

## Comparison between Chained CPI-U and Regular CPI-U All-US Indexes in the Housing Sector (2000-2004) December 2006

Owen J. Shoemaker  
U.S. Bureau of Labor Statistics, 2 Mass Ave., NE, Rm 3655  
Washington, DC 20212  
[shoemaker\\_o@bls.gov](mailto:shoemaker_o@bls.gov)

**Key Words:** Tornqvist; Superlative Index

*Any opinions expressed in this paper are those of the author and do not constitute policy of the Bureau of Labor Statistics.*

In February, 2006, the BLS calculated and published its fourth annual set of C-CPI-U indexes --- for the 12 months of 2004. The C-CPI-U (Chained Consumer Price Index – Urban) is calculated and published every year, with a one year lag, using a Tornqvist formula, which is a “Superlative” index formula, and which is supplied with *unique* and *timely* weights for times  $t$  and  $t-1$  as called for in the Tornqvist formula. By contrast, the regular CPI-U uses a Laspeyres formula as its final estimator, using expenditure weights that are, on average, three years old. Both indexes draw on the same set of lower level price relatives, which are calculated using an econometrically appropriate combination (Hybrid) of Geomeans and Laspeyres formulas. For 12-month price changes, the All-US–All-Items chained C-CPI-U index results continue to diverge from regular CPI-U index results, although half of the differences are no longer *significant*. Moreover, in the Housing sector and in the Micropolitan areas (C-Size Cities – All-Items), beginning with the 2003 results, the Superlative results have begun to track consistently *higher* than their Regular CPI results. In our new 2004 comparative data results, this unexpected result is also evident in three other Major Group sectors as well: Apparel, Recreation, and Transportation. We

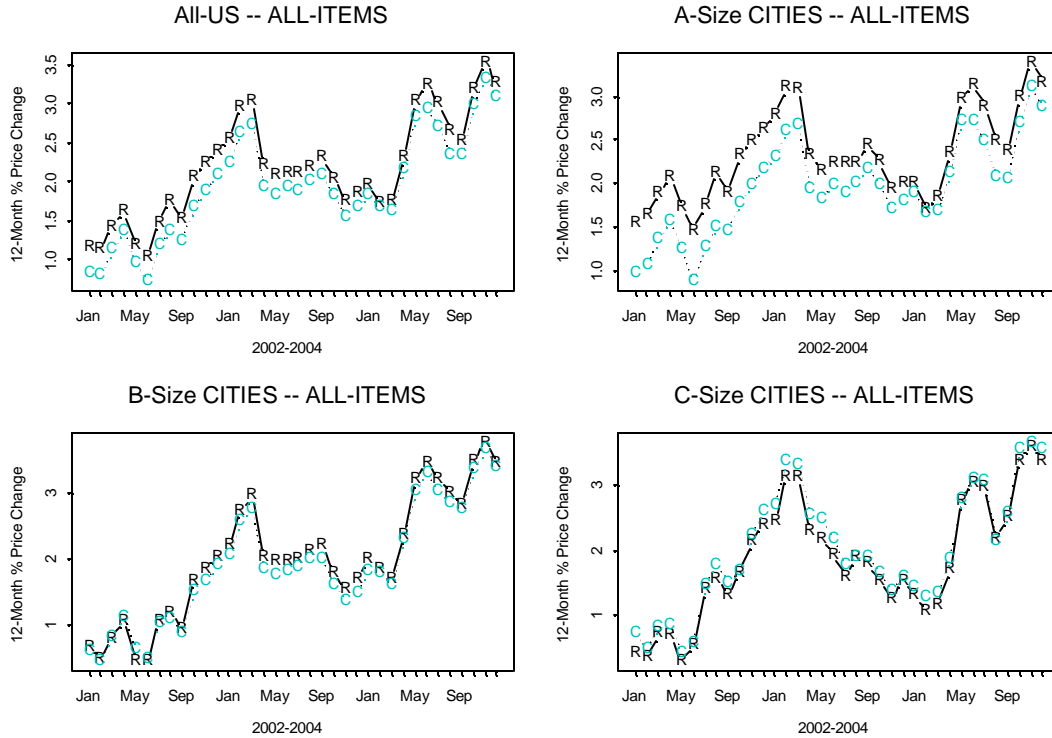
investigate the anomalous nature of these results, and suggest that a mathematical result rather than index theory may be determinative.

### 1. Chained CPI vs. Regular CPI

A Superlative formula, like the Tornqvist, is generally expected to produce a lower index than an index that uses a Laspeyres formula. According to classical price index theory, the Laspeyres formula, under homothetic assumptions, will provide an *upper* bound for a Konus (Cost of Living) Index — with the Paasche formula providing a lower bound. The Tornqvist formula, like the Fisher Ideal or a CES formula, provides a close approximation to a true cost-of-living index (i.e., closest to a Konus), and as such is expected to produce a consistently lower index than an index employing a Laspeyres formula. The Boskin Commission’s 1996 “Final Report on the Advisory Commission to Study the Consumer Price Index” estimated the (upper-level) substitution bias between a Superlative and a Laspeyres index at approximately 0.15 percent points per annum for an All-US–All-Items index. While the differentials between a 12-month superlative (C-CPI-U) and a Laspeyres (CPI-U) index remain at least as much as 0.15 percent points apart, the differences seem to be diminishing and, in selected lower level aggregations, even producing index results where superlative is *greater* than its Laspeyres counterpart. It is these anomalies that we want to investigate more closely.

## 2. All-US and City-Size Comparisons

**Fig 1** **12-Month % Price Changes**  
(R = Regular CPI C = Chained CPI)



**Table 1 Mean Yearly Per Cent Differences --- City Size Classes**

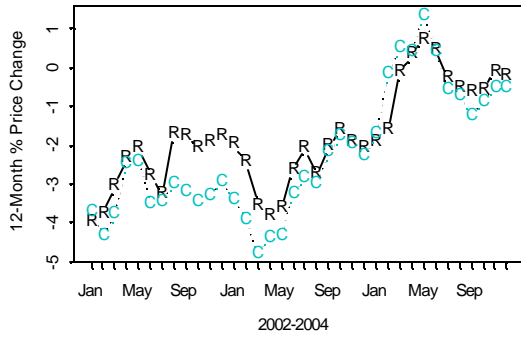
	<b>Yr2002</b>	<b>Yr2003</b>	<b>Yr2004</b>	<b>RELIMP</b>
<b>ALL US</b>	+0.312	+0.247	+0.192	100.0%
<b>A-Size CITIES</b>	+0.524	+0.326	+0.270	57.5%
<b>B-Size CITIES</b>	+0.036	+0.169	+0.120	36.5%
<b>C-Size CITIES</b>	-0.142	-0.177	-0.113	6.0%

We will be looking at the latest three years' worth of 12-month price change data for all of our work here: 2002, 2003, and 2004. **Fig 1** displays the 36 months of 12-month price changes for the All-US–All-Items indexes along with the three City-Size–All-Items indexes, with “R” designating the Regular CPI results and “C” designating all Chained (Superlative) CPI results. **Table 1** presents their average yearly per cent differences. Not surprisingly, all three of the City-Size–All-Items indexes track very closely with their respective All-US–All-Items indexes, even

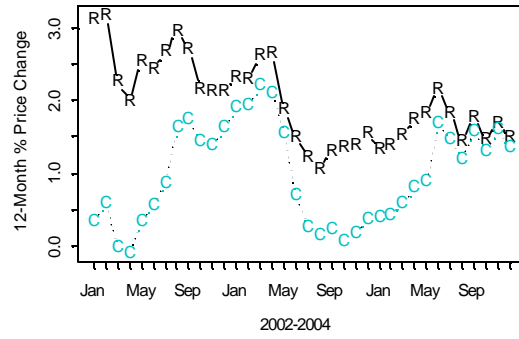
C-Size. However, in the GSize Cities–All-Items indexes, the differences between the two types of indexes – Regular versus Superlative – have not just diminished, but have actually reversed direction. For all three consecutive years, indeed for every one of the 36 months shown, the Chained CPI is higher than its Regular CPI counterpart. At the end of the presentation of all our relevant data and graphs, we will attempt to account for this unanticipated reversal, but, for the moment, we will only note this anomaly and move on to the rest of our results.

### 3. All-US by Major-Group Comparisons

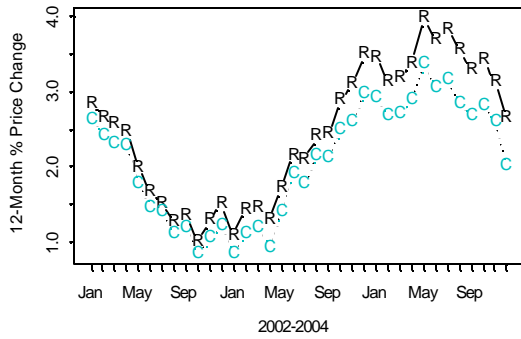
APPAREL



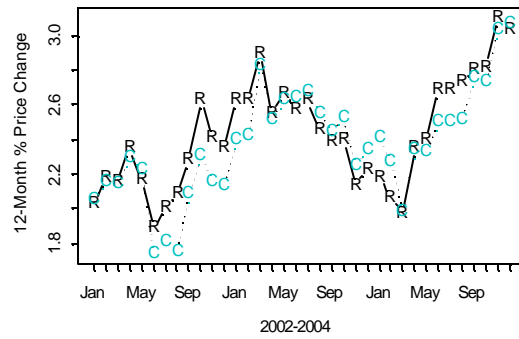
EDUC & COMM



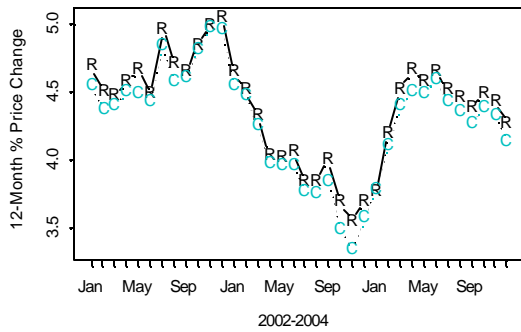
FOOD & BEVEGS



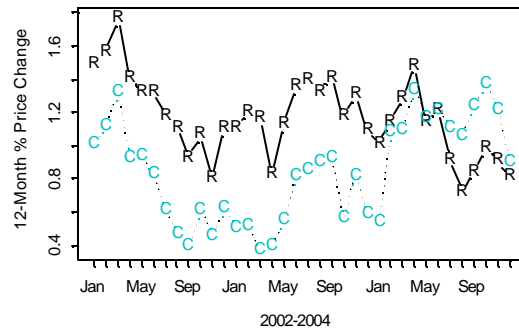
HOUSING



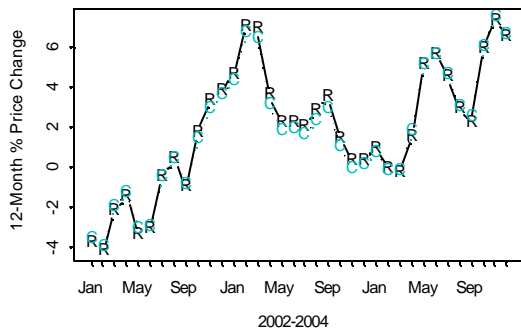
MEDICAL



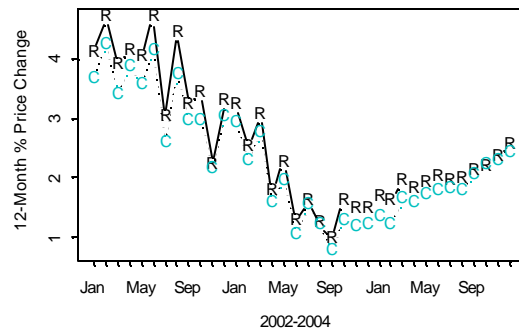
RECREATION



TRANSPORTATION



OTHER



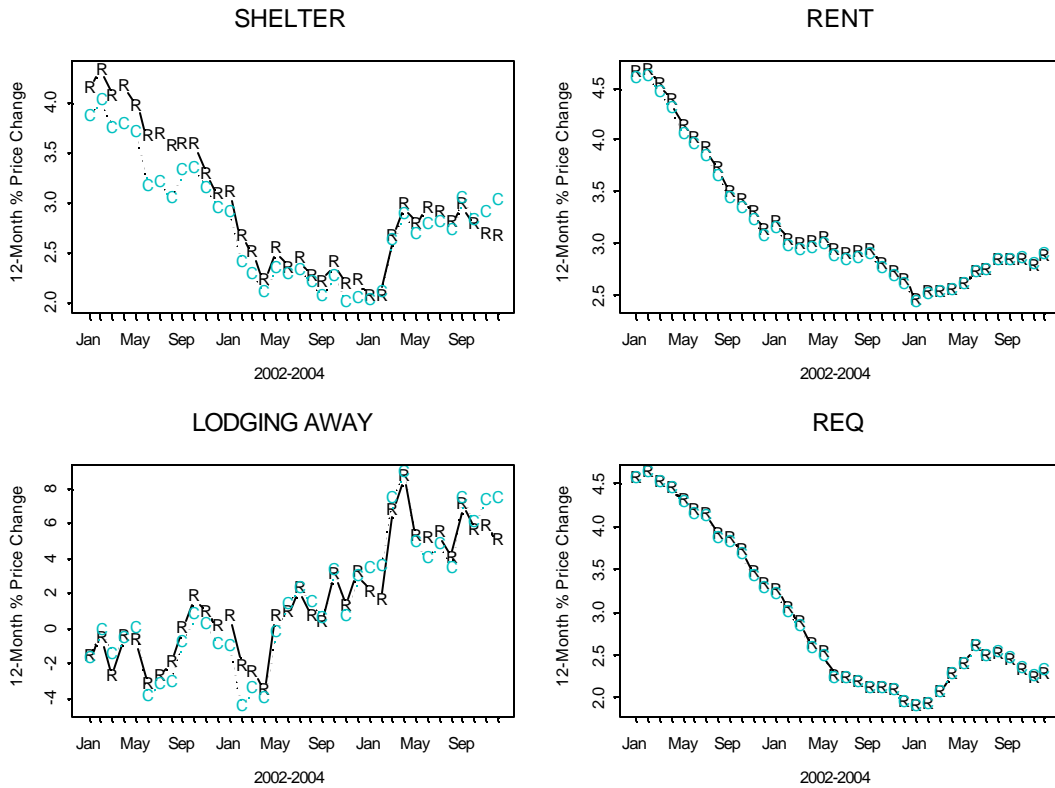
**Table 2. Mean Yearly Per Cent Differences --- By Major Group (All-US)**

	Yr2002	Yr2003	Yr2004	RELIMP
APPAREL	+0.762	+0.642	-0.063	4%
EDUC & COMM	+1.654	+0.797	+0.536	6%
FOOD & BEVG	+0.208	+0.337	+0.585	16%
HOUSING	+0.142	-0.004	+0.034	41%
MEDICAL	+0.092	+0.109	+0.094	6%
RECREATION	+0.479	+0.559	-0.072	6%
TRANSPORTN	-0.042	+0.418	-0.068	17%
OTHER	+0.406	+0.226	+0.190	4%

When we break down the differences by Major Group, we again encounter a certain significant set of index comparisons where Superlative 12-month price changes are tracking higher than its Regular CPI counterparts. In 2002 the anomaly showed up only in Transportation, but in 2003 it showed up in the much more influential Major Group of Housing (representing 41% of the entire CPI). The importance of the anomaly occurring in Housing in 2003

largely motivated this study, but when the 2004 Superlative date came in (in early 2006) the picture again muddled. This time the sign reversal occurred in Apparel, Recreation and once again in Transportation. Housing regained a mean positive differential for 2004, but for four of those months the anomaly was present. As such, a closer look at the Housing sector still seemed in order.

**4. Housing Sector Comparisons**



**Table 3. Mean Yearly Per Cent Differences --- Housing Sector (All-US)**

	Yr 2002	Yr 2003	Yr 2004	RELIMP
<b>SHELTER</b>	+0.320	+0.155	-0.009	32%
<b>RENT</b>	+0.072	+0.050	+0.001	6%
<b>REQ</b>	+0.025	+0.024	-0.014	23%
<b>LODGING</b>	+0.271	+0.440	-0.511	3%

The main component of the Housing Major Group is Shelter (32% out of its 41% total), with its Rent and REQ (Owners' Equivalent Rent) sub-components accounting for a full 29% out of that 32%. The initial intuition motivating this study — which is exploring the instances where Chained CPI results track unexpectedly *higher* than Regular CPI — was that Rent and REQ within the large Housing sector would be the main sources of this anomalous behavior. However, such is not exactly the case. As the results in **Table 3** show — when viewed against the Housing results from **Table 2** — Shelter and its main constituent parts show negative differentials in Yr2004 when the full Housing sector itself is not negative, and then positive differentials in Yr2003 when the Shelter differential *is* negative. Clearly, some of the anomalous price change behaviors are occurring within the two other smaller components sectors of Housing: Fuels & Utilities (5%) and Household Furnishings (4%). Also the largest negative differential in these tables is evidenced in Lodging Away From Home in Yr2003. This mixed bag of results shows no clear-cut pattern of *misbehavior* within the various sub-indexes.

### 5. Four Possible Explanations

First of all, unless a Laspeyres price relative calculation is expressed precisely as a weighted arithmetic mean, no exact mathematical comparison can be made between a Laspeyres and a Geomeans formulation (which mathematically is what a Tornqvist is). Moreover, only if the respective weights of the two types of index calculation are equal will the Geomeans index be guaranteed to come in lower than a Laspeyres index. Such a structure is not the case here. The Laspeyres is not exactly an arithmetic mean and the two set of weights for the two indexes are certainly not identical. So,

while the presumption, both econometrically and mathematically, remains strongly in favor of a consistently *lower* Chained CPI versus Regular CPI at all levels, there is no assurance that the pattern will always hold.

The first possible explanation for the anomalies observed in the data above is, of course, the weights. The two sets of weights do come from the same Consumer Expenditure (CE) Survey. The Regular CPI weights, however are on average three years old when they are used, while the Chained (Superlative) CPI weights are expenditure estimates specific to the exact times (months  $t$  and  $t-1$ ) of its index calculations. Moreover, the individual weights in the two sets of weights are developed in slightly different manners and generally at different aggregate levels. A lengthy study might be possible which could compare the expenditure *shares* that two comparable index calculations are employing, and possibly some sort of pattern might emerge from that study, but that will not be our focus here. The variability and comparative shares of the weights in the respective indexes may be important to study, but the constituent price relatives themselves drive the final price change results so much more than do the weight differentials that we find this line of data analysis largely unproductive. However, the clearly anomalous behavior shown in the C-Size Cities-All-Items graph in **Fig 1**, and punctuated in the **Table 1** results, just might be attributable to unstable or erratic or less than robust weights in these areas.

A second possible explanation might come from a closer investigation of substitution behaviors in these instances where Chained CPI is observed to be higher than Regular CPI. Perhaps the actual elasticity of substitution behavior is driving the unexpected differentials between the two indexes. But we will have to leave that for

economists to determine, one way or the other.

A third area of explanation might lie in examining inflation levels, particularly where and when the constituent price relatives for their larger price change calculations are in “negative inflation” territory. Comparative index behaviors may indeed vary greatly at significantly different levels of inflation. With relatively high inflation levels, where literally all prices are constantly and consistently going up every month, the comparative behavior may follow one pattern, whereas at low inflation levels or at deflationary levels, the comparative behaviors may be quite different.

A fourth possible explanation is mostly mathematical and draws on a small result from an earlier paper (“Performance Comparisons of Laspeyres Indexes with Geometric Mean Indexes in the U.S. Consumer Price Index” – ASA 1998) that investigated this very same anomaly when it occurred within Rent and REQ at the *lower* level of the CPI’s price relative calculations. The formulas and data we are observing are *upper* level operations. At the lower level price relative calculation, the formulas work directly with prices, and with individual quote weights for those prices. At the *higher* level our respective formulas work with *indexes* and with higher-level weights for those indexes. At the lower level, for Rent and REQ anyway, price *levels* turned out to be determinative. At the higher level of aggregation and index calculation, however, prices are replaced by *index* levels. At the lower level when *lower* Rent (and thus REQ) prices *moved* at a consistently higher rate of inflation than *higher* Rent (and REQ) prices, a proven mathematical result prevailed and Geomeans Indexes for Rent (and REQ) did indeed end up higher than their Laspeyres-

calculated counterparts. It is this mathematical paradigm that we would like to offer to the current data as a possible explanation for the anomalous behavior we have been observing.

## 6. A Mathematical Explanation

In my 1998 ASA paper, “Performance Comparisons of Laspeyres Indexes with Geometric Mean Indexes in the U.S. Consumer Price Index”, it was shown that a Geomeans index would necessarily be higher than a Laspeyres index if the lower priced items’ rate of increase was higher than the higher priced items’ rate of increase. To be precise, assuming equal weights, and a simple two-partition model:

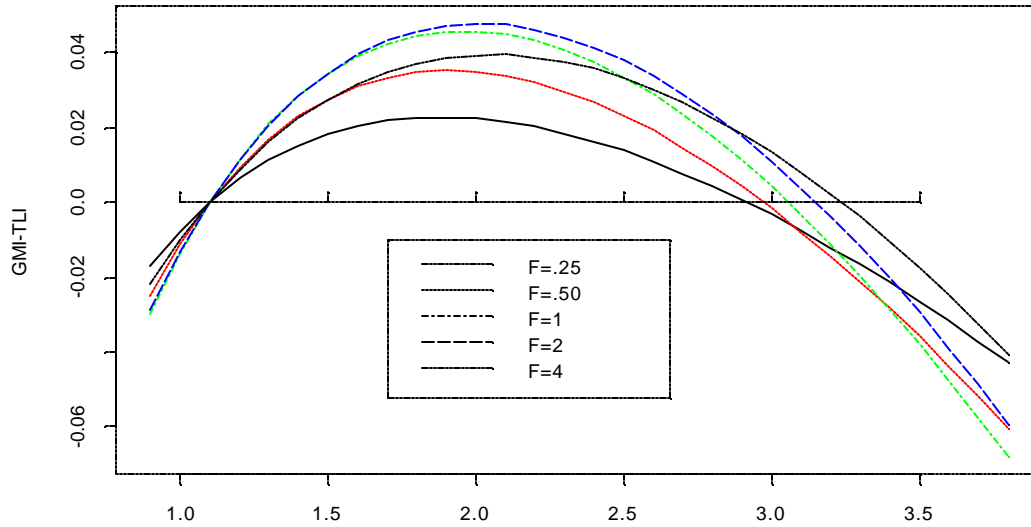
(1) Geomeans > Laspeyres if (a)  $B/A > D/C$  and (b)  $0 < A < B < C < D$

with Geomeans =  $[(BD)/(AC)]^{1/2}$  and Laspeyres =  $(B+D)/(A+C)$

In the area of Rent prices, where this pattern was shown to be the case most of the time, the lower-level Rent (and REQ) price relatives calculated using a Geomeans formula *were* consistently higher (75% of the time) than the Rent price relatives using a Laspeyres formula. For the rent data we analyzed, the lower priced rents’ rate of increase averaged a full 1.5 percentage points higher than the higher rents’ rates of increase. Thus, in the Rent (and REQ) sector the stage was set for producing the unexpected result of having Geomeans higher than Laspeyres.

Even if unequal weight are applied to this two-partition model, the range of annual price relatives where the Geomeans > Laspeyres result obtains is fairly wide and includes most observed price relatives.

### Geo vs Laspeyres with 550, 500, 300



Note: GMI = Geomeans price relative, and TLI = Test Laspeyres price relative. F is a variable weight factor. The two higher rents in the simulations are the 550 and 500; the lower rent at time  $t_0$  is the 300, and we let the time  $t_1$  lower rent vary from 200 to 1200. The x-axis is the unweighted ratio of the variable lower rent at time  $t_1$  to the lower rent (300) at time  $t_0$ . The results clearly show the faster moving lower rents producing a higher Geomeans price relative.

Moving this simple two-item paradigm to the more complicated, and more realistic, multi-quoted (or multi-price) level we can no longer assure a proven result, but, in general, if the price increases of all the lower half of prices (or rents) are greater than the price increases of all the higher half of prices, then the proven result will hold --- assuming equal weights across the board.

But when we move this paradigm to the higher-aggregate level, does this mathematical result continue to apply? The answer would have to be: mostly not. However, this result *may* be the reason we are observing higher Tornqvist price changes than Laspeyres price changes in certain sectors of the CPI at various times. It cannot easily be shown to be *the* reason for the observed anomalies, particularly

since the index comparisons we have been looking at involve the higher levels of index aggregates. (Analyzed from the Superlative perspective, for instance, we are looking at 8,018 separate index ratios, each with a unique expenditure share, that go into the Tornqvist formula for the All-US-All-Items price relative calculation.) But the critical difference between lower-level usage of this sharp mathematical result and higher-level usage is that for our higher-level calculations the ingredients in the respective index formulas are *indexes* and not raw prices.

The two upper-level formulas use indexes (and not prices), and effectively the *same* indexes, in their calculations.

$$RPREL_{I,A,t}^R = \frac{\sum_{i \in I, a \in A} IX_{i,a,t}^R * W_{i,a,t_0}^R}{\sum_{i \in I, a \in A} IX_{i,a,t-1}^R * W_{i,a,t_0}^R}$$

$$PREL_{I,A,t}^C = \prod_{i \in I, a \in A} \left( \frac{IX_{i,a,t}^C}{IX_{i,a,t-1}^C} \right)^{\left( \frac{W_{i,a,t}^C + W_{i,a,t-1}^C}{2} \right)}$$

where R = Reg CPI C = Chained CPI

At the lower-level price relative calculations we have actual *prices* (or rents) used in the index formulas. Oddly enough, the indexes we use do, over time, range about as widely as do rents themselves (from 90 to 600 or more, in general). However, these are unit-less index numbers and have no real-world referents. Lower rents moving faster than higher rents makes perfect real-world sense. Lower index numbers moving faster (or slower) than higher index numbers makes no real-world sense at all. But, with a similar range of numbers in the respective calculations, these higher-level formula calculations *may* be similarly impacted by the cited mathematical result -- provided the same assumptions would hold.

In addition, there is a corollary to the above mathematical result, which may also apply in certain deflationary instances. When, within a given area-item combination, all or almost all of its constituent prices (or indexes) are falling, then this corollary of (1) may apply:

(2) Geomeans > Laspeyres if (a)  $B/A > D/C$  and (b)  $0 < B < A < D < C$

Thus, if the rate of *decrease* is *greater* for the higher prices (or indexes), then a Geomeans formula, like the Tornqvist, will produce a higher percent price change than a Laspeyres formula. In **Fig 3** and **Table 3**, for Lodging Away From Home in 2004, we can observe, in the twelve months of 2004, that that particular sub-index was clearly deflating somewhat uniformly across the year's time. The result in (2) could be a valid explanation for the observed negative differentials between its Chained CPI results and its Regular CPI results for that year.

The case for using these two mathematical results to explain the anomaly of a given set of Superlative price changes tracking higher than their Laspeyres counterparts is more based on a process of elimination than by direct inference. The mathematical results are informative and *may* be applicable as valid explanations for our observed anomalous behaviors, but the true explanation may lie elsewhere.

## 7. Conclusions

- Unexpected negative differentials between Regular CPI and Chained CPI results at some aggregate levels below the All-US-All-Items level continue to be observed.
- The non-uniformity of the observed anomaly in the new 2004 data makes conclusions about why certain Superlative indexes are producing higher 12-month price changes highly problematical.
- Several explanations for the anomaly are offered for consideration.
  - Variability of weights, which might explain the situation in the C-Size Cities sector.
  - The actual elasticity of substitution behavior may be driving the unexpected differentials between the two indexes.
  - Lower or even negative inflation levels may be influencing the observed anomalies.
  - Given certain price (or index) levels and certain rates of increases, there is a mathematical result which finds an index using a Geomeans formula (like the Superlative's Tornqvist) perform *higher* than an index using a Laspeyres formulation, no matter what the weights themselves might be.
- No *one* explanation has been found to account for all the occurrences of the observed anomaly.