When Numbers Don't Add Up

The Statistical Discrepancy in GDP Accounts

David Rosnick and Dean Baker

August 2011

Center for Economic and Policy Research
1611 Connecticut Avenue, NW, Suite 400
Washington, D.C. 20009
202-293-5380
www.cepr.net

Contents

About the Authors

David Rosnick is a Senior Economist at the Center for Economic and Policy Research in Washington, D.C. Dean Baker is an economist and the Co-director of CEPR.

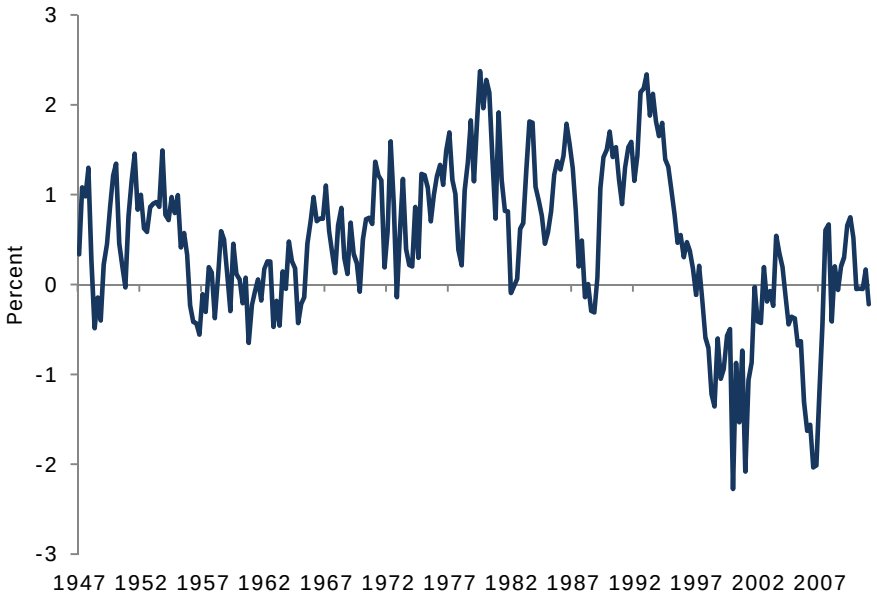
Introduction

Economists have two main ways of estimating the amount of real economic activity in a given period. First, there is gross domestic product (GDP), which totals up the market value of production of goods and services offered within the period. Second, there is gross domestic income (GDI), which represents the earnings associated with the production of the same goods and services. In theory, these two measures are identical, but in practice they are not, since they are based on different sources of data.

The output measure (GDP) is generally regarded as more reliable than the income measure (GDI) because there is a good deal of self-reporting of income.[1] In most years the GDP measure has exceeded the GDI measure. The average difference between GDP and GDI, known as the "statistical discrepancy," over the years 1947 through the first quarter of 2011 is 0.5 percent of GDP. The most obvious explanation for this gap is that the Bureau of Economic Analysis' GDI data depend on income tax data. Insofar as income is under-reported to avoid taxes, it will lead to an under-reporting of GDI.

In a reversal of this longstanding pattern, the income side exceeded the output side by large amounts in the late 1990s and again in the years 2004-07, as can be seen in **Figure 1**.

FIGURE 1
The Statistical Discrepancy between Gross Domestic Product and Gross Domestic Income, 1947-2011 (percent of GDP)



Source: Bureau of Economic Analysis, NIPA Table 1.10, Lines 1 and 26.

1 There is less variance in the quarterly data in the output measure than income measure. In addition, the mean revision from first report to the latest revision is smaller in the case of the output measure than the income measure, both factors that might suggest the superiority of the output measure. See Fixler, Dennis J. and Bruce T. Grimm. 2008. "The Reliability of GDP and GDI Estimates." Washington, D.C.: Bureau of Economic Analysis, Survey of Current Business, February, pp. 16-34. Available at http://www.bea.gov/scb/pdf/2008/02%20February/0208_reliable.pdf.

This means that the income measure provides a somewhat different picture of economic growth than the more commonly used output measure. Compared to the GDP measure, the GDI measure shows more growth in the late 1990s, a steeper falloff in the 2001 recession, and then more rapid growth in the 2000s cycle. In order to get a better assessment of the movements in the economy and the impact of policy, it is worth determining the basis for the fluctuations in the size and sign of the statistical discrepancy.

The Capital Gains Explanation for the Fluctuations in the Statistical Discrepancy

One obvious place to look for an explanation for the fluctuations in the statistical discrepancy is capital gains.[2] In principle, capital gains should not be counted in GDP. They reflect the change in the price of existing assets, not the value of newly created goods and services. However, as a practical matter, it is virtually inevitable that some amount of capital gains income will be misclassified as wage or profit income.

While long-term capital gains are given favorable tax treatment, short-term capital gains are taxed as ordinary income. Therefore, it is likely that some number of taxpayers will carelessly record their short-term capital gains income as ordinary income. Clearly, the IRS has no incentive to pursue this kind of error, but any such errors would increase the amount of ordinary income reported. Such errors would most likely appear in the household sector, as the business sector has—as a rule—the necessary experience and accountants to avoid any accidental misreporting of income.

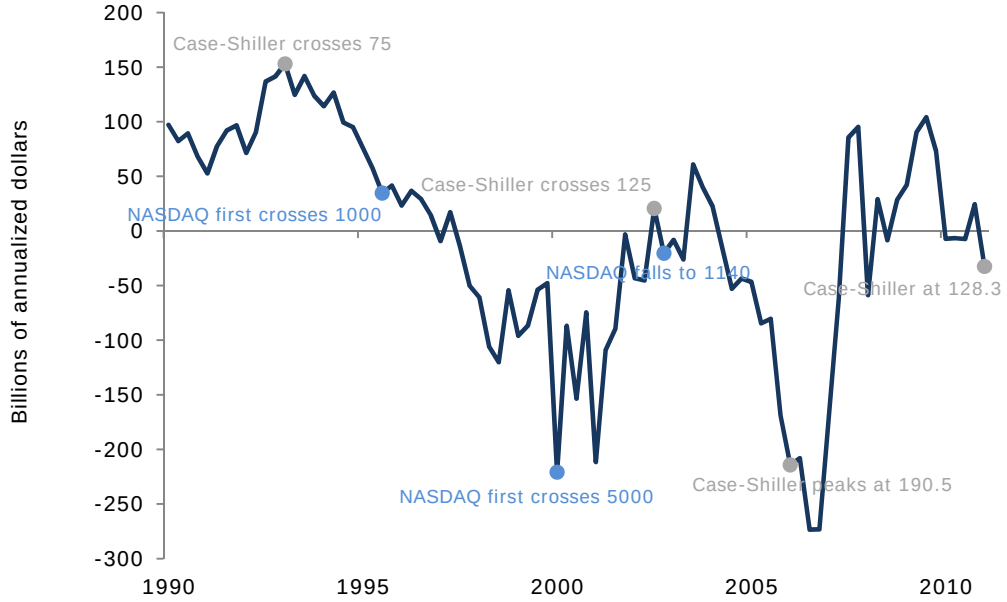
It is also likely that some amount of long-term capital gains income shows up as normal income as well, despite the tax benefits given to long-term gains. Inevitably there will be some misclassification. If the misclassification for both reasons is a constant portion of capital gains, then it would imply that the greater the amount of capital gains in the economy relative to GDP, the greater will be the overstatement of income relative to output.[3]

Figure 2 shows the pattern in the statistical discrepancy (in billions of annualized dollars) over the last two decades. As can be seen, the peak negative values of the statistical discrepancy coincided with the two most significant asset bubbles since the Great Depression—the dot-com bubble of the late 1990s and the real-estate bubble of the 2000s.

2 As an alternative explanation, Mark Thoma links the statistical discrepancy to increases in the government share of output. See Thoma, Mark. 2005. "Explaining the Discrepancy Between GDP and GDP with Non-Defense Related Government Consumption." Economist's View, August 26. Available at http://economistsview.typepad.com/economistsview/2005/08/explaining_the_.html.

3 There undoubtedly is misstatement in the other direction with people hiding ordinary income as capital gains income. However, misclassification in this direction would presumably be roughly constant as a share of GDP, except in response to changes in the tax code.

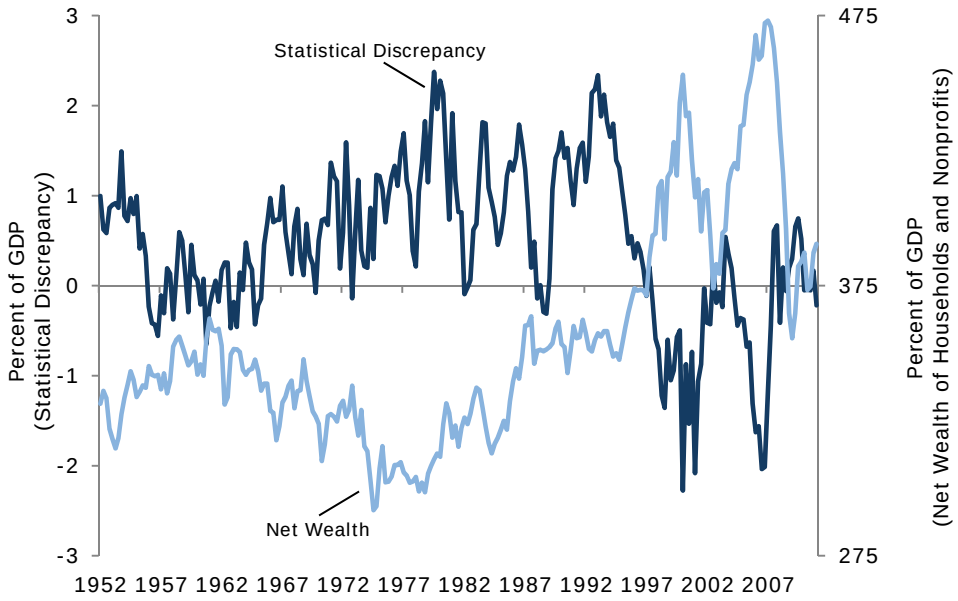
FIGURE 2
The Statistical Discrepancy between Gross Domestic Product and Gross Domestic Income, 1990-2011 (billions of annualized dollars)



Source: Bureau of Economic Analysis, NIPA Table 1.10, Line 26

Figure 3 shows the inverse relationship between the statistical discrepancy and household net worth.[4] The statistical discrepancy turns sharply negative as households become unusually wealthy and reverts as household assets fall.

FIGURE 3
The Statistical Discrepancy and Net Wealth, 1952-2011



Source: Bureau of Economic Analysis, NIPA Table 1.10, Lines 1 and 26; Fed Flow of Funds Table B.100, Line 42.

4 The Federal Reserve Flow of Funds reports consolidated numbers for households and nonprofit organizations.

A simple regression gives us an idea of how much of an effect household net worth might have on the statistical discrepancy. As seen in **Table 1**, an increase in household net worth corresponding to 100 percentage points of quarterly GDP (25 percentage points annualized) is associated with a fall in the statistical discrepancy of perhaps 0.32-0.35 percentage points of GDP.

TABLE 1
The Impact of Net Worth on the Statistical Discrepancy

	1990Q1		1952Q1	
Net worth	-0.0058 (0.0005)**	-0.0035 (0.0010)**	-0.0035 (0.0003)**	-0.0032 (0.0006)**
constant	0.092 (0.008)**	0.056 (0.016)**	-0.054 (0.004)**	0.050 (0.008)**
rho	--[(a)]	0.766	--[(a)]	0.767
R^2	0.63	0.13	0.42	0.12
Durbin-Watson	0.64	2.34	0.47	2.15

** 1% level of significance. (a) Not corrected for AR(1) errors.

The exact timing of the relationship between misreporting of a capital gain and when the gain took place is not obvious. After a large run-up in asset prices, wealth holders may realize a large gain as the market itself is falling. Similarly, many may wait to realize capital gains on assets that were purchased at market lows. Thus, it is unclear whether the actual capital gains would be better than the valuation level in explaining the statistical discrepancy.

Therefore, in order to test whether misreporting of capital income drives the statistical discrepancy, we regress the statistical discrepancy (as a share of GDP) against lags of household-owned corporate equities and real estate (also as shares of GDP). The results are shown in **Table 2**.

TABLE 2
The Impact of Capital Gains in the Household Sector on the Statistical Discrepancy

		Cumulative Effect of Faster Capital Gains over N Quarters[a]		Cumulative Effect of Capital Gain over N Quarters[b]				
Start	Maximum Lag (N)	Corporate Equities	Real Estate	Corporate Equities	Real Estate	Constant	R^2	Durbin-Watson (transformed)
		-0.1667	-0.0709	-0.0130	-0.0060	0.064		
1990Q1	16	(0.0366)**	(0.0309)*	(0.0028)**	(0.0025)*	(0.014)**	0.61	1.95
		-0.0653	-0.0526	-0.0117	-0.0036	0.048		
1990Q1	8	(0.0205)**	(0.021)*	(0.0031)**	(0.0023)	(0.015)**	0.46	2.09
		-0.0222	-0.0325	-0.0118	-0.0044	0.053		
1990Q1	4	(0.0114)#	(0.0171)#	(0.0032)**	(0.0018)*	(0.011)**	0.40	2.13
		-0.1210	-0.0806	-0.0072	-0.0042	0.039		
1952Q1	16	(0.0274)**	(0.0246)**	(0.0017)**	(0.0014)**	(0.008)**	0.36	2.05
		-0.0406	-0.0639	-0.0084	-0.0040	0.041		
1952Q1	8	(0.0125)**	(0.0186)**	(0.0017)**	(0.0013)**	(0.007)**	0.28	2.03
		-0.0134	-0.0471	-0.0079	-0.0038	0.039		
1952Q1	4	(0.0061)*	(0.0138)**	(0.0016)**	(0.0011)**	(0.007)**	0.25	2.06

Notes: Levels of significance: * 5%, ** 1%. Prais-Winsten regression with statistical discrepancy (divided by GDP) as dependent variable. Semi-robust standard errors in parenthesis. Dependent variables are concurrent and N lags of change in assets divided by GDP and Nth lag of asset valuation divided by GDP. (a) Sum of coefficients on changes in assets divided by GDP. (b) Coefficient on Nth lag of asset valuation divided by GDP.
Sources: Bureau of Labor Statistics, NIPA Table 1.10, Lines 1 and 26; Federal Reserve, Flow of Funds Table B.100, Lines 3 and 24.

In all these regressions, assets had significant negative effects on the statistical discrepancy. A single-quarter increase in the value of household real estate corresponding to 10 percentage points of (quarterly) GDP is associated with a decrease in the statistical discrepancy by 0.04-0.06 percentage points in subsequent quarters. The effect of a 10 percentage-point-of-GDP increase in the value of household ownership of corporate equities was larger—a 0.07-0.13 percentage-point-of-GDP decrease in the statistical discrepancy in later quarters.

The story does not substantially change in moving to lags of capital gains. For corporate equities, a single-quarter's increase in capital gains of 10 percentage points of GDP is associated with a decrease in the statistical discrepancy of 0.03-0.10 percentage points. For real estate, the effect was somewhat larger than in considering levels— a 0.04-0.12 percentage-point drop in the statistical discrepancy.[5]

5 The authors also ran regressions that included capital gains and asset levels for assets of nonfarm, nonfinancial business. These regressions showed little or no statistically significant relationship with the statistical discrepancy. See Appendix.

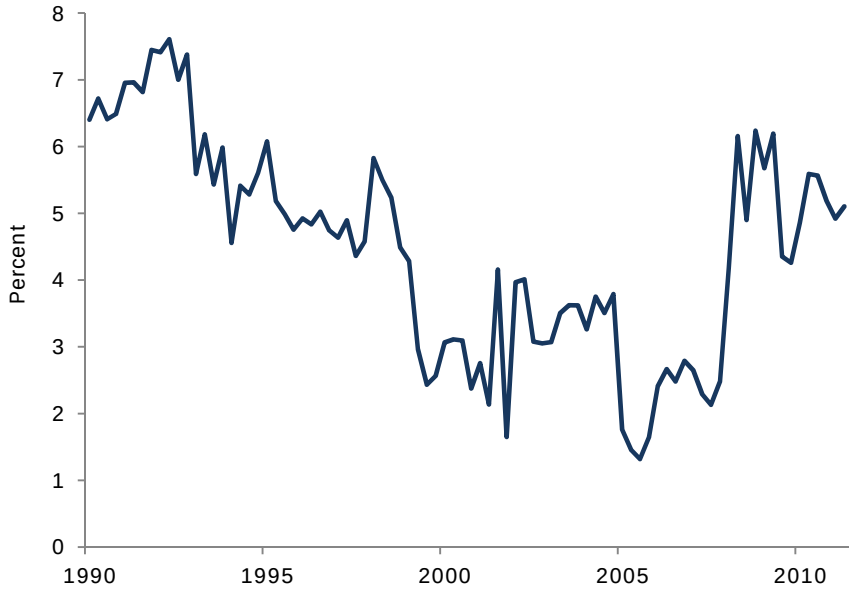
The Statistical Discrepancy and Capital Gains: Why It Matters

There are two main reasons why it matters if the fluctuations in the size and direction of the statistical discrepancy are driven by capital gains. First, there have been some economists who have viewed the income side as providing additional and possibly better information on GDP growth than the output side. If the fluctuations in the size of the statistical discrepancy are primarily explained by misclassified capital gains, then the income side is not providing additional information on GDP growth. There is no reason to move away from the output measure or to take an average of the two.

The other reason that this capital gain explanation would be important is that it would have a profound effect on our understanding of savings rates in the last twenty years. Private spending and taxation are well-measured, so any increase in measured income corresponds to an increase in measured private savings as well. Thus, in normal economic times, the savings rate is generally a bit higher than reported. During an asset boom, however, the actual savings rate may be considerably lower.

Figure 4 shows the measured personal (household) savings rate as a share of disposable personal income since 1990.

FIGURE 4
Measured Household Savings Rate, 1990-2011



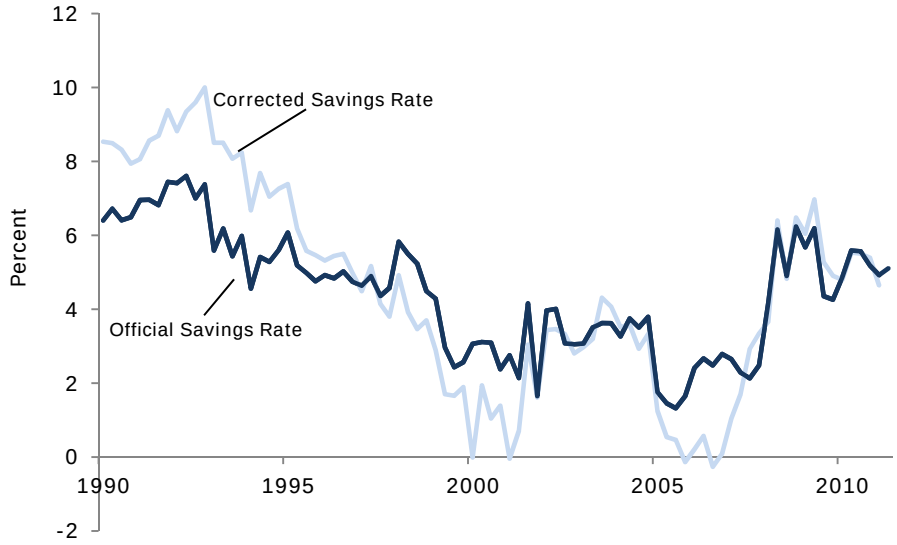
Source: Bureau of Economic Analysis, NIPA Table 2.1 Lines 26 and 33.

From a high of 7.6 percent in the second quarter of 1992, the personal savings rate fell through the rest of the decade—reaching a low of 2.1 percent nine years later. After recovering for a few years

in the early 2000s, the savings rate plummeted in early 2005 to 1.3 percent as the housing bubble neared its peak. Following the collapse of the housing bubble, households began to resume saving and the rate again topped 6.2 percent by the second quarter of 2008.

If, however, we believe that the statistical discrepancy represents misreported personal income, then the actual falls in savings (and subsequent recoveries) were considerably more pronounced. **Figure 5** shows the official and corrected measures of personal savings under this assumption.

FIGURE 5
Official and Corrected Household Savings Rates, 1990-2011



Source: Bureau of Economic Analysis, NIPA Table 2.1 Lines 26 and 33; Table 5.1, Lines 9 and 43.

As seen in the figure, savings rates were significantly higher in the early 1990s than reported incomes would indicate. Savings may have reached 10 percent of disposable income by the end of 1992 and fallen all the way to zero by the first quarter of 2000.

As a result, misreporting of income masked the depth of the fall in the personal savings rate. Rather than an official fall of 5.5 percentage points, the savings rate fell by 10 percentage points. For two years from mid-1999 to mid-2001, the savings rate stayed below 2.0 percent and stayed well below 1.0 percent for nearly two years starting in the second quarter of 2005. In fact, the corrected savings rate fell to a low of -0.3 percent in mid-2006 before climbing to 7.0 percent in the middle of 2009.

Because household wealth appears to increase during asset bubbles, failure to recognize that the rapid appreciation in housing prices was unsustainable led to the failure to recognize that households were saving insufficiently for their future needs. As the bubbles burst, households' need to resume saving was assured. Insofar as some of the gain in wealth attributable to these bubbles appeared in GDP accounts as misclassified disposable income, the impact of the bubbles in discouraging saving is even larger than the official data indicate.

Conclusion

At the peak of both the stock and housing bubbles, there were extraordinary shifts in the statistical discrepancy between the national output and income accounts. The statistical discrepancy fell from its normal range of 0.5 – 1.0 percent of GDP to levels below -1.0 percent of GDP. The analysis in this paper suggests that this reversal was directly related to these bubbles, with the likely explanation that a portion of the capital gains from these bubbles being misclassified in national income accounts as ordinary income. If this is the case, then the drops in household saving during the bubbles and the subsequent rises following their collapse were even larger than the official data show.

Appendix

Appendix Table 1 shows a set of regression results that used the value of all corporate equity and real estate, including commercial real estate, as independent varaibles.

APPENDIX TABLE 1
The Impact of Capital Gains on the Statistical Discrepancy

		Cumulative Effect of Faster Capital Gains over N Quarters[(a)]		Cumulative Effect of Capital Gain over N Quarters[(b)]				
Start	Maximum Lag (N)	Corporate Equities	Real Estate	Corporate Equities	Real Estate	constant	R^2	Durbin-Watson (transformed)
		-0.0702	-0.0259	-0.0051	0.0005	0.032		
1990Q1	16	(0.0166)**	(0.0163)	(0.0009)**	(0.0015)	(0.014)*	0.69	2.21
		-0.0271	-0.0269	-0.0052	-0.0001	0.029		
1990Q1	8	(0.0079)**	(0.0097)**	(0.0008)**	(0.0011)	(0.010)**	0.61	2.24
		-0.0083	-0.0240	-0.0052	-0.0003	0.031		
1990Q1	4	(0.0046)#	(0.0070)**	(0.0009)**	(0.0010)	(0.008)**	0.58	2.23
		-0.0610	-0.0331	-0.0032	0.0001	0.017		
1952Q1	16	(0.0157)**	(0.0168)*	(0.0008)**	(0.0012)	(0.011)	0.39	2.14
		-0.0178	-0.0317	-0.0037	-0.0004	0.021		
1952Q1	8	(0.0068)**	(0.0111)**	(0.0009)**	(0.0012)	(0.010)*	0.29	2.12
		-0.0049	-0.0256	-0.0035	-0.0004	0.020		
1952Q1	4	(0.0036)	(0.0068)**	(0.0008)**	(0.0010)	(0.009)*	0.24	2.15

Notes: Levels of significance: # 10%, * 5%, ** 1%. Prais-Winsten regression with statistical discrepancy (divided by GDP) as dependent variable. Semi-robust standard errors in parenthesis. (a) Sum of coefficients on changes in assets divided by GDP. Dependent variables are concurrent and N lags of change in assets divided by GDP and Nth lag of asset valuation divided by GDP. (b) Coefficient on Nth lag of asset valuation divided by GDP.
Sources: Bureau of Economic Analysis, NIPA Table 1.10, Lines 1 and 26; Federal Reserve, Flow of Funds Table B.100, Lines 3 and 24; Table B.102, Line 3; Table L.4, Line 11.