The Personal Saving Rate: Data Revisions and Forecasts

Dean Croushore*, Pedro Del Monaco Santos

Economics Department, 102 UR Drive, University of Richmond, VA 23059, USA

Abstract

Revisions to the U.S. personal saving rate are very large and may be predictable. We decompose the revisions of the personal saving rate into those caused by revisions to income and those caused by revisions to household outlays. We find that the main source of the revisions to the personal saving rate has shifted over time. We use our findings to explore the forecastability of future revisions of the personal saving rate.

Keywords: data revisions, saving, real-time data
JEL codes: E01, E21

1. Introduction

Data are revised as new source data become available to government data agencies. The revisions make the data more accurate and complete. However, some variables are revised more than others, as Croushore (2011) illustrates. Depending on how the data agency operates, revisions may be characterized as adding news, or reducing noise, or both; see Jacobs and van Norden (2011) for a discussion.

Data revisions to the personal saving rate have not been examined in much detail except by Nakamura and Stark (2005), Nakamura and Stark (2007), and Nakamura (2008). They use the personal saving rate and its revisions in several different contexts, with the focus on forecasting. They document that the personal saving rate is revised dramatically over time and suggest that its use in forecasting other variables is questionable.

In a 2004 speech, Fed vice-chair Roger Ferguson argued that "The fall in the personal saving rate could have important implications for the ability of the country to finance investment in plant and equipment, for future growth in productivity and real incomes, and for our growing economic dependence on other countries to finance our spending patterns." (Roger W. Ferguson (2004)). Further, Jagadeesh Gokhale, Laurence Kotlikoff, and John Sabelhaus

*Corresponding author: Dean Croushore, dcrousho@richmond.edu

examine cohort data on personal saving in the United States from 1960 to 1990 and conclude that "Anemic rates of saving will spell anemic rates of domestic investment, labor productivity growth, and real wage growth. This is the legacy of the uncontrolled intergenerational redistribution from young savers to old spenders that has been fueling ever-higher rates of consumption in the United States" (Gokhale et al. (1996), p. 382).

However, the personal saving rate might not be as low as its initial releases suggest. Nakamura and Stark (2007) highlight the revisions to the personal saving rate and show that the revisions are mainly a function of upward revisions to personal income, so the personal saving rate is generally much higher than is first reported.

This unreliability of data on the personal saving rate might create significant issues for economists, policymakers, and households. For instance, economists investigating the relationship between the personal saving rate and the health of the economy might draw the wrong conclusions if their data is incorrect. These researchers might also fail to diagnose if households are saving enough to sustain economic growth and prevent drops in consumption as the population grows older. Furthermore, artificially-low levels of the personal saving rate might prompt policymakers to wrongly incentivize higher savings, leading households to oversave. Therefore, improving our measures of the personal saving rate may yield substantial benefits to society. For a more detailed discussion, see Gokhale et al. (1996), Roger W. Ferguson (2004).

Our focus in this paper is on the nature and sources of the data revisions themselves and whether an analyst might be able to forecast revisions to the personal saving rate. We begin by examining the data and illustrating the size of the revisions. Then we show how revisions to the personal saving rate can be broken down into revisions to income and to outlays. Finally, we investigate whether the revisions are forecastable.

2. Data

The Bureau of Economic Analysis (BEA) defines the personal saving rate (expressed as a percentage) as

$$PSR = \frac{I - O}{I} \cdot 100$$

where I is income, defined as disposable personal income, and O is personal outlays (outlays, for short).

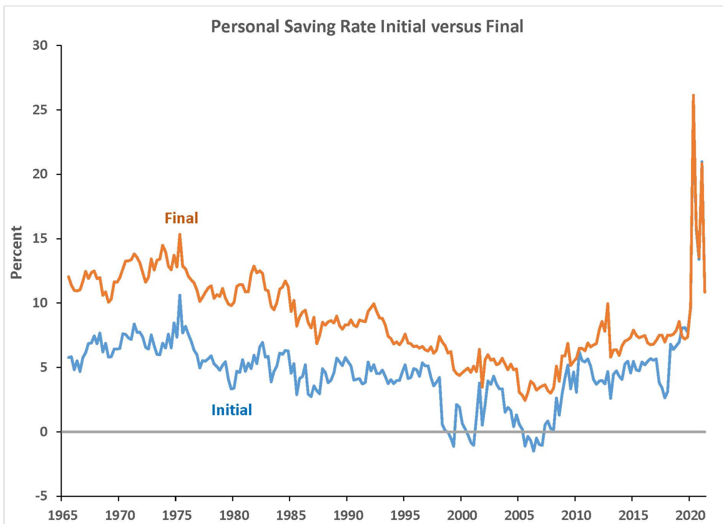
Data on the personal saving rate, outlays, and personal income are available from the Real-Time Data Set for Macroeconomists (RTDSM) at the Federal Reserve Bank of Philadephia[1], described in Croushore and Stark (2001). We use quarterly vintages of the quarterly variables for the personal saving rate,

[1] www.philadelphiafed.org/surveys-and-data/real-time-data-research/real-time-data-set-for-macroeconomists

nominal personal saving, nominal personal disposable income, and the PCE price deflator, all of which have vintages going back to 1965Q4.

Figure 1 plots the personal saving rate from 1965Q3 to 2021Q2, both as it was initial reported and as it is reported in today's dataset.

Figure 1: Personal Saving Rate Initial versus Final



As Figure 1 shows, the personal saving rate is revised substantially from its value when it is first released (initial) to the final value (which is the vintage of 2021Q3). In some cases, the revision is much larger than the initial release. In a number of quarters, the initial release showed a negative personal saving rate but in the final data, the smallest personal saving rate is 2.4 percent.

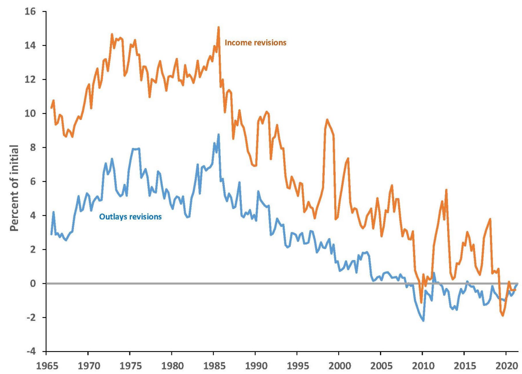
What is causing the personal saving rate to be revised so much? In order to answer this question, we investigate the variables that comprise this metric, since revisions to the personal saving rate could arise from revisions to income or to outlays.

Income and outlays are both nominal variables that grow over time because of both inflation and growth in the size of the economy, so examining their levels and revisions is not fruitful. We eliminate the effect of inflation by deflating each variable by the PCE price index.[2] Revisions to income and outlays reveal some answers as to why the personal saving rate is generally revised upwards, as

[2]Because of changes in base years over time, we deflate using the PCE price index in our last vintage, 2021Q3, so that the data are scaled properly.

Figure 2 shows.

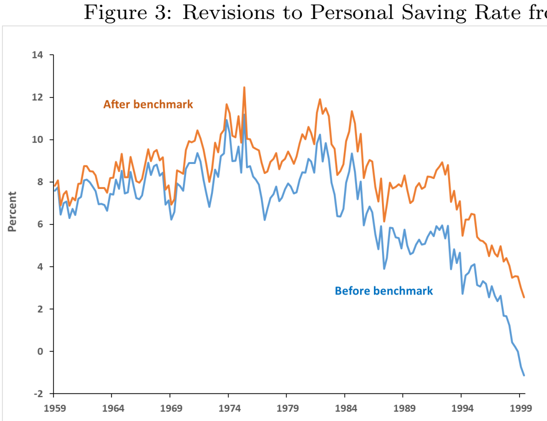
Figure 2: Revisions to Income and Outlays as a Percentage of Initial Release



Note that both income and outlays are revised up substantially, especially from 1965 to 2003 for outlays and 1965 to 2010 for income. But the percentage upward revision to income is substantially larger than the percentage upward revision to outlays, suggesting that income revisions play a more important role than outlay revisions in affecting the personal saving rate.

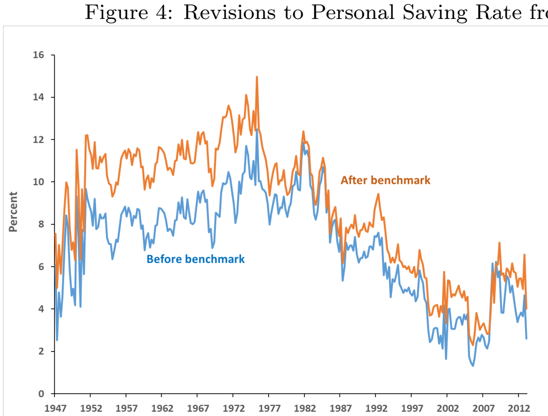
To investigate this more fully, we examine the benchmark revisions that occur roughly every 5 years or so, to see how those benchmark changes led to revisions of the personal saving rate. Looking at the data just prior to the benchmark revisions compared to just after the benchmark revisions, we see that two of the benchmark revisions, in 1999 and 2013, had particular large impacts on the personal saving rate, as Figure 3 and Figure 4 illustrate.

Figure 3: Revisions to Personal Saving Rate from 1999 Benchmark Revision



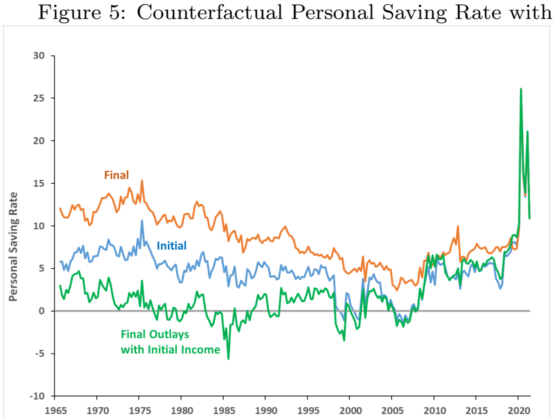
How can we tell if revisions to income or outlays are mainly responsible? Both outlays and income are revised up, so which matters more? To investigate

Figure 4: Revisions to Personal Saving Rate from 2013 Benchmark Revision



this, we create counterfactual graphs of the personal saving rate. First, we calculate a counterfactual personal saving rate, using final outlays with rescaled initial income. We plot that counterfactual personal saving rate against both the initial personal saving rate and the final revised version, as shown in Figure 5. Second, we calculate a counterfactual personal saving rate, using final income with rescaled initial outlays.

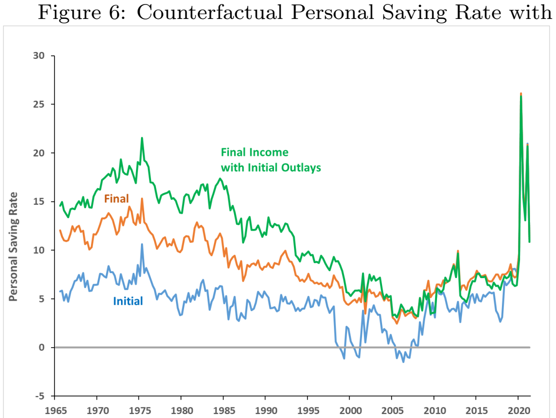
Figure 5: Counterfactual Personal Saving Rate with Final Outlays and Initial Income



Using final outlays with initial income shows a lower PSR than for either initial or final, so goes the wrong way. Using final income with initial outlays shows a higher PSR than for either initial or final, so goes too far. Both income revisions and outlays revisions matter, in opposite directions, but income revisions dominate.

Why are income revisions so important in determining the personal saving rate? According to the BEA, personal interest received is often revised upward, which could explain a significant portion of the revisions to income. Investigat-

Figure 6: Counterfactual Personal Saving Rate with Final Income and Initial Outlays



ing this in greater detail is a topic for future research.

3. JvN Statistics

Next, we look at a set of statistics that reveals the contribution of revisions of each component series to the revisions of the personal saving rate. This approach allows us to identify the sources of the revisions to the personal saving rate, helping elucidate the root cause of the unreliability of this variable.

We modify the statistics developed by Jacobs and van Norden (2011) and call them the JvN statistics. Jacobs and van Norden note that if we calculate the overall noise-signal ratio (where noise means revisions, and signal means variance, as measured using the latest vintage) of one variable, we can decompose it into the contribution of its two component variables. We modify their statistics to allow for removal of the trend in the income and consumption series (see the Appendix for details).

In order to calculate these statistics, we start by processing quarterly RTDSM data on the personal saving rate, nominal outlays, and nominal disposable personal income. Then, using PCE price index data from the vintage 2021Q3, we deflate the outlay and income numbers. Then, to reproduce the statistics developed by Jacobs and van Norden (2011), we calculate $ln(1 - PSR)$, $ln(OUTL)$, and $ln(NDPI)$, where PSR is the personal saving rate, $OUTL$ is real outlays, and $NDPI$ is real personal income. Lastly, we de-trend outlays and income, assuming a long-term growth rate of 3% per year. For the calculations below, consider PSR_{JvN} to be the de-trended real value of $ln(1-PSR)$, $OUTL_{JvN}$ to be the de-trended real value of $ln(OUTL)$, and $NDPI_{JvN}$ to be the de-trended real value of $ln(NDPI)$.

Afterwards, we calculate the noise-to-signal ratio of $OUTL_{JvN}$ and $NDPI_{JvN}$. To calculate the noise of $OUTL_{JvN}$, we take the variance of revisions to $OUTL_{JvN}$. To calculate the signal of $OUTL_{JvN}$, we take the variance of $OUTL_{JvN}$, as measured using the latest vintage available (21Q3). By dividing these two calculated numbers, we get the noise-to-signal ratio of $OUTL_{JvN}$. Repeating these steps with $NDPI_{JvN}$ data gives us the noise-to-signal ratio of $NDPI_{JvN}$.

$$\text{Noise-to-signal ratio of outlays} = \frac{Var(Revisions_{OUTL_{JvN}})}{Var(OUTL_{JvN})}$$

$$\text{Noise-to-signal ratio of income} = \frac{Var(Revisions_{NDPI_{JvN}})}{Var(NDPI_{JvN})}$$

Lastly, we calculate the noise-to-signal ratio of PSR_{JvN}. To this end, we calculate the sum of cross-products of revisions to $OUTL_{JvN}$ and $NDPI_{JvN}$, calculate the variance of PSR_{JvN} (as measured using the latest vintage available, 21Q3), and divide both the variance of revisions to $OUTL_{JvN}$ and $NDPI_{JvN}$ by the variance of PSR_{JvN}. By adding these numbers, we have the noise-to-signal ratio of PSR_{JvN}:

$$\text{Noise-to-signal ratio of } PSR_{JvN} =$$

$$\frac{Var(Revisions_{OUTL})}{Var(PSR)} + \frac{Var(Revisions_{NDPI})}{Var(PSR)} + \frac{\sum \text{Cross-Product of Revisions}}{Var(PSR)}$$

Intuitively, the first two terms represent the direct contribution of $OUTL_{JvN}$ or $NDPI_{JvN}$ revisions (respectively) to PSR_{JvN} revisions, while the third term represents the contribution of the interaction between $OUTL_{JvN}$ and $NDPI_{JvN}$ revisions to PSR_{JvN} revisions.

Over the full sample, PSR_{JvN} has a noise-to-signal ratio of 2.19, $OUTL_{JvN}$ has a noise-to-signal ratio of 0.53, and $NDPI_{JvN}$ has a noise-to-signal ratio of 5.87. Our measure of the direct contribution of $OUTL_{JvN}$ revisions to PSR_{JvN} revisions is 1.47, significantly less than the corresponding metric for $NDPI_{JvN}$, 6.75. This is consistent with our hypothesis that income revisions are more significant than outlay revisions in driving revisions to the personal saving rate.

Splitting the sample into four roughly-equal periods (65Q3 through 78Q3, 78Q4 through 91Q4, 92Q1 through 04Q4, 05Q1 through 17Q4), we notice that the second interval has the highest PSR noise-to-signal ratio, 7.04 (compared to 2.13 in the first period, 1.78 in the third, and 0.60 in the fourth). Since revisions tend to increase as the time since the initial release increases, we would expect the noise-to-signal ratio to go down as we move from the first period to the last. However, the period from 78Q4 through 91Q4 deviates from this pattern, as shown above.

To explain this phenomenon, we hypothesize that revisions may be higher in recessionary periods. Since the second split of our data set encompasses the highest number of recessions, 30 (versus 27 in the first period, 8 in the third, and 18 in the last), we would expect its noise to be particularly high. However, further research is needed to investigate this relationship.

4. Forecast Improvement Exercises

Given what we have found so far, we proceed to test the bias in the personal saving rate with simulated real-time out-of-sample forecasting exercises, which are "Forecast Improvement Exercises".[3] In each of the forecasting exercises below, we use only data that an analyst would have had in real time to simulate whether it would have been possible to forecast revisions to the PSR. First, we just use the estimated bias in the PSR over time to forecast the revisions. Second, we use the estimated bias in the growth rates of income and outlays to forecast revisions to PSR. For both types of forecasts, we consider an expanding sample beginning in 1965Q3, as well as a rolling ten-year sample.

Forecast Improvement Exercise Based on Bias in PSR Consider a forecaster standing in 1976Q1, with data from 1965Q3 to 1975Q4 on PSR. Data for the past two years have not had much chance to be revised, so suppose the forecaster ignores those data and just uses the average revision in the PSR from 1965Q3 to 1973Q4, then adds that to the initial release of PSR to estimate the revised PSR. Save that forecast, then step forward one quarter at a time and repeat the process, adding one quarter of new data each time; and in each case using only the vintage of data that a forecaster would have had at the time. Continue this process through 2021Q2. Then, we have forecasts of the revised PSR from 1976Q1 to 2021Q2. Using the latest vintage from 2021Q3 as our measure of "actuals", calculate the $RMSFE$ of the forecast of the revision versus using the initial release of PSR, which implicitly assumes that the revisions to PSR are not forecastable. Results are shown in the column labeled "Expanding window" of Table 1. The DM row shows the p-value of the Diebold-Mariano test with a null of equal forecast accuracy.

Table 1: Out-of-Sample Results for Using Bias to Forecast Revisions

Version	Expanding window	Ten-year rolling window
RMSFE initial	3.96	3.92
RMSFE bias	3.22	3.27
Percent difference	19%	17%
DM	0.00	0.00

Because the bias might change over time, a second experiment is to use ten-year rolling windows in the bias estimation. The results of that exercise are shown in the column labeled "Ten-year rolling window" in Table 1. For both the expanding window and the rolling window, the results suggest that the bias in the initial release is exploitable in real time. Based on past bias, revisions to the PSR can be forecasted and doing so reduces the $RMSFE$ significantly.

[3] See Croushore (2010).

Forecast Improvement Exercise Based on Bias in Income and Outlays Now consider the same type of exercise, but suppose that instead of forecasting the revision of PSR by using the average past bias of PSR, the forecaster instead forecasts revisions to income and outlays first, then uses those forecasts to estimate the revised PSR. Results are shown in Table 2.

Table 2: OOS Results Based on Income and Outlay Revisions

Version	Expanding window	Ten-year rolling window
RMSFE initial	3.96	3.92
RMSFE using I & O	4.00	3.99
Percent difference	-1%	-2%
DM	0.45	0.16

In this case, using bias in the income and outlays series to forecast the revision to PSR fails to reduce the $RMSFE$. It may be that because of the real-time complications of estimated real income and outlays when the base year changes leads causes an inconsistent pattern in the estimated bias. This may be worthy of further investigation.

5. Conclusions

In this paper, we have shown that data on the personal saving rate, disposable personal income, and personal outlays all tend to be revised up. The bias in the initial estimates of the personal saving rate is the clearest, and we can use the average of past revisions to forecast future revisions. Both income and outlays also tend to be revised up over time, but we are not able to use forecasts of those revisions to forecast revisions to the personal saving rate.

These results have implications for current policy analysis and future research. Policymakers may place too much emphasis on the signal coming from recent data on the personal saving rate but they need to understand that the data are likely to be revised in the future. Any alarm about a low rate of personal saving is likely misplaced. Researchers may find these results of interest and may wish to explore the broader implications of revisions to the personal saving rate. For example, if the personal saving rate is biased down initially, then why is gross domestic income considered to be useful in estimating revisions to GDP, as in Nalewaik (2010)?

6. Appendix: Derivation of the JvN statistic with trends

For this paper, consider Z as
$PSR_{JvN} = ln(1 - PSR_{de-trended})$,
Y as $OUTL_{JvN} = ln(OUTL_{de-trended})$,
and L as $NDPI_{JvN} = ln(NDPI_{de-trended})$

The original JvN stat is this:
$\phi^2 = \sum (R_t^Z)^2 / \sum (Z_t - \overline{Z})^2$
Now use the decomposition of Z_t:
$Z_t = Y_t - L_t$
$\phi^2 = \sum (R_t^Y - R_t^L)^2 / \sum (Z_t - \overline{Z})^2$
$\phi^2 = \sum [(R_t^Y)^2 + (R_t^L)^2 - 2(R_t^Y)(R_t^L)] / \sum (Z_t - \overline{Z})^2$
Break this down into A, B, C, D, E, and F components
$\phi^2 = (A \cdot B) + (C \cdot D) + (E \cdot F)$
$A = \sum (R_t^Y)^2 / \sum (Y_t - \overline{Y})^2$
$B = \sum (Y_t - \overline{Y})^2 / \sum (Z_t - \overline{Z})^2$
$C = \sum (R_t^L)^2 / \sum (L_t - \overline{L})^2$
$D = \sum (L_t - \overline{L})^2 / \sum (Z_t - \overline{Z})^2$
$E = -2 \sum (R_t^Y)(R_t^L) / \sum (Y_t - \overline{Y})(L_t - \overline{L})$
$F = \sum (Y_t - \overline{Y})(L_t - \overline{L}) / \sum (Z_t - \overline{Z})^2$
Now suppose that Y and L have a common trend
$Y_t = T_t + \widetilde{Y}_t$
$L_t = T_t + \widetilde{L}_t$
$Z_t = Y_t - L_t = [T_t + \widetilde{Y}_t] - [T_t + \widetilde{L}_t] = \widetilde{Y}_t - \widetilde{L}_t$
Assuming no trend revisions, then revision calculations remain unchanged but now use detrended Y and L instead of levels.
$\phi^2 = (A \cdot B) + (C \cdot D) + (E \cdot F)$
$A = \sum (R_t^Y)^2 / \sum (\widetilde{Y} - \overline{\widetilde{Y}})^2$
$B = \sum (\widetilde{Y} - \overline{\widetilde{Y}})^2 / \sum (Z_t - \overline{Z})^2$
$C = \sum (R_t^L)^2 / \sum (\widetilde{L} - \overline{\widetilde{L}})^2$
$D = \sum (\widetilde{L} - \overline{\widetilde{L}})^2 / \sum (Z_t - \overline{Z})^2$
$E = -2 \sum (R_t^Y)(R_t^L) / \sum (\widetilde{Y}_t - \overline{\widetilde{Y}})(\widetilde{L} - \overline{\widetilde{L}})$
$F = \sum (\widetilde{Y}_t - \overline{\widetilde{Y}})(\widetilde{L} - \overline{\widetilde{L}}) / \sum (Z_t - \overline{Z})^2$

References

Croushore, D., 2010. An evaluation of forecasts from surveys using real-time data. B.E. Journal of Macroeconomics 10, 1–30.

Croushore, D., 2011. Frontiers of real-time data analysis. Journal of Economic Literature 49, 72–100.

Croushore, D., Stark, T., 2001. A real-time data set for macroeconomists. Journal of Econometrics 105, 111–130.

Gokhale, J., Kotlikoff, L., Sabelhaus, J., 1996. Understanding the postwar decline in U.S. saving: A cohort analysis. Brookings Papers on Economic Activity 27, 315–407. URL: `https://EconPapers.repec.org/RePEc:bin:bpeajo:v:27:y:1996:i:1996-1:p:315-407`.

Jacobs, J.P., van Norden, S., 2011. Modeling data revisions: Measurement error and dynamics of 'true' values. Journal of Econometrics 161, 101–109. URL: http://neumann.hec.ca/pages/simon.van-norden/wps/CCSO%20WP%202006-07.pdf.

Nakamura, L., 2008. The mismeasured personal saving rate is still useful: Using real-time data to improve forecasting. Federal Reserve Bank of Philadelphia Business Review , 9–20.

Nakamura, L.I., Stark, T., 2005. Benchmark revisions and the u.s. personal saving rate. URL: http://www.phil.frb.org/files/wps/2005/wp05-6.pdf. federal Reserve Bank of Philadelphia working paper 05-6.

Nakamura, L.I., Stark, T., 2007. Mismeasured personal saving and the permanent income hypothesis. URL: http://www.philadelphiafed.org/files/wps/2007/wp07-8.pdf. federal Reserve Bank of Philadelphia working paper 07-8.

Nalewaik, J.J., 2010. The income- and expenditure-side estimates of U.S. output growth. Brookings Papers on Economic Activity , 71–127.

Roger W. Ferguson, J., 2004. Questions and reflections on the personal saving rate. URL: https://www.federalreserve.gov/boarddocs/speeches/2004/20041006/default.htm. remarks to the National Bankers Association, Nashville, Tennessee.