Survey Methodology, December 1986
Vol. 12, No. 2, pp. 107–120
Statistics Canada

Estimating a Monthly Index Based on Trimestrial Data

JOHN G. KOVAR[1]

ABSTRACT

A problem of estimating monthly movements in rents based on data collected every four months is explored. Five alternative composite estimators of the rent index are presented and justified, both from an intuitive as well as theoretical point of view. An empirical study testing and comparing the proposed methods is described and summarized. Recommendations are put forth.

KEY WORDS: Index numbers; Rotating samples; Composite estimation.

1. INTRODUCTION

The rent component of the Consumer Price Index is based on data collected on a six month rotating basis using a Labour Force Survey Supplement. Since changes in rents generally occur on an annual basis, the effective sample size of the Labour Force Survey design is reduced. Furthermore, special annual benchmarks, which are obtained by revisiting the June sample of dwellings one year later, indicate that the rent component can suffer from varying degrees of bias (Dolson 1982). To ameliorate the situation, several data collecting schemes were proposed in order to combine the monthly data with the yearly benchmarks in a continuous and timely fashion. One of these methods, which collects data every four months, was selected for practical application.

The proposed design consists of four sets of four rotation groups of rented dwellings, each set of which is to be surveyed in one of four consecutive months, on a rotating basis. Each month, one rotation group is surveyed for the first time and the other three are those that rotated in four, eight and twelve months ago respectively. Each group would thus be surveyed four times over a period of thirteen months,before rotating out of the sample. Every month, data on current rents, as well as matched rents collected four months ago, are available from exactly three rotation groups (the fourth group is new and thus has no matching "backrents"). Yearly benchmarks can be calculated monthly based on one rotation group. This paper discusses several methods of estimating a monthly index based on such trimestrial data.

In estimating the indices, the constraints of the Consumer Price Index publication policy must be kept in mind. In other words, it must be practically as well as technically possible to produce the indices on a monthly basis for each of the index cities. The estimates must be timely: produced no later than mid-month following the reference month. Furthermore, no revisions can be made once the indices are published. While not entirely essential, it would be desirable that any proposed estimator be able to reflect (real) sudden changes in trend very quickly. On the other hand, in order to remain credible,the indices must be relatively stable: volatile, saw-toothed indices are to be avoided.

[1] John G. Kovar, Business Survey Methods Division, Statistics Canada, 11th floor, R.H. Coats Building, Tunney's Pasture, Ottawa, Ontario K1A 0T6.

In Section 2, five estimators will be presented, justified, and compared on a theoretical basis. Some empirical adjustments to these indices will be discussed in Section 3. In order to compare the performance of these estimators over time and between locations, a simulation study involving eight cities with observations over a period of 48 months was performed. The results of the study are presented in Section 4. The conclusions and recommendations can be found in Section 5.

2. INDEX ESTIMATORS

In this paper, only matched indices will be considered. While relative changes could easily be derived by comparing independent (unmatched) estimates of rent levels at distinct time points, such estimates of levels would have to be very reliable, necessitating prohibitively large sample sizes. Moreover, past studies indicate that such direct estimators tend to be volatile, upwardly biased and generally not practical in use (Szulc 1983). In what follows, therefore, an estimate of relative change between two time points will be based only on those units that report rents for both of these time points.

We will denote by x_m the total rent paid, in the current month m, by a certain subset s of dwellings in a given city. Thus, more rigorously,

$$x_m = \sum_{i \in s} x_{mi}, \tag{2.1}$$

where x_{mi} denotes the rent paid by the i-th dwelling in month m. The rent index is customarily estimated by chaining one month relatives, that is, the ratios of average rents between two consecutive months denoted by r^m_{m-1}. In other words, the index in month m, I_m, over a base period zero, is estimated recursively by

$$\hat{I}_m = \hat{I}_{m-1} \times \hat{r}^m_{m-1} = 100 \times \hat{r}^1_0 \times \hat{r}^2_1 \times \ldots \times \hat{r}^{m-1}_{m-2} \times \hat{r}^m_{m-1}. \tag{2.2}$$

where 100 is the (abritrary) level of the index at time zero. The difficulty then rests only in estimating the relatives.

In general, consider the relative change in rent in month m over month 1, denoted by r^m_1. This "m over 1 relative" can be estimated by

$$\hat{r}^m_1 = x_m / x_1. \tag{2.3}$$

However, if one considers matched indices only, the only estimable relatives under the proposed design are the four-month relatives, in other words, those of the form r^m_{m-4j}, $j = 1, 2, 3$, because it is only in these cases that there are common units between the two months. These relatives are estimated by

$$\hat{r}^m_{m-4j} = x_m / x_{m-4j}, \tag{2.4}$$

where the set s of dwellings consists of only those units that report rents at both time m and $m-4j$. Unfortunately, the interest lies in estimating monthly relatives of the form r^m_{m-1}. On the positive side, the rotation scheme ensures that a four-month relative is available every month. It is also assumed that units rotating out of the sample are replaced by equivalent units rotating into the sample. As such, the set s of common dwellings in (2.1) depends on the time m only and any future reference to it, while implicitly retained, can thus be suppressed in what follows. For a rigorous discussion of these assumptions and the effect on the index if the assumptions fail, the reader is invited to consult Szulc (1983) and Kovar (1984).

In the following paragraphs, five methods of estimating monthly relatives from four-month relatives will be described. Each will be justified intuitively as well as theoretically, and its advantages and disadvantages will be pointed out. The first three methods are derived on a theoretical basis alone while the fourth attempts to exploit the rotation pattern of the survey. All four assume that at least a four month back history of data is available. The last approach takes advantage of prior empirical knowledge: that of high probability of observing one change in rent per year. Methods two and four have been discussed earlier by Kovar (1984).

2.1 Interpolated Index (Additive Index)

One way of estimating the relative r^m_{m-1} is to estimate the previous month's rent, x_{m-1}. This can be accomplished, among other methods, by linearly interpolating the observed rents at time m and $m - 4$, that is, by assuming that the rents increase (decrease) linearly over time. Note that this assumption does not require each individual rent to increase every month by a fixed amount, but merely that the sum of all the rents does. In general, to describe linear interpolation briefly, consider two measurements of the same quantity at two distinct time points, say y_t and y_{t-s}. Suppose that we wish to estimate the value of y at some point between the times $t - s$ and t, say at time $t - u$ $(u < s)$. Assuming that the measurements increase linearly in time, y_{t-u} can be estimates from y_t and y_{t-s} by

$$y_{t-u} = (1 - \frac{u}{s}) y_t + \frac{u}{s} y_{t-s} \qquad (2.5)$$

or in the case at hand, where $s = 4$ and $u = 1$, by

$$y_{t-1} = (3/4) y_t + (1/4) y_{t-4}. \qquad (2.6)$$

Thus the previous month's total rent can be estimated by

$$x_{m-1} = (1/4) x_{m-4} + (3/4) x_m \qquad (2.7)$$

and consequently, the monthly relative for month m by

$$r^m_{m-1} = \frac{x_m}{x_{m-1}} = \frac{4x_m}{x_{m-4} + 3x_m}. \qquad (2.8)$$

The index is then derived by chaining the relatives as in (2.2) above.

Provided that the rents follow the linear interpolation model, that is, provided that we can write the current month's rent as a recursive function of previous months' rents, namely, as

$$x_m = x_{m-1} + d = x_0 + md, \qquad (2.9)$$

then it can be shown that the index at time m is given by $I_m = x_m / x_0$, as is desired. In other words, if the data follow the model in (2.9), the index will suffer no time lags. But, of course, if the model were true at all times, the index would be fixed for all time points, based on any two observations. Since this is clearly not the case, one can at best use (2.8) as an approximation over short periods of time only. In that case, however, if the relationship in (2.9) is not exact, the index at time m will depend on all the rents between time -4 and m. In other words, the index is then susceptible to accumulating various biases over time.

Note that the same index would be derived by assuming that the four-month increment, $x_m - x_{m-4}$, occurred in 4 equal additive steps: $(x_m - x_{m-4})/4$. Since then, the previous month's rent would be estimated by

$$x_{m-1} = x_m - (x_m - x_{m-4})/4, \tag{2.10}$$

which is the same as (2.7); hence the alias: additive index.

2.2 Geometric Index

In this section, in contrast to the above, we will attempt to estimate the relative directly. We first note that

$$\begin{aligned} r^m_{m-4} = \frac{x_m}{x_{m-4}} &= \frac{x_m}{x_{m-1}} \frac{x_{m-1}}{x_{m-2}} \frac{x_{m-2}}{x_{m-3}} \frac{x_{m-3}}{x_{m-4}} \\ &= r^m_{m-1}\, r^{m-1}_{m-2}\, r^{m-2}_{m-3}\, r^{m-3}_{m-4}. \end{aligned} \tag{2.11}$$

We then assume that the four relatives on the right hand side of (2.11) are equal, or equivalently, that the four-month movement is due to four equal movements which act multiplicatively (Kosary *et al.* 1982). Under this assumption, the relationship (2.11) can be written as

$$r^m_{m-1} = (r^m_{m-4})^{1/4}. \tag{2.12}$$

From (2.2) and (2.3), assuming that there are no sample changes or that units rotating out of the sample are replaced by equivalent units rotating into the sample, the index in month m over the base period zero becomes

$$\begin{aligned} I_m &= I_0 \times r^1_0 \times r^2_1 \times \ldots \times r^m_{m-1} \\ &= I_0 \times (r^1_{-3})^{1/4} \times (r^2_{-2})^{1/4} \times \ldots \times (r^m_{m-4})^{1/4} \\ &= I_0 \frac{(x_{m-3}\, x_{m-2}\, x_{m-1}\, x_m)^{1/4}}{(x_{-3}\, x_{-2}\, x_{-1}\, x_0)^{1/4}} \end{aligned} \tag{2.13}$$

In other words, the index is a ratio of two geometric averages; hence the name geometric index. We note that at any time, assuming the panels are stationary, the index depends on eight months worth of data only, and thus is independent of any movements between time 0 and $m-4$, though in practice matched sets contributing to each r^m_{m-4} are different, so the cancellation is only theoretical. By contrast the index suffers from one-month to three-month lags and will thus tend to dampen true sudden changes. These changes, however, will be reflected eventually, that is, the index will selfcorrect (Kovar 1984).

As a point of clarification, note also that the relatives in (2.12) can be rewritten as

$$\frac{x_m}{x_{m-1}} = \left[\frac{x_m}{x_{m-4}}\right]^{1/4}$$

or as

$$x_{m-1} = (x_{m-4})^{1/4}\,(x_m)^{3/4}$$

or finally as

$$\log(x_{m-1}) = (1/4)\log(x_{m-4}) + (3/4)\log(x_m). \tag{2.14}$$

The geometric index is therefore equivalent to an index derived by estimating the previous month's rent by linearly interpolating the logarithms of the observed rents at time m and $m - 4$. (See (2.6) with $y_m = \log x_m$.)

2.3 Incremental Index

Analogous to the above geometric index, here we assume that the four consecutive monthly relative net increments are equal and acting additively. More precisely, we can write r_1^m as

$$r_1^m = 1 + i_1^m$$

where i_1^m is the relative net increment in month m over month 1. To estimate r_{m-1}^m we need therefore i_{m-1}^m. Assuming that the available $i_{m-4}^m = 4\,i_{m-1}^m$, the relative r_{m-1}^m can be estimated. Namely, we will estimate i_{m-1}^m by

$$i_{m-1}^m = (1/4)\,i_{m-4}^m = (1/4)\,(r_{m-4}^m - 1) = (1/4)\left(\frac{x_m}{x_{m-4}} - 1\right), \tag{2.15}$$

and r_{m-1}^m by

$$r_{m-1}^m = 1 + i_{m-1}^m = \frac{x_m + 3x_{m-4}}{4x_{m-4}}. \tag{2.16}$$

We note that $r_{m-1}^m = x_m/x_{m-1}$ and thus (2.16) can be written as

$$\frac{x_m}{x_{m-1}} = \frac{x_m + 3x_{m-4}}{4x_{m-4}}$$

or as

$$\frac{1}{x_{m-1}} = (1/4)\frac{1}{x_{m-4}} + (3/4)\frac{1}{x_m}. \tag{2.17}$$

In other words, the incremental index corresponds to one which would be derived by estimating the previous month's rent by linearly interpolating the reciprocals of the observed rents at time m and $m - 4$. (See (2.6) with $y_m = x_m^{-1}$.)

As is the case with the interpolated index, the incremental index will be independent of the intermediate observations only under the restrictive condition that the interpolation model be followed. In this case, analogous to (2.9), the model is

$$\frac{1}{x_m} = \frac{1}{x_0} + md. \tag{2.18}$$

However, in most real situations, the chained incremental index will depend on all the data between times -4 and m and therefore will be susceptible to various accumulating biases.

Since all three indices discussed to this point can be described in terms of linear interpolation of various functions of the observed rents, it is also possible to compare them theoretically. It can in fact be shown that the three indices are ordered in magnitude, from smallest to largest in the order of their presentation. That is, in an inflationary situation the interpolated index will always be smaller in absolute value than the geometric index which in turn will always be dominated by the incremental index. The reverse holds true when the trend is downward, that is, when prices are decreasing. As one referee pointed out, this phenomenon can be explained by noting that "the interpolated, geometric and incremental relatives are respectively the weighted arithmetic, geometric, and harmonic means of rent quotations four months apart. The standard relationship between these means explains the behaviour of the estimates in inflationary or deflationary times".

2.4 Carried Index (Arithmetic Index)

The carried index is constructed by taking advantage of the rotating sample at hand. Noting that all units reappear periodically in the sample, we construct the index by simply carrying each unit's rent value forward until a new observation is recorded. In this way all units on the file have a matching previous month's rent and thus the monthly relative, r^m_{m-1}, can be constructed in a straightforward manner. The obvious drawback is that the rent increases (decreases) are not recorded until observed. However, since all changes are eventually recorded, the index will selfcorrect (Kovar 1984) but will suffer from a mixture of one to three-month lags. Just as for the geometric index, sudden (real) changes will be dampened but the carried index will reflect them eventually.

On the technical side, we note that in computing the carried index for any given month one quarter of the observations on the file reflect a four-month movement, whereas three quarters of the observations are carried for one to three months and reflect no change. In fact, in month m we observe x_m and carry x_{m-1}, x_{m-2} and x_{m-3}. Similarly, in month $m-1$ we observe x_{m-1} and carry x_{m-2}, x_{m-3} and x_{m-4}. The monthly relative is therefore given by

$$r^m_{m-1} = \frac{x_m + x_{m-1} + x_{m-2} + x_{m-3}}{x_{m-1} + x_{m-2} + x_{m-3} + x_{m-4}}. \tag{2.19}$$

Chaining the relatives as in (2.2), and assuming again that the samples are stationary, we obtain the index for month m over the base period zero as

$$I_m = I_0 \frac{x_{m-3} + x_{m-2} + x_{m-1} + x_m}{x_{-3} + x_{-2} + x_{-1} + x_0}. \tag{2.20}$$

In other words, the index is a ratio of two arithmetic averages. Analogous to the geometric index, the carried index depends on eight months worth of data only, and thus is independent of the movements between time 0 and $m - 4$. As mentioned above, it too suffers from one to three-month lags, and therefore dampens sudden changes.

2.5 Annual Index

Empirical observations suggest that most units change rent once a year. One could therefore argue that yearly relatives are more stable than monthly relatives, since the distribution of individual monthly relatives will necessarily demonstrate two spikes, one around the annual relative and the other at 1. The rotation pattern of the proposed rent pilot (Kovar 1984) ensures that an annual relative be estimable every month, that is that r^m_{m-12} be available. To compute the annual index on a monthly basis, we note that for any chained index the following relationships hold:

$$I_m = r^m_{m-1} I_{m-1} \tag{2.21}$$

and

$$I_m / I_{m-12} = r^m_{m-12}. \tag{2.22}$$

From these relationships we obtain an expression for a monthly relative r^m_{m-1} as

$$r^m_{m-1} = r^m_{m-12} I_{m-12} / I_{m-1}. \tag{2.23}$$

These relatives can then be chained as above to produce an index. Since such a relationship is recursive, we need 12 months worth of indices to be able to "start up". One possibility that exists, is to define the index for the first 12 months, by analogy to the geometric index, as

$$I_k = (r^k_{k-12})^{k/12}, k = 1, 2, \ldots, 12. \tag{2.24}$$

As defined, the annual index is independent of intermediate changes. On the other hand it will be saw-toothed unless individual monthly sample sizes are large. This is due to the fact that consecutive monthly estimates are totally independent. Moreover, it must be noted that the lagging problem will be at least as serious in the case at hand as it is for the indices presented earlier.

3. ADJUSTMENTS

In this section, two adjustment procedures for the above indices will be discussed. First, because the first four indices suffer from one to three month lags, they will smooth out true, sharp peaks. From prior data, it has been observed that rent indices do exhibit sharp rises, in certain cities, with some regularity. To "correct" the smoothed out index, an empirical adjustment will be proposed. By contrast, due to the volatility of the annual index, a smoothing adjustment will also be proposed.

3.1 Empirical Adjustments

It is known, for example, that most rents in Montreal change in July. The first four indices discussed in the previous section would distribute this July change over July, August, September and October. One could however adjust the index in July to reflect a larger change and counter adjust it in the following three months. More precisely, the index could be multiplied by r^* in the reference month and then by $(r^*)^{-1/3}$ in each of the following three months. Since all the proposed indices are chained indices, in the third month after the reference month the four multipliers will offset each other, leaving no trailing biases. As for the choice of r^*, this will depend on continued empirical observations in each particular city.

It is to be noted that such adjustments must be performed in rare situations only and with great care. It is imperative that the particular situation be monitored, for it is not uncommon for such aberrations to disappear suddenly.

3.2 Smoothing

As a last effort in redeeming a volatile, saw-toothed index, one could consider smoothing it. Like the above adjustments, smoothing should be considered in rare and extreme situations only: in cases where no other alternative exists. The smoothing procedure we consider here involves averaging the index at time m with a linear extrapolation to time m of the smoothed index from time $m - 1$ and $m - 2$. One possible choice of the smoothed index at time m, S_m, is then given by

$$\begin{aligned} S_m &= I_m/2 + (2S_{m-1} - S_{m-2})/2 \\ &= S_{m-1} + (I_m - S_{m-2})/2. \end{aligned} \tag{3.1}$$

Since the smoothing operation basically projects past data into the future, the smoothed index will extend past trends and therefore introduce some lags. Moreover, the method is recursive and consequently could also introduce unwanted biases. Other smoothing methods could be considered, although the utility of smoothing an index that suffers from serious lags is questionable.

4. EMPIRICAL STUDY

The study described in the following paragraphs was initiated in order to test the performance over time of the proposed indices and adjustments. The study provides quantitative information on the ability of the indices to track the true index accurately. It supports the mostly heuristic observations made above and reinforces the theoretical ones.

4.1 The Population

The population of rented dwellings used in this study was designed to duplicate the real situation as closely as possible. For this purpose, the cities,their sizes, and their sample sizes were selected to correspond to those used by the Rent Component of the CPI. Since all real data on rents is available for periods of six months only, the needed thirteen months of data had to be simulated. Eight cities were chosen for this purpose. Some are large, some are small, some have periodic jumps in their indices, but all are CPI index cities and have sufficient amount of rent data available. Moreover, while some of the indices in these cities are strictly increasing, others are both increasing and decreasing.

Only the initial rents of all units (those collected when the unit rotated in) on the CPI rent database for the years 1979 to 1984 inclusive, for the eight cities mentioned above, were

Table 1

Average Sample Sizes (Distinct Units) and the Index at 8401 for Eight Cities Based on the Simulated Population

City	Average Monthly Sample Size	Index at 8401 (8001 = 100)
Halifax	51	144.3
Montreal	268	136.6
Ottawa	35	130.0
Toronto	170	130.4
Winnipeg	105	132.0
Edmonton	112	125.2
Calgary	97	123.5
Vancouver	105	130.5

retrieved. For each unit, twelve additional months worth of data were then simulated using the observed parameters. (This approach is operationally easier then simulating seven months of data in addition to the existing six.) More precisely, for each unit, first a decision was made whether or not a change in rent will occur sometime in the next twelve months. The probability of this event was set to be equal to the observed probability of a rent change in that particular city and year. Then, given that a change was to occur,the appropriate month was selected proportional to the observed incidence of rent changes, again specific to the city and month at hand. The actual amount of the rent change was assumed to be distributed normally with a fixed mean and variance. Robust estimates of these two parameters were obtained from the existing data for each city and each month.

All programming was done in SAS (Statistical Analysis System). The random numbers were generated using the routines RANUNI and RANNOR. The resulting population consists of eight cities and four years of fully rotated data (that is,discarding start up months). The average monthly sample sizes and the value of the simulated index for January 1984 (with Jan 1980 = 100) can be seen for each city in Table 1. The indices, calculated for each of the cities, resemble very closely those observed originally. In the following comparisons, the indices of the simulated population were taken to be the true reference points to be reproduced.

4.2 Comparison of Indices

For the purpose of calculating the indices, it was assumed that of the 13 available observations for each unit, only those for months 1, 5, 9 and 13 were actually observed. All calculations were then based on this (4/13) subsample.The five indices described above were calculated for each city and compared to the true index. All indices are fixed at 100 in January 1980. The empirical adjustment was tested with the Montreal, Halifax and Winnipeg data, for the month of July, January and October respectively. While the results for all the possible combinations of cities and indices are too numerous to include herein, they are available from the author. Some selected highlights will be put forth in the following paragraphs. While not exhaustive, they are hoped to be representative as well as indicative of the situation at hand.

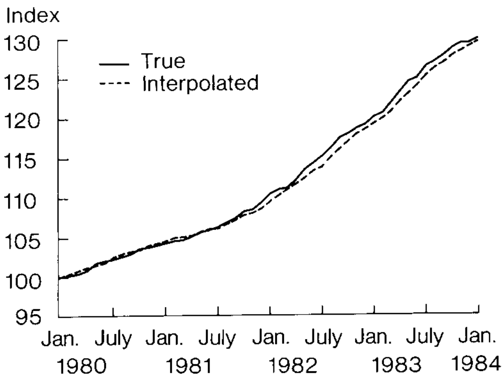

Figure 1. Plot of the True Index and the Interpolated Index for the City of Ottawa

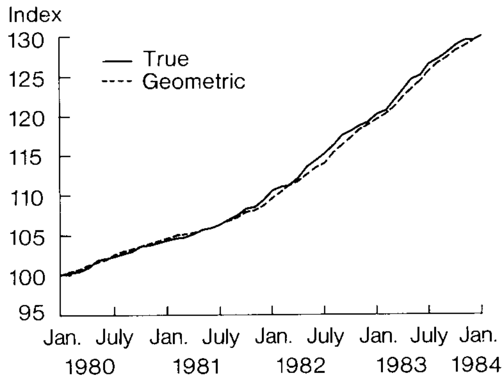

Figure 2. Plot of the True Index and the Geometric Index for the City of Ottawa

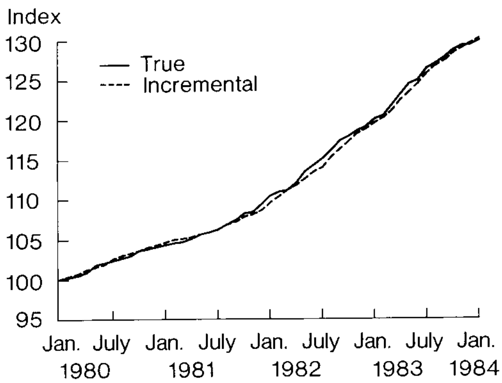

Figure 3. Plot of the True Index and the Incremental Index for the City of Ottawa

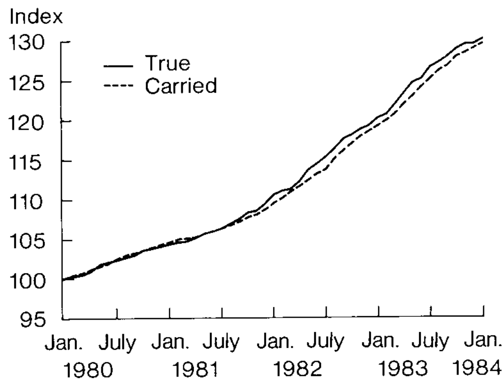

Figure 4. Plot of the True Index and the Carried Index for the City of Ottawa

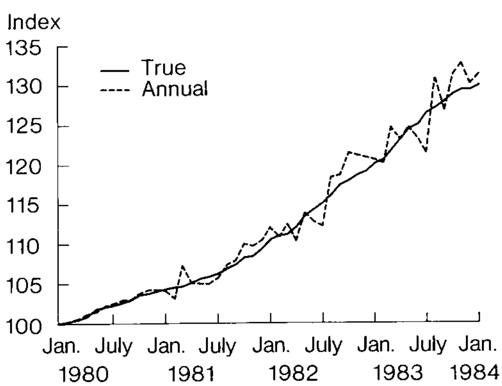

Figure 5. Plot of the True Index and the Annual Index for the City of Ottawa

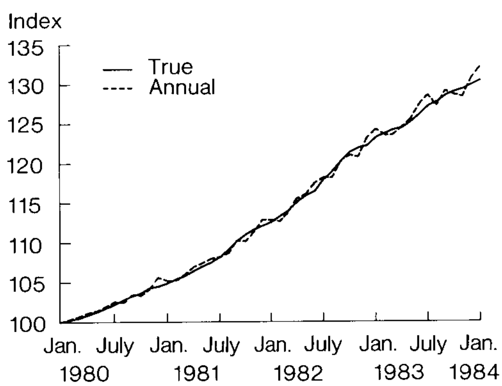

Figure 6. Plot of the True Index and the Annual Index for the City of Toronto

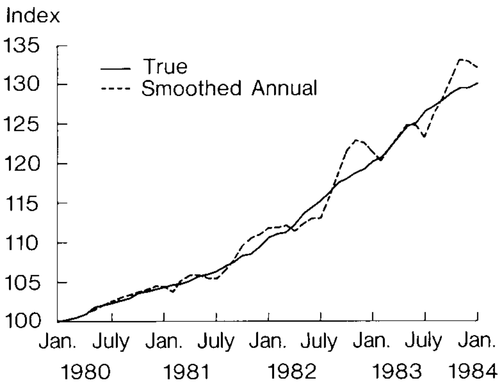

Figure 7. Plot of the True Index and the Smoothed Annual Index for the City of Ottawa

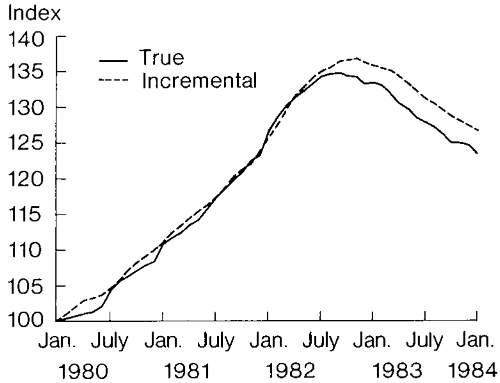

Figure 8. Plot of the True Index and the Incremental Index for the City of Calgary

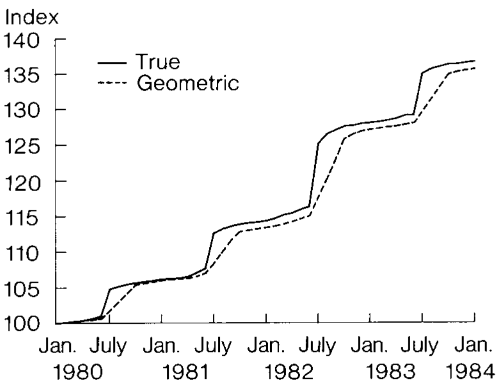

Figure 9. Plot of the True Index and the Geometric Index for the City of Montreal

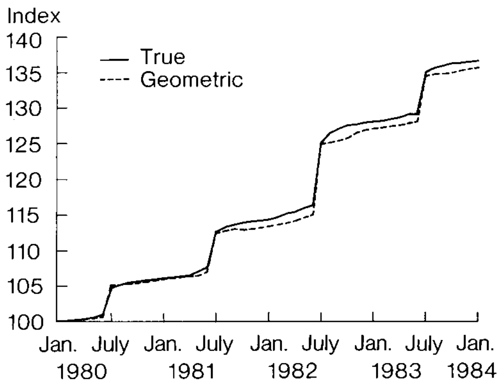

Figure 10. Plot of the True Index and the Adjusted Geometric Index for the City of Montreal

As can be seen in Figures 1-5, all five indices track the true index reasonably well, even in the case of small sample sizes such as in the city of Ottawa. As expected, the first four indices show some lags, those being more pronounced in the carried and interpolated index. (Note that the lagging problem could likely be accentuated by generating the population with exponentially increasing prices). Not surprisingly, the annual index is rather volatile. For cities with large sample sizes however,(e.g. Toronto), the annual index performs well (see Figure 6). While the smoothing adjustment of Section 3.2 does indeed smooth the index, the results are less than satisfactory as can be seen in Figure 7 (c.f. Figure 5). Perhaps a larger number of points should be used for the extrapolation but then the lagging problem would be even more pronounced. Figure 8 further demonstrates how sudden unexpected changes in trends are reported with a delay. However, expected jumps in the index (as in July in Montreal, Figure 9) can be adjusted successfully using the adjustment procedure of Section 3.1 (Figure 10).

Table 2

Mean Square Errors of Five Indices in Eight Cities

City	Interpolated		Geometric		Incremental		Carried		Annual	
Halifax	30*	(3)	19*	(2)	12*	(1)	48	(4)	74	(5)
Montreal	48*	(3)	24*	(2)	9*	(1)	160	(5)	82	(4)
Ottawa	17	(3)	12	(2)	8	(1)	22	(4)	95	(5)
Toronto	36	(4)	27	(3)	20	(2)	29	(5)	13	(1)
Winnipeg	27*	(3)	17*	(2)	10*	(1)	66	(5)	41	(4)
Edmonton	46	(1)	64	(4)	88	(5)	55	(3)	50	(2)
Calgary	56	(2)	81	(4)	121	(5)	64	(3)	46	(1)
Vancouver	70	(5)	53	(2)	39	(1)	64	(4)	60	(3)

Note: 1. Bracketed figures indicate ranking within cities.
2. Starred figures are results of adjusted indices as per Section 3.1.

Mean square errors of the five indices away from the true index have been calculated for each city (Table 2). The three interpolation based indices (interpolated, geometric and incremental) have been adjusted for the cities of Montreal, Halifax and Winnipeg. Table 2 also presents the rankings (from smallest to largest) of the mean square errors of the five indices within each city. The carried and the annual index tend to perform the worst. The three interpolation-based indices perform relatively alike. In general, in cities where the index is climbing consistently, the performance of these three indices worsens in the order: incremental, geometric, interpolated. The order is reversed in cities where sharp decreases in the index have been observed. It is unlikely, however, that the strategies could be interchanged based on observed behaviours only.

5. SUMMARY

Both the theoretical as well as the empirical observations suggest that the yearly index is too volatile in cities where sample sizes are not large enough. Smoothing, at least of the type described, has proven fruitless. For this reason the annual index should be reserved only for those rare cases where sample sizes permit. On the other hand, the annual index could be used in conjunction with one of the more stable four-month indices to produce a composite estimate analogous to that proposed by Kosary *et al.* (1982). However, empirical observations would be needed to determine the appropriate weights to be used in averaging the two indices.

By contrast, the carried, and to some degree, the interpolated index tend to be too smooth. That is they tend to smooth out all peaks in addition to demonstrating a one or two (index) point lag. While the incremental and geometric indices are not entirely free of these lags, they tend to track the true index a little more closely. The incremental index performs the best overall, however, because of the mathematical ''cleanliness''of the geometric index (i.e. its theoretical independence of its history and its correspondence to the chaining structure), it is the latter that is recommended here. In other words, the geometric index does not retain terms that could cause biases in the long run.

It is also apparent that whenever possible, prior knowledge can be used to improve the index. Empirical adjustments as described in Section 3.1 can be useful, provided that they are well founded. If their use is contemplated, it is imperative that the empirical knowledge that leads to their application be monitored and its continued existence verified.

ACKNOWLEDGEMENT

Thanks are due to Mr. George Sampson for all of his help and to the referees and the editorial staff for their constructive comments.

REFERENCES

DOLSON, D.D. (1982). Rent status survey: Analysis. Technical Report, Statistics Canada.

KOSARY, C.L., BRANSCOME, J.M. and SOMMERS, J.P. (1982). Evaluating alternatives to the rent estimator. Technical Report, Bureau of Labor Statistics.

KOVAR, J. (1984). Note on calculating the rent index. Technical Report, Statistics Canada.

SZULC, B. (1983). Linking price index numbers. Technical Report, Statistics Canada.