Is the ECI sensitive to the method of aggregation?

Research by BLS economists indicates that the ECI is not particularly sensitive to the methodology used in constructing the index; also, because changes in the price of labor can be due to shifts in supply, demand, or both, caution should be used when applying standard index number analysis to the ECI

Michael K. Lettau, Mark A. Loewenstein, and Aaron Cushner

Michael K. Lettau and Mark A. Loewenstein are economists, and Aaron T. Cushner is a computer programmer in the Compensation Research and Program Development Group, Bureau of Labor Statistics.

The Employment Cost Index (ECI), published quarterly by the U.S. Bureau of Labor Statistics (BLS), measures changes in the price of labor—defined as total compensation per employee hour worked. As a fixed-weight or Laspeyres index, the ECI controls for changes occurring over time in the industrial–occupational composition of employment. The ECI is computed from survey data on compensation by occupation, collected from a sample of establishments and occupations, weighted to represent the universe of establishments and occupations in the economy.[1]

The cost to employers for employee compensation has two components: wages or salaries paid to employees, and the cost of all nonwage benefits. The wage and salary component of the ECI is represented by straight-time average hourly earnings in an occupation, whether or not the employees are actually paid by the hour. Nonwage benefits, which account for about 30 percent of total compensation costs, include such things as employer contributions to employees' health and other insurance, pension plans, and Social Security, as well as paid vacations and sick leave, premium pay for overtime, and nonproduction bonuses. As with the wages portion of compensation, benefit cost data are converted to an hourly basis in the ECI.

In computing the national ECI, the myriad wage quotes from the sample of individual jobs must somehow be aggregated into a single index number. This aggregation process involves two key steps. Each establishment surveyed for the ECI is placed within 1 of 73 two-digit SIC industries, and each surveyed job is placed within 1 of 10 major occupation groups.[2] Because all of the occupations are not represented in all of the industries, only 720 industry–occupation cells exist (as opposed to 730). Each job quote in the survey falls into exactly one of these cells. The first step in the aggregation process involves combining all of the job quotes within a given cell to obtain a cell average. The second step involves aggregating across the cell averages to obtain the ECI.[3]

Another index produced by BLS using similar methodology to that used in the ECI is the Consumer Price Index (CPI). The CPI has received a great deal of attention recently concerning its alleged upward bias. Because the two indexes are constructed similarly, some of the alleged problems in the CPI also may exist in the ECI. This article examines the sensitivity of the ECI to the method of aggregation.

There are two distinct aggregation issues. The first involves how the various job quotes within a given cell should be combined to obtain a cell average. In light of recent work by Marshall Reinsdorf and Brent R. Moulton on the elementary aggregation of price quotations in the CPI, it seems prudent to evaluate the possible benefits of using geometric rather than arithmetic means to calculate the proportionate change in employee compensation each period. Indeed, the analysis in this article shows that, as with the CPI, ECI wage and benefit indexes constructed using geometric cell means tend to have lower growth rates than those using arithmetic means. In the case of the ECI, however, the difference in growth rates is negligible.

The second aggregation issue concerns the way that cell averages are aggregated to obtain the actual index. The ECI is obtained by taking the weighted sum of the proportionate changes in compensation costs for the various categories of labor, where the weight for each category is simply its share of total labor compensation. Thus, like the CPI, the ECI is a Laspeyres index. To check for a bias in the ECI resulting from the fact that it does not take into account employers' ability to substitute toward labor that has become relatively less costly, this article compares three different indexes—fixed weight, current weight, and superlative—over the period from September 1981 to December 1994. The results indicate that estimates of changes in employee compensation over time are not very sensitive to the choice of index used. An additional finding is that standard index number analysis may not always be adequate for analyzing the ECI (or even the CPI). Changes in the price of labor (or commodities) can be caused by shifts in demand, supply, or both. But the standard interpretation implicitly assumes that only one of these curves is shifting.

The remainder of this article is organized as follows. The next section describes the current process by which individual job quotes are aggregated to obtain the ECI. The next two sections compare fixed weight, current weight, and superlative indexes. The analysis concludes by comparing geometric and arithmetic means.

Current calculation of the ECI

As noted above, the first step in calculating the ECI is to estimate the mean compensation for each category of labor, and the second is to aggregate the cell averages to obtain a final index. Because it is simpler to do so, we begin by discussing the second step in the process. The ECI is designed to indicate how the average compensation paid out by employers would have changed over time if the industrial–occupational composition of employment had not changed from the base period. It is calculated as the weighted sum of the changes in compensation costs for all two-digit industry–major occupation job cells, where the weighting factor for each cell is its share of total labor compensation in the base period. The employment numbers used in the construction of the weights are obtained from the Census of Population.[4]

Now consider the first step in the calculation of the ECI, namely, the estimation of mean compensation for each of the industry–major occupation job cells. Although this step might seem straightforward at first glance, without a costly census, the compensation received by all workers cannot be directly measured. Thus, calculating the mean wages for the various estimating cells is partly a statistical estimation problem. This problem is complicated by the fact that one must estimate mean wages over a number of periods rather than at a single point in time. This is difficult because the ECI sample changes over time. After a job is initially surveyed, it remains in the ECI sample for 16 to 20 quarters, after which the job is deleted from the sample and replaced by a new one.

The simplest way to estimate the mean change in compensation for a category of labor between period 0 (the base period) and period *t* (the reference period) would be to compare the average compensation for that category in the reference period and in the base period. Because the ECI sample changes over time, however, this would involve comparing averages across jobs that may not be homogenous. Thus, the ECI takes a different approach. To start, the mean change in an estimating cell's compensation between period 0 and period 1 is estimated as the ratio of the average compensation for that category's jobs in period 1 to that in period 0.[5] To ensure that this estimate is not affected by changes in the sample, only those in the sample in both periods are used in the calculation. A similar procedure is then used to calculate the mean change in compensation between periods 1 and 2, 2 and 3, and so on. The proportionate change in mean compensation from time 0 to *t* can then be calculated as the product of the individual per-period changes.

Alternative index numbers

Having reviewed the way that the ECI is presently calculated, we now turn to alternative formulas for aggregating the cell compensation relatives to obtain a single index number. Recall that the ECI is calculated as the weighted sum of the compensation relatives for the various categories of labor, where the weight for a given category is simply that category's share of total labor compensation in the base period. The CPI uses a similar weighting scheme. Specifically, the CPI is a weighted sum of price changes, where a particular component's weight is equal to its share of total household spending in the base period. An index using this kind of weighting scheme is known as a *Laspeyres index*. The Laspeyres index is not without its problems. To see why, suppose that the price of a good rises relative to that of another good. Economic theory holds that utility-maximizing consumers would tend to purchase less of the first good, and more of the second. A Laspeyres price index, however, does not allow for this *substitution ef-*

fect, and hence tends to overstate the increase in the cost of living.[6] Similarly, a Laspeyres compensation index may be expected to overstate increases in the price of *labor* because it does not allow for the fact that employers may be able to substitute one type of worker for another in response to changes in relative compensation.

Other weighting schemes exist besides that used in the Laspeyres index. A *Paasche index*, for example, uses current period quantities to aggregate across the various price relatives.[7] Thus, if the ECI were computed as a Paasche index, it would be calculated as the weighted sum of the changes in compensation costs for all of the industry–occupation job cells, where the weighting factor for each cell is what that cell's share of total compensation would have been in the base period had employers purchased current period inputs in the base period.

While it is not clear on purely theoretical grounds whether the Laspeyres or Paasche index is preferable, something can be said about the potential biases in the two indexes. Other things remaining equal,[8] if the cost of employing one type of labor rises relative to that of another type, employers will tend to substitute in favor of the less costly type of labor. Thus, by ignoring this substitution effect, the Laspeyres index will tend to overstate employers' labor cost in the *reference period*, while the Paasche index will tend to overstate employers' labor cost in the *base year*.[9]

Because the Laspeyres index tends to overstate increases in labor costs, and the Paasche index tends to understate them, it might seem sensible to take an average of the two indexes. In fact, the *Fisher ideal* index does precisely that, being simply a geometric average of the Laspeyres and Paasche indexes. The Törnquist index—which is simply a weighted geometric mean of the price relatives, where the weights are the average shares of spending on the various inputs in the two years—is yet another appealing type of index. Not only are the Fisher ideal and Törnquist indexes intuitively appealing, they also have some desirable theoretical properties. For this reason, Diewert calls these indexes superlative.[10]

One last type of index that requires some discussion is the chained index. The chained Laspeyres index is constructed by chaining together the series of one-period Laspeyres indexes, each of which has a different base and thus uses different weights. The chained Paasche, chained Fisher ideal, and chained Törnquist indexes are all defined similarly. A chained index has the advantage that no single period is singled out to play an asymmetric role. Consequently, chaining tends to reduce the discrepancy between the Paasche and Laspeyres indexes.

Comparing alternative index numbers

We now consider how the various indexes compare in practice. For convenience, only indexes for private industry employment are computed. To focus on the effects of the different index number formulas, all of the calculations use the same official estimates for the cell means.

The available data make it possible to compute the alternative index numbers from September 1981 to December 1994. One practical difficulty in calculating the various indexes is that one needs employment counts for every category of labor in every quarter. (The exception is the Laspeyres index, for which these counts are required only for the base period.) Because the Census of Population only collects employment counts every 10 years, the quarterly employment counts are estimated using industry employment counts from the BLS Current Employment Statistics program in conjunction with the ECI sample weights.[11] The same method for constructing current weights is used in the official calculation of the Employer Cost for Employee Compensation each March.

Two further complications, though minor, should be noted. First, although the ECI is a fixed-weight or Laspeyres index, the weights are revised periodically. For the period through March 1986, ECI weights are constructed using employment counts from the 1970 Census of Population. Cost changes beyond March 1986 are measured using 1980 weights.[12] The official index in periods after March 1986 thus chains together the index in March 1986 with the average change in the cost of compensation between March 1986 and the reference period (using 1980 weights in the latter period).

The second complication to note is that the ECI cell definitions have undergone some minor changes over time. For the present analysis, it is necessary to aggregate across some cells to obtain consistency over time. The effect of these changes turns out to be minor, however, as the indexes calculated here are nearly identical to the published indexes.

Table 1 presents quarterly Laspeyres, Paasche, Fisher ideal, and Törnquist indexes for total compensation. As expected, the Fisher ideal and Törnquist indexes lie between the Laspeyres and Paasche indexes. Contrary to initial expectations, however, the Paasche index appears to yield higher estimates of the increase in the cost of compensation than does the Laspeyres index.[13] There is a plausible explanation for this.

In arguing that the Laspeyres index should be higher than the Paasche index, it was implicitly assumed that changes in relative wages were the only cause of changes over time in employers' relative demands for the various types of labor. However, there are two other factors that may affect employers' relative labor demands. First, technological change may be non-neutral, in that worker productivity may grow at different rates in different industries and occupations. Second, the relative demands for the products of the various industries may change over time. If one type of labor becomes more productive relative to a second type, then

Table 1. Total compensation indexes, September 1981 to December 1994

[September 1981=100]

Quarter	Index			
	Laspeyres	Paasche	Fisher	Törnquist
1981:				
September	100.0	100.0	100.0	100.0
December	102.1	102.1	102.1	102.1
1982:				
March	103.4	103.4	103.4	103.4
June	105.0	105.1	105.0	105.0
September	106.9	107.1	107.0	107.0
December	108.2	108.5	108.3	108.3
1983:				
March	110.2	110.5	110.3	110.3
June	111.6	111.9	111.8	111.8
September	113.1	113.3	113.2	113.2
December	114.2	114.5	114.4	114.4
1984:				
March	116.1	116.5	116.3	116.3
June	117.4	117.7	117.5	117.5
September	118.6	118.8	118.7	118.7
December	119.9	120.2	120.1	120.1
1985:				
March	121.3	121.7	121.5	121.5
June	122.5	122.7	122.6	122.7
September	124.0	124.4	124.2	124.2
December	124.5	125.0	124.8	124.8
1986:				
March	125.9	126.6	126.2	126.3
June	126.9	127.7	127.3	127.4
September	127.7	127.8	127.8	127.8
December	128.5	128.6	128.5	128.5
1987:				
March	129.6	129.9	129.8	129.8
June	130.8	131.0	130.9	130.9
September	132.0	132.3	132.1	132.2
December	133.1	133.4	133.2	133.2
1988:				
March	135.1	135.8	135.4	135.5
June	136.7	137.2	136.9	136.9
September	138.2	138.8	138.5	138.5
December	139.3	139.7	139.5	139.5
1989:				
March	141.2	141.6	141.4	141.3
June	142.6	143.2	142.9	142.9
September	144.6	145.2	144.9	144.8
December	146.0	146.7	146.3	146.3
1990:				
March	148.2	149.3	148.7	148.7
June	150.2	151.1	150.6	150.6
September	151.6	152.7	152.2	152.1
December	152.7	153.9	153.3	153.3
1991:				
March	154.7	156.4	155.6	155.5
June	156.5	158.0	157.3	157.2
September	158.2	161.0	159.6	159.8
December	159.4	162.2	160.8	161.0
1992:				
March	161.4	164.4	162.9	163.1
June	162.8	165.7	164.2	164.5
September	164.0	166.8	165.4	165.6
December	165.4	168.2	166.8	167.0
1993:				
March	167.7	170.6	169.1	169.3
June	169.2	172.3	170.7	171.0
September	170.4	173.6	172.0	172.3
December	171.6	174.8	173.2	173.4
1994:				
March	173.0	176.6	174.8	175.0
June	174.5	177.3	175.9	176.2
September	175.9	178.7	177.3	177.6

employers will demand relatively more workers of the first type and relatively fewer workers of the second type. This will cause the compensation paid the first category of labor to rise relative to that paid the second category of labor, which in turn will induce a supply response, with the number of workers of the first type increasing, and the number of the second type falling. Thus, technological change will lead to a *positive* correlation between changes in relative wages and employment across sectors. By the same reasoning, changes in relative demands for the products of the various industries also will lead to a positive correlation between changes in relative wages and employment across sectors.[14] This will cause the Paasche index to be *higher* than the Laspeyres index.

Table 2 provides additional insight into the demand effect referred to above. For each of the major industrial divisions, the first column of the table indicates compensation in December 1994 relative to that in September 1981. The second column indicates the difference between an industry's (average) December 1994 Paasche weight and its (average) Laspeyres weight.[15] Finally, the third column indicates each industry's December 1994 employment relative to its September 1981 employment. Note that the greatest growth in compensation occurred in the service industries. Services also had very high employment growth. This is consistent with the hypothesis that an increase in the demand for the service industry's output has caused an increase in relative compensation in the service industry, which in turn has induced a response in the amount of labor supplied.

Perhaps the most important observation to be made concerning table 1 is that the differences among the various indexes are quite small. The increase in the Laspeyres index from 100 in September 1981 to 176.5 in December

1994 corresponds to an average annual growth rate of 4.38 percent. During the same period, the Paasche total compensation index grew at an average annual rate of 4.5 percent. Thus, the average annual growth rate of the Paasche total compensation index exceeds that of the Laspeyres index by only 0.12 percent. Apparently, the substitution and demand effects largely offset each other, although there is no way of telling whether these effects individually are large or small.

To complete the analysis, chained indexes are presented in table 3. Note that the same basic pattern found in table 1 is also found in table 3—the Paasche index yields slightly higher estimates of compensation growth than the Laspeyres index, and these two indexes bound the Fisher and Törnquist indexes. However, the difference between the chained Paasche and Laspeyres indexes is smaller than the difference between the unchained Paasche and Laspeyres indexes. This last result should not be surprising because the weights of the chained Laspeyres index change over time, unlike those of the unchained Laspeyres index.

Chained geometric cell means

As discussed above, aggregating individual job quotes to obtain the ECI involves two key steps. The first step is to obtain cell means and the second step is to aggregate over the cell means to obtain a single index number. The previous section analyzed the sensitivity of the ECI to the method chosen to aggregate over the various industry–occupation cells. This section focuses on the process by which individual job quotes are aggregated to obtain cell means.

Recall that the current quarter's average compensation for a given category of labor is estimated by chaining together the proportionate changes in average compensation for that category in all previous quarters, the proportionate change in compensation in each previous quarter being calculated as the ratio of mean compensation in that quarter to mean compensation in the prior quarter. Instead of using arithmetic means to calculate the proportionate changes in compensation each quarter, one might use geometric means.

Previous work has shown that chained Laspeyres price indexes (which are themselves weighted arithmetic means) are subject to upward drift when prices oscillate.[16] Brent R. Moulton has recently calculated Laspeyres arithmetic and geometric mean indexes for the most important items in the CPI for the period June 1992 to June 1993.[17] (Spending on these goods and services accounts for 70 percent of spending on all CPI items.) While the arithmetic mean index exceeds the geometric mean index for every item, the difference between the indexes varies greatly across the various goods and services, with the difference between the indexes being especially large for goods that have highly variable prices. For example, the arithmetic and geometric mean indexes for fruits and vegetables differ by 3 percent and the indexes for women's and girls' apparel differ by 2.87 percent. For all items together, the arithmetic mean index exceeds the geometric mean index by half a percent.

There are some unique features in the old CPI methodology that made the CPI especially sensitive to the use of geometric versus arithmetic means.[18] These are not relevant to the ECI. However, following up on Moulton's (1996) suggestion,[19] one consideration that needs to be taken into account when choosing between an arithmetic and geometric mean is that the distribution of wages within a cell may be changing over time. This can cause the geometric and arithmetic means to diverge over time. Which one is preferable will depend on the actual form of the compensation distribution. As demonstrated by Michael K. Lettau, Mark A. Loewenstein, and Aaron Cushner,[20] if the compensation within each cell is dis-

Table 2. Relative compensation and employment growth by major industry

Industry	December 1994 total compensation divided by September 1981 total compensation	Difference between Paasche and Laspeyres weights	December 1994 employment divided by September 1981 employment
Mining	1.61	–0.012	0.49
Construction	1.68	–.008	1.13
Manufacturing	1.83	–.098	.89
Durables	1.82	–.064	.85
Nondurables	1.83	–.034	.95
Transportation and public utilities	1.68	–.013	1.14
Transportation	1.67	–.008	1.18
Public utilities	1.75	–.005	.98
Wholesale and retail trade	1.68	–.011	1.32
Wholesale trade	1.77	–.009	1.14
Retail trade	1.64	–.002	1.39
Finance, insurance, and real estate	1.76	.030	2.18
Banking, savings and loan, and other credit agencies	1.72	.031	8.32
Insurance	1.86	–.002	1.26
Other finance, insurance, and real estate	1.69	.002	1.35
Services	2.03	.111	1.88
Business services	1.78	.061	3.42

tributed log normally,[21] arithmetic mean compensation for a cell will be growing more quickly (slowly) than geometric mean compensation for the cell if the variance of log compensation is increasing (decreasing) over time. Interestingly, the inequality literature seems to indicate that the variance of log compensation has been rising in recent years, which suggests that the arithmetic cell means will be growing more quickly than the geometric cell means.[22]

Is the ECI sensitive to the use of arithmetic versus geometric cell means? The available data make it possible to compute the alternative indexes from March 1986 to December 1994. An analysis of the data indicates that the average within-cell variance in log compensation grew at an annual rate of 0.9 percent from March 1986 to December 1994. This translates into a small predicted difference in the annual growth rate between the Laspeyres index based on arithmetic means and the Laspeyres index based on geometric means of 0.06 percent.

The actual effect of calculating average compensation using geometric as opposed to arithmetic means can be seen in table 4. Column 1 of the table presents the index that results when one uses Laspeyres September 1981 employment weights and arithmetic cell means. The index obtained when one uses Laspeyres weights and geometric cell means appears in column 2. According to the table, the Laspeyres index calculated using arithmetic cell means increased by 39.8 percent from March 1986 to December 1994, while the Laspeyres index constructed using geometric cell means increased by 39.4 percent. Thus, the geometric mean index does indeed grow at a slower rate than the arithmetic mean index. However, the difference in annual growth rates is only 0.02 percent, which is even smaller than the predicted difference of 0.06 percent. For completeness, the Paasche index has also been calculated

Table 3. Chained total compensation indexes, September 1981 to December 1994

[September 1981=100]

Quarter	Index			
	Laspeyres	Paasche	Fisher	Törnquist
1981:				
September	100.0	100.0	100.0	100.0
December	102.1	102.1	102.1	102.1
1982:				
March	103.4	103.4	103.4	103.4
June	105.0	105.1	105.1	105.1
September	107.0	107.0	107.0	107.0
December	108.3	108.3	108.3	108.3
1983:				
March	110.2	110.3	110.3	110.3
June	111.7	111.7	111.7	111.7
September	113.2	113.2	113.2	113.2
December	114.4	114.4	114.4	114.4
1984:				
March	116.4	116.3	116.3	116.3
June	117.6	117.6	117.6	117.6
September	118.8	118.7	118.8	118.8
December	120.1	120.1	120.1	120.1
1985:				
March	121.5	121.5	121.5	121.5
June	122.7	122.6	122.6	122.6
September	124.2	124.2	124.2	124.2
December	124.8	124.8	124.8	124.8
1986:				
March	126.2	126.2	126.2	126.2
June	127.2	127.4	127.3	127.3
September	128.0	128.2	128.1	128.1
December	128.8	128.9	128.9	128.9
1987:				
March	130.1	130.3	130.2	130.2
June	131.2	131.3	131.3	131.3
September	132.5	132.7	132.6	132.6
December	133.6	133.7	133.7	133.7
1988:				
March	135.6	135.8	135.7	135.7
June	137.1	137.4	137.2	137.2
September	138.6	138.9	138.8	138.8
December	139.8	140.1	140.0	140.0
1989:				
March	141.6	141.9	141.8	141.8
June	143.3	143.6	143.4	143.4
September	145.2	145.5	145.4	145.4
December	146.6	146.9	146.8	146.8
1990:				
March	149.0	149.3	149.1	149.1
June	150.9	151.3	151.1	151.1
September	152.5	152.9	152.7	152.7
December	153.5	153.9	153.7	153.7
1991:				
March	155.6	156.0	155.8	155.8
June	157.3	157.7	157.5	157.5
September	159.0	159.5	159.2	159.2
December	160.1	160.6	160.3	160.3
1992:				
March	162.0	162.5	162.2	162.2
June	163.4	164.0	163.7	163.7
September	164.6	165.2	164.9	164.9
December	166.1	166.6	166.3	166.3
1993:				
March	168.2	168.7	168.4	168.4
June	169.7	170.2	170.0	170.0
September	171.0	171.5	171.2	171.2
December	172.0	172.6	172.3	172.3
1994:				
March	173.7	174.2	174.0	174.0
June	174.8	175.3	175.1	175.0
September	176.2	176.7	176.4	176.4
December	176.9	177.4	177.2	177.2

Table 4. Total compensation indexes, 1986–94

[March 1986=100]

Quarter	Index			
	Laspeyres		Paasche	
	Arithmetic cell means	Geometric cell means	Arithmetic cell means	Geometric cell means
1986:				
March	100.0	100.0	100.0	100.0
June	100.8	100.7	100.8	100.7
September	101.5	101.5	101.5	101.4
December	102.2	102.1	102.1	102.1
1987:				
March	103.2	102.9	103.1	103.0
June	103.9	103.8	104.0	104.0
September	104.9	104.7	105.0	104.9
December	105.7	105.6	105.9	105.9
1988:				
March	107.2	107.3	107.6	107.7
June	108.5	108.5	108.8	108.8
September	109.6	109.6	110.1	110.1
December	110.7	110.6	111.0	111.1
1989:				
March	112.1	112.0	112.4	112.5
June	113.4	113.1	113.8	113.7
September	114.8	114.5	115.4	115.3
December	115.9	115.6	116.5	116.4
1990:				
March	117.7	117.4	118.4	118.3
June	119.2	119.0	119.9	119.8
September	120.4	120.2	121.1	121.1
December	121.3	121.0	122.0	122.0
1991:				
March	123.0	122.5	123.8	123.6
June	124.4	124.0	125.2	125.1
September	125.7	125.3	126.7	126.6
December	126.6	126.2	127.7	127.6
1992:				
March	128.1	127.8	129.2	129.2
June	129.0	128.8	130.2	130.2
September	130.0	129.8	131.2	131.1
December	130.9	130.7	132.3	132.1
1993:				
March	132.6	132.3	133.8	133.7
June	133.7	133.4	134.9	134.7
September	134.9	134.4	136.0	135.7
December	135.7	135.2	136.9	136.7
1994:				
March	137.1	136.6	138.3	138.1
June	138.1	137.7	139.2	138.9
September	139.3	138.9	140.4	140.2

using geometric cell means. As can be seen from columns 3 and 4 of table 4, the Paasche arithmetic and geometric mean indexes grew by nearly identical amounts over the period.[23]

THE RESULTS OF THIS STUDY INDICATE that the estimation of compensation growth is not very sensitive to the choice of the index formula employed. An additional finding is that one has to be careful in applying standard index number analysis to the ECI or even the CPI. Changes in prices can be caused by shifts in demand, sup-, ply, or both, but the standard interpretation presumes implicitly that only supply curves are shifting. Interpretation of the CPI is made more difficult if changes in income lead to changes in consumer demands, which in turn cause changes in relative output prices. Similarly, changes in labor demand will make the fixed-base ECI somewhat difficult to interpret from the standpoint of standard index number analysis. Changes in technology and changes in the relative demands for industries' products will both lead to relative changes in the demand for labor, which will in turn lead to changes in relative wages. The changes in relative wages in turn induce shifts in the amounts of labor workers supply to the various industries. Indeed, the finding that the Paasche version of the ECI is actually slightly higher than the Laspeyres version indicates that this effect more than offsets the standard substitution effect. The fact that the superlative indexes can accommodate a changing technology and output mix is an additional argument for their use (although in the case of the ECI this advantage may be offset by the fact that current employment counts must be partially estimated from the ECI sample weights and thus may suffer from measurement error).[24] Intuitively, if the cost index is to accommodate a technology and output mix that is changing over time, then the expenditure weights must take into account prices and quantities in both the base year and the current year. □

Footnotes

ACKNOWLEDGEMENT: The views expressed here are those of the authors and do not necessarily reflect the views of the U.S. Department of Labor or the Bureau of Labor Statistics. The authors are grateful to Al Schwenk for answering their many questions about the ECI data, and for providing them with estimates of current employment counts. They have also received helpful suggestions from Katharine Abraham, Anthony Barkume, Brent Moulton, Marshall Reinsdorf, John Ruser, and Kimberly Zieschang.

[1] The ECI covers all establishments and occupations in both the private nonfarm and public sectors. As such, private sector coverage is limited to the private nonfarm economy, excluding private household workers. Public sector coverage includes employees of State and local governments, but ex-

cludes Federal workers. Finally, the self-employed, owner-managers, and unpaid family workers are excluded from coverage in the ECI.

For a more complete description of the ECI, see *BLS Handbook of Methods*, Bulletin 2414 (Bureau of Labor Statistics, 1992), ch. 8, "The Employment Cost Index." Also, see the following two articles from the *Monthly Labor Review*: Victor Sheifer, "Employment Cost Index: a measure of change in the 'price of labor'," July 1975, pp. 3–12; and G. Donald Wood, "Estimation procedures for the Employment Cost Index," May 1982, pp. 40–42.

[2] Prior to March 1995, only nine major occupational groups were used.

[3] Note that an analogous procedure is used to calculate the CPI. Specifically, the goods and services that consumers purchase are classified into 207 groups and urban areas are divided into 44 areas. The first step in constructing the CPI involves estimating the mean price of each of the 9,108 (207*44) basic CPI components. The second step involves aggregating the basic components to obtain a single index number. For a more detailed description, see the following articles in the December 1993 issue of the *Monthly Labor Review*: Dennis Fixler, "The Consumer Price Index: underlying concepts and caveats," pp. 3–12; Brent R. Moulton, "Basic components of the CPI: estimation of price changes," pp. 13–24; and Ana M. Aizcorbe and Patrick C. Jackman, "The commodity substitution effect in CPI data, 1982–91," pp. 25–33.

[4] Beginning in March 1995, the employment numbers used for the weights are from the BLS Occupational Employment Statistics (OES) survey, rather than the Census of Population.

[5] The ECI sample weights indicate exactly how much weight should be given to each job quote. Once a job is surveyed, its sample weight is held fixed during the entire time that it remains in the sample. As a consequence, the sample weights in period t will not strictly reflect current employment in period t. The importance of this consideration is limited, however, because 20 to 25 percent of the sample is replenished every year.

[6] For a textbook discussion of the Laspeyres and Paashe indexes and the substitution effect, see John M. Barron, Mark A. Loewenstein, and Gerald J. Lynch, *Macroeconomics* (Reading, Addison–Wesley, 1988), pp. 34–9. See also Hal R. Varian, *Intermediate Microeconomics*, 3rd ed. (New York, W.W. Norton and Company, 1993), pp. 129–33 (index numbers) and pp. 136–41 (substitution effect).

[7] The Gross Domestic Product deflator is an example of a Paasche index.

[8] Of course, one of the things that could change is output. Scale will not affect relative factor demands if the aggregate production function is homothetic. Another thing that might change is technology. If technological change is neutral, then it too will not affect relative factor demands. We return to this point below.

[9] Aizcorbe and Jackman, in "Commodity substitution effect," estimate that by ignoring the substitution effect, the CPI overstates the annual increase in the cost of living by about 0.2 percent during the 1982–91 period. Marilyn E. Manser and Richard J. MacDonald, in "An Analysis of Substitution Bias in Measuring Inflation, 1959–85," *Econometrica*, July 1988, pp. 909–30, obtain a similar result for the 1959–85 period. Analyzing the 1958–73 period, Steven D. Brathwait, in "Substitution Bias in the Laspeyres Price Index: An Analysis Using Estimated Cost-of-Living Indexes," *American Economic Review*, March 1980, pp. 64–77, estimates an annual substitution effect of only 0.1 percent. As noted by Aizcorbe and Jackman, Braithwait's smaller estimated effect probably results from using more-aggregate cells to construct the index.

[10] For further discussion of this point, see W.E. Diewert, "Index Numbers," in *The New Palgrave: A Dictionary of Economics* (New York, MacMillan Press, 1987), pp. 767–79. Also, see the authors' longer version of the present article, same title, Mimeo. (Bureau of Labor Statistics, 1996).

[11] The authors thank Al Schwenk, of the Office of Compensation and Working Conditions, Bureau of Labor Statistics, for deriving the employment estimations.

[12] In March 1995, the ECI began using 1990 weights. For further discussion of the reweighting of the ECI, see Albert E. Schwenk, "Introducing new weights for the Employment Cost Index," *Monthly Labor Review*, June 1985, pp. 22–27; and Albert E. Schwenk, "Introducing 1990 weights for the Employment Cost Index," *Compensation and Working Conditions*, June 1995, pp.1–5.

[13] BLS also publishes separate indexes for wages and salaries and for benefits. The relative ordering of the Laspeyres, Paasche, and superlative indexes is the same for the separate indexes.

[14] Note that the above arguments do not imply that there will be relatively greater employment growth in sectors with high wage *levels* than in sectors with low wage *levels*.

[15] More precisely, the second column is the difference between the Paasche and Laspeyres weights that result when cells are defined only by major industry.

[16] See F. G. Forsyth and R. F. Fowler, "The Theory and Practice of Chain Price Numbers," *Journal of the Royal Statistical Society*, Series A, 1981, pp. 224–46; Bohdan J. Szulc, "Linking Price Index Numbers," in W. E. Diewert and C. Montmarquette, eds., *Price Level Measurement: Proceedings from Conference Sponsored by Statistics Canada* (Ottawa, Minister of Supply and Services, 1983), pp. 537–66; and Jorgen Dalen, "Computing Elementary Aggregates in the Swedish Consumer Price Index," *Journal of Official Statistics*, 1992, pp. 129–47.

[17] See Brent R. Moulton, "Basic components of the CPI: estimation of price changes," *Monthly Labor Review*, December 1993, pp. 25–33.

[18] The following three articles discuss rotation bias in the old CPI: Marshall Reinsdorf, "The Effect of Outlet Price Differentials in the U.S. Consumer Price Index," in M.F. Foss, M.E. Manser, and A.H. Young, eds., *Price Measurements and their Uses*, NBER Studies in Income and Wealth (Cambridge, MA, National Bureau of Economic Research, 1993) vol. 57, pp. 227–54; Marshall Reinsdorf and Brent R. Moulton, "The Construction of Basic Components of Cost of Living Indexes," in T. Bresnahan and R.J. Gordon, eds., *New Goods*, Proceedings of the NBER Conference on New Products (Cambridge, MA, National Bureau of Economic Research, 1996); and Brent R. Moulton, "Bias in the Consumer Price Index: What is the Evidence?" *Journal of Economic Perspectives*, Fall 1996, pp. 159–79. For a summary of BLS research on the issue, see Brent R. Moulton, "Estimation of Elementary Indexes of the Consumer Price Index," Mimeo. (Bureau of Labor Statistics, May 1996). To eliminate rotation bias in the CPI, BLS has recently changed the way it calculates the weights of replacement items. See Bureau of Labor Statistics, "Extending improvements in CPI sample rotation procedures for substitute items," Mimeo., March 29, 1996.

[19] See Brent R. Moulton, "Estimation of Elementary Indices of the Consumer Price Index," Mimeo. (Bureau of Labor Statistics, May 1996).

[20] See Michael K. Lettau, Mark A. Loewenstein, and Aaron Cushner, "Is the ECI Sensitive to the Method of Aggregation?" Mimeo. (Bureau of Labor Statistics, August 1996).

[21] This assumption is common in the literature. The distribution of wages in the economy tends to be skewed rightward. The log normal distribution has this property. In contrast, the normal distribution is symmetric.

[22] See, for example, Frank Levy and Richard Murnane, "U.S. Earnings Levels and Earnings Inequality: A Review of Recent Trends and Proposed Explanations," *Journal of Economic Literature*, 1992, pp. 1333–81.

[23] A comparison of indexes for wages and salaries and for benefits using arithmetic and geometric means yields very similar results.

[24] For further discussion of this point, see Michael K. Lettau, Mark A. Loewenstein, and Aaron Cushner, "Is The ECI Sensitive to the Method of Aggregation?" Mimeo. (Bureau of Labor Statistics, August 1996).

Appendix: Index calculation

This appendix describes the calculation of the ECI and the alternative indexes more formally. Let W_{it} denote the mean compensation paid to category i workers in period t and let E_{it} denote the number of category i workers employed in period t. Letting 0 denote the base period, the Employment Cost Index in year t, ECI_t, is calculated as

$$(1) \quad ECI_t = \sum_i \alpha_i^L \frac{W_{it}}{W_{i0}} \cdot 100,$$

where

$$(2) \quad \alpha_i^L = \frac{E_{i0} W_{i0}}{\sum_i E_{i0} W_{i0}}$$

Out of the ECI sample in period t, let I_t denote the subsample of jobs corresponding to labor category i. In addition, let W_{ijt} denote the period t compensation for the j^{th} job quote in cell i, and let W_{ijt-1} denote the corresponding compensation in period t–1. Finally, let s_{ij} denote the sample weight corresponding to the j^{th} job quote in cell i. Letting τ be an arbitrary period between periods 0 and t, the proportionate change, $r_{i\tau}$, in the average compensation paid type i workers between period 1 and τ is estimated as

$$1 + r_{i\tau} = \frac{\sum_{j \in I_\tau} S_{ij\tau} W_{ij\tau}}{\sum_{j \in I_\tau} S_{ij\tau} W_{ij\tau-1}}$$

$$(3) \quad = \sum_{j \in I_\tau} S_{ij\tau}' \frac{W_{ij\tau}}{W_{ij\tau-1}},$$

where

$$S_{ij\tau}' = \frac{S_{ij\tau} W_{ij\tau-1}}{\sum_{j \in I_\tau} S_{ij\tau} W_{ij\tau-1}}$$

is the implicit expenditure weight for the j^{th} job quote in cell in i in period τ. The proportionate change in compensation for category i from period 0 to period t is then calculated as

$$(4) \quad \frac{W_{it}}{W_{i0}} = (1 + r_{i1})(1 + r_{i2}) \ldots (1 + r_{it})$$

If the ECI were computed as a Paasche index, one would use an equation like (1), but with weights defined by

$$(5) \quad \alpha_{it}^p = \frac{E_{it} W_{i0}}{\sum_i E_{it} W_{i0}}$$

The Fisher ideal index, F_t, and the Törnquist index, T_t, are given by

$$(6) \quad F_t = L_t^{1/2} P_t^{1/2}$$

and

$$(7) \quad T_t = \prod_{j=1}^{N} \left(W_{jt} / W_{j0} \right)^{\alpha_{jt}^T} \cdot 100$$

where

$$(8) \quad \alpha_{jt}^T = (1/2) W_{j0} E_{j0} / \sum_{k=1}^{N} W_{k0} E_{k0} + (1/2) W_{jt} E_{jt} / \sum_{k=1}^{N} W_{kt} E_{kt}$$

To obtain current employment weights, the fraction of industry m employment in major occupation n is estimated by

$$(9) \quad S_{mnt} = \frac{\sum_{j \in (MN)_t} s_{mnj}}{\sum_n \sum_{j \in (MN)_t} s_{mnj}}$$

where $(MN)_t$ denotes the set of industry m –occupation n jobs for which compensation quotes are available in periods t–1 and t. An estimate of industry m employment E_{mt} is obtained from the Current Employment Statistics (CES) survey. Total period t employment in industry m and occupation n is estimated as $E_{mnt} = E_{mt} \cdot S_{mnt}$.

The proportionate change in geometric mean compensation between period τ–1 and τ is given by

$$(10) \quad \ln(1 + r_{it}^g) = \sum_{j \in I_t} S_{ijt} \ln(W_{ijt}) - \sum_{j \in I_t} S_{ijt} \ln(W_{ijt-1}) \equiv \gamma_{it}$$

and the ratio of geometric mean period t compensation to mean period 0 compensation for cell i is calculated using the formula

$$(11) \quad \frac{W_{it}^g}{W_{i0}^g} = \exp\left(\sum_{\tau=1}^{t} \gamma_{i\tau}\right)$$