A Note on Estimating Small Sample Bias in Two Price Index Formulas

DRAFT: PLEASE DO NOT QUOTE WITHOUT PERMISSION

Robert McClelland
Marshall Reinsdorf

December 31 1996

In this analysis we compare the finite sample bias (aka small sample bias) of the geometric means index and the 'seasoned modified Laspeyres index ' (hereafter referred to as the seasoned index.) Our tentative conclusion is that the geometric means index has a larger, growing rate of small sample bias while the seasoned index has a smaller, fluctuating bias.

Given an index $I_n(P_t,P_{t+k})$ of a sample of n quotes using price vector P_t in time t and price vector P_{t+k} in time t+k, small sample bias occurs when

$$E\left[I_{n^*}(P_t,P_{t+k})\right] \neq I_n(P_t,P_{t+k}),$$

where n* is less than the population of quotes n. This tends to happen in price indexes because they are often ratios of the form $i(P_{t+k})/i(P_t)$, and by Jenson's inequality, $E[i(P_{t+k})/i(P_t)] < E[i(P_{t+k})]/ E[i(P_t)]$.

The first of the two indexes considered here is the seasoned index

$$S_n(P_{t+k},P_t) = \frac{\sum_j^n (E_{j,B}/P_{j,I})P_{j,t+k}}{\sum_j^n (E_{j,B}/P_{j,I})P_{j,t}},$$

where $P_{j,t}$ is the price of good j at time t. E_{jB} is the total expenditures on good j in the base period B, the time period during which the expenditures are made. Rather than occurring in a single month, this period varies from one week to five years, depending upon the item stratum. In addition, the expenditures need to be multiplied by several adjustment factors, such as a correction for the difference in the definition of goods between the survey which collects expenditures and the survey which collects the prices. If time period I equals B, then we have a modified Laspeyres index of the form

$$\frac{\sum_j^n Q_{j,B}P_{j,t+k}}{\sum_j^n Q_{j,B}P_{j,t}}.$$

Because prices are not collected by the Bureau of Labor Statistics (BLS) in the base period, prices in some other time period must be used. From 1978 until 1996, this period was the first month in which the prices were used in the index. To correct 'formula bias', the BLS either uses prices when they are 'initiated', which occurs approximately three or four months before the prices are used in an index or simply holds the prices for three or four months before using them in the index. These processes are collectively referred to as "seasoning", and an index using the procedure is a "seasoned" index.

The second index to be considered is the geometric means index

$$G_n(P_{t+k}, P_t) = \prod_j^n \left[P_{j,t+k} / P_{j,t} \right]^{E_{j,B}/\sum E_{j,B}} .$$

One problem with estimating either index is that for each item stratum the BLS collects only a small proportion of all possible price quotes in a given area. To represent this, we can replace $E_{j,B}$ with $\bar{E}_{j,B} \equiv W_{j,B} * N_{j,B}$, with $N_{j,B}$ being a random variable equal to unity if a particular outlet and item is chosen and zero otherwise and $W_{j,B}$ being the product of the adjustment factors mentioned above. The BLS sets the probability of choosing a given outlet proportional to the outlet's expenditures. Within each outlet, the price of a specific good is 'initiated' by being randomly selected in a similar manner. This ensures that the expected value of $\bar{E}_{j,B}$ equals E_{jB}.

To simulate the measurement of small sample bias, we pool quotes nationally into a population from which indexes could be formed using the two formulas. We then calculate the expected value of the index formed with a subset of the population. For the geometric means formula, this expectation is found by exhaustively forming indexes with all possible combinations of quotes. For the seasoned index, this expectation was formed in a similar manner for small populations and through Monte Carlo simulations for larger populations.[1] The link month for all quotes in all strata is assumed to be March, 1993 and indexes compare this month with prices in the 18 month period starting April 1993 and ending September 1994. The 'initiation period' is assumed to be January, 1993. Please see the appendix for additional information on the data.

Our results are in figures one, two and three. In the first two figures, we show the small sample bias of the geometric means and seasoned index using one, two, three, four and the average number of quotes over the 18 months in our sample. This average is the average number of quotes in each of the 44 geographic areas in which indexes are formed. As expected, the bias falls as the number of quotes increases. Seasonality appears to strongly affect the bias of both indexes. However, while the bias of the geometric means index shows a definite upward

[1] For the geometric means formula it is always computationally tractable to use all possible combinations of quotes. The seasoning formula does not allow the same kinds of manipulations possible in the geometric means, so that exact calculation of the expectation is possible for only small populations.

trend, the seasoned index appears to have only a cyclical component. The difference is most striking in figure three, where we compare indexes with the two formula, using the average number of quotes.

One possible explanation for this is that the two indexes formed with the population are moving away from each other, as shown in figure four. Because both formula yield essentially identical indexes when only a single quote is used, the geometric mean must, in some sense, fall farther than the seasoned index.

APPENDIX
DATA SET CONSTRUCTION

The prices used in this analysis are the prices of items in the commodities and services section that are used in calculating the CPI. Item classes in which the average number of quotes was less than two were deleted -- calculated by dividing the total number of quotes by 44, the total number of geographic areas for which indexes are created-- as were strata with an unusually large number of missing prices or with anomalous price behavior. The total weight of these three classes in the CPI is less than 0.5 percent. The number of remaining item strata is 162.

Any quote that was missing in January 1993 or for which the number of quotes in the area was unknown was deleted. If the price for a given quote was missing for every month from January to November of 1994, then the quote was deleted. Because approximately 20 percent of the sample rotates every year this left about 60 percent of the original sample. In addition, any quote in which a noncomparable substitution was made was deleted. If an index was missing because all of the values were missing, the index was imputted using the average of all other indexes.

For the 82 item groups where prices are collected monthly, the index is calculated from March 1993 to later months ending in September 1994. For the 80 groups where prices were collected bimonthly, even-month indexes run from February through August and odd-month indexes run from January through September. The even-numbered months were treated as odd-month quotes (i.e., quotes from February were combined with the January quotes.) There were also several large metropolitan areas that collect the bimonthly items on a monthly basis. Odd-month quotes of this type were used and even-month quotes were ignored. To avoid any potential problem from chaining, indexes between March 1994 and later months were formed by comparing the prices in March 1994 with that month. As with BLS policy, if at least one quote existed in both periods of the index, then missing values were imputed using the index formed by the existing quotes.

The indexes were aggregated using the relative importance weights for December 1994. These weights are the product of the index in a given period and the item's expenditure share estimated from the 1982-4 Consumer Expenditure Survey. While this method is slightly inaccurate, the inaccuracy is only a function of how the relative inflation rates across goods has changed from January to December 1994. The outlet expenditures used for the expenditure weights were taken to be the average expenditures in the area from which a price quote was taken. All

the adjustment factors in the weights, such as the percent of POPS correction, were ignored. For the Monte Carlo simulations, quotes were drawn from the population 20,000 times, with the probability of an outlet being selected was proportional to its expenditures, as defined above.

Official Bureau procedure is to select randomly one entry level item (ELI) to represent the item strata of goods and then randomly select outlets from that ELI proportional to their expenditures. Here, we do not separate the quotes by ELI, but simply sample outlet/quote combinations from the entire item stratum.

FIGURES

Figure 1
Small Sample Bias for Geomeans Method

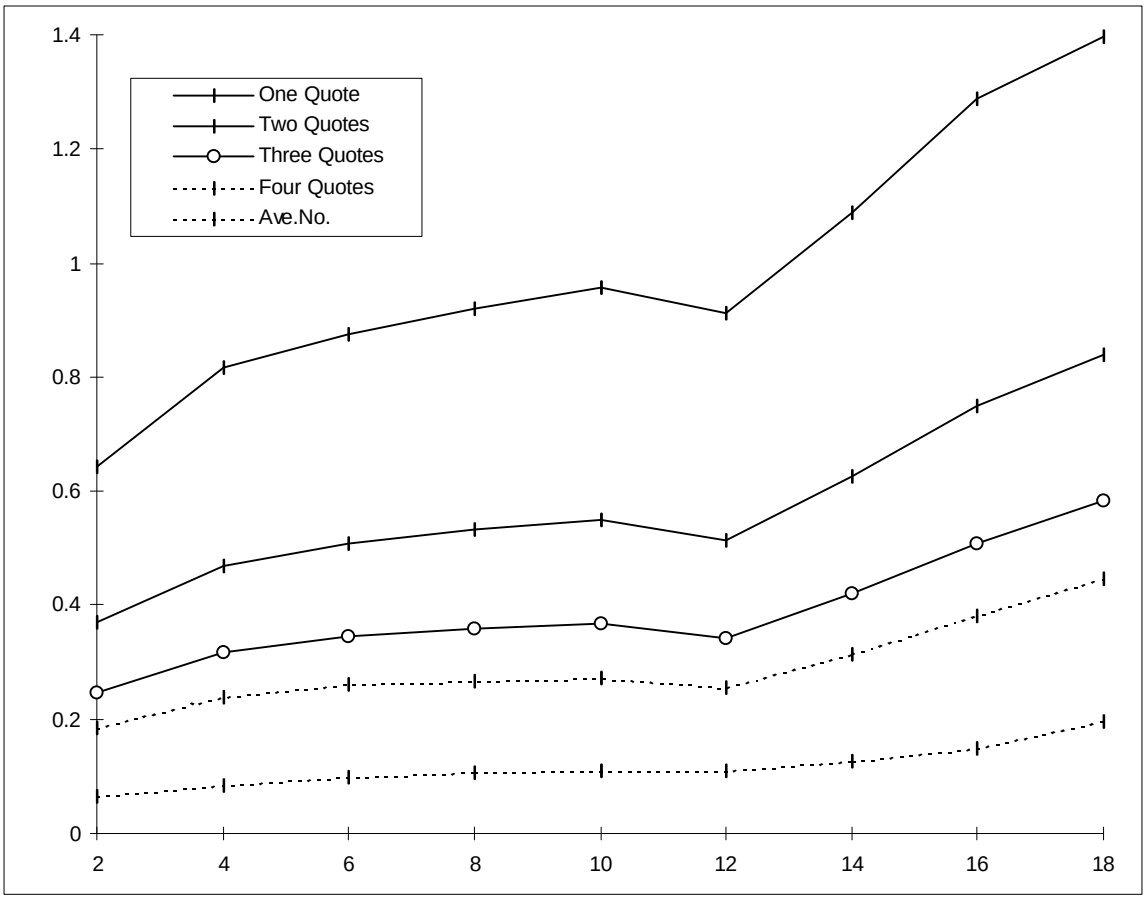

Figure 2
Small Sample Bias for Seasoned Method

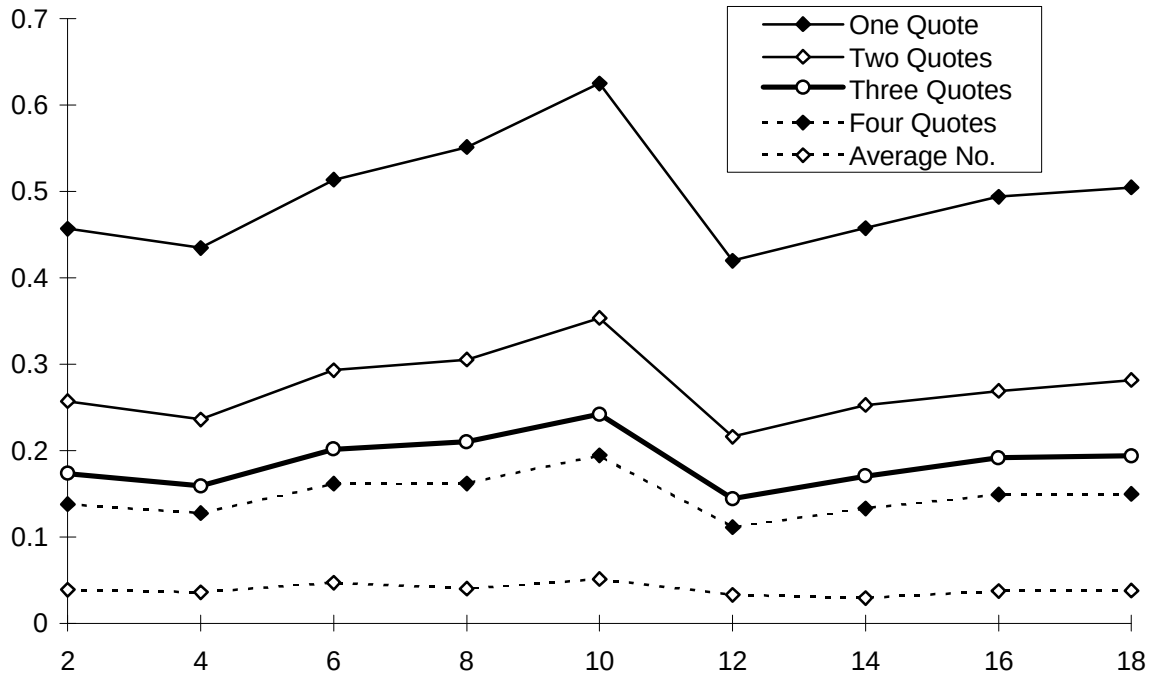

Figure 3

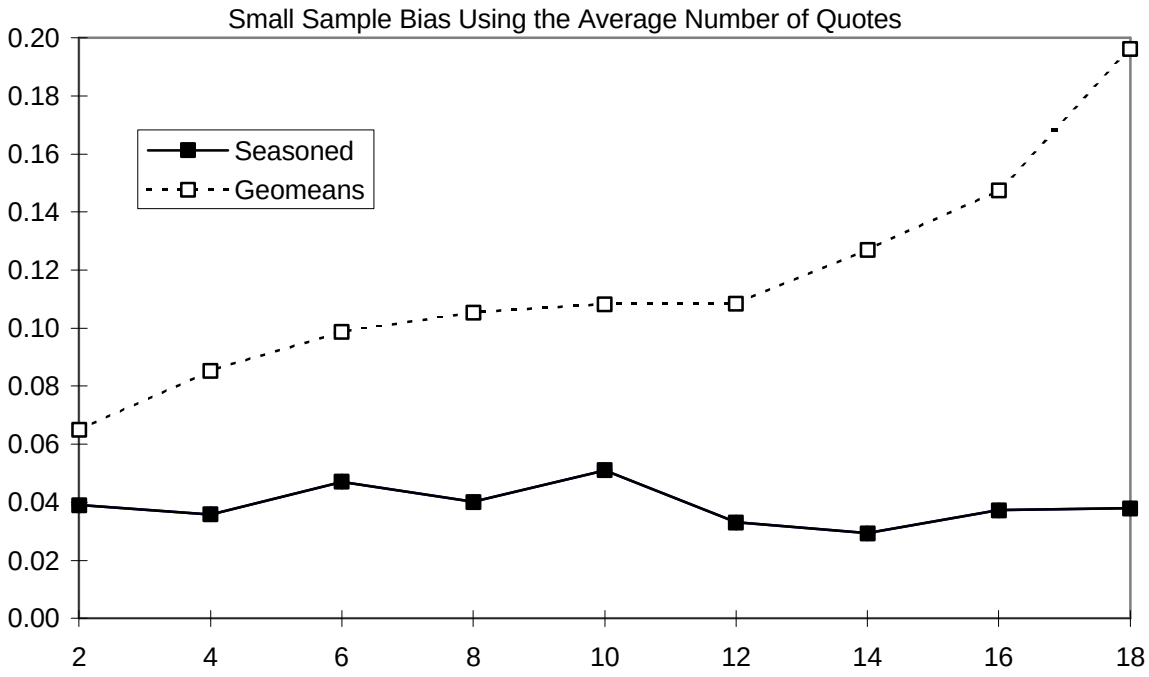

Figure 4
Population Indexes