RAS Updates in a Commodity-by-Industry Setting:

What Does the Extra Information Buy You?

By

Michael L. Lahr
Center for Urban Policy Research
Rutgers University
33 Livingston Avenue, Suite 400
New Brunswick, NJ 08901-1982

lahr@rci.rutgers.edu

1. Introduction.

US national tables still take at least seven years from survey to production. As late as 1996 anyone performing work on American economy was doing so with a 1987 model. That is, we were operating with a base-year model that was over nine-years old. This is despite the decreased costs and increased speed of computation. Interestingly, the US is not alone; many countries are not able to spin out input-output tables any faster than they used to.

The literature for updating input-output tables is rather rich. The biproportional technique known as RAS is probably the most widely used approach, although others have been put forward (e.g., Snower, 1990). Recently, Polenske (1997) thoroughly reviewed the literature on this technique, so I will not do so here. Suffice it to say, the technique conventionally required that intermediate (primary) input and output totals of a recent year are applied to the industry-by-industry transactions matrix of a base year (or in unison with the total gross output totals of the recent year and the direct matrix of a base year) to create an updated direct coefficients matrix.[1] A main benefit of the technique is its simplicity. Nonetheless, many researchers have pointed out that RAS tends to result in updated matrices that are "more accurate" than the alternative—the direct coefficients matrices from which the updates were derived.

The literature on this subject died out after a rapid exchange across many journals and books from 1961 to 1980. The reasons why this literature petered out are not entirely clear. Certainly, Barker's (1975) suggestion that it would be more fruitful to get partial

[1] Szyrmer (1984) developed a similar technique that he called RTS in which he extended the transactions matrix by final-demand columns and value-added rows and only n elements of margin information, as opposed to the $3n$ in standard RAS, as mentioned above.

information on particular flows resonated throughout the I-O community.[2] Miernyk's (1977) statement that the technique was overly mechanical and that it lacked sufficient economic foundations may have weighed in heavily as well.

Nonetheless, RAS remains very effective technique, as I pointed out earlier. Plus it is a relatively simple and inexpensive, requiring only data generally readily estimable, if not readily available, from annual national accounts. It is still widely used.

Another event may have served to prevent some publications—the Make-Use system of national accounts (SNAs) was established as a UN standard in 1968. Many nations began to produce tables in that form soon after the standard was set. The US, for example, released its first tables in that form in 1979 —the 1972 US national I-O tables. Indeed, the demise of RAS updating tests and the rise of the new UN standard for SNAs seems more than just a coincidence. Certainly the updating procedure seemed to have become more complex, requiring at least the RASing of two tables—the Make and the Use matrices, not just one. And more commodity-based marginal data were needed as well. St. Louis (1989) and Gilchrist and St. Louis (1999) are among the only papers published on the subject of RASing I-O tables specified in the currently used SNA.

Is there a reason why such literature has not been forthcoming? Well, one reason may be that updates can be performed on the SNA standard tables by converting them to industry-by-industry form, that in which most national I-O tables existed in years gone

---

[2] Although he does not acknowledge Barker (1975), Szyrmer (1989) shows that in a dense 79-by-79 I-O setting that individual margin data are far more important than are values of individual coefficients or transactions in obtaining an accurate model, fueling a potential dispute with Barker's suggestion. Conversely, essentially in support of Barker, Israilevich (1986) developed a modification of RAS, which he labeled "ERAS" (Extended RAS), for which subtotals of certain blocks in the transaction matrix, rather than individual cells, are known. Gilchrist and St. Louis (1999) note the significance of an approach similar to Israilevich's that they called TRAS (three-stages RAS) and obtained results superior to those gotten through conventional RAS, supporting Barker's notion

by. This approach enables analysts to use conventional RAS-updating techniques. But, at least at the national level, there is much appeal to using the industry-by-commodity form since final demand is typically expressed in commodity terms.[3] Hence, another tack must be taken. Such a tack was proposed as early as St. Louis (1989). A sketch of part of St. Louis's approach follows using the notation of Jackson (1998), which is outlined in Table 1 below.

Table 1: Schematic of Commodity by Industry Accounts
with Imports Negative Entry in Final Demand

	Commodities	Industries	Final Demand	Total Output
Commodities		**U**	**F|x| (-m)**	**q**
Industries	**V**			**g**
Value Added		**W**		
Total Inputs	$\mathbf{q}'$	$\mathbf{g}'$		

Note: **F** is the matrix of domestic final demand and **x** is a vector of exports, so that **F|x** is matrix formed by catenating **x** to **F**. The vector **i** is a summation vector that consists strictly of ones.

Thus, the fixed relationship between commodity inputs and industry output can be expressed as

$$\mathbf{B} = \mathbf{U}\hat{\mathbf{g}}^{-1} \tag{1}$$

where ^ above a vector denotes a matrix with that vector on the diagonal. The fixed relationship between industry inputs and commodity outputs as

$$\mathbf{D} = \mathbf{V}\hat{\mathbf{q}}^{-1} \tag{2}$$

[3] Regional I-O analysts, in the US at least, remain tied to the industry-by-industry model since they typically apply final demands that are more readily expressed in terms of industry demand. See Jackson (1998) for some further discussion on this.

Let us now distinguish the elements of a two-point time series of each of these vectors and matrices by subscripts. That is, let $\mathbf{B_0}$, $\mathbf{D_0}$, U_0, $\mathbf{V_0}$, $\mathbf{g_0}$, and $\mathbf{q_0}$ be the corresponding components of a country's SNA for the base year, $\mathbf{B_1}$, $\mathbf{D_1}$, U_1, $\mathbf{V_1}$, $\mathbf{g_1}$, and $\mathbf{q_1}$ be those same components for the next period, and so on. Then according to St. Louis (1989) in order to derive $\dot{\mathbf{B}}_1$ (the $\dot{}$ denotes an estimate), in addition to $\mathbf{B_0}$ one needs values for $\mathbf{i}'\mathbf{U}_1$ or alternatively $\left(\mathbf{g}_1 - \mathbf{i}'\mathbf{W}_1\right)$, the vector of intermediate (primary) industry inputs, and for $\mathbf{U}_1\mathbf{i}$ or alternatively $\left(\mathbf{q}_1 - \mathbf{F}_1 \,|\, \mathbf{x}_1 \,|\, (-\mathbf{m}_1)\right)$, the vector of intermediate (primary) commodity outputs. To derive $\dot{\mathbf{D}}_1$, one needs not only $\mathbf{D}_0$ but also values for $\mathbf{q_1}$ and $\mathbf{g_1}$. Further, according to Jackson (1998), to gain an industry-by-industry direct coefficients matrix, the conventional $\mathbf{A}$ matrix, which is derived from $\dot{\mathbf{B}}$ and $\dot{\mathbf{D}}$ by the following formula:[4]

$$\dot{\mathbf{A}}_1 = \dot{\tilde{\mathbf{D}}}_1\dot{\mathbf{B}}_1 \tag{3}$$

where

$$\dot{\tilde{\mathbf{D}}}_1 = \begin{pmatrix} \dot{\mathbf{V}}_1 - \dot{\mathbf{D}}_1\hat{\mathbf{x}}_1 \\ \mathbf{m_1}' \end{pmatrix}\left(\mathbf{q}_1 - \mathbf{x}_1 + \mathbf{m_1}\right)^{-1} \tag{4}$$

and

$$\dot{\mathbf{V}}_1 = \dot{\mathbf{D}}_1\hat{\mathbf{q}}_1 \tag{5}$$

[4] This formulation is quite different than that proposed by St. Louis (1989, p. 377), who Used Miller and Blair's (1985) formulation that $\dot{\mathbf{A}}_1 = \dot{\mathbf{D}}_1\left(\mathbf{I} - \hat{\mathbf{m}}_1\hat{\mathbf{q}}_1^{-1}\right)\dot{\mathbf{B}}_1$. Interestingly, Jackson's formula closely parallels Miller and Blair's (1985) for regional supply percentages, which are Used to regionalize national direct requirements matrices.

1. Information and RAS Updating in a Make-Use World

Clearly the operations described in Equations (3) to (6) require two more *m*-element vectors those for exports, $\mathbf{x}$, and imports, $\mathbf{m}$. According to the procedure above then, a total of two base matrices—$\mathbf{B}_0$ and $\mathbf{D}_0$—and six vectors— $\mathbf{x}_1$, $\mathbf{m}_1$, $\mathbf{q}_1$, $\mathbf{g}_1$, $\mathbf{i}'\mathbf{W}_1$, and $\mathbf{F}_1'\mathbf{i}$ —should be used to update an I-O model via the 1968 UN SNA. If there are m commodities and n industries and if there are k more commodities than are industries (i.e., $m=n+k$), this means that $2(n+k)n + 2n + 4m$ pieces of information can be applied to derive a solution to which previously only $n^2 + 2n$ were brought to bear. Hence, now $n^2+2nk+4m$ more pieces of information can be used in the I-O updating process when RAS is employed. Indeed, more than twice as much information (precisely $2nk$+2n+ $4k$ more pieces than twice as much) can be used now for updating than was needed prior to the 1968 UN to produce the industry-by-industry direct requirements matrix.

The rub is, however, that only $4m$ of the $n^2+2nk+4m$ extra pieces of information are "actual" data: the rest are the extra base-year data that are the objects of the updating. Nonetheless, all evidence point to a hypothesis the model resulting from the newer format and updating procedures outlined above should be more accurate. This is shown by the pair of ratios on either side of the inequalities in (6). The ratio of "actual" to base information for the traditional case in shown on the left-hand-side of (6): that for the 1968 UN SNA as outlined above is on the right-hand side of (6). When k=0, the right-hand side is one-and-a-half times the value of the left-hand side of (6). When k=n then the right-hand side of (6) becomes $\frac{5n}{2n^2}$ or $\frac{2.5}{n}$, and thus remains greater than $\frac{2}{n}$, the value of the right-hand-side.

$$\left.\begin{array}{l}\dfrac{2n}{n^2}\overset{?}{<>}\dfrac{2n+4m}{2nm}\\ \\ \text{or}\\ \\ \dfrac{2}{n}\overset{?}{<>}\dfrac{2(3n+2k)}{2n(n+k)}\\ \\ \text{or}\\ \\ \dfrac{2}{n}<\dfrac{3n+2k}{n^2+nk}\end{array}\right\}\tag{6}$$

Regardless, the right-hand side of (6) is a decreasing function of *k*, since it's derivative with respect to k is $\frac{-n^2}{\left(n^2+nk\right)^2}$, a negative number since *n* and *k* must both be positive and typically $k<<n$.

Let us now assume that no piece of information is any more important than another [perhaps a rather strong assumption according to Szyrmer (1989)]. In turn, this finding supports the argument that the double commodity-by-industry updating approach should provide twice the accuracy of the traditional strict application of RAS to **A**. Indeed, it probably should be far better since in very industry-detailed I-O tables, the Make matrices **V**, not in the traditional format, are particularly sparse, severely reducing the number of cells across which the vectors of "actual" information are spread, i.e., far less than the *nm* elements accounted for in (6) above .

As suggested earlier, however, one could argue that to derive intermediate inputs and intermediate outputs that three, rather than two, vectors are needed to RAS I-O tables of pre-1970 vintage. This is because intermediate inputs and outputs typically can be derived from information on net final demand, $(\mathbf{F}_1 \mid \mathbf{x}_1 \mid (-\mathbf{m}_1))'\mathbf{i}$, and value added, $\mathbf{i}'\mathbf{W}_1$, only when outputs, $\mathbf{g}_1$, are known In this case, where $3n$ rather than $2n$ pieces of information are employed, the left-hand side of (6) is altered to $\frac{3}{n}$, giving it the maximum of the range of values for the right-hand side (when k=0). If this is valid reasoning (and I put forward that it is) the hypothesis for the method sketched out above is indeterminate. That is, it not clear that the proposed method of updating to an industry-by-industry table in a commodity-by-industry accounting frame should result in more accurate updated tables when compared to RASing in the traditional industry-by-industry accounting environment.[5]

2. The Research Approach

With this in mind, I set out to make an initial foray to discover whether, as expected, enough accuracy ***is*** gained in an I-O update from the extra information and calculation time required to apply it to a system of accounts in a commodity-by-industry format. To do this, I opted to use I-O tables from two successive time periods: the US tables for 1987 and 1992.[6] I also opted to investigate the performance of the approach with two different levels of aggregation: one with as much detail as possible, about 498 sectors, and the standard 97-sector version published in the *Survey of Current Business*.

---

[5] Interestingly, St. Louis (1989) eliminated the indeterminacy of the hypothesis applying RAS, yet again, to the $\dot{\mathbf{A}}_1$ matrix in (3) in the conventional manner.

[6] In experiments to be conducted after the conference, I will add an extra five-year period, the idea being that the period from 1982 to 1992 is more like the amount of time that elapses until a new tables are available in the US. Further, finding in past studies showed that technology changed much more during this kind of time frame making updating paramount, as opposed to simply using an old one.

Some amount of aggregation on each of the two detailed models was required in order to obtain the same set of industries.[7]

The original design of the experiment was to apply RAS to both $\mathbf{B}_{87}$ and $\mathbf{D}_{82}$ and subsequently to form $\dot{\mathbf{A}}_{92}$. This estimated matrix was then to be compared to the actual one, $\mathbf{A}_{92}$, using several comparison measures established in the I-O literature. As a control, I form $\mathbf{A}_{87}$ directly and then update it using the conventional RAS approach. The resulting $\dot{\mathbf{A}}_{92}$ then also was to be compared to the actual $\mathbf{A}_{92}$. The rationale behind using the two different level of aggregation was to see if the density of the matrix does, indeed, affect the relative accuracy of the new approach.

After developing the Make and Use coefficients matrices at both the 468 and 95 sector levels,[8] it became clear that the resulting estimates of their corresponding direct matrices were invalid. The sums of a small number of columns wound up being either negative or more than 1.0. In an attempt to avoid an extremely short paper, I opted to follow the example of St. Louis (1989) and RAS the invalid direct matrices in the same fashion that I did for the traditional approach. The results of the RASing process are summarized in Tables 1 and 2. In the case of the detailed 468-sector table, I still obtained columns sum larger than 1.0. Hence, the results for the resulting matrix are not reported for this level of sectoral detail.

Table 1: Results for Updating of US 468-Sector Tables[9]

[7] For this, I gratefully acknowledge effort made by Alexandru Voicu.

[8] Aggregation was undertaken to Make tables form the two timeframes commensurate. Some non-existent sectors and others with negative total gross outputs were also dropped from the analysis.

[9] Note that error for multipliers was estimated by first subtracting unity from them before performing the requisite calculations.

Table 1.1: Difference for Straight Update of Direct Coefficients: 468 Sectors

	MAD	WAPE	STPE
Estimated vs. Actual	0.00051217245	1.8332388	47.817551
Original vs Actual	0.00052918782	2.0474196	49.406143
Estimated/Original	0.96784626	0.2141808	1.588592

Table 1.2: Difference between Leontief Multipliers of Straight Update: 468 Sectors

	MAD	MAPE	STPE
Estimated vs. Actual (%)	4.8725237e-005	3.2373529	2.3737836
Original vs. Actual (%)	0.00027208511	13.636368	13.255373
Estimated/Original	0.17908086	10.262849	10.8815894

Table 1.3: Difference for Update of Use Coefficients: 468 Sectors

	MAD	WAPE	STPE
Estimated vs. Actual (%)	0.0016660504	7.1331315	147.14770
Original vs. Actual (%)	0.00057612088	2.3769714	50.883733
Estimated/Original	2.8918417	-4.7561601	-96.2639670

Table 1.4: Difference for Multipliers of Make Coefficients Matrix: 468 Sectors

	MAD	MAPE	STPE
Estimated vs. Actual (%)	0.00041666985	6.7421700	19.625912
Original vs. Actual (%)	0.00016588631	2.3314353	7.8135485
Estimated/Original	2.5117797	-4.4107347	-11.812363

Table 2: Results for Updating of US 95-Sector Tables

Table 2.1: Difference for Straight Update of Direct Coefficients: 95 Sectors

	MAD	WAPE	STPE
Estimated vs. Actual (%)	0.0016422833	1.2558907	30.632332
Original vs. Actual (%)	0.0017330452	1.4031619	32.325247
Estimated/Original	0.94762868	0.1472712	1.692915

Table 2.2: Difference between Leontief Multipliers of Straight Update: 95 Sectors

	MAD	MAPE	STPE
Estimated vs. Actual (%)	0.00017088952	1.8027387	1.6832332
Original vs. Actual (%)	0.0014500123	15.391527	14.282378
Estimated/Original	0.11785384	13.5887883	12.5991448

Table 2.3: Difference for Update of Use Coefficients: 95 Sectors

	MAD	WAPE	STPE
Estimated vs. Actual (%)	0.00056684636	2.3417753	50.059611
Original vs. Actual (%)	0.00057600816	2.3769805	50.868712
Estimated/Original	0.98409432	0.0352052	0.8091010

Table 2.4: Difference for Multipliers of Make Coefficients Matrix: 95 Sectors

	MAD	MAPE	STPE
Estimated vs. Actual (%)	0.00055688000	2.1905797	5.4597494
Original vs. Actual (%)	0.00044158474	1.2589032	4.3293744
Estimated/Original	1.2610943	-0.9316765	-1.1303750

Table 2.5: Difference for Update of Make-Use Formed Direct Coefficients: 95 Sectors

	MAD	WAPE	STPE
Estimated vs. Actual (%)	0.0015556931	1.4799634	33.015939
Original vs. Actual (%)	0.0017330452	1.4031619	32.325247
Estimated/Original	0.8976645	-0.0768015	-0.690692

Table 2.6: Difference for Multipliers of Make-Use Direct Coefficients Matrix: 95 Sectors

	MAD	MAPE	STPE
Estimated vs. Actual (%)	0.00078978078	9.1750048	8.8514722
Original vs. Actual (%)	0.0014500123	15.391527	14.282378
Estimated/Original	0.40851563	6.2165222	5.4309058

3. Analysis of RAS Results

The most prominent result is that the Make-Use format for updating via RAS is not found here to be superior to that gained by means of direct coefficients matrices only. While (at the 95-sector level only) it yielded results that were superior to those of the base matrix alone, traditional RAS (applied strictly to the industry-by-industry matrix) yielded error on the order of only 11percent, compared to the 40 percent figure for the Make-Use format.

Next note from Table 1.1 that much of the difference between the two detailed direct matrices occurs in smaller elements. This is identified by the fact that the WAD is small (2%) and the STPE is rather large (49%). This jibes with other studies, which have found that there tends not to be much substantive difference in technology during a five-year period. But Table 1.2 shows that these minor perturbations in technology do result in a relatively significant difference in Leontief multipliers (on average about 13 percent). Thus, the table further shows that while the RAS-based improvements were rather insignificant in improving the direct coefficients themselves, that a full 10 percent reduction in error was obtained as a result of the process, certainly making this effort worthwhile in any case.

There is one other significant result to be gained from information in Table 1. That is, from Table 1.6 it is relatively clear that the coefficients of the Make table has not changed much at all during the period of study, just under 8 percent as measured by STPE during the five-year period.

In general, the results in Table 2 mimic those in Table1. A main difference is that the Use matrix is improved by RASing at this level of aggregation where it was not at the more detailed level. In combination with the last finding on the lack of change in the Make matrix, this points to the prospect that it may be that updating of the Use matrix only is necessary during intervening years—certainly a topic worthy of further study!

4. Conclusions

From the experiments conducted here it is clear that that not only is the additional information and effort required for a commodity-by-industry (Make-Use) update via RAS not worthwhile, but it tends to yield results that are inferior to those of a traditional update of the industry-by-industry base table via RAS. But there appeared to be some promise. That is it was clear that, during the period of study, the coefficients of the Make matrix did not change much. Further we found that at one level of aggregation, in any case, that RASing yielded a superior Use matrix. Thus it may be that by updating the Use matrix alone that superior results may be obtained. But the prospects are not all that promising given that this result was not obtained for the more detailed Use matrix.

References

Barker, T.S. 1975. “An Analysis of the Updated 1963 Input-Output Transactions table,” in R.I.G. Allen and W.F. Gossling (eds.), Estimating and Projecting Input-Output Coefficients (London: Input-Output Publishing Company), pp. 57-67.

Gilchrist, Donald A. and Larry V. St. Louis. 1999. “Completing Input-Output Tables using Partial Information, with an Application to Canadian Data,” *Economic Systems Research*, 11, 185-191.

Israilevich, Philip R. 1986. *Biproportional Forecasting of Input-Output Tables* (Ann Arbor, MI: University Microfilms).

Jackson, Randall W. 1998. “Regionalizing national Commodity-by-Industry Accounts,” *Economic Systems Research*, 10, 223-238.

Lynch, R.G. 1986. “An Assessment of the RAS Method for Updating Input-Output Tables,” In: Ira Sohn (ed.), *Readings in Input-Output Analysis* (New York: Oxford University Press), pp. 271-284.

Miernyk William H. (1977) “The Projection of Technical Coefficients for Medium-Term Forecasting,” In: W.F. Gossling (ed.) *Medium-Term Dynamic Forecasting, The 1975 London Input-Output Conference* (London: Input-Output Publishing Company), pp. 29-42.

Polenske, Karen R. 1997. “Current Uses of the RAS Technique: A Critical Review,” In Andr·s Simonovits and Albert E. Steenge (eds.), *Prices, Growth and Cycles: Essays in Honour of Andr·s Brody* (Basingstoke, England: Macmillan Press Ltd.), pp. 58-88.

Snower, Dennis J. 1990. “New Methods of Updating Input-Output Matrices,” *Economic Systems Research*, 2, 27-37.

St. Louis, Larry V. 1989. “Empirical Tests of Some Semi-Survey Update Procedures Applied to Rectangular Input-Output Tables,” *Journal of Regional Science*, 29, 373-385.

Szyrmer, Janusz M. 1984. “Estimating Interregional Input–Output Tables with RAS,” In: *Commodity Flow Estimation*, report edited by W. Bruce Allen et al. for U.S. Department of Transportation.

_____. 1989. “Trade-Off between Error and Information in the RAS Procedure,” In: Ronald E. Miller, Karen R. Polenske and Adam Z. Rose (eds.) *Frontiers of Input-Output Analysis* (New York: Oxford University Press), pp. 258-278.