Journal of Monetary Economics 24 (1989) 31–51. North-Holland

PRODUCTION, SALES, AND THE CHANGE IN INVENTORIES
An Identity That Doesn't Add Up*

Jeffrey A. MIRON

University of Michigan, Ann Arbor, MI 48109, USA
National Bureau of Economic Research, Cambridge, MA 02138, USA

Stephen P. ZELDES

The Wharton School, University of Pennsylvania, Philadelphia, PA 19104, USA
National Bureau of Economic Research, Cambridge, MA 02138, USA

Received June 1987, final version received March 1989

We examine two measures of monthly manufacturing production. The first is the index of industrial production; the second is constructed from the accounting identity that output equals sales plus the change in inventories. We show that the means, variances, and serial correlation coefficients of the log growth rates differ substantially between the two series, and the cross-correlations are in most cases less than 0.4. A model of classical measurement error indicates that in 15 of 20 two-digit industries measurement error accounts for over 35% of the variation in the monthly growth rates of seasonally adjusted industrial production.

1. Introduction

In this paper we examine two measures of monthly production that have been used by economists. The first measure, which we refer to as *IP*, is the index of industrial production constructed by the Board of Governors of the Federal Reserve. This measure is used extensively in empirical work on the business cycle, as well as by policymakers and others to assess the current state of the economy. The second measure, which we refer to as *Y4*, is constructed from the accounting identity that output equals sales plus the change in

*We are grateful to Joe Beaulieu, Edward Gold, and Amy Lipton for research assistance. We thank Joe Altonji, Larry Christiano, Marty Eichenbaum, Alan Greenspan, Jim Kahn, Craig MacKinlay, Christina Romer, David Romer, Ruth Runyon, Chris Sims, Mark Watson, David Wilcox, a referee, members of the Penn Macro Lunch Group and workshop participants at Brown, Harvard, Michigan, Northwestern, and the University of Texas at Austin for useful comments. We presented an earlier version of this paper at the Federal Reserve Board and are appreciative of particularly helpful discussions with Zoltan Kenessey (head of the section that produces the industrial production series), Richard Raddock, Dixon Tranum, and others at the FRB, as well as with Frank de Leeuw and Robert Parker of the BEA. Both authors are thankful for support from John M. Olin Fellowships at the National Bureau of Economic Research. Miron thanks NSF Grant SES-8710140 for financial support.

0304-3932/89/$3.50

inventories. Sales and inventory data are reported by the Department of Commerce. This measure of output is frequently used to estimate models of inventory accumulation. Theoretically, these two series measure the same underlying economic variable – the production of goods by manufacturing firms during the month.

We show here that the time series properties of these two series are radically different. We examine means, variances, and serial correlation coefficients of the log growth rates and show that these statistics differ substantially between the two series. Generally, *IP* is a less volatile and more persistent series than *Y4*. In addition, the cross-correlations between the two seasonally adjusted series range from 0.6 to 0.0 and are in most cases less than 0.4.[1] We then demonstrate the significance of these differences in two ways. First, we show that the variance bounds results of Blinder's (1986) study of inventory behavior are partially reversed when the *IP* rather than the *Y4* output measure is used. Second, we examine two specific models of the measurement error in the series. The estimates under one of them (classical measurement error) indicate that in 15 out of 20 two-digit industries measurement error accounts for over 35% of the variation in the monthly growth rates of seasonally adjusted industrial production data.

These results are important for all those who use the *IP* or *Y4* data. This includes researchers on inventories, since some studies use the *IP* measure while others use the *Y4* measure.[2] More generally, many studies of the business cycle employ *IP* as a measure of economic activity. Our results supplement the work of Lichtenberg and Griliches (1989), who show that substantial measurement error exists in industry level price indexes.

The remainder of the paper is organized as follows. Section 2 describes how the two data series are constructed. Section 3 presents summary statistics that demonstrate the differences between the two series, and section 4 gives an example of the economic significance of the discrepancies. In section 5 we model the measurement error and estimate its importance under alternative sets of assumptions. Section 6 concludes the paper.

2. Data construction

In this section we describe how the data released by the relevant government agencies are constructed and how we use these data to construct *Y4*.

[1] The correlations between the growth rates of the raw seasonally unadjusted series are always higher, ranging from 0.4 to 0.9.

[2] Blinder (1986) and West (1986) use the *Y4* measure, while Maccini and Rossana (1984) and Reagan and Sheehan (1985) use the *IP* measure. Miron and Zeldes (1988b) report two sets of results: one using *IP* and the other using *Y4*. West points out in his footnote 13 that he estimated his equations for a few of the industries using the *IP* measure as well. He found that the parameters were uniformly nonsensical and therefore did not report them.

2.1. Construction of IP

The Federal Reserve Board's (FRB) index of industrial production is available monthly, both seasonally adjusted (SA) and seasonally unadjusted (NSA), at the two-digit level, from 1959 to the present. The two-digit series and the more aggregated series are constructed from disaggregated data using value-added weights.

The disaggregated *IP* indexes are constructed from three types of data: physical product measures, kilowatt-hours of electrical power input, and man-hours of labor input. Each of these is collected at either the establishment (plant) level or at the more specific product level. The input measures are used in cases where the physical product numbers are not available or would not make sense because of heterogeneity in the product. For the physical product measures, the FRB uses series from the Department of Energy, the Bureau of the Census, and other public and private sources. Most of these are counts of output goods, although occasionally (e.g., steel) they are constructed as the sum of sales and inventory changes.[3] For the kilowatt-hour data, the FRB asks utility companies their sales of kilowatt-hours of electric power to firms in manufacturing.[4,5,6] For the man-hours series, the Bureau of Labor Statistics provides data from its payroll reports.[7] Virtually all of the three- or four-digit level series used to construct the two-digit level series are based on only one of these sources (i.e., not on a combination of these sources).

Table 1 gives the fraction of output in each industry that is based on physical product data, electricity use, and labor use respectively.[8] Some industries are based almost entirely on electricity use (e.g., furniture), while others are based predominantly on man-hour use (e.g., apparel). Overall, the series that are based solely on kilowatt-hours represent 35% of the output covered by the industrial production manufacturing index, and those based solely on production-worker-hours represent 29%. Only about one third of the industrial production index for manufacturing is actually calculated directly from data on physical units of output.

[3] The physical product numbers are divided by the number of working days in the reporting period in order to put each series on a per-working-day basis.

[4] Unfortunately, the reports do not measure electricity use on a calendar month basis, because billing dates fall throughout the month and thus cover different month-long periods for different customers [Federal Reserve Board (1986, p. 42)].

[5] Monthly movements in the data are reviewed to eliminate 'abrupt movements that cannot be accounted for by such developments as work stoppages, power shortages, or cyclical movements' and are presumed due to inappropriate reports [Federal Reserve Board (1986, p. 42)].

[6] The FRB also asks 'self-generators' of electricity in the manufacturing industry to report power used in manufacturing.

[7] Only one week of data (the week containing the 12th day of the month) is used to estimate the monthly labor input [Federal Reserve Board (1986, p. 42)].

[8] These fractions were aggregated, using value-added weights, from the numbers for the three- and four-digit industries given in Federal Reserve Board (1986, pp. 133–148).

Table 1

Composition of IP, by source.[a,b,c]

	SIC code	Nondurable/ durable	Fraction of index based on			
			Physical output	Kilowatt-hours	Production-worker-hours	Other data
Food	20	N	0.409	0.390	0.200	0.000
Tobacco	21	N	0.903	0.097	0.000	0.000
Textiles	22	N	0.671	0.307	0.022	0.000
Apparel	23	N	0.000	0.165	0.835	0.000
Lumber	24	D	0.539	0.035	0.313	0.113
Furniture	25	D	0.000	0.953	0.047	0.000
Paper	26	N	0.990	0.000	0.010	0.000
Printing	27	N	0.297	0.703	0.000	0.000
Chemicals	28	N	0.334	0.376	0.227	0.062
Petroleum	29	N	0.924	0.076	0.000	0.000
Rubber	30	N	0.234	0.695	0.000	0.071
Leather	31	N	0.547	0.151	0.302	0.000
Stone, Clay, Glass	32	D	0.255	0.628	0.117	0.000
Primary Metal	33	D	0.908	0.021	0.071	0.000
Fab. Metal	34	D	0.000	0.510	0.490	0.000
Machinery	35	D	0.005	0.673	0.240	0.082
Elec. Machinery	36	D	0.134	0.271	0.554	0.041
Trans. Equip.	37	D	0.418	0.005	0.577	0.000
Instruments	38	D	0.000	0.173	0.711	0.117
Other	39	D	0.000	1.000	0.000	0.000
Nondurables	—	N	0.446	0.363	0.171	0.020
Durables	—	D	0.242	0.348	0.375	0.034
Total	—	T	0.328	0.355	0.289	0.028

[a]*Source:* Federal Reserve Board, *Industrial Production, 1986*.

[b]The entries in the last four columns are the fraction of the industrial production index of each industry that is based on physical-output data, kilowatt-hours data, production-worker-hours data, and other data, respectively.

[c]Includes Federal Reserve estimates and combined kilowatt-hour and production-worker-hour data.

The FRB constructs production factor coefficients (PFC's) to convert input data into estimates of monthly output.[9] The PFC's are a weighted average of the ratio of annual Census output data to annual input data, adjusted to incorporate trends and cyclical movements in productivity.[10] These productivity adjustments are 'based on the historical behavior of the series in earlier cycles and on an assessment of the position of the series in the current cycle'

[9]The FRB also applies analogous PFC's to some of the physical product series, to account for incomplete sampling in the monthly numbers.

[10]Census calculates both annual and quinquennial indexes of production, calculated as the deflated sum of shipments and the change in inventories (work in progress plus finished goods) using data from the Annual Survey of Manufactures and the quinquennial Census of Manufactures, respectively [Federal Reserve Board (1986, p. 46), U.S. Department of Commerce (1983)]. The FRB combines these to get a benchmarked annual index of production.

[Federal Reserve Board (1986, p. 49)]. The *PFC*'s do not vary seasonally. The estimate of monthly output is equal to the product of monthly input and the monthly *PFC*. Because of this procedure, using industrial production data to analyze short-run productivity changes [e.g., Stockman (1988)] may be misleading, since estimating productivity from the *IP* data will to a large extent result in backing out the FRB estimates of the *PFC*'s.

2.2. Construction of Y4

The *Y4* measure of production is defined as $y_t^{Y4} \equiv x_t + n_t - n_{t-1}$, where y_t^{Y4} is real production during period t, x_t is the real value of shipments during period t, and n_t is the sum of the real values of the stock of finished goods and work in progress inventories at the end of period t.[11] Constant dollar shipments and inventories are provided by the Bureau of Economic Analysis (BEA) of the Department of Commerce and are available monthly from 1959 to the present at the two-digit level. We adjust the finished goods and work in progress inventory series from cost to market by multiplying each by an industry-specific constant, as described in West (1983) and Holtz-Eakin and Blinder (1983), respectively.

To arrive at the constant dollar inventory series, the BEA begins with data on the book value of inventories collected by the Bureau of the Census at the Commerce Department and adjusts these for differences between book and current dollar values and also for differences between current and constant dollar values. This complicated procedure incorporates information about whether firms use LIFO or non-LIFO accounting methods and involves estimating the accounting age structure of the existing stock of goods. The conversion procedures are described in Hinrichs and Eckman (1981), Foss et al. (1980), and Miron and Zeldes (1988a).

The book value data are collected by the Census through the monthly M3 (Manufacturers' Shipments, Inventories, and Orders), the Annual Survey of Manufactures, and the quinquennial Census of Manufactures. The M3 is a voluntary survey of large companies. There are a total of only 4500 reporting units, made up of 3400 companies and 1100 divisions of 450 companies.[12] On

[11] There is some debate as to whether to include work in progress inventories in the definition of *Y4*. Although the use of only finished goods inventories is more standard in work on inventories, Blinder (1986) argues that the definition including work in progress is a more desirable measure of production, and he suggests that the use of only finished goods inventories might be the source of the discrepancy between *Y4* and *IP* [see West (1986, fn. 13) and Blinder (1986)]. We have also calculated the statistics in tables 2a and 2b using the finished goods only definition of *Y4*. The results, presented in Miron and Zeldes (1988a), are very similar to those in tables 2a and 2b and do not yield consistently higher or lower correlations between *IP* and *Y4*.

[12] Reporting units often produce more than one type of good, and sometimes these goods fall into different industry classifications. In this case, all of the inventories and shipments of the reporting unit are lumped into the primary industry classification. Units report total book value inventories, and then a breakdown into three stages of fabrication: materials and supplies, work in progress, and finished goods.

Table 2a

Summary statistics, seasonally adjusted data, 1967:5–1984:12.[a,b]

	Mean			Std. dev.			Autocorrelation			Cross-correlation
	IP	*Y4*	*t*-stat.	*IP*	*Y4*	*t*-stat.	*IP*	*Y4*	*t*-stat.	
Food	0.0025	0.0014	1.77	0.009	0.021	12.14	−0.27	−0.34	0.82	0.18
Tobacco	0.0004	0.0002	0.12	0.046	0.098	7.60	−0.52	−0.45	−0.64	0.19
Textiles	0.0014	0.0016	−0.16	0.021	0.035	3.54	0.31	−0.39	5.08	0.25
Apparel	0.0010	0.0010	0.03	0.024	0.057	8.28	−0.23	−0.35	1.10	0.09
Lumber	0.0015	0.0023	−0.55	0.026	0.051	8.28	0.04	−0.32	2.55	0.32
Furniture	0.0030	0.0031	−0.10	0.020	0.061	15.24	0.05	−0.51	4.96	0.16
Paper	0.0028	0.0023	0.67	0.019	0.024	2.62	−0.02	−0.34	1.72	0.38
Printing	0.0030	0.0021	1.04	0.012	0.033	13.14	−0.13	−0.51	3.58	−0.02
Chemicals	0.0041	0.0030	1.54	0.015	0.026	8.42	0.10	−0.23	2.32	0.17
Petroleum	0.0009	0.0017	−0.60	0.021	0.031	3.25	−0.12	−0.37	2.45	0.08
Rubber	0.0054	0.0023	2.29	0.031	0.043	2.80	0.09	−0.27	3.61	0.37
Leather	−0.0028	−0.0032	0.21	0.029	0.075	13.19	−0.21	−0.44	2.28	0.09
Stone, Clay, Glass	0.0022	0.0009	1.71	0.020	0.034	6.99	−0.01	−0.34	3.31	0.33
Primary Metal	−0.0007	−0.0007	0.04	0.041	0.042	0.10	0.19	0.10	0.86	0.64
Fab. Metal	0.0013	0.0011	0.15	0.014	0.051	22.63	0.41	−0.44	7.79	0.25
Machinery	0.0035	0.0028	0.86	0.015	0.039	24.99	0.30	−0.35	5.39	0.34
Elec. Machinery	0.0049	0.0045	0.40	0.016	0.040	16.03	0.16	−0.39	5.67	0.29
Trans. Equip.	0.0016	0.0016	0.05	0.031	0.058	11.95	0.29	−0.07	3.56	0.61
Instruments	0.0048	0.0041	0.48	0.011	0.066	50.41	0.09	−0.47	5.58	0.25
Other	0.0014	0.0013	0.03	0.020	0.057	13.46	−0.23	−0.38	1.52	0.10
Nondurables	0.0028	0.0018	2.47	0.009	0.014	6.30	0.30	−0.25	3.28	0.45
Durables	0.0024	0.0021	0.64	0.014	0.025	14.51	0.47	−0.09	5.36	0.59
Total	0.0026	0.0019	1.88	0.011	0.017	8.49	0.44	−0.10	4.91	0.61

[a]The statistics in the table are computed for monthly logarithmic growth rate.
[b]The *Y4* results are based on the finished goods plus work-in-progress definition of output. The *t*-statistics are for tests of the hypotheses that the relevant moments are the same.

each monthly survey, units get an opportunity to revise the previous two months' information.

The BEA reports only SA data, and therefore the above procedure gives seasonally adjusted *Y4*. We create NSA shipments and inventories data using the procedures in Reagan and Sheehan (1985), West (1986), and Miron and Zeldes (1988b). For both the level of shipments and the level of inventories, the technique is to multiply the real seasonally adjusted series by a seasonal factor equal to the ratio of the seasonally unadjusted to the seasonally adjusted nominal (shipments) or book value (inventories) data. This procedure is appropriate as long as there is relatively little seasonality in prices or in the factors used to convert from book to nominal.[13] In Miron and Zeldes (1988a)

[13]The book/nominal distinction is only relevant for inventories (not for shipments).

we find that the seasonality in producer prices, while statistically significant, is sufficiently small relative to the seasonality in output that it is probably not an important factor in the reseasonalization of the data.

3. The time series properties of the two measures of production

In this section, we quantify the extent to which *IP* and *Y4* diverge and show that the differences are significant. The analysis is carried out for all 20 two-digit manufacturing industries, as well as for three aggregates of these industries (durables, nondurables, and total). We consider first the seasonally adjusted data, since these are the ones most familiar to a majority of readers. We also present results for seasonally unadjusted data, however, and we examine the seasonal movements themselves. With the exception of the variance bounds tests, the results presented below all focus on the logarithmic growth rates of the relevant series. We employ growth rates because the resulting series are likely to be stationary whether the secular growth is generated by a unit root or by a deterministic time trend.[14]

3.1. Descriptive statistics

Table 2a presents the means, standard deviations, and first-order autocorrelation coefficients of the log growth rates of the monthly seasonally adjusted *IP* and *Y4* series, as well as tests of the hypotheses that these statistics are equal for *IP* and *Y4*.[15] The sample period is May 1967 through December 1984.[16]

The results in the table indicate that the time series properties of *IP* and *Y4* are substantially different. The cross-correlations between the growth rates range from a low of -0.02 for Printing to a high of 0.64 for Primary Metals.

[14] In Miron and Zeldes (1988a) we present results of Dickey–Fuller tests of the hypothesis of no unit root in the autoregressive representation of these series. In almost all of the two-digit industries, we do not reject the null hypothesis of a unit root at the 95% level of significance.

[15] We compute these test statistics as follows. For the means, we regress the difference between the log growth rates of *IP* and *Y4* on a constant and test the hypothesis that the constant term is zero. For the variances, we regress the log growth rate of *IP* on the difference between the log growth rates of *IP* and *Y4* and test the hypothesis that the coefficient on the difference in growth rates is equal to 0.5. For the autocorrelations, we stack the *IP* and *Y4* observations and regress the growth rate of output on the lagged growth rate, a dummy that is 1 for the *IP* observations and 0 for the *Y4* observations, *and* this dummy multiplied by the lagged growth rate. The test statistic is the *t*-statistic on this last variable. For the seasonal patterns, we regress the difference in log growth rates on a constant and eleven seasonal dummies and test the hypothesis that the coefficients on the eleven dummies are jointly equal to 0. In all of these tests, we use the Hansen and Hodrick (1980) and Newey and West (1987) procedure to estimate the standard errors (with the lag length set to 12 and the damping factor set to 1.0). These test procedures therefore allow for general serial correlation and/or heteroscedasticity in the log growth rates of *IP* and *Y4*.

[16] We use only post-1967 data because there were changes in the definition of the SIC codes in 1967 that make the pre-1967 data not completely compatible with post-1967 data.

Eighteen of the twenty-three correlations are less than 0.4.[17,18] The correlations are higher for the aggregates than for the individual series.

Examination of the first-order autocorrelations reveals the surprising result that in 13 out of 23 cases the autocorrelation is positive for *IP* but negative for *Y4*. For example, for nondurables as a whole, the first-order serial correlation of growth rates equals 0.30 for *IP* and -0.25 for *Y4*. The difference in the autocorrelation coefficients is statistically significant in 17 of 23 cases. These differences are generally not eliminated over longer horizons; we find that the sum of the first 24 autocorrelations is almost always higher for *IP* than for *Y4*.[19] Thus the *IP* measure exhibits significantly more persistence than does *Y4*.

Turning to the standard deviations, the results indicate that the *IP* measure is much less volatile than the *Y4* measure. In all cases the standard deviation is higher for the *Y4* measure than for the *IP* measure, and in 12 of the industries the point estimates indicate it is more than twice as large. The differences are statistically significant in all but one case. Finally, in a few cases the mean growth rate is twice as high for one measure as for the other. The differences in means, however, are in most cases not statistically significant.[20]

In table 2b we present summary statistics and hypothesis tests for the seasonally unadjusted data. The results are similar to those in table 2a. The correlations between the two series are in every case higher than with adjusted data, reflecting the comovements due to seasonality, but the correlations are nevertheless well below 1 in most cases. The seasonal patterns are in most cases similar with respect to the timing of the peaks and troughs. In several industries, however, the magnitude of the peaks and/or troughs is substan-

[17]Harrison and Stewart (1986) report similar results for the two corresponding Canadian series. They report correlation coefficients between the detrended seasonally adjusted levels (rather than growth rates) as low as 0.56, with the majority of industries between 0.7 and 0.8. Fair (1969, p. 128) estimates the correlation coefficient between log growth rates in physical units production series and Department of Commerce shipments plus the change in inventories production series for three three-digit U.S. manufacturing industries. He finds correlation coefficients that range from 0.03 to 0.59. Sims (1974, p. 704), using U.S. manufacturing data, finds that labor input is estimated as a one-sided distributed lag of *IP* but a two-sided distributed lag of the BEA's measure of shipments or shipments plus the change in finished goods inventories (see Sims' footnote 20). Sims interprets this as evidence that shipments (as a proxy for output), or shipments plus the change in inventories, may be measured with greater error than *IP*.

[18]The correlations between *IP* and shipments (unadjusted for inventory changes) are actually greater than those between *IP* and *Y4* in 14 of 20 industries in the SA data (10 of 20 in the NSA data).

[19]Campbell and Mankiw (1988) explain why the sum of the autocorrelations is a useful, nonparametric measure of persistence.

[20]Dickey–Fuller tests on the difference in the logs of the two series indicate that in almost all cases we cannot reject the hypothesis of a unit root in this difference. The fact that the difference between the log levels of the two series is positively autocorrelated explains why we cannot usually reject the hypothesis that the growth rates are the same even though plots of the log levels in some cases diverge substantially over time.

Table 2b

Summary statistics, seasonally unadjusted data, 1967:5–1984:12.[a,b]

	Mean			Std. dev.			Autocorrelation			Seasonals	Cross-correlation
	IP	*Y4*	*t*-stat.	*IP*	*Y4*	*t*-stat.	*IP*	*Y4*	*t*-stat.	χ^2	
Food	0.0025	0.0013	1.55	0.031	0.047	8.58	0.17	−0.10	4.24	98.0	0.73
Tobacco	−0.0009	−0.0006	−0.09	0.144	0.150	0.48	−0.44	−0.37	−1.15	96.9	0.61
Textiles	0.0008	0.0011	−0.19	0.082	0.110	8.46	−0.32	−0.41	1.62	133.7	0.87
Apparel	0.0009	−0.0004	0.51	0.074	0.124	13.50	−0.32	−0.27	−0.76	133.7	0.63
Lumber	0.0009	0.0016	−0.45	0.054	0.086	11.77	−0.01	−0.16	2.35	75.3	0.72
Furniture	0.0030	0.0027	0.16	0.059	0.115	11.04	−0.39	−0.39	0.19	403.6	0.71
Paper	0.0022	0.0017	0.65	0.063	0.059	1.57	−0.30	−0.39	1.67	94.1	0.89
Printing	0.0030	0.0020	0.82	0.038	0.052	4.94	0.44	−0.29	11.88	656.4	0.37
Chemicals	0.0040	0.0025	1.30	0.027	0.057	13.33	0.06	−0.11	2.06	806.1	0.52
Petroleum	0.0011	0.0015	−0.27	0.032	0.037	2.18	0.13	−0.33	4.38	72.8	0.37
Rubber	0.0051	0.0015	2.33	0.063	0.084	5.78	−0.10	−0.25	2.10	113.4	0.79
Leather	−0.0031	−0.0035	0.19	0.082	0.110	4.53	−0.39	−0.42	0.62	66.8	0.62
Stone, Clay, Glass	0.0019	0.0003	1.68	0.043	0.065	10.65	0.10	−0.13	4.01	205.4	0.74
Primary Metal	−0.0014	−0.0017	0.26	0.064	0.072	2.10	0.15	0.02	1.91	20.7	0.85
Fab. Metal	0.0012	0.0004	0.47	0.026	0.098	45.34	−0.07	−0.37	4.49	326.1	0.63
Machinery	0.0033	0.0027	0.37	0.029	0.095	34.34	0.00	−0.31	5.53	761.6	0.65
Elec. Machinery	0.0049	0.0043	0.46	0.032	0.091	22.81	0.03	−0.28	4.30	375.5	0.70
Trans. Equip.	0.0014	0.0008	0.34	0.070	0.120	20.62	0.03	−0.01	0.82	350.3	0.84
Instruments	0.0049	0.0040	0.50	0.020	0.101	45.07	−0.09	−0.39	4.56	304.0	0.45
Other	0.0012	0.0009	0.17	0.049	0.103	14.75	−0.07	−0.19	1.86	134.9	0.62
Nondurables	0.0026	0.0014	2.27	0.035	0.045	8.88	−0.08	−0.19	1.89	244.8	0.91
Durables	0.0022	0.0016	0.61	0.032	0.078	24.62	0.01	−0.14	3.08	766.4	0.89
Total	0.0024	0.0015	1.25	0.032	0.060	21.52	−0.05	−0.16	2.07	819.7	0.92

[a] The statistics in the table are computed for monthly logarithmic growth rates.

[b] The *Y4* results are based on the finished goods plus work-in-progress definition of output. The *t*-statistics are for tests of the hypotheses that the relevant moments are the same. The χ^2-statistic is for the test of the hypothesis that the seasonal patterns in the log growth rates of *IP* and *Y4* are the same. The 95% critical value of the $\chi^2(11)$ is 19.67.

tially greater for *Y4* than for *IP*.[21] Hypothesis tests indicate that the seasonal coefficients are statistically different in all 23 cases.

In Miron and Zeldes (1988a) we present results analogous to those in table 2a for growth rates of quarterly and annual averages of monthly data, respectively. The cross-correlations of the quarterly growth rates are higher than those of the monthly data, but still less than 0.7 in half of the industries. The correlations of annual growth rates are significantly higher, being greater than 0.9 in 15 out of 20 industries. The generally high correlations of the annual growth rates are consistent with the fact that the information in the

[21] For plots of the seasonal patterns, see Miron and Zeldes (1988a).

Annual Survey of Manufactures and the quinquennial Census of Manufactures is in most cases used to benchmark both *IP* and *Y4*.

The evidence presented above demonstrates that there are dramatic differences between the time series properties of the *IP* and *Y4* measures of production. The standard deviations and autocorrelations of the two series differ systematically, and the cross-correlations between the two series indicate that there is remarkably little variation that is common to both series. We have discussed the differences with researchers at the BEA and FRB, and, while they are aware of the problem and of numerous differences in the construction of the data, they are not able to offer a definitive explanation.[22]

4. The variance of production and the variance of sales

In this section we underscore, by way of an example, the economic importance of the discrepancies between the two measures of production. We show that the results of Blinder's (1986) widely cited study of firms' inventory behavior are at least partially sensitive to the choice of output measure. Blinder (1986) emphasizes that, in the absence of cost shocks, the production smoothing model implies that the variance of production should be less than the variance of sales (shipments).[23] Using the *Y4* measure of output, Blinder shows that the variance of production is greater than the variance of shipments for all but one of the industries examined, and he interprets this as strong evidence against the production smoothing model.

In table 3 we present the ratio of the variance of output to the variance of shipments based on each of the two output measures. The sample period, inventory definition, and detrending techniques were all chosen to correspond as closely as possible to Blinder (1986). Thus, unlike the data in the previous tables, these data are levels (not growth rates), detrended with an exponential trend, and cover the period 1959:2–1981:7.[24] We convert the *IP* measure from

[22] One possible source of discrepancy is that FRB uses value-added weights to aggregate the individual series, while the BEA, by adding constant dollar series, effectively uses gross value weights. We examine an alternate *IP* series calculated by the FRB using gross value weights (for seasonally adjusted total manufacturing). Over the period 1972:3–1984:12, we find that the correlation of growth rates between this series and the standard *IP* is equal to 0.83. The correlation between this series and *Y4* is equal to 0.59, compared to a correlation of the standard *IP* and *Y4* of 0.65 over the same time period. Thus, the use of value-added weights does not appear to be a quantitatively important source of the difference between *IP* and *Y4*.

[23] If cost shocks are present, then this inequality need not hold [Eichenbaum (1984), Blinder (1986)]. Kahn (1987) argues that, if production for the period must be chosen before sales are known and if stockouts are possible, then this inequality can be violated even in the absence of cost shocks.

[24] For a detailed description, see Blinder (1986) and Miron and Zeldes (1988a).

Table 3

Variance of production over variance of sales, seasonally adjusted data, 1959:2–1981:7.[a,b]

	IP	*Y4*
Food	0.62	1.20
Tobacco	0.54	2.43
Textiles	1.17	1.06
Apparel	0.91	1.38
Lumber	0.94	1.12
Furniture	0.96	1.24
Paper	1.42	1.02
Printing	1.15	1.18
Chemicals	0.82	1.01
Petroleum	0.59	1.06
Rubber	1.12	1.13
Leather	1.09	1.36
Stone, Clay, Glass	1.08	1.12
Primary Metal	0.98	0.96
Fab. Metal	0.59	1.13
Machinery	1.28	1.35
Elec. Machinery	1.13	1.26
Trans. Equip.	0.84	1.23
Instruments	0.75	1.81
Other	0.81	1.42
Nondurables	1.52	1.05
Durables	1.13	1.19
Total	1.31	1.14

[a]The statistics in the tables are computed for deviations from exponential trend; see text for details.

[b]The *Y4* results are based on the finished goods plus work-in-progress definition of output.

an index into a constant dollar figure by multiplying it by the ratio of average *Y4* to average *IP*.[25]

The results for the *Y4* measure match Blinder's results almost exactly. For all but one industry, the variance of output is greater than the variance of shipments. The results for the *IP* measure, however, are quite different. The variance ratio is in most cases less than the one based on *Y4*, and for 11 industries the variance inequality is actually reversed.[26] It is especially noteworthy that reversals occur in five of the six industries identified by Belsley (1969) as production to stock; these are the industries for which the produc-

[25]Because the ratio of average *Y4* to average *IP* is different for different averaging periods, the choice of base period for conversion of *IP* sometimes affects the resulting variance bounds ratio for a few industries. Results for different base periods consistently show, however, that the variance ratio using *IP* data is less than the ratio using *Y4* data in a significant number of cases.

[26]This reversal of the variance bounds inequality was first pointed out by West (1986, fn 13).

tion smoothing model is the most plausible theoretically.[27] Had Blinder originally chosen to use *IP* instead of *Y4*, he would have reached substantially different conclusions about the empirical validity of the production smoothing model.[28]

5. The sources and importance of measurement error

The fact that *IP* and *Y4* differ means that at least one of them is measured with error. So far, none of our results has required making assumptions about the types of measurement error present in the data. In this section we model the measurement error in the series. We begin with a general model and then gauge the importance of different types of measurement error under alternative assumptions about the type of measurement error present. We also attempt (only partially successfully) to determine which types of measurement error are most likely to be present.

We consider two types of measurement error, as in Mankiw, Runkle, and Shapiro (1984) and Mankiw and Shapiro (1986). The first type, classical measurement error, is uncorrelated with the true underlying series. The second type is correlated with the true series but uncorrelated with the observed series. This second type of measurement error could arise for two reasons. If the announced series are rational forecasts of the underlying series, then the measurement error will be a rational expectations forecast error and thus uncorrelated with the forecast itself. In addition, if there are productivity changes (true productivity shocks or, e.g., changes in productivity due to labor hoarding) that are not captured by the measured series, then the measurement error will include the productivity change, which will be correlated with the true output but may be uncorrelated with the measured figures.

We consider the following model:

$$y_t^{IP} + u_t^{IP} = y_t^* + e_t^{IP}, \qquad y_t^{Y4} + u_t^{Y4} = y_t^* + e_t^{Y4}, \tag{1}$$

where y_t^* is the log of the true series; y_t^{IP} and y_t^{Y4} are the logs of the two measured series; e_t^{IP} and e_t^{Y4} are measurement errors that are assumed uncorrelated with y_t^* (and y_{t-1}^*, y_{t+1}^*); and u_t^{IP} and u_t^{Y4} are measurement errors that are assumed uncorrelated with y_t^{IP} ($y_{t-1}^{IP}, y_{t+1}^{IP}$) and y_t^{Y4} ($y_{t-1}^{Y4}, y_{t+1}^{Y4}$), respectively. The cross-correlations of the errors $e_t^{IP}, e_t^{Y4}, u_t^{IP}$, and u_t^{Y4} are

[27] These six industries are food, tobacco, apparel, chemicals, petroleum and rubber.

[28] While this may be interpreted as support for the production smoothing model, Miron and Zeldes (1988b) present additional tests and find that a generalized production smoothing model is rejected for both the *Y4* data and the *IP* data (although the rejections are not as strong based on *IP* data).

assumed to be 0 at lags 0 and 1.[29,30] The model in first differences is then

$$\Delta y_t^{IP} + \Delta u_t^{IP} = \Delta y_t^* + \Delta e_t^{IP}, \qquad \Delta y_t^{Y4} + \Delta u_t^{Y4} = \Delta y_t^* + \Delta e_t^{Y4}, \tag{2}$$

where the contemporaneous orthogonality conditions assumed above hold for these first differences.

Ideally, we would like to be able to answer the following questions. First, how poorly do the two series measure true production, i.e., how large is the variance of the measurement error relative to the variance of the true series? Second, which of the output measures is a better proxy for output, i.e., for which series is the variance of the measurement error smaller?

Each of these questions could be answered if we could estimate the population moments $V(\Delta u^{IP})$, $V(\Delta e^{IP})$, $V(\Delta u^{Y4})$, $V(\Delta e^{Y4})$, and $V(\Delta y^*)$. Under the assumptions given so far, these moments are related to the moments of the measured series as follows:

$$\operatorname{var}(\Delta y_t^{IP}) = \operatorname{var}(\Delta y_t^*) + \operatorname{var}(\Delta e_t^{IP}) - \operatorname{var}(\Delta u_t^{IP}), \tag{3}$$

$$\operatorname{var}(\Delta y_t^{Y4}) = \operatorname{var}(\Delta y_t^*) + \operatorname{var}(\Delta e_t^{Y4}) - \operatorname{var}(\Delta u_t^{Y4}), \tag{4}$$

$$\operatorname{cov}(\Delta y_t^{IP}, \Delta y_t^{Y4}) = \operatorname{var}(\Delta y_t^*) - \operatorname{var}(\Delta u_t^{Y4}) - \operatorname{var}(\Delta u_t^{IP}). \tag{5}$$

Unfortunately, obtaining three sample moments does not enable us to estimate five population moments (although we can estimate certain combinations of the moments).

As an example of the difficulties with inference, note that a large variance of e-type measurement error tends to increase the variance of measured output, but a large variance of u-type measurement error tends to decrease the variance of measured output. Thus, observing (as we have) that the log growth rate of *IP* has a smaller variance than *Y4* could mean that *IP* is a better series (smaller classical type measurement error) or that *IP* is a worse series (larger

[29] We make no assumptions about the autocorrelation structure of the individual measurement errors.

[30] The assumption that e_{t+1} and e_{t-1} are uncorrelated with y_t^* seems plausible in the case of pure sampling error. Likewise, if the u's are unmeasured productivity shocks, then the assumption that u_{t-1} and u_{t+1} are uncorrelated with the corresponding y_t seems no stronger than the assumption of contemporaneous orthogonality. If u_t is a rational forecast error based only on time t information, then the assumption that u_{t-1} is uncorrelated with y_t is not in general valid. Since the reported numbers are revised, however, they should properly be viewed as forecasts based on information through the end of the sample. In this case the forecast errors will be orthogonal to forecasts at all leads and lags in the sample.

artificial smoothing or larger rational forecast error).[31] Distinguishing between these possibilities and estimating the relevant parameters requires either additional identifying assumptions or additional information.

In section 5.1 below, we calculate the importance of measurement error under the assumption that all measurement error is of the e-type. In the following section, we make the calculations under the other extreme assumption that all measurement error is of the u-type. In each case, we use the sample moments to estimate the fractions (κ^{IP}, κ^{Y4}) of the total variance of each series due to measurement error and the ratio (λ^{IP}) of the variance of the measurement error in IP to the sum of the variances of the measurement errors in IP and $Y4$. Under the assumptions in this paper, the optimal indicator of the true series based solely on the contemporaneous observations is a linear combination of the two series, and the weight on IP is equal to λ^{IP}.[32] Finally, in section 5.3 we attempt to use information about the differences in construction of IP across industries to shed light on which type of measurement error is likely to be most important.

5.1. Classical measurement error (e)

Assume for the moment that $V(\Delta u^{IP}) = V(\Delta u^{Y4}) = 0$ so that all measurement error is the classical type. Under this assumption, $\mathrm{cov}(\Delta y^{IP}, \Delta y^{Y4}) = V(\Delta y^*)$.[33] We define $\kappa^{IP} = \mathrm{var}(\Delta e^{IP})/\mathrm{var}(\Delta y^{IP})$ (κ^{Y4} is defined analogously). In the absence of measurement error, a regression of one series on the other would yield a coefficient of unity. With e-type measurement error, the coefficient will be biased downward, and the bias (the difference between unity and the coefficient) will be a consistent estimate of κ for the right-hand-side

[31]As long as the serial correlation in the growth rate of the measurement errors is less than that in the growth rate of the true series, higher measurement error *decreases* the autocorrrelation of the growth rate of the measured series in a model with only e-type measurement error, while higher measurement error variance *increases* the autocorrelation of the growth rate of the measured series with only u-type measurement error. Thus, the fact that the growth rate of $Y4$ has a lower autocorrelation than IP could again be evidence of a large e-type measurement error in $Y4$ or a large u-type measurement error in IP.

[32]See de Leeuw and McKelvey (1983) for the case of e-type measurement error. An analogous argument carries through for u-type measurement error. These each assume that the optimal weights sum to 1, which might not be the case if some weight is put on the unconditional mean. Also, if the first differences of the measurement error are serially correlated, a superior indicator could be constructed using both contemporaneous and past values of the two series.

[33]Prescott (1986) makes this observation and estimates the variance of true hours of employment based on household and firm measures of hours. Lichtenberg and Griliches (1989) also assume that only classical measurement error is present and estimate the same variance ratios as we do in this section for two measures of prices. They examine long-run inflation rates and base their measurement error estimates on sample moments computed across industries for a single time period, rather than across time for a single industry as is done here.

variable.[34] Thus, regressing *IP* on *Y4* gives us information about the measurement error in *Y4*, and regressing *Y4* on *IP* gives us information about the measurement error in *IP*.[35]

The results are presented in table 4. Looking at the seasonally adjusted data, these estimates indicate that in all but one industry at least 60% of the variation in the growth rate of *Y4* is due to measurement error, and in 16 out of 20 industries it is over 80%. Looking at *IP*, we find that in 15 out of 20 industries measurement error accounts for over 35% of the variation in the monthly growth rate. The estimated standard errors of the ratios indicate that in most cases they are estimated precisely. When we turn to the seasonally unadjusted data, we find a different set of results. Relative to the seasonally adjusted data, the measurement error shares are estimated to be smaller for *Y4* and *IP*, and often negative for *IP*. The ratios should be smaller if seasonality in the measurement error is small relative to the seasonality in the true series. However, the negative estimates suggest a misspecification, to which we return in section 5.3.[36]

The seasonally adjusted results in table 4 indicate that the optimal weight λ^{IP} $(= V(\Delta e^{Y4})/[V(\Delta e^{Y4}) + V(\Delta e^{IP})])$ is in all but one case significantly greater than 0.5.[37,38] This indicates that under the assumption that all measurement error is of the classical (e) type, *IP* is the better measure of output: the variance in its measurement error is less, and an optimal indicator would place more weight on it. However, there is evidence of substantial measurement error in both series.

5.2. u-type measurement error

In this section, we make the opposite polar assumption: $V(\Delta e^{IP}) = V(\Delta e^{Y4}) = 0$, so that all measurement error is orthogonal to the measured series. Here

[34] Call $\hat{\beta}_{Y4,\,IP}$ the estimated coefficient when Δy^{Y4} is regressed on Δy^{IP}. Then

$$\begin{aligned}\text{plim}(1 - \hat{\beta}_{Y4,\,IP}) &= 1 - \text{cov}(\Delta y^{IP}, \Delta y^{Y4})/\text{var}(\Delta y^{IP}) \\ &= 1 - \text{var}(\Delta y^*)/[\text{var}(\Delta y^*) + \text{var}(\Delta e^{IP})] \\ &= \text{var}(\Delta e^{IP})/[(\text{var}(\Delta y^*) + \text{var}(\Delta e^{IP})] \\ &= \text{var}(\Delta e^{IP})/\text{var}(\Delta y^{IP}) = \kappa^{IP}.\end{aligned}$$

[35] Using this simple regression technique has the advantage of enabling us to calculate in a straightforward way the standard errors of these variance ratios. We employ the Hansen and Hodrick (1980) procedure, as modified by Newey and West (1987), to calculate standard errors that are consistent given the serial correlation in the residuals.

[36] We also calculated the tables using seasonal dummy adjusted data, and the results were similar to the SA results in table 4.

[37] We again use a simple regression technique to calculate an estimate of λ^{IP} and its standard error. The coefficient in the regression of Δy^{Y4} on $\Delta y^{Y4} - \Delta y^{IP}$ is a consistent estimate of λ^{IP}.

[38] The seasonally unadjusted results indicate that the optimal weight on *IP* is greater than 1, again suggesting the possibility of misspecification.

we define κ as the variance of the measurement error as a fraction of the variance in the *true* series,[39] and λ^{IP} is defined as $V(\Delta u^{Y4})/[V(\Delta u^{IP})+V(\Delta u^{Y4})]$. These numbers are presented in table 5.[40]

Looking at the SA data, the κ's indicate that measurement error continues to represent a substantial part of the variation in each series. The numbers for *IP* are in all but three cases greater than 60%. The ratios for *Y4* are generally in the neighborhood of 10–20%. Note that λ^{IP} in this table is exactly 1 minus the value of λ^{IP} in table 4. In other words, while the numbers in table 4 indicate that *Y4* contains more measurement error than *IP*, the numbers in table 5 calculated under the alternative assumption about the type of measurement error suggest the opposite conclusion: that the measurement error in *IP* is worse than that in *Y4*.

The seasonally unadjusted results are also in table 5. Unfortunately, the negative estimates for κ^{Y4} and λ^{IP} again suggest a misspecification in the NSA data.

5.3. *Including both types of measurement error and attempting to distinguish between them*

What can be said if we are not willing to take as strong a stand as to which type of measurement error is present in the two series? We make two observations. First, by combining sample moments we can estimate the sum of the variances of Δe^{Y4} and Δu^{IP} and also the sum of the variances of Δe^{IP} and Δu^{Y4}. The ratio of the κ's in table 5 indicates that the former sum is dramatically greater than the latter.[41] Thus, e^{Y4} and u^{IP} together appear to constitute the primary sources of measurement error. However, we cannot tie down the importance of e^{Y4} relative to that of u^{IP}.

Second, the fact that there is no issue in producing *Y4* about adjustments for missing productivity shocks suggests that u^{Y4} may be an unimportant source of measurement error. In this case, the estimates of the size of e-type measurement error in *IP* in table 4 are accurate estimates, whether or not there is u-type measurement error in *IP*.

[39] Thus, for example, $\kappa^{IP}=V(\Delta u^{IP})/V(\Delta y^*)$. We scale by the variance of Δy^* rather than Δy, so that the ratio will be interpretable as a fraction between 0 and 1 [recall that in this model, $V(\Delta y^*)=V(\Delta y)+V(\Delta u)$].

[40] We estimate κ^{IP} by estimating $1-[V(\Delta y^{IP})/(V(\Delta y^{IP})+V(\Delta y^{Y4})-\text{cov}(\Delta y^{IP},\Delta y^{Y4}))]$, which, under these assumptions, is equal to $V(\Delta u^{IP})/V(\Delta y^*)$. Reversing *IP* and *Y4* gives the analogous ratio for *Y4*.

[41] Under the model with both e- and u-type errors, $V(\Delta y_t^{IP})-\text{cov}(\Delta y_t^{IP},\Delta y_t^{Y4})=V(\Delta e_t^{IP})+V(\Delta u_t^{Y4})$ and $V(\Delta y_t^{Y4})-\text{cov}(\Delta y_t^{IP},\Delta y_t^{Y4})=V(\Delta e_t^{Y4})+V(\Delta u_t^{IP})$. The ratio of κ^{IP} to κ^{Y4} in table 5 is equal to $[V(\Delta e^{Y4})+V(\Delta u^{IP})]/[V(\Delta e^{IP})+V(\Delta u^{Y4})]$.

Table 4

Estimates of e-type measurement errors, 1967:5–1984:12.[a,b]

	Seasonally adjusted						Seasonally unadjusted					
	$\hat{\kappa}_{IP}$	se	$\hat{\kappa}_{Y4}$	se	$\hat{\lambda}_{IP}$	se	$\hat{\kappa}_{IP}$	se	$\hat{\kappa}_{Y4}$	se	$\hat{\lambda}_{IP}$	se
Food	0.59	0.16	0.92	0.03	0.89	0.03	−0.10	0.09	0.52	0.04	1.09	0.07
Tobacco	0.60	0.15	0.91	0.04	0.87	0.05	0.37	0.07	0.42	0.09	0.55	0.11
Textiles	0.59	0.11	0.84	0.06	0.79	0.08	−0.17	0.05	0.35	0.03	1.37	0.10
Apparel	0.79	0.21	0.96	0.03	0.87	0.04	−0.06	0.07	0.62	0.06	1.03	0.04
Lumber	0.38	0.19	0.83	0.07	0.89	0.05	−0.15	0.07	0.55	0.04	1.12	0.05
Furniture	0.52	0.22	0.95	0.02	0.94	0.03	−0.39	0.14	0.63	0.02	1.19	0.06
Paper	0.52	0.10	0.70	0.11	0.68	0.07	0.15	0.03	0.06	0.04	0.26	0.15
Printing	1.07	0.25	1.01	0.03	0.87	0.03	0.51	0.07	0.73	0.03	0.72	0.05
Chemical	0.71	0.10	0.90	0.04	0.79	0.03	−0.11	0.14	0.76	0.02	1.03	0.04
Petroleum	0.88	0.09	0.95	0.04	0.70	0.06	0.58	0.06	0.68	0.05	0.60	0.05
Rubber	0.49	0.09	0.74	0.07	0.75	0.09	−0.06	0.03	0.41	0.04	1.08	0.10
Leather	0.78	0.21	0.97	0.03	0.89	0.03	0.17	0.09	0.54	0.03	0.85	0.08
Stone, Clay, Glass	0.43	0.10	0.81	0.04	0.85	0.05	−0.13	0.07	0.51	0.03	1.12	0.06
Primary Metals	0.35	0.10	0.36	0.10	0.51	0.13	0.05	0.06	0.24	0.05	0.85	0.17
Fab. Metals	0.09	0.25	0.93	0.02	0.99	0.02	−1.33	0.26	0.83	0.02	1.13	0.01
Machinery	0.12	0.12	0.87	0.02	0.98	0.02	−1.10	0.10	0.80	0.02	1.15	0.02
Elec. Machinery	0.32	0.17	0.88	0.04	0.94	0.03	−0.97	0.13	0.76	0.02	1.19	0.03
Trans. Equip.	−0.13	0.10	0.67	0.07	1.06	0.05	−0.44	0.06	0.51	0.03	1.42	0.04
Instruments	−0.54	0.42	0.96	0.01	1.01	0.01	−1.27	0.28	0.91	0.01	1.06	0.01
Other	0.72	0.21	0.96	0.02	0.91	0.03	−0.32	0.14	0.71	0.02	1.11	0.04
Nondurables	0.31	0.07	0.71	0.08	0.84	0.05	−0.17	0.05	0.29	0.03	1.55	0.12
Durables	−0.07	0.08	0.68	0.04	1.03	0.04	−1.16	0.08	0.63	0.02	1.46	0.04
Total	0.02	0.09	0.63	0.05	0.99	0.06	−0.75	0.07	0.52	0.02	1.66	0.05

[a]The statistics in the table are computed for monthly logarithmic growth rates.

[b]The *Y4* results are based on the finished goods plus work-in-progress definition of output.

λ^{IP} is the weight on *IP* in an optimal forecast of the true series. We estimate $\lambda^{IP} = \mathrm{cov}(\Delta y^{Y4}, \Delta y^{Y4} - \Delta y^{IP})/\mathrm{var}(\Delta y^{Y4} - \Delta y^{IP})$. Under the assumption that $\mathrm{var}(\Delta u^{Y4}) = \mathrm{var}(\Delta u^{IP}) = 0$, $\lambda^{IP} = \mathrm{var}(\Delta e^{Y4})/(\mathrm{var}(\Delta e^{Y4}) + \mathrm{var}(\Delta e^{IP}))$.

κ^{IP} is the fraction of the variation in *IP* due to measurement error. We estimate $\kappa^{IP} = 1 - \mathrm{cov}(\Delta y^{IP}, \Delta y^{Y4})/\mathrm{var}(\Delta y^{IP})$. Under the assumption that $\mathrm{var}(\Delta u^{Y4}) = \mathrm{var}(\Delta u^{IP}) = 0$, $\kappa^{IP} = \mathrm{var}(\Delta e^{IP})/\mathrm{var}(\Delta y^{IP})$. κ^{Y4} is defined similarly.

While we have exhausted all of the information in the sample moments within each industry, we next consider using information about variations in moments and construction of the *IP* data *across* industries. An important part of the measurement error in *IP* may be due to the fact that a large number of the individual *IP* series are constructed from input data. The use of input data could add either e-type measurement error because it involves added noise, or u-type error because the use of inputs to measure output may omit productiv-

Table 5

Estimates of u-type measurement errors, 1967:5–1984:12.[a,b]

	Seasonally adjusted				Seasonally unadjusted			
	$\hat{\kappa}_{IP}$	$\hat{\kappa}_{Y4}$	$\hat{\lambda}_{IP}$	se	$\hat{\kappa}_{IP}$	$\hat{\kappa}_{Y4}$	$\hat{\lambda}_{IP}$	se
Food	0.83	0.10	0.11	0.03	0.54	−0.05	−0.09	0.07
Tobacco	0.81	0.12	0.13	0.04	0.31	0.25	0.45	0.11
Textiles	0.69	0.18	0.21	0.06	0.38	−0.10	−0.37	0.10
Apparel	0.84	0.12	0.13	0.03	0.63	−0.02	−0.03	0.04
Lumber	0.76	0.09	0.11	0.07	0.59	−0.06	−0.11	0.05
Furniture	0.90	0.05	0.06	0.02	0.71	−0.11	−0.19	0.06
Paper	0.53	0.25	0.32	0.11	0.05	0.14	0.74	0.15
Printing	0.88	0.13	0.13	0.03	0.57	0.22	0.28	0.05
Chemicals	0.73	0.19	0.21	0.04	0.78	−0.02	−0.03	0.04
Petroleum	0.68	0.29	0.30	0.04	0.47	0.31	0.40	0.05
Rubber	0.59	0.20	0.25	0.07	0.42	−0.03	−0.08	0.10
Leather	0.86	0.11	0.11	0.03	0.49	0.09	0.15	0.08
Stone, Clay, Glass	0.70	0.13	0.15	0.04	0.54	−0.06	−0.12	0.06
Primary Metals	0.27	0.26	0.49	0.10	0.23	0.04	0.15	0.17
Fab. Metals	0.92	0.01	0.01	0.02	0.92	−0.11	−0.13	0.01
Machinery	0.86	0.02	0.02	0.02	0.89	−0.12	0.15	0.02
Elec. Machinery	0.84	0.05	0.06	0.04	0.86	−0.14	−0.19	0.03
Trans. Equip.	0.70	−0.04	−0.06	0.07	0.60	−0.18	−0.42	0.04
Instruments	0.97	−0.01	−0.01	0.01	0.96	−0.05	−0.06	0.01
Other	0.88	0.08	0.09	0.02	0.76	−0.08	−0.11	0.04
Nondurables	0.63	0.12	0.16	0.08	0.32	−0.12	−0.55	0.12
Durables	0.69	−0.02	−0.03	0.04	0.79	−0.25	−0.46	0.04
Total	0.62	0.01	0.01	0.05	0.65	−0.26	−0.66	0.05

[a]The statistics in the table are computed for monthly logarithmic growth rates.

[b]The $Y4$ results are based on the finished goods plus work-in-progress definition of output.

λ^{IP} is the weight on IP in an optimal forecast of the true series. We estimate $\lambda^{IP} = 1 - \text{cov}(\Delta y^{Y4}, \Delta y^{Y4} - \Delta y^{IP})/\text{var}(\Delta y^{Y4} - \Delta y^{IP})$. Under the assumption that $\text{var}(\Delta e^{Y4}) = \text{var}(\Delta e^{IP}) = 0$, $\lambda^{IP} = \text{var}(\Delta u^{Y4})/(\text{var}(\Delta u^{Y4}) + \text{var}(\Delta u^{IP}))$.

κ^{IP} is the fraction of the variation in IP due to measurement error. We estimate $\kappa^{IP} = 1 - \text{var}(\Delta y^{IP})/(\text{var}(\Delta y^{IP}) + \text{var}(\Delta y^{Y4}) - \text{cov}(\Delta y^{IP}, \Delta y^{Y4}))$. Under the assumption that $\text{var}(\Delta e^{Y4}) = \text{var}(\Delta e^{IP}) = 0$, $\kappa^{IP} = \text{var}(\Delta u^{IP})/\text{var}(\Delta y^{*})$. κ^{Y4} is defined similarly.

ity changes.[42] If the error induced by the use of input data is primarily of the smoothing variety (u^{IP}), then industries based most on input data should exhibit the most smoothing, and thus should have the lower variance (relative to that of $Y4$). However, if the use of input data simply causes more classical measurement error (e^{IP}), the high input data industries should have relatively high variance. We estimate the Spearman rank correlation between the use of input data (as reported in table 1) and the difference of the variances of IP

[42]As indicated previously, the FRB attempts to correct for this by using cyclically adjusted PFC's. Here we allow for the possibility that this adjustment does not fully capture productivity changes.

and *Y4* to be -0.48 and -0.78 for SA and NSA data respectively, each significant at the 5% level.[43]

This negative correlation provides evidence that *u*-type measurement error is present in the *IP* data and is related to the use of inputs. The interpretation is that the use of inputs for *IP* serves to artificially smooth the data and therefore that the relatively low standard deviation of *IP* should not necessarily be taken to mean that it is a better measure of true output.

One final point, suggested by the NSA results, concerns the possibility that the true coefficients relating y^{IP} and y^{Y4} to y^* are not equal to 1. Recall that in the model above the regression of Δy^{IP} on Δy^{Y4} and the regression of Δy^{Y4} on Δy^{IP} each give a coefficient that is biased downward from the true coefficient of 1. However, in the NSA results, the regression of Δy^{Y4} on Δy^{IP} in many cases gave a coefficient greater than 1, causing the negative estimates of κ^{IP} in table 4 and κ^{Y4} in table 5. When the FRB does productivity adjustments within the year, they assume (approximately) that the elasticity of output with respect to inputs is equal to 1. They then modify this judgementally to account for cyclical factors but make no adjustment for seasonal factors. If the elasticity of output with respect to inputs is in fact greater than 1, due to either labor hoarding or productivity shocks correlated with input use, then the FRB procedure biases upward the coefficient in the regression of Δy^{Y4} on Δy^{IP}. In fact, there is a strong positive rank correlation (0.65) between the use of inputs and the regression coefficients of Δy^{Y4} on Δy^{IP} using NSA data.[44] Users of NSA data should thus be aware that in the industries based primarily on inputs, the seasonal movements in *IP* appear to significantly understate the true seasonal variation in output.

6. Concluding remarks

In this paper we have documented the radically different time series properties of two different measures of monthly manufacturing output. Under specific assumptions about the nature of the measurement error, we estimate that a large fraction of the variation in the observed growth rate of both measures of output is due to measurement error. The results suggest that empirical analyses that rely heavily on the time series properties of the month-to-month variation in either of these measures may give very misleading results.

[43] The Spearman correlations between the ratio of variances and the use of inputs are even stronger: -0.75 and -0.82 for SA and NSA data, respectively.

[44] This problem may be present in the SA data as well, although the fact that the coefficients are almost never greater than 1 in the SA data suggests that this may be less of a problem in SA data. Even if it is present in the SA data, this will bias *downward* the estimates of *e*-type measurement error in *IP* in table 4, i.e., even if this is an issue, we can still conclude that measurement error is an important part of industrial production data under an *e*-type model of measurement error.

There are ample reasons for the presence of noise in both series. For $Y4$, real monthly inventory changes are difficult to estimate from book value data. For IP, the use of inputs to proxy for output involves strong assumptions about productivity. In both cases, there is likely to be standard sampling error.

We find evidence that the error is concentrated in the sum of u^{IP} and e^{Y4}, and that the use of inputs adds to u^{IP}, thus artificially smoothing the data. Unfortunately, we are unable to estimate the importance of u^{IP} relative to e^{Y4}, so we cannot unambiguously recommend one measure over the other. The finding that most of the measurement error is predominantly in u^{IP} or e^{Y4}, however, is likely to be useful in certain contexts in choosing between these two measures of production.

References

Belsley, David A., 1969, Industrial production behavior: The order-stock distinction (North-Holland, Amersterdam).

Blinder, Alan, 1986, Can the production smoothing model of inventory behavior be saved?, Quarterly Journal of Economics 101, 431–453.

Campbell, John Y. and N. Gregory Mankiw, 1988, International evidence on the persistence of economic fluctuations, NBER working paper no. 2498.

de Leeuw, Frank and Michael J. McKelvey, 1983, A 'true' time series and its indicators, Journal of the American Statistical Association 7, 37–46.

Eichenbaum, Martin S., 1984, Rational expectations and the smoothing properties of inventories of finished goods, Journal of Monetary Economics 14, 71–96.

Fair, Ray C., 1969, The short-run demand for workers and hours (North-Holland, Amersterdam).

Federal Reserve Board, 1986, Industrial production (Federal Reserve Board, Washington, DC).

Foss, Murray F., Gary Fromm, and Irving Rottenberg, 1980, Measurement of business inventories (Government Printing Office, Washington, DC).

Hansen, Lars P. and Robert J. Hodrick, 1980, Forward rates as optimal predictors of future spot rates: An econometric analysis, Journal of Political Economy 88, 829–853.

Harrison, Alan and Mark Stewart, 1986, Conditional strike-settlement probabilities and the cost of a strike, Manuscript (McMaster University, Hamilton, Ont.).

Hinrichs, John C. and Anthony D. Eckman, 1981, Constant dollar manufacturing inventories, Survey of Current Business 61, Nov., 16–23.

Holtz-Eakin, Douglas and Alan Blinder, 1983, Constant dollar manufacturers' inventories: A note, Manuscript (Columbia University, New York, NY).

Kahn, James A., 1987, Inventories and the volatility of production, American Economic Review 77, 667–679.

Lichtenberg, Frank R. and Zvi Griliches, 1989, Errors of measurement in output deflators, Journal of Business and Economic Statistics, forthcoming.

Maccini, Louis J. and Robert Rossana, 1984, Joint production, quasi-fixed factors of production, and investment in finished goods inventories, Journal of Money, Credit and Banking 16, 218–236.

Mankiw, N. Gregory, David E. Runkle, and Matthew D. Shapiro, 1984, Are preliminary announcements of the money stock rational forecasts?, Journal of Monetary Economics 14, 15–27.

Mankiw, N. Gregory and Matthew D. Shapiro, 1986, News or noise: An analysis of GNP revisions, Survey of Current Business, May, 20–25.

Miron, Jeffrey A. and Stephen P. Zeldes, 1988a, Production, sales, and the change in inventories: An identity that doesn't add up, NBER working paper no. 2765.

Miron, Jeffrey A. and Stephen P. Zeldes, 1988b, Seasonality, cost shocks, and the production smoothing model of inventories, Econometrica 56, 877–908.

Newey, Whitney K. and Kenneth D. West, 1987, A simple, positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix, Econometrica 55, 703–708.
Prescott, Edward C., 1986, Theory ahead of business cycle measurement, Carnegie–Rochester Conference Series 25, 11–44.
Reagan, Patricia and Dennis P. Sheehan, 1985, The stylized facts about the behavior of manufacturers' inventories and backorders over the business cycle: 1959–80, Journal of Monetary Economics 15, 217–246.
Sims, Christopher A., 1974, Output and labor input in manufacturing, Brooking Papers on Economic Activity 3, 695–735.
Stockman, Alan C., 1988, Sectoral and national aggregate disturbances to industrial output in seven European countries, Journal of Monetary Economics 21, 387–409.
U.S. Department of Commerce, Bureau of the Census, 1983, Census of manufactures, 1977, indexes of production (Government Printing Office, Washington, DC).
West, Kenneth D., 1983, A note on the econometric use of constant dollar inventory series, Economics Letters 13, 337–341.
West, Kenneth D., 1986, A variance bounds test of the linear quadratic inventory model, Journal of Political Economy 94, 374–401.