V

Expenditures Estimates Compared with Income Estimates

Accuracy of Successive Estimates

One conclusion of the preceding analysis is that the provisional estimates can be viewed as relying partly on extrapolations, and that the revisions are primarily a measure of the extrapolation error. Since less extrapolation would be required for the income side of the accounts, we might then expect that the estimates of GNP derived from income data would be revised less and that they might give a better indication of the final expenditures estimates than the initial expenditures estimates do. For example, the errors of benchmark extrapolation that are included in the expenditures estimates have few counterparts in the income estimates. Moreover, there is no body of data for the expenditures estimates that is comparable to the highly reliable wage and salary component of the income estimates. On the other hand, the data for individual proprietors' income are relatively unreliable, and there are no early data on profits.

Table 12 compares the accuracy of the two sets of estimates. Not surprisingly, it shows that the income estimates are revised less than the expenditures estimates (compare lines 1 through 4 with the corresponding entries in lines 5–8).

If the income estimates (A_j^*) contained less error than the corresponding expenditures estimates (A_j), they would be better predictions of A_n. With only one exception, the estimates of quarterly levels and changes in GNP based on income (A_j^*) do show slightly greater over-all accuracy than A_j (columns 3 and 6, lines 1–4 compared with lines 9–12). They are less biased initially, though the advantage decreases as

TABLE 12. Errors in Expenditures Estimates Compared with Errors in Income Estimates of Quarterly Levels and Changes in Gross National Product, 1947 II–1961 IV (billion dollars)

		Quarterly Levels			Quarterly Changes		
Line	Code of Estimate[a]	Mean Error (1)	Standard Deviation of Error (2)	Root Mean Square Error (3)	Mean Error (4)	Standard Deviation of Error (5)	Root Mean Square Error (6)
		EXPENDITURES ESTIMATES (GNP)[b]					
1	A_0	−8.9	5.6	10.5	−0.6	3.2	3.2
2	A_1	−6.1	4.8	7.7	−0.4	2.9	2.9
3	A_2	−5.3	4.2	6.7	−0.1	2.4	2.4
4	A_3	−5.0	3.6	6.1	−0.0	2.0	2.0
		INCOME ESTIMATES (GNP, exclusive of the statistical discrepancy)[c]					
5	A_0^*	−5.4	6.3	6.6	−0.3	2.2	2.2
6	A_1^*	−4.0	5.0	6.3	−0.4	1.8	1.8
7	A_2^*	−3.4	4.4	5.5	−0.3	1.4	1.4
8	A_3^*	−3.0	3.4	4.5	−0.1	1.3	1.3
		INCOME ESTIMATES AS PREDICTIONS OF FINAL EXPENDITURES ESTIMATES[d]					
9	A_0^*	−6.8	6.3	9.2	−0.3	3.0	3.0
10	A_1^*	−5.4	4.9	7.3	−0.4	2.7	2.7
11	A_2^*	−4.8	4.5	6.6	−0.2	2.3	2.3
12	A_3^*	−4.4	3.8	5.8	−0.1	2.3	2.3

[a]The estimates refer to quarter t of year T.

Estimate	Date of Publication
A_0	$t + 2$ months
A_0^*	$t + 3$ months
A_1 and A_1^*	July, $T + 1$
A_2 and A_2^*	July, $T + 2$
A_3 and A_3^*	July, $T + 3$

[b]Errors are computed as $A_j - A_n$, where A_n denotes the 1965 statistically revised estimates.

[c]Errors are computed as $A_j^* - A_n^*$.

[d]Errors are computed as $A_j^* - A_n$.

the estimates are revised and in fact vanishes in the change estimates with the second July revision (columns 1 and 4). Income estimates of levels are less efficient than the expenditures estimates, but in the case of changes, they are initially more efficient (columns 2 and 5).

In sum, Table 12 suggests there is no great difference between the accuracy of the early expenditures estimates and that of the early income estimates of GNP. Any differences in the primary data which would favor the accuracy of the income estimates are apparently offset by the errors arising from the lack of early data on profits.

Use of the Statistical Discrepancy to Measure Error

Another implication of the conclusion that revisions are primarily a measure of extrapolation errors is that the revisions may give almost no indication of the magnitude or the behavior of total measurement error.[38] Indeed, extrapolation error may be a very small element of the total error ($E_0 = \epsilon + E_n$).

Some indirect evidence on total error is provided by the statistical discrepancy, the item in the national accounts which reconciles the income with the product, or expenditures, estimates. The discrepancy between the two sets of GNP estimates (D_j) can be written

$$D_j = A_j - A_j^* = E_j - E_j^* \quad (j = 0, \cdots, n),$$

where A_j stands for the product estimates, A_j^* for the income estimates, and E_j and E_j^* for their respective errors.

Under strong but not unreasonable assumptions, the discrepancy can be used in conjunction with the revisions to obtain two types of error estimates: (1) rough estimates of the fraction of error eliminated by successive revisions and (2) rough estimates of the ratio of measurement error variance to the variance of the final series.

If extrapolation errors were important components of the initial errors, E_0 and E_0^*, respectively, their elimination through successive revisions of the estimates would result in a substantial reduction in the discrepancy. The variance of the initial discrepancy (D_0) would be

$$\sigma^2(D_0) = \sigma^2(E_0 - E_0^*) = \sigma^2(\epsilon - \epsilon^*) + \sigma^2(E_n - E_n^*),$$
$$\text{assuming Cov}\,(\epsilon - \epsilon^*, E_n - E_n^*) = 0,$$

and the variance of the final discrepancy (D_n) would be

$$\sigma^2(D_n) = \sigma^2(E_n - E_n^*).$$

[38] However, the revisions would give a rough index of accuracy among the components of GNP if there were a positive correlation between ϵ_i and E_{ni}, the error in measuring the ith component.

We would then expect $\sigma^2(D_n) < \sigma^2(D_0)$ and, indeed, the following tabulation shows that the revisions reduced the variance of the initial discrepancy by more than half:

Statistical Discrepancy Between Estimates of GNP Based on Expenditures and Income Data, 1947 II–1961 IV

		$\sigma^2(D_j)$
Provisional Estimates	$D_0 = E_0 - E_0^*$	7.95
First July Revised Estimates	$D_1 = E_1 - E_1^*$	6.10
Second July Revised Estimates	$D_2 = E_2 - E_2^*$	4.84
Third July Revised Estimates	$D_3 = E_3 - E_3^*$	4.41
1965 Revised Estimates	$D_n = E_n - E_n^*$	3.46

It might be tempting to conclude that the revisions eliminated somewhat over one-half of the initial measurement error. However, this conclusion would be correct only under certain conditions. To illustrate what is involved, assume that:

(1) The revisions are an exact measure of extrapolation errors and that they are independent of other types of measurement errors. Then,

$$\mathrm{Cov}(\epsilon, E_n) = \mathrm{Cov}(\epsilon^*, E_n^*) = 0,$$

such that

$$\sigma^2(E_0) = \sigma^2(\epsilon) + \sigma^2(E_n) \text{ and } \sigma^2(E_0^*) = \sigma^2(\epsilon^*) + \sigma^2(E_n^*).$$

(2) Errors in the income and in the expenditures estimates are of the same magnitude, such that

$$\sigma^2(\epsilon) = \sigma^2(\epsilon^*) \text{ and } \sigma^2(E_0) = \sigma^2(E_n^*).$$

The ratio of the variances of the initial and the final discrepancy could then be written

$$\frac{\sigma^2(D_0)}{\sigma^2(D_n)} = \frac{2\sigma^2(\epsilon)(1 - r_{\epsilon\epsilon^*}) + 2\sigma^2(E_n)(1 - r_{E_nE_n^*})}{2\sigma^2(E_n)(1 - r_{E_nE_n^*})}$$

$$= 1 + \frac{\sigma^2(\epsilon)(1 - r_{\epsilon\epsilon^*})}{\sigma^2(E_n)(1 - r_{E_nE_n^*})}$$

In the special case in which the correlations between the two extrapolation errors and between the two remaining errors are equal ($r_{\epsilon\epsilon^*} = r_{E_nE_n^*}$),

$$\frac{\sigma^2(D_0)}{\sigma^2(D_n)} - 1 = \frac{\sigma^2(\epsilon)}{\sigma^2(E_n)}.$$

Since $\frac{\sigma^2(D_0)}{\sigma^2(D_n)} = 2.298$, $\frac{\sigma^2(\epsilon)}{\sigma^2(E_n)} = 1.298$ and since $\sigma^2(\epsilon) = (5.6)^2 = 31.36$, $\sigma^2(E_n) = 31.36 \div 1.298 = 24.16$, then $\frac{\sigma^2(\epsilon)}{\sigma^2(E_0)} = \frac{31.36}{31.36 + 24.16} = .554.$ Thus under these special conditions, extrapolation error would amount to about 55 per cent of the initial measurement error, or put differently, the revisions eliminated about 55 per cent of the initial error.

It could be objected, however, that the reduction in the discrepancy may overstate the error eliminated by the revisions (or in other words, the importance of extrapolation error). The statistical discrepancy, as noted earlier, serves as a tool for controlling error. An unusually large discrepancy suggests the presence of unusual error in either the income or in the expenditures estimates and an attempt is generally made to trace the error and eliminate it before the figures are released. Hence the discrepancy as published is not a pure residual and the revisions, it might be contended, reflect in part an allocation of the discrepancy.

It is clear, however, that the revisions of the expenditures estimates do not *primarily* reflect an allocation of the discrepancy—or the revisions of GNP components, as well as of the aggregate, would not have resembled extrapolation errors. It is nonetheless possible that the revisions *in part* reflect an effort to bring the income and product estimates together. It would therefore be well to consider the estimate of the fraction of error eliminated by the revisions (.554) an upper limit and the estimate of the variance of E_n (24.16), a lower limit.

Alternative estimates can be obtained from the final discrepancy in the following way: Assume, as before, that the errors remaining in both sets of estimates are of the same magnitude such that

$$\sigma^2(E_n) = \sigma^2(E_n^*).$$

The variance of the final discrepancy could then be written

$$\sigma^2(D_n) = 2\sigma^2(E_n)(1 - r_{E_nE_n^*}),$$

and

$$\sigma^2(E_n) = \frac{\sigma^2(D_n)}{2(1 - r_{E_nE_n^*})}.$$

Estimates of $\sigma^2(E_n)$ can be computed for given values of $r_{E_nE_n^*}$. For example, if $r_{E_nE_n^*}$ were $+.9$, $\sigma^2(E_n)$ would be five times $\sigma^2(D_n)$, or, if

$r_{E_n E_n^*}$ were equal to $+.95$, $\sigma^2(E_n)$ would be ten times $\sigma^2(D_n)$. The following tabulation shows estimates of $\sigma^2(E_n)$ for arbitrary values of $r_{E_n E_n^*}$. These estimates are added to the variance of the revisions to obtain crude estimates of $\sigma^2(E_0)$, the initial error. The last two columns show the variance of the revisions as a fraction of the variance of the initial error and the initial error as a fraction of the variance of the final series.

Estimates of Error

Assumed Value of $r_{E_n E_n^*}$	$\sigma^2(E_n) = \dfrac{\sigma^2(D_n)}{2(1 - r_{E_n E_n^*})}$	$\sigma^2(E_0) = \sigma^2(\epsilon) + \sigma^2(E_n)$	$\dfrac{\sigma^2(\epsilon)}{\sigma^2(E_0)}$	$\dfrac{\sigma^2(E_0)}{\sigma^2(A_n)}$
.950	34.6	66.0	.475	.008
.975	69.2	100.6	.313	.012
.990	173.0	204.6	.153	.025
.995	346.0	377.4	.083	.045

These error estimates are of course crude and they are based on strong assumptions, which may not be fulfilled. Nonetheless they provide some indication of orders of magnitude. Taking them at face value, they would suggest that the revisions eliminated anywhere from 8 to nearly 50 per cent of the initial error and that the initial measurement error could range from roughly 1 to 5 per cent of the variance of GNP for the period covered, 1947 II–1961 IV.