

# PERFORMANCE COMPARISONS OF LASPEYRES INDEXES WITH GEOMETRIC MEAN INDEXES IN THE U.S. CONSUMER PRICE INDEX

Owen J. Shoemaker, U.S. Bureau of Labor Statistics  
2 Mass Ave., NE, Room 3655, Washington, DC 20212

KEY WORDS: Quantity weight, price relative

The official U.S. Consumer Price Index (CPI) is a Laspeyres index. An alternative form of the index known as the geometric mean (GeoMeans) index is scheduled to be published in 1997. In this paper, rental housing data have been studied with regard to rent levels, to the average weights attributed to these levels and to the average rate of rental price change at these levels. These weights are quantity-based. A two-partition model is constructed, with "low" rents compared with "higher" rents. Simulations are run comparing Laspeyres and GeoMeans index results. The differences between the two indexes are plotted against the rate of price change in the "low" rent sector. In the simulations, all the other factors are treated as fixed. A smooth (quadratic) curve results, and the GeoMeans indexes are shown to score consistently higher than the Laspeyres indexes.

## 1. Background

When the GeoMeans Analysis Team at BLS investigated the CPI's Rent and Rent Equivalence (REQ) price relatives, they discovered that the GeoMeans' Rent/REQ indexes were running consistently higher than their Test Laspeyres counterparts. Rent/REQ appeared to be the exception to the rule that said that GeoMeans indexes will generally come in lower than Test Laspeyres indexes. In fact, for nearly all other basic-level indexes in the CPI (roughly 75% of the Index), the reverse is the case. There, GeoMeans run consistently lower than Laspeyres. Providing an explanation for the higher GeoMeans in Rent/REQ is the subject of this paper.

## 2. Possible Explanations

First of all, unless the Laspeyres' price relative calculation (PRC) is expressed precisely as a weighted arithmetic mean, no exact mathematical comparison can be made between the Laspeyres and its GeoMeans counterpart. Moreover, only if their respective weights are equal will the GeoMeans index be guaranteed to come in lower than a Laspeyres index. Such a structure is not the case here. (Nor is it the case for C&S data, that other 75% of the CPI, but that issue is for a separate investigation.) For Rent/REQ data, the

two BLS production PRCs have the following general forms:

$$\begin{aligned} \text{Laspeyres PRC} = R_{T,T-12}^L &= \frac{\sum_i W_i P_{T,i}^*}{\sum_i W_i P_{T-12,i}}, \\ &= \frac{\sum_i (W_i P_{T-12,i}) \left( \frac{P_{T,i}^*}{P_{T-12,i}} \right)}{\sum_i (W_i P_{T-12,i})}, \end{aligned}$$

where  $W_i = Q_i$  = a quantity weight,  
and  $P_{\bullet,i}$  is a rent price ( $P_{T,i}^* = P_{T,i} + r$ ).

$$\text{GeoMeans PRC} = R_{T,T-12}^G = \prod_i \left( \frac{P_{T,i}^*}{P_{T-12,i}} \right)^{\frac{W_i}{\sum W}}$$

$$= \text{EXP} \left[ \frac{\sum_i W_i \log \left( \frac{P_{T,i}^*}{P_{T-12,i}} \right)}{\sum_i W_i} \right]$$

Neither of these PRCs exactly duplicates BLS production, but the two weight structures do duplicate BLS production. (It should be noted that quantity weights were used in the experimental GeoMeans because more appropriate weights were not available.) Currently,  $S^L = \text{Laspeyres Weights} = W * P_{T-12}$ , while  $S^G = \text{GeoMeans Weights} = W$ . Thus, the two weight structures are not the same and there is no mathematical guarantee that one or the other of these two PRCs will always be lower than the other.

(T-12 is the time designation for previous rents, since that is the time differential that the empirical results in this paper uses. Actual BLS production uses six-month rent changes – plus a depreciation adjustment factor  $r$  in the new rents – for its rent indexes. BLS takes the sixth root of these six-month relatives to calculate one-month price relatives.)

The explanation that we will pursue in this paper focuses on (a) the weighting structure of the rents as it relates to rent-level, and (b) the relative percentage increases in the rent-levels. We will use simulations to compare the relative performances of the two indexes, by making the difference between the two indexes a function of the variables outlined in (a) and (b).

### 3. BasicWeights as Related to Rent-Level

The following table gives summary statistics for BasicWeights (1987 Revision) as they are distributed across three rent levels — low, medium, high. The results confirm that low-rent units do have significantly higher weights than medium- or high-rent units. (Three pair-wise comparison tests –  $H_0$ : LOW=MEDIUM,  $H_0$ : LOW=HIGH, and  $H_0$ : MEDIUM=HIGH – all give P-values less than 0.0001.)

#### BasicWeight Averages By Strata Levels

RENT LEVEL	FREQ	WEIGHT MEAN	WEIGHT STDERR
LOW	7728	4842.6	42.33
MED	13753	4485.4	32.72
HIGH	5803	3767.5	50.06
ALL	27284	4433.9	23.12

It has been suggested that this particular weight distribution, with the higher weights in the lower rent levels, is contributory to producing higher GeoMeans indexes (see internal memoranda between Louise Campbell and Dave Richardson, September 1996). However, as it turns out, the weights have little or nothing to do with the issue at hand. As our simulations will show, the GeoMeans (Rent/REQ) index will come in higher than its Test Laspeyres counterpart whatever the respective weights are.

### 4. Percentage Increases in Rent Levels

In the previously referenced memoranda, it was postulated (1) that low rent units are more likely to have larger percentage increases than higher rent units and (2) that the impact of these larger percentage increases is, in some measure, responsible for

producing a GeoMeans (Rent/REQ) index that tracks higher than its Test Laspeyres counterpart. Both statements are true, as we shall demonstrate.

To test the first supposition, rent data from the Research Analysis Database (RAD), from 1988 through 1995, were analyzed. New construction units, rent-control units and special rent-reduction units were excluded. The idea was to analyze “regular” rental units in regard to their average percentage increases across rent levels. The medium and high rents were collapsed into one higher rent-level and then compared to low rents. Testing the hypothesis that the percentage increases for low rents were the same as the increases for higher rents produced the following results.

#### Percentage Rent Increases Year to Year

	PCT Increase (Low rents)	PCT Increase (Higher rents)	P-value
1988-89	6.85%	4.47%	0.0005
1989-90	5.78%	4.36%	0.0186
1990-91	5.96%	5.20%	0.1812
1991-92	4.99%	3.66%	0.0233
1992-93	4.81%	3.45%	0.0078
1993-94	5.55%	3.50%	0.0001
1994-95	4.38%	3.54%	0.0927

Low P-values reflect significant differences. An  $\alpha = .05$  significance level can be assumed appropriate here. Variances are assumed to be unequal, using regular two-sample t-tests. Sample size for low rents run just under 4000, with higher rents running around 14000. The six January-June data panels are used in the analysis.

Low rent percentage increases track consistently higher across the years studied. In two of the yearly time frames ('90-'91 and '94-'95) the differences are not significant (at an  $\alpha = .05$  level), but the differences in the other five yearly time frames are clearly significant. The question is whether these larger percentage rate increases for the low rents are affecting the relative performances of the two indexes. The answer is they are.

### 5. The Effect on GeoMeans vs. Test Laspeyres Indexes

To test the proposition that higher percentage rate increases for low rents result in the GeoMeans index coming in higher than Laspeyres, we turn to

simulations, an appropriate paradigm and one general proof.

The simulation program need not be very complicated. We are interested in a mix of higher and lower percentage rent increases for low rents in tandem with appropriate weights that may themselves range higher or lower. This mixture is the essential paradigm.

Using appropriate weights and appropriate rents for the higher rental units, we can simulate these indexes without too much complication. For one thing, we can view the varying weights of the low rents as simple linear combinations of the fixed basic weights of the higher rents. This allows the basic weights themselves to cancel out in both indexes.

We simulate an index consisting of one low rent unit and one higher rent unit from time  $t_0$  to time  $t_1$  (think of 12 months as an appropriate time interval). The higher rent unit and its appended weight will be treated as fixed, along with the low rent unit's previous rent price. The new low rent, along with its weight, will be allowed to vary considerably. A comparison of the two index calculations (GeoMeans and Test Laspeyres) should tell us something useful, one way or the other.

The two index formulas applicable here are:

TLI (Test Laspeyres Index) =  $I_p * R_L$  and  
 GMI (GeoMeans Index) =  $I_p * R_L$ ,  
 where (WLOG), we can set  $I_p = 1$ ,  
 and where  $R_i$  = a rent price relative.

$$\text{Then, } TLI = R_L = \frac{B(wF) + c_1w}{b_0(wF) + c_0w}$$

$B$  = low rent at time  $t_1$   
 $b_0$  = low rent at time  $t_0$   
 $w$  = base weight for higher rent unit  
 $F$  = a variable weight factor ( $>0$ )  
 $c_1$  = higher rent at time  $t_1$   
 $c_0$  = higher rent at time  $t_0$

$$\text{thus, } TLI = \frac{B \cdot F + c_1}{b_0 \cdot F + c_0},$$

since the weights cancel.

$$\begin{aligned} GMI = R_G &= \left(\frac{B}{b_0}\right)^{\frac{wF}{wF+w}} \left(\frac{c_1}{c_0}\right)^{\frac{w}{wF+w}} \\ &= \left(\frac{B}{b_0}\right)^{\frac{F}{F+1}} \left(\frac{c_1}{c_0}\right)^{\frac{1}{F+1}} \end{aligned}$$

With these two index formulas, the simulation program sets up easily enough (using Splus). The previous low rent ( $b_0$ ) is set at 300. The previous higher rent ( $c_0$ ) is set at 500. The current higher rent ( $c_1$ ) is set first at 550 and then at 600. The current low rent is allowed to vary from 200 to 1000 in increments of 20. The weight factors are  $\frac{1}{4}$ ,  $\frac{1}{2}$ , 1, 2 and 4.

Any number of settings can be viewed (with the same paradigmatic results each time). We can set the low rents as low as 200 and as high as 500, and set the higher rents as low as 300 and as high as 800, and still we will get the same pattern. The paradigm holds steady and clear in any and all plausible settings. A (smooth) quadratic curve defines the function. At one clear-cut mathematical point, both indexes are equal. At another less definable point, the two indexes are again equal. Rent price reality exists, in the main, between these two points — i.e., precisely where GeoMeans indexes always run higher than its Laspeyres counterpart. The difference between GeoMeans and Test Laspeyres (GMI-TLI) is plotted against the ratio of new to old low-rents ( $B / b_0$ ). The three fixed numbers in the header refer to  $c_1$ ,  $c_0$ , and  $b_0$ , respectively.

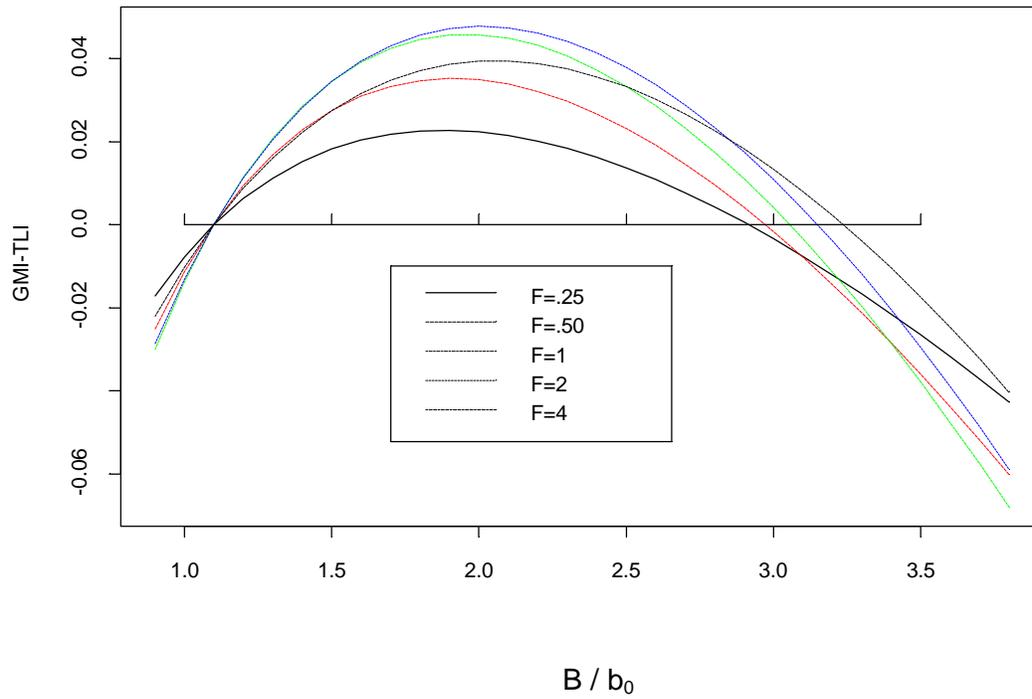
The simulations point to a clear change of direction as the two indexes pass their point of intersection, i.e., at the point where the two price relatives are equal. Thus, in the locality just past that intersection point, GeoMeans runs higher than Test Laspeyres. Only when  $B$  starts to range out of the low rent level does the inequality begin to reverse itself. Thus, for all practical purposes, when low rents have higher percentage increases than their higher rent counterparts, the GeoMeans index comes in higher than Test Laspeyres. Moreover, this higher GeoMeans index begins tracking higher as soon as that intersection point has been crossed, no matter what the respective weights are.

Furthermore, the weights just do not influence the direction of this paradigm. Lower low-rent weights make for lower differences between the two indexes,

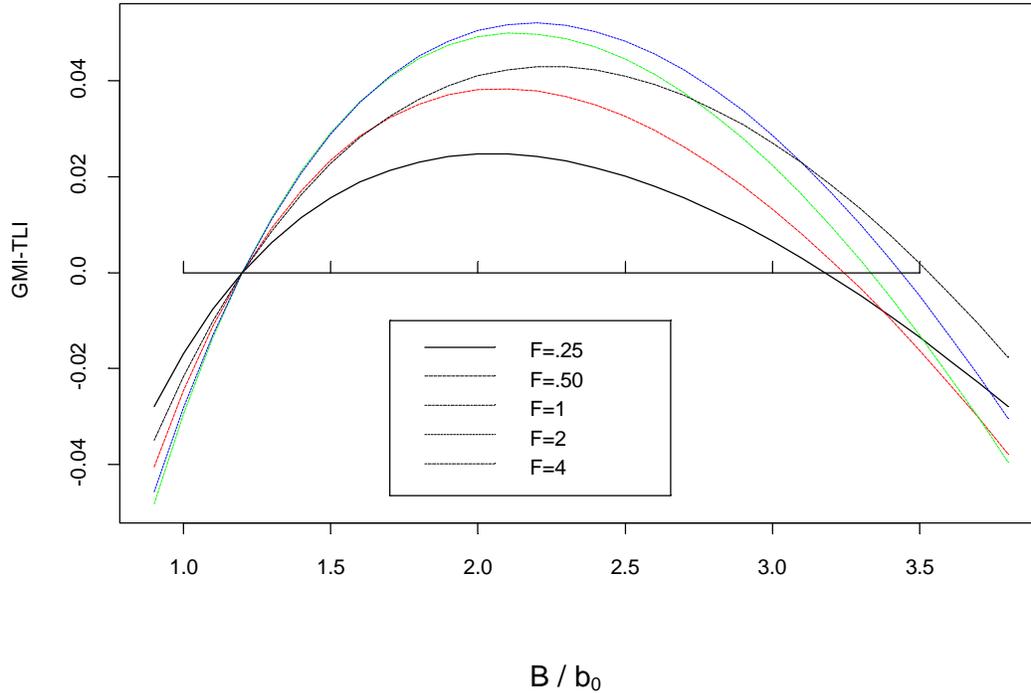
but the signed differences themselves are never changed (under the conditions of the paradigm as denoted above). So long as low rent units on average exhibit higher percentage rate increases than their

higher rent counterparts, the GeoMeans index will track higher than Test Laspeyres.

Geo vs Laspeyres with 550, 500, 300



### Geo vs Laspeyres with 600, 500, 300



The simulations bear out the following results:

- (1)  $GMI = TLI$  whenever  $\frac{B}{b_0} = \frac{c_1}{c_0}$ ,  
no matter what F is.

Let  $\frac{B}{b_0} = \frac{c_1}{c_0} = \frac{x}{y}$ , then,

$$GMI = \left(\frac{x}{y}\right)^{\frac{F}{F+1}} \left(\frac{x}{y}\right)^{\frac{1}{F+1}}$$

$$= \left(\frac{x}{y}\right)^{\frac{F}{F+1} + \frac{1}{F+1}} = \frac{x}{y}$$

and likewise,  $TLI = \frac{x \cdot F + x}{y \cdot F + y} = \frac{x}{y}$ .

- (2)  $GMI > TLI$  if (a)  $\frac{B}{b_0} > \frac{c_1}{c_0}$  and  
(b)  $0 < b_0 < B < c_0 < c_1$

for all reasonable

F.

If we let F=1 (i.e., equal weights), we can prove that  $GMI > TLI$ , provided conditions (a) and (b) are true.

Given (1)  $B/A > D/C$  and (2)  $0 < A < B < C < D$ ,

Prove:  $[(B/A)(D/C)]^{1/2} > (B+D)/(A+C)$ .

**Proof 1.1:**  $BC > AD$  [ from (1) ]  
 $CD > AB$  [ from (2) ]

Then  $BC - AD > 0$   
 $CD - AB > 0$   
 $(BC - AD)(CD - AB) > 0$   
 $BDC^2 - ACB^2 - ACD^2 + BDA^2 > 0$   
 $BDA^2 + BDC^2 > ACB^2 + ACD^2$   
 $BDA^2 + 2ABCD + BDC^2 >$   
 $ACB^2 + 2ABCD + ACD^2$   
 $BD \cdot (A^2 + 2AC + C^2) >$   
 $AC \cdot (B^2 + 2BD + D^2)$   
 $BD / AC > (B^2 + 2BD + D^2) /$   
 $(A^2 + 2AC + C^2)$   
 $BD / AC > (B + D)^2 / (A + C)^2$

$$[(BD) / (AC)]^{1/2} > (B+D) / (A+C)$$

□

In the extended multi-quote setting, the following condition should be sufficient (albeit as yet unproven) to insure the claim that (2) is still true.

$$\left[ \frac{\prod_i^m B_i}{\prod_j^m b_{0,j}} \right]^{1/m} > \left[ \frac{\prod_i^n c_{1,i}}{\prod_j^n c_{0,j}} \right]^{1/n},$$

$$\text{where } \begin{cases} m = n, & \text{if } (m+n) \text{ is even.} \\ m = n - 1, & \text{if } (m+n) \text{ is odd.} \\ m + n = \# \text{ of quotes.} \end{cases}$$

$$\begin{aligned} b_{0,1} < b_{0,2} < \dots < b_{0,m} < c_{0,1} < c_{0,2} < \dots < c_{0,n}, \\ \forall (B/b_0)_k, B_i = B_j \text{ and} \\ \forall (c_1/c_0)_k, c_{1,i} = c_{1,j}. \end{aligned}$$

The extended multi-quote version of (2), where  $GMI > TLI$  would then be

$$\prod_i^m \left( \frac{B_i}{b_{0,i}} \right)^{\frac{1}{m}} \prod_i^n \left( \frac{c_{1,i}}{c_{0,i}} \right)^{\frac{1}{n}} > \frac{\sum_i^m B_i + \sum_i^n c_{1,i}}{\sum_i^m b_{0,i} + \sum_i^n c_{0,i}}.$$

Since it has been demonstrated that the weights matter little in regard to the direction of this inequality, this version is quite close to production reality. I.e., when the percentage rate of increase is greater across the lower rents than it is across the higher rents, then this inequality holds and holds strong.

## 6. Conclusion

The somewhat simplistic two-partition model that we have used may not always reflect an actual Rent or REQ index, but it is not purely hypothetical either. If the average low rent percentage increase is higher than its higher rent counterpart, then the inequality as expressed in (2) will almost always hold. (An average low-rent percentage increase is directly analogous to the left-hand side of the multi-quote condition above.)

BLS statistician, Sylvia Leaver, working independently of this paper, calculated 6-month price changes in rents over this same 1988-1995 period and found that the GeoMeans index came in higher in three out the four major geographic regions of the CPI. Only in the South was the direction reversed, with GeoMeans there coming in lower than Test Laspeyres. When I then separated out my own rent numbers by region and fitted these averages into the above paradigm, I found that only in the South did the first “average” rent price relative come in lower than its higher rent counterpart, thus corroborating the paradigm.

## 7. Acknowledgements

The author would like to thank Dave Richardson and Louise Campbell for their extensive analyses on performance comparisons of the Laspeyres and GeoMeans indexes, and in particular for their raising the main issue of this paper. The author would also like to thank Janet Williams and especially Robert Baskin for their help and encouragement in developing these issues. The author also wishes to thank Dave Swanson, Shawn Jacobson, Sylvia Leaver, and Janet Williams for their careful reading of this paper and for their helpful comments.

## 8. References

Bureau of Labor Statistics, *BLS Handbook of Methods* (1992), Washington, DC. U.S. Government Printing Office, 180-181, 190-191, 196-198.

Campbell, L. and Richardson, D. (1996), “The Comparison of the Geometric Mean Index (GMI) with the Test Laspeyres Index (TLI) on the Basis of 1990-1991 Data”, *BLS Memorandum*, Bureau of Labor Statistics, Washington, DC.

Leaver, S.G. (1997), “Variance Estimates for Alternative Estimators of Price Change for Rent and Owners’ Equivalent Rent, for U.S. All Cities and Regions, 1990-1995”, *BLS Memorandum*, Bureau of Labor Statistics, Washington, DC.

Stewart, K., Smedley, K. (1997), “The Experimental CPI using Geometric Means (CPI-U-XG)”, *BLS Memorandum*, Bureau of Labor Statistics, Washington, DC.