**Comparing Input-Output Tables of Different Economies: A New Application for Decomposition Techniques?**

Michael L. LAHR & Ling YANG

# Introduction.

In 1986 Janusz Szyrmer and the senior author of this paper teamed up to produce a paper “On Measures for Comparing Input-Output matrices.” They presented it later that year at the North American Meeting of the Regional Science Association (NA RSA Meetings) and, in hindsight, naturally without much fanfare. The paper was a natural for the two. Lahr was in his formative years as a PhD candidate, and Szyrmer was a fresh Adjunct Assistant Professor: Frankly they both needed publications.

Research-wise, Lahr needed some sort of measure to enable a comparison of two input-output tables, and Szyrmer (1984) already had some experience in the matter, having published an appendix to a report that included a short list of measures. The two rolled up their sleeves and delved in. Their roles were clearly delineated. Lahr was to comb the literature for measures used by input-output analysts since the beginning of history, and Szyrmer would pull together the data to test and categorize “the most promising and most used” of those measures, as identified by Lahr. Szyrmer also analyzed the results.

The two started out their paper with the premise that “the results of comparisons [of input-output tables] have been, at best, inconclusive with regard to goodness of fit (or should we say badness of fit?) of the estimates.” They leaned heavily on Knudsen and Fotheringham (1986), henceforth K-F, who covered similar material for spatial interaction modeling to make certain statements about the means of comparisons that had been used, largely because the two modeling arenas share many of the same measures. K-F noted that such comparisons largely employ measures of distance and association, which serve two main purposes; (1) to measure the ability of models to produce accurate results and (2) to determine the statistical significance of the difference between two tables. Interestingly in the intervening quarter century plus, little new has sprouted forth on the subject.

There were some consequences from our presentation at the NA RSA meetings though. Szyrmer (1989) did produce a piece that dealt with similar material in “The Trade-Off between Error and Information in the RAS Procedure,” and Lahr (1992) completed a dissertation in which much of his part of the investigation became an appendix to a chapter. An extract of which is in Lahr (2001), shows the appendix culminates in the development of the Weighted Absolute Deviation, which has since been used occasionally in subsequent research by others.

Lahr and Szyrmer (1986) noted, while many similarities exist between spatial interaction model and input-output tables, following K-F approach would not be appropriate for the needs they outlined at the outset of their paper. Moreover, they found it unsatisfactory to categorize the various measures as “1) those that performed quite well, 2) those that appear quadratically related to induced error, and 3) those that performed erratically,” although some satisfaction was derived from culling the last group from the chaff. They identified the “best” measures for direct coefficients as absolute psi and the weighted absolute difference. The two best for transactions tables and multipliers they identified as root square-error and total absolute deviation. Still, the two knew they were both on to something and missing something. In their suggestion for future research, they proposed normalizing industries’ the input mixes, excluding imports and value added, and then comparing. This is clear recognition that the basic structures of input-output tables were disabling their ability to compare tables themselves when limited to the vast array of measures they had at hand.

# Research Approach.

An axiomatic approach like that used by Asami and Smith (1995) undoubtedly could be applied to uncover the attributes of a proper measure for comparing input output matrices. We jump the gun and dive right into an approach that has all of the qualities that were lacking in the measures reviewed by Szyrmer and Lahr (1986). The key reference to the line of literature that we employ is the review of structural decomposition analysis (SDA) by Casler and Rose (1996) who liken it to a combination of shift-share, growth accounting, and index number analysis. For this reason, it has been used successfully to analyze the causes of change in economies. The factors analyzed and focuses of any investigation are limited only by available data and the creativity of analysts.

The key appeal of SDA in the case of identifying error or bias for input-output tables is that they can explicitly account for their key structural attributes. As Casler and Rose (1996) note they can account for differences between an actual and estimated table in technology (material input mix), industry mix, productivities of capital and labor, trade, consumption and export patterns, and more generally the composition of final demand. Indeed, these attributes of SDA, which enable its capability in growth accounting analysis, are also essential in the examination of error and bias in estimated tables compared to benchmark equivalents. Hence, SDA there is no reason to believe it should be any less effective in comparing an estimated table to the “real thing.” Similarly since SDA can inform how two different snapshots of an economy compare, it follows that it should also be able to inform how snapshots of two very different economies compare. Ultimately the issue is not whether SDA is viable in other input-output table comparisons but more simply a matter of identifying what specific SDA should be applied.

The selection of the “right’ SDA, of course, depends on the bases of the test to be undertaken. Using principles discussed in Lahr (1993, 2001), we propose using a pre-existing decomposition for the present paper—one detailed in Dietzenbacher, Lahr and Los (DLL, 2004).

The overriding principle guiding our selection of a specific decomposition is that the input-output tables, which we use as targets of our later examination in this paper, will be (have been) used for purposes as yet known—i.e., they are general-purpose input-output tables. A single-purpose table would eliminate the need for holistic accuracy of the table examined and, therefore, militates against any a detailed analysis of table comparisons.

The second set of principles is grounded in the relative quality of data in the target tables. The industry-by-industry tables that we examine are the U.S. annual table for 2002 (US BEA, 2003) and the Washington State table for 2002 (Beyers, 2008): we compare both to the 2002 U.S. benchmark tables (Stewart, Stone, and Streitwieser, 2007). For one, data related to compensation in an estimated I-O table in the U.S. should be quite accurate since robust annual data are available on the aggregate payroll of almost all industries (government and railroads are exceptions) at both the national and regional levels in the U.S. through state-level reporting via the U.S. Bureau of Labor Statistic’s Quarterly Census of Employment and Wages. To a lesser extent, same goes for household consumption (assuming that researchers have developed a bridge from the U.S. Bureau of Labor Statistic’s Consumer Expenditure Survey to sectors of the input-output table) and for major components of gross product originating (GPO, gross domestic product by industry) other than compensation, for which annual estimates are available for states from the U.S Bureau of Economic Analysis. At the national level trade by commodity is also available as is information on government spending (The U.S. Census bureau), albeit the latter in rather aggregated forms.

In summary the differing reliabilities of data on compensation, value added (GPO), and final demand suggest that we scrutinize these table aspects. In DLL, we added hours worked to facilitate a close examination of labor productivity, which also is widely reported. The final version of the paper likely will demand we eliminate this aspect and pay more attention to the composition of the final demand array as well as trade. Still, the DLL composition is interesting, instructive, and convenient for present purposes. A main feature is that it focuses on decomposing compensation’s share of value added (GPO) by industry. This target ratio, as suggested above, relies on data for which national and regional input-output analysts in the US should have rather good data on an annual basis, at least at a typical level of aggregation of 85 sectors or so. This makes it an excellent target for our tests since any deviations from the benchmark would be surprising unless we compare very different economies. Moreover, we expect also to detect greater variances in the components (terms) of the ratio’s decomposition than in the ratio itself.

The following description of the DLL decomposition is largely extracted from that paper. The definitions, all for industry *i*, are:

 = value added (in 2002 dollars)

 = labor compensation (in 2002 dollars)

 = labor inputs (in hours worked)

 = labor productivity

= compensation per hour worked

 = hours worked per 2002 dollar of gross output

  = labor compensation’s share of value added.

The overall labor productivity is the weighted average of industry labor productivities, again using industry labor input shares (in total labor inputs) as weights. This implies for the ratio of aggregate labor compensation to value added

 .

The two polar decompositions now yield that  equals

 (1a)

and

 (1b)

Each of the two polar decompositions breaks down the changes in the aggregate labor share into five effects. The five Fisher-type indexes are obtained by taking the geometric average of the two corresponding effects. Note that in the polar decompositions, the first factor relates to changes in the real compensation per hour worked. The second factor indicates how labor productivity changes affect the aggregate labor share. The first two factors together are the shift effects as discussed in Section 2. The other three factors relate to changes in economic structure and together form the share effects. These three factors are: changes in the labor input coefficients; changes in the intermediate input coefficients; and changes in the final demands, respectively. The third and fourth factors reflect technological changes in the production structure. The fifth factor covers changes in the preferences of consumers. Note that the last two terms are common determinants in structural decomposition analyses.

# Discussion of Results.

To simplify presentation of the material and still make it distinctive, in each of the two core tables of factor-based results (one for the annual-benchmark comparison, the other for the US-Washington State comparison) we first present the various factors economywide for the two polar decompositions. We follow that with their geometric mean or the Fisher Index. Subsequently we present our findings for major sectors of the economy delineating services into Producer services (those services that almost exclusively deliver to businesses), Distributive services (Finance, Insurance, Real Estate, Transport, Information, and Wholesale Trade—sectors that serve both to businesses and households), and Consumer services.

***The US Annual Table.***

We opt first to discuss results of our decomposition that compares the national annual table to the national benchmark for the same year—2002. From a macro perspective, it is gratifying to note that the two polar decompositions yield very similar results. This suggests our analysis of the Fisher Index both by sector and economywide are robust.

Ultimately bias economywide is generally rather small (See Table 1). That is, in producing the annual 2002 table, which was released about five years prior to the benchmark, the BEA team did a remarkable job. As expected, in the main, they had the compensation per hour (**α**) fairly precisely pegged. More surprisingly, perhaps, they estimated input mix (**L**) fairly precisely too. Estimates of hours worked per unit of output (**λ**) and composition of final demand (**f**) were rather as well: each was not much more than about a half percentage point off, on average. They do contribute bias but in conflicting directions: they generally under estimated labor per unit output and over-estimated final demand’s influence.

**Table 1: Decomposition Comparing the U.S. Annual to the U.S. Benchmark Table, 2002**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Factor values*** |  | Fisher |  |  | Services |
| Factor | Total 1 | Total 2 | Index | Primary | Mfg | Producer | Distrib | Consumer |
| α | 1.0000 | 1.0000 | 1.0000 | 0.9491 | 1.1260 | 0.8977 | 0.9837 | 1.0202 |
| π | 1.0210 | 1.0210 | 1.0210 | 0.9787 | 0.9404 | 1.1016 | 1.0703 | 0.9772 |
| λ | 0.9946 | 0.9936 | 0.9941 | 1.0073 | 1.0027 | 1.0119 | 0.9902 | 1.0023 |
| L | 0.9997 | 1.0013 | 1.0005 | 1.0031 | 1.0047 | 0.9868 | 1.0009 | 0.9991 |
| f | 1.0058 | 1.0051 | 1.0054 | 0.9896 | 0.9926 | 1.0015 | 1.0090 | 0.9986 |
| ***Factor bias (%)*** |  | Fisher |  |  | Services |
| Factor | Total 1 | Total 2 | Index | Primary | Mfg | Producer | Distrib | Consumer |
| α | 0.00% | 0.00% | 0.00% | -5.22% | 11.87% | -10.79% | -1.64% | 2.00% |
| π | 2.08% | 2.08% | 2.08% | -2.15% | -6.14% | 9.68% | 6.79% | -2.31% |
| λ | -0.54% | -0.64% | -0.59% | 0.73% | 0.27% | 1.18% | -0.98% | 0.23% |
| L | -0.03% | 0.13% | 0.05% | 0.31% | 0.47% | -1.33% | 0.09% | -0.09% |
| f | 0.58% | 0.51% | 0.54% | -1.05% | -0.74% | 0.15% | 0.90% | -0.14% |

The main point of deviation, based on economywide numbers only, appears to have been in their measurement of labor productivity, which on average they overestimated by about 2.0 percent. Since compensation per worker and hours per unit output were well measured, this suggests that BEA team must have tended to err in their estimate of returns to capital, which, of course, is specifically what benchmark input-output tables are designed to address.

The macro view of the error structure masks apparent problems in BEA’s procedures, however. It turns out that compensation per hour was not as well estimated for all sectors as we were led to believe from the economywide results. Moreover, overestimation in Manufacturing recompenses for underestimates of the rate of compensation for both the Primary (agriculture, mining, and construction) and Producer service sectors. Indeed, bias in both Manufacturing and Producer services averaged in double digits, although in compensating directions. That is the compensation rate for Producer services was underestimated in the annual table and that for Manufacturing overestimated. Interestingly the outcomes for labor productivity in those two sectors were nearly opposite. That is, for labor productivity Producer services was overestimated and underestimated for Manufacturing. This suggests that labor likely was regaining ground lost during the 1990s, This loss was a main point of Dietzenbacher, Lahr, and Los (2004).

While bias in compensation per hour was evident in Distributive and Consumer services too, it was not nearly so strong, hovering close to 2 percent in both sectors. Distributive services followed the general path of Producer services with respect to compensation per hour and labor productivity. That is, its compensation was under-estimated but its labor productivity overestimated, following the general pattern of convergence toward a long-run mean for labor’s share of GDP.

In summary, based an evaluation of the U.S. annual estimate using the DLL decomposition suggests that it is rather good vis-à-vis the benchmark. This suggests that the approach outlined in Planting and Guo (2004) amd Lawson et al. (2006) is a least a solid starting point.

**Table 3: Decomposition Comparing the Washington State
to the U.S. Benchmark Table, 2002**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Factor values*** |  | Fisher |  |  | Services |
| Factor | Total 1 | Total 2 | Index | Primary | Mfg | Producer | Distrib | Consumer |
| α | 1.06100 | 1.07990 | 1.07041 | 1.12381 | 1.31993 | 1.14279 | 0.95772 | 0.91533 |
| π | 1.49470 | 1.23260 | 1.35734 | 1.38423 | 0.93421 | 1.75017 | 1.13466 | 1.26546 |
| λ | 0.73660 | 0.92807 | 0.82681 | 0.85271 | 1.04671 | 0.62359 | 0.99510 | 0.99600 |
| L | 1.05760 | 1.01780 | 1.03751 | 1.08177 | 0.96260 | 1.11771 | 1.00140 | 1.00502 |
| f | 1.04650 | 1.02820 | 1.03731 | 0.95017 | 1.04240 | 1.13347 | 0.97958 | 0.99972 |
| ***Factor difference (%)*** | Fisher |  |  | Services |
| Factor | Total 1 | Total 2 | Index | Primary | Mfg | Producer | Distrib | Consumer |
| α | 5.92% | 7.69% | 6.80% | 11.67% | 27.76% | 13.35% | -4.32% | -8.85% |
| π | 40.19% | 20.91% | 30.55% | 32.51% | -6.81% | 55.97% | 12.63% | 23.54% |
| λ | -30.57% | -7.46% | -19.02% | -15.93% | 4.57% | -47.23% | -0.49% | -0.40% |
| L | 5.60% | 1.76% | 3.68% | 7.86% | -3.81% | 11.13% | 0.14% | 0.50% |
| f | 4.55% | 2.78% | 3.66% | -5.11% | 4.15% | 12.53% | -2.06% | -0.03% |

***The 2002 Washington State Table.***

With regard to a comparison of two different economies, it was difficult to be prescient. We expected the economies to differ somewhat on all factors: compensation per hour, productivity, hours worked, input-mix, and final demand. We knew the differences would be markedly different compared to those for between U.S. benchmark and the annual estimate. We, in fact, expected Washington State to have higher rates of pay, productivity, and more efficient input mix and final demand. This is because workers in coastal states of the U.S. west and northeast tend to have higher accumulations of human capital and, hence, yield higher productivity to the industries there. We therefore also expected the number of hours worked per unit of output to be lower, however. But we could not pinpoint an expectation of the magnitude of the differences nor of the relative factor contributions, even across sectors.

As Table 3 reveals the results of the comparative decomposition. As expected final demand and input mix were generally more efficient. Sectoral exceptions exist. The decomposition suggests that manufacturing in Washington State on average has a slightly less efficient input mix than the U.S. Washington State is entrenched in aerospace equipment and both paper and wood products manufacturing. Indeed, since 2002, the state has been suffering major declines in aerospace contracts. In the same vein, final demand for primary goods (exceptionally agriculture and logging) and distributive services (accentuated by the state’s air and water transportation services and ) are less productive than they are for other sectors.

Producer services, led by software and waste remediation services, are the most different of all sectors compared to the national economy and positively affected the overall share of GDP that compensation comparatively owns in Washington State. Indeed, Producer services generally gave the state’s economy a most marked advantage in all factors except for the hourly compensation rate, where manufacturing led.

**Conclusions.**

In summary, it would seem that using decomposition techniques to measure differences in economies holds promise. The U.S. annual table for 2002 fairly close matched its benchmark equivalent but the Washington State table for the same year showed considerable differences. Most interestingly those differences occurred in directions and in magnitudes for sectors that might have been expected, given Washington’s industries with extraordinary concentrations compared to the nation.

If regional input-output table producers have access to reasonably accurate secondary data on jobs, aggregate compensation, and GDP by producing sector, it would seem that most of the differences between national and Washington State data detected could be anticipated without survey work. This is generally the case for states in the U.S. as noted by Lahr (2001). Still some error will persist due to any differences in input mix and the composition in final demand. This suggests more applicable decompositions undoubtedly exist for regional accounts. Of course, it depends on the detail that the regional accounts enable. For example, it is likely that greater analytical detail on both international and interregional trade and on the various final demand subsectors would be beneficial.

Asami, Yasushi and Tony E. Smith. (1995) “Additive-Ratio Measures of Interactivity in Input-Output Systems, *Journal of Regional Science*, 35, 85-115.

Dietzenbacher, Erik,  Michael L. Lahr, and Bart Los, (2004) “The Decline in Labor Compensation's Share of GDP: A Structural Decomposition Analysis for the US, 1982-1997,” in Erik Dietzenbacher and Michael L. Lahr (eds.), *Wassily Leontief and Input-Output Economics.* Cambridge University Press: New York, pp. 188-212.

Lawson, Ann, Brian Moyer, Sumiye Okubo, and Mark Planting. (2006) “Integrating Industry and National Economic Accounts: First Steps and Future Improvements,” Chapter 6 in Dale W. Jorgenson, J. Steven Landesfeld, and William D. Nordhaus (eds) *A New Architecture for the U.S..National Accounts.* National Bureau of Economic Research and the University of Chicago Press: Chicago,pp. 215-262.

Planting, Mark and Jiemin Guo. (2004) “Increasing the Timeliness of U.S. Annual I-O Accounts,” *Economic Systems Research*, 16.

Stewart, Ricky L., Jessica B. Stone, and Mary L. Streitwieser. (2007) *U.S. Benchmark Input-Output Accounts, 2002.*