividing by 3.[5] The problem is that these data are revised (rebenchmarked) each January, and the revisions are not carried back. There are thus spikes, either positive or negative, each January, or for the quarterly variable each first quarter. Figure 4 shows the percentage change in quarterly population for the 1994:1–2018:4 period (at quarterly rates). The spikes in the first quarters are evident. This problem was first pointed out in Edge and Gürkaynak (2010, p. 218). They discovered this problem too late to revise the results in their paper, and

---

[5]The BLS notation for civilian noninstitutional population ages 16 and over is LFU800000000 before January 1976 and LNS10000000 from January 1976 on.

others have also not adjusted for this.[6] The variables that are divided by population in the SW model to put them in per capita terms are real output, real consumption, real investment, and hours. Aside from the mismeasurement problems discussed above, the per capita variables are more variable than they should be.

3 The Papers

Smets and Wouters (2007), Edge and Gürkaynak (2010), Kolasa, Rubaszek, and Skrzypczyński (2012), Wolters (2013), and Del Negro, Giannoni, and Schorfheide (2015) all use the basic SW setup and measurement procedures and so suffer from the above problems. One exception, as noted in footnote 6, is that Del Negro, Giannoni, and Schorfheide (2015) now use the HP filter to adjust for the population spikes. Another exception is that Kolasa, Rubaszek, and Skrzypczyński (2012) do not use population data.

Wolters (2018) uses the SW measurements except for hours. For hours he uses the correct data except that he divides by the population data with the spikes. He shows that using better hours data has a large effect on the estimated output gap in DSGE models.

Kaplan, Moll, and Violante (2018) in thier HANK model use the SW model for comparison purposes, which means that the comparisons may be off, but their model does not use aggregate data on consumption, investment, and output and so there is no deflation issue. In their work with the HANK model Auclert, Rognlie, and Straub (2020) deflate both nominal consumption and nominal investment by

---

[6]In private correspondence Marco del Negro has informed me that the population data are now smoothed using the HP filter in the New York Fed DSGE model. In my macroeconometric model I have always adjusted for this, not using the HP filter, but linearly interpolating the January adjustments back for 40 quarters.

the GDP deflator, and they use the wrong hours variable. Anzoategui, Comin, Gertler, and Martinez (2019) use the SW model as a base and so are subject to the same mismeasurement issues.

Beraja, Hurst, and Ospina (2019) use the population data with spikes. They also deflate service consumption, nondurable consumption, durable consumption, private investment, and nominal GDP by the consumer price index (CPI). This is, of course, wrong for all five variables. The CPI is a particularly bad deflator to use for nominal GDP since the CPI includes import prices and excludes export prices. GDP is domestically produced output.

4 Evaluation

It is discouraging that the influential macro research cited here has not taken more care with the aggregate data. This is more than a decade of research. I first pointed out the SW measurement problems for consumption, investment, and hours in a paper in 2009, which was eventually published as Fair (2012), and, as noted above, Edge and Gürkaynak (2010) pointed out the population problem in their paper. Unfortunately the problems remain.

Are the measurement errors serious? As noted above, Wolters (2018) shows that for hours it is serious regarding the measurement of the output gap in DSGE models. As discussed above, the fluctuations in Figures 1, 2, and 3 appear large. Another way to gauge magnitudes is to consider the ratio of the variables to real GDP. The ratio of C to GDP, where C is the correctly measured variable, has a mean of 0.651 with a standard deviation of 0.0249 over the sample period. For the ratio of CZ to GDP the mean is 0.640 with a standard deviation of 0.0312, 25

percent larger. This is a large difference in the relationship between consumption and GDP. Similar stories hold for investment and hours. The ratio of I to GDP has a mean of 0.144 with a standard deviation of 0.0196. The ratio of IZ to GDP has a mean of 0.172 with a standard deviation of 0.0138, 30 percent smaller. The ratio of L to GDP has a mean of 0.0224 with a standard deviation of 0.00523. The ratio of LZ to GDP has a mean of 0.0223 with a standard deviation of 0.00552, 3.6 percent larger. Note also that some of the fluctuations in Figures 1–3 are correlated with business cycle fluctuations, and some show considerable noise in the 2008-2009 recession, an important period to analyze. .

It is not feasible in a study like this to redo the models using the correct data. There are too many details. Also, it is not clear how more than one price should be handled in DSGE models with Calvo pricing. Using, say, the price deflator for consumption, the price deflator for investment, and the GDP deflator requires non trivial changes to the models.

Say it is the case that the parameters of a model and its properties are not sensitive to the use of the correctly measured variables. What is one to make of this? If the data are quite different, the lack of sensivitity may suggest that the model is not tied to the data much, that it is not really making empirical statements. This could be because of tight theoretical restrictions or a heavy use of calibration. One purpose of this comment is to point out that the data do seem quite different.

Figure 1
Real Consumption Mismeasurement
Ratio of CZ to C
1966:1--2018:3

Figure 2
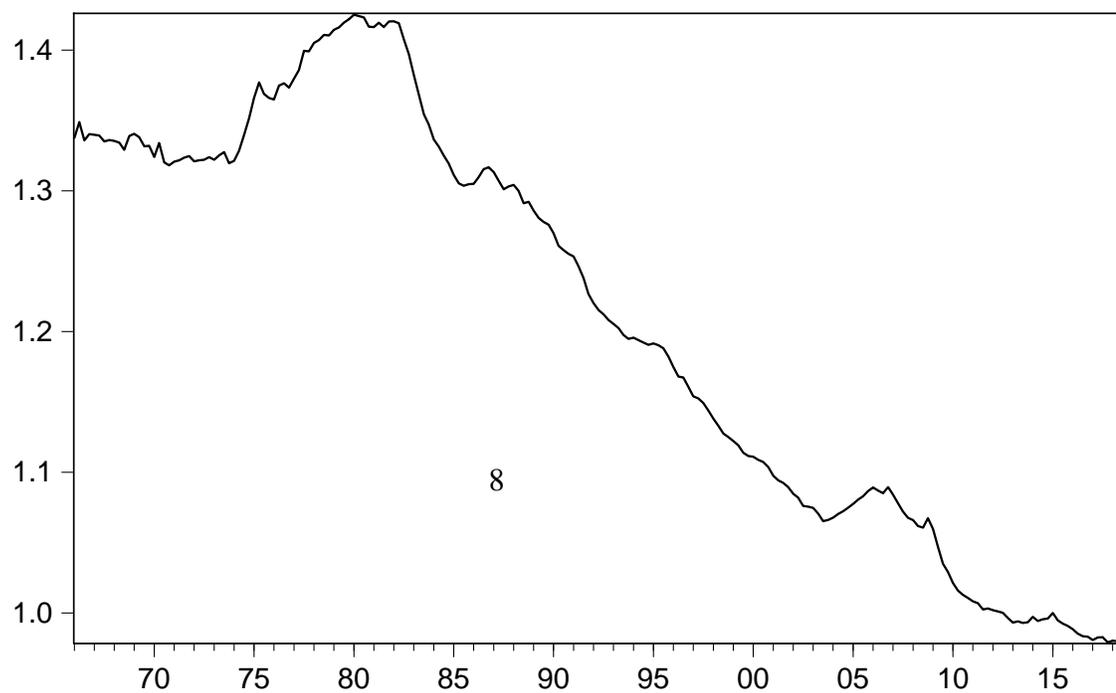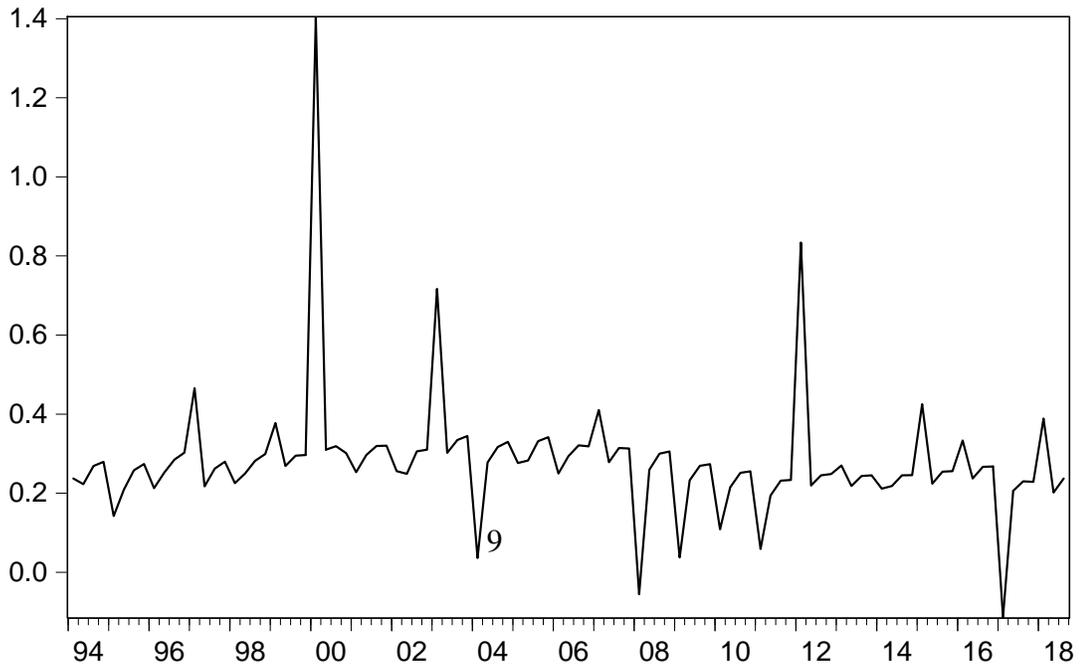## Real Investment Mismeasurement
Ratio of IZ to I
1966:1--2018:3

8

Figure 3
Hours Mismeasurement
Ratio of LZ to L
1966:1--2018:3

Figure 4
Percentage Change in Population, Quarterly Rate
1994:1--2018:3

References

[1] Auclert, Adrien, Matthew Rognlie, and Ludwig Straub, 2020, "Micro Junps, Macro Humps: Monetary Policy and Business Cycles in an Estimated HANK Model," NBER Working Paper No. 26647, March.

[2] Anzoategui, Diego, Diego Comin, Mark Gertler, and Joseba Martinez, 2019, "Endogenous Technology Adoption and R&D as Sources of Business Cycle Persistence," *American Economic Journal: Macroeconomics*, 11, 67–110.

[3] Beraja, Martin, Erik Hurst, and Juan Ospina, 2019, "The Aggregate Implications of Regional Business Cycles," *Econometrica*, 87, 1789-1833.

[4] Del Negro, Marco, Marc P. Giannoni, and Frank Schoreheide, 2015, "Inflation in the Great Recession and New Keynesian Models," *American Economic Journal: Macroeconomics*, 7, 168–196.

[5] Edge, Rochelle M., and Refet S. Gürkaynak, 2010, "How Useful are Estimated DSGE Model Forecasts for Central Bankers?" *Brookings Papers on Economic Activity*, Fall, 209–259.

[6] Fair, Ray C., 2012, "Has Macro Progressed?" *Journal of Macroeconomics*, 2–10.

[7] Fair, Ray C., 2018, *Macroeconometric Modeling: 2018*, *fairmodel.econ.yale.edu/mmm2/mm2018.pdf*.

[8] Kaplan, Greg, Benjamin Moll, and Giovanni L. Violante, 2018, "Monetary Policy According to HANK," *American Economic Review*, 108, 697-743.

[9] Kolasa, Marcin, Michal Rubaszek, and Pawel Skrzypczyński, 2012, "Putting the New Keynesian DSGE Model to a Real-Time Forecasting Test," *Journal of Money, Credit, and Banking*, 44, 1301–1324.

[10] Smets, Frank, and Rafael Wouters, 2007, "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97, 586–606.

[11] Wolters, Maik H., 2013, "Evaluating Point and Density Forecasts of DSGE Models," *Journal of Applied Econometrics*, 30, 74–96.

[12] Wolters, Maik H., 2018, "How the Baby Boomers' Retirement Wave Distorts Model-Based Output Gap Estimates," *Journal of Applied Econometrics*, 33, 680-689.