Different Approaches to the Treatment of Seasonal Products: Tests on the Israeli CPI

Yoel Finkel
Anna Rakhmilevich
Victoria Roshal

Central Bureau of Statistics, Israel

Abstract:

Seasonal products are either (1) not available during certain seasons of the year, and are termed *strong seasonality products,* or (2) have regular fluctuations in prices or quantities according to different seasons and are termed *weak seasonality products*. In this paper we analyze various approaches to the treatment of seasonal products according to Chapter 22 of the CPI/PPI Manuals which use an artificial dataset to present them. We use the real data from the Israeli CPI to compare different methods of seasonality treatment, to test the conclusions made in Chapter 22, and to reveal some problems arising in calculation of the CPI on a practical basis, especially for products that have very strong seasonality.

This paper is based on the one published by Artsev and Finkel (ECE 21(2) who make a similar analysis for the years 1997-2001. In this paper, the price and quantity dataset for fresh fruits (strong seasonality) is extended for the years 1997-2006. We test whether the contradictory findings of the CPI Manual and the paper are due to the data structure and its sensitivity to the months at which the commodity enters the market.

I. Introduction

In their paper from 2004, Diewert, Finkel and Artsev summarized the main methods of dealing with seasonal products in the CPI, presented in Chapter 22 of the CPI Manual. They implemented many of the solutions proposed, by inserting real market data from the Israeli CPI for the years 1997-2002. While some of the conclusions that they reached when exploiting this true market data confirmed those of the CPI Manual, others contradicted them.

In the present paper we aim to extend the data used in order to check whether the inference still holds. We use the data for the same commodities and extend them up until December 2006. We also focus on some of the problems raised by the previous paper, which were never explained, and propose several solutions.

In Section II we start with the comparison of the databases of the artificial dataset, used in the CPI Manual, and the real market data from the Israeli CPI. This comparison is important, as the differences in the data structure can explain some of the contradictory conclusions.

Section III presents the results of the main methods that deal with seasonal products proposed by the CPI Manual. We use the Israeli data extended up to the year 2006.

In Section IV we attempt to explain some of the main challenges that a real dataset poses, in contrast to the artificial one. The sensitivity of the data to the length of the period through which the commodity is consumed is examined and some solutions are proposed.

Section V presents the conclusions.

II. Comparing the databases

Data structure can be a key factor in determining the performance of an index. For example, Turvey's artificial dataset produces a Rothwell price index that is less volatile than the Lowe index. Using Israeli real market data, however, leads to a totally opposite conclusion. The sensitivity of the specific dataset to different calculation methods can result in different decisions as to the best way to compile a price index. Hence, it is important to make a detailed description of the dataset.

In section B of Chapter 22 of the CPI Manual, a modified version of Turvey's artificial data set for computing seasonal items is introduced. This set includes data for 5 seasonal commodities (apples, peaches, grapes, strawberries and oranges) over 4 years (48 months). Modifications of this set were made as follows: the data for grapes were adjusted to emphasize the differential between annual Laspeyres and Paasche indices; the monthly inflation rate for the data in the fourth year was doubled compared to the average of the first three years.

Three of these commodities are present throughout the year (first, third and fifth). The second item appears for 5 months from June to October every year. The fourth appears from May to July. These commodities always appear at high initial prices that decrease in the subsequent months.

The Israeli "true" dataset includes 7 commodities of **fresh fruits** – lemons, apricots, avocado, watermelon, persimmon, grapefruits and bananas. The fresh fruits have *strong seasonality* (many months with zero prices). As strong seasonality presents more challenges for the index compilers, we will restrict our analysis to these products.

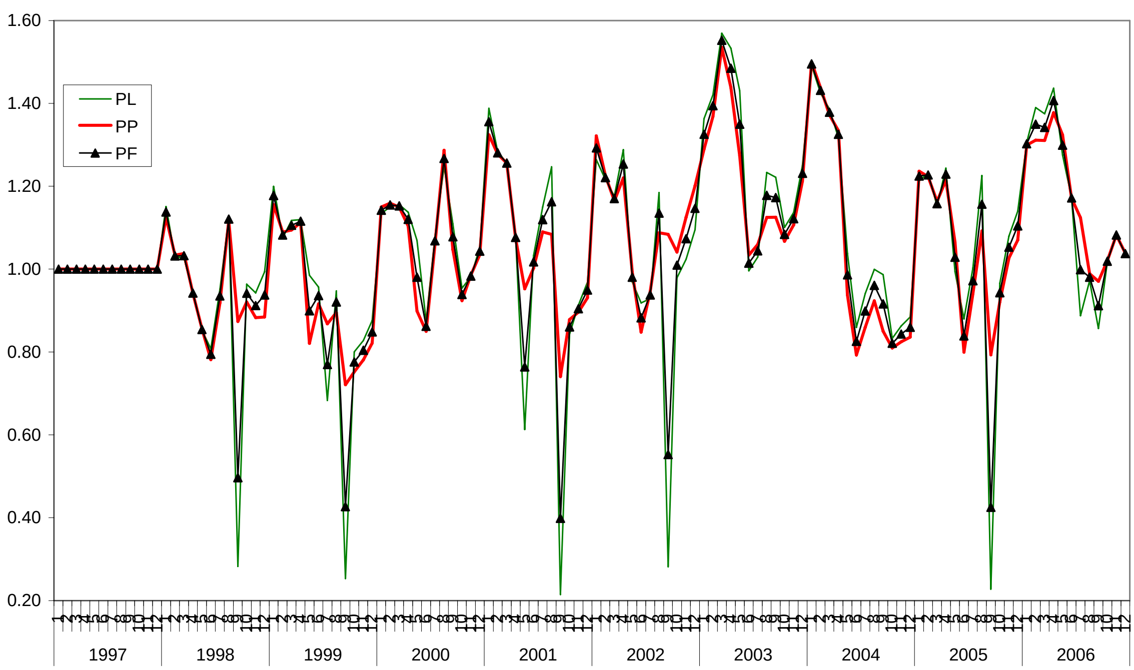
Out of the group of 7 commodities, 2 are present throughout the year (lemons and bananas). 2 more commodities (avocados and grapefruits) disappear from the market only for only two or three months, in the summer, and only in some of the years. Apricots, watermelons and persimmons have the strongest seasonality of all – they are present in the market for 1 to 5 months in the year. The period during which the items appear on the market is not the same every year, which poses the greatest challenge on the compilation of indices of several types.

Our sources for testing the results of the CPI Manual are: (a) an annual Household Expenditure Survey (for the years 1997-2006), consisting of over 6,200 households and provides the quantity data; (b) the monthly CPI, which includes data on prices. We built a dataset for the ten-year period of prices and quantities for the commodities mentioned above.

III. The performance of different methods based on the Israeli market data

The concept of comparing the prices of the same month in different years is the first being presented in Chapter 22 of the CPI Manual. Strongly seasonal commodities are likely to reappear in the same month, thus increasing the overlapping of commodities. These indices perform reasonably well for the Turvey's artificial dataset, but they experience huge peaks and troughs, as can be seen from Figure 1, especially for the case of the Laspeyres index.

Figure 1: Year Over Year Monthly Fixed Base Laspeyres, Paasche and Fisher Indices



The 2004 paper tests some of the conclusions presented in Chapter 22 of the CPI Manual. We would like to check whether these conclusions still hold for our extended dataset. The first test checked whether "approximate" methods, which use data normally at the disposal of statistical agencies at the time of computation, can replace the "current" methods, as was stated in the CPI Manual. Approximate methods usually use the base-year expenditure weights, instead of current weights (thus the

Laspeyres true index and its approximation share the same formula, in opposite to the Paasche and Fisher indices – see appendix 2). In **Table 1** below we compare the year over year "current month" fixed base Fisher index (P_F) with the approximate monthly fixed base Fisher index (P_{AF}).

In 1997 the same data are used for both methods. Therefore, the relevant comparison is for 1998-2006. In 17 out of the 108 months compared, the differential is 5 percent or more. However, only 8 of these months are really extreme (five of them in September). Only the Fisher formula is compared, since if approximate methods may be used, Fisher is preferable to Laspeyres or Paasche in order to reduce the upward or downward bias.

Table 1: Ratio Between Year over Year "Current Month" Fixed Base Fisher and Approximate Fixed Base Fisher Indices (P_{AF}/P_F)

Month/Year	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
1	1.00	1.01	0.99	0.99	1.02	0.98	1.01	0.98	0.99	0.97
2	1.00	1.00	1.01	1.00	1.00	1.00	1.00	0.99	1.00	0.98
3	1.00	1.00	1.00	1.01	1.01	1.01	1.00	1.00	1.00	0.98
4	1.00	1.00	0.98	1.01	1.00	1.02	0.99	0.98	1.01	0.98
5	1.00	0.99	1.04	**1.05**	**1.27**	0.97	1.01	0.99	**0.95**	**0.94**
6	1.00	1.01	1.00	1.01	1.01	1.03	1.02	1.02	1.03	1.03
7	1.00	1.01	**1.16**	1.01	**1.06**	1.02	1.04	1.04	1.03	**1.14**
8	1.00	0.99	1.02	1.00	1.04	**1.05**	1.04	1.04	1.05	1.04
9	1.00	**1.77**	**1.83**	1.04	**1.82**	**1.58**	1.04	**1.07**	**1.81**	**1.06**
10	1.00	1.01	1.00	1.01	1.01	0.96	1.00	0.99	1.00	0.99
11	1.00	0.99	0.98	0.99	0.98	**0.94**	0.98	0.98	0.99	1.00
12	1.00	1.01	0.99	1.00	1.00	**0.93**	0.99	0.98	0.98	1.00

In **Table 2** below the same comparison is made, this time for year over year monthly chained Fisher indices.

Table 2: Ratio Between Year over Year "Current" Monthly Chained Fisher and Approximate Monthly Chained Fisher Indices (P_{AF}/P_F)

Month/Year	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
1	1.00	1.01	0.98	0.99	**1.05**	1.01	0.98	**0.95**	0.96	**0.94**
2	1.00	1.00	1.03	0.99	1.00	1.00	**0.95**	**0.94**	0.97	**0.95**
3	1.00	1.00	1.00	1.02	1.03	1.04	**0.95**	**0.95**	**0.94**	**0.92**
4	1.00	1.00	0.98	1.03	1.01	**1.07**	**0.91**	**0.92**	0.97	**0.92**
5	1.00	0.99	**1.09**	**1.12**	**1.58**	**1.32**	**1.20**	**1.19**	**1.21**	**1.17**
6	1.00	1.01	1.00	1.00	1.00	**1.06**	**1.17**	**1.17**	**1.20**	**1.24**
7	1.00	1.01	**1.14**	1.01	**1.06**	1.02	**1.05**	1.04	**1.05**	**1.17**
8	1.00	0.99	1.01	1.00	**1.05**	1.04	1.04	1.01	**1.06**	1.03
9	1.00	**1.77**	**1.82**	**1.23**	**1.75**	**1.77**	**1.18**	**1.22**	**1.62**	**1.61**
10	1.00	1.01	1.02	1.01	1.02	**1.07**	0.99	1.00	0.99	0.99
11	1.00	0.99	0.99	1.02	1.01	0.99	0.97	0.97	0.98	0.97
12	1.00	1.01	1.01	1.01	1.01	0.99	0.98	0.97	0.97	0.97

It seems that the chained indices are not approximated as well as the base period indices, especially if the base year is far from the current year, which fits the conclusion of the Manual. In forty-four observations the differences between the true index and its approximate counterpart is 5 percent or more. May and especially September seem to be “outlier” months in both comparisons. More on the causes to this are discussed in section IV of the paper. As we move farther away from the base year, the number of months with high differentials between the "current" chained and approximate chained indices grows.

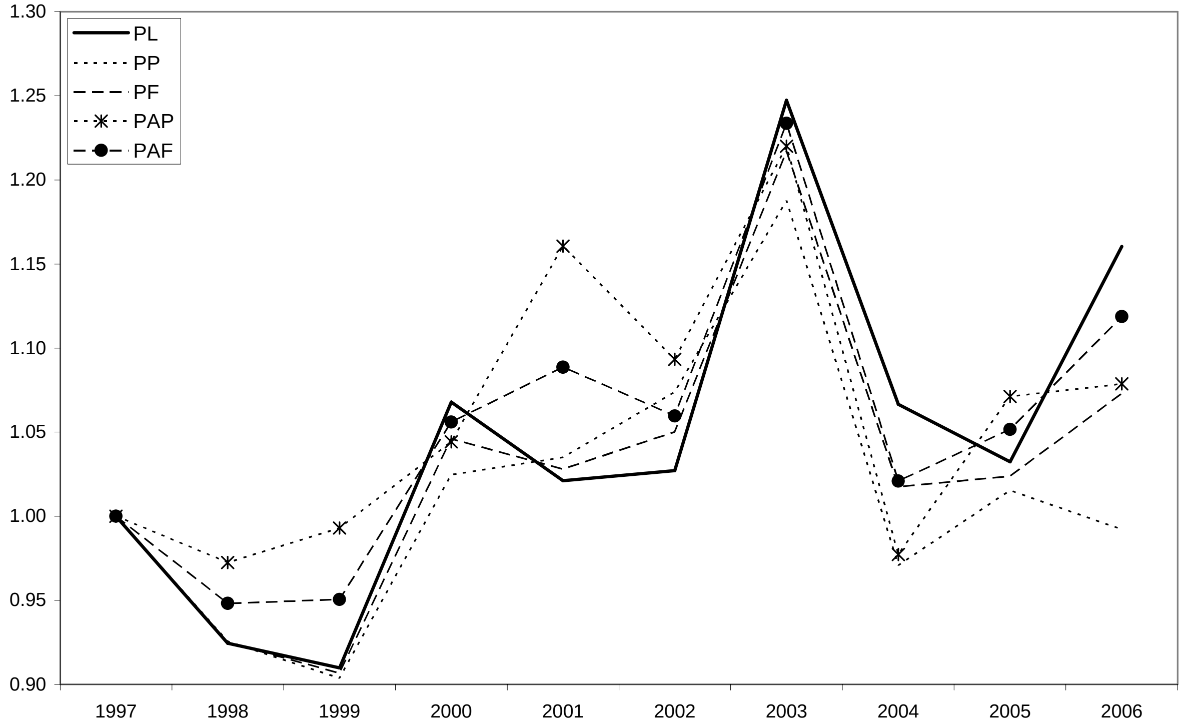
An additional conclusion from Chapter 22 was that chained indices would usually reduce the spread between the Laspeyres (P_L) and Paasche (P_P) indices. In **Table 3** below we compare the spread between these two formulas, for the fixed-base indices and their chained counterparts. The 2004 paper stated that for the years 1998-2002, the average spread for chained indices was lower for the fixed-base indices than for the chained ones. The results seem to contradict the conclusions of the Manual. After extending the database to include 4 more years, we claim that there is no clear pattern in the difference of approximation of fixed base or chained indices. For short-time periods they perform poorly, but for a period of 10 years, the average spread ratio is almost 1 in both cases.

Table 3: Mean Annual Ratio Between Year over Year Monthly Laspeyres and Paasche Indices (P_L/P_P)

Year	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**Average**
Fixed Base	1.00	0.97	0.974	1.03	0.94	0.93	1.04	1.05	0.97	0.98	1.00
Chained	1.00	0.97	0.972	0.98	0.93	0.94	1.04	1.05	1.06	1.07	0.99

Annual year-over-year approximate indices do not perform better, as they are only compilations of the monthly indices. In Figure 1 the dotted line represents the Paasche index, and the dashed line represents the Fisher annual index. It can be seen that only for four years (2000, 2002, 2003 and 2004) are the approximate indices (defined by marks) close to their true counterparts.

Figure 2: Annual Fixed Base and Fixed Base Approximate Laspeyres, Paasche and Fisher Indices



Rolling year indices, based on the data of the preceding 12 months, are expected to be smooth and free from seasonal fluctuations and are regarded as seasonally adjusted CPIs. There are also fixed-base and chained versions of this index. However, one has to be cautious when using approximated versions of this index (which use the base year as a reference period): as we move farther from the base period, Laspeyres and Paasche approximate chained indices diverge considerably from their true counterparts. Fixed base approximation works much better, and so does the Fisher index (Figures 3 and 4).

Figure 3: Rolling Year Fixed Base and Chained Laspeyres, Paasche, and Fisher Indices

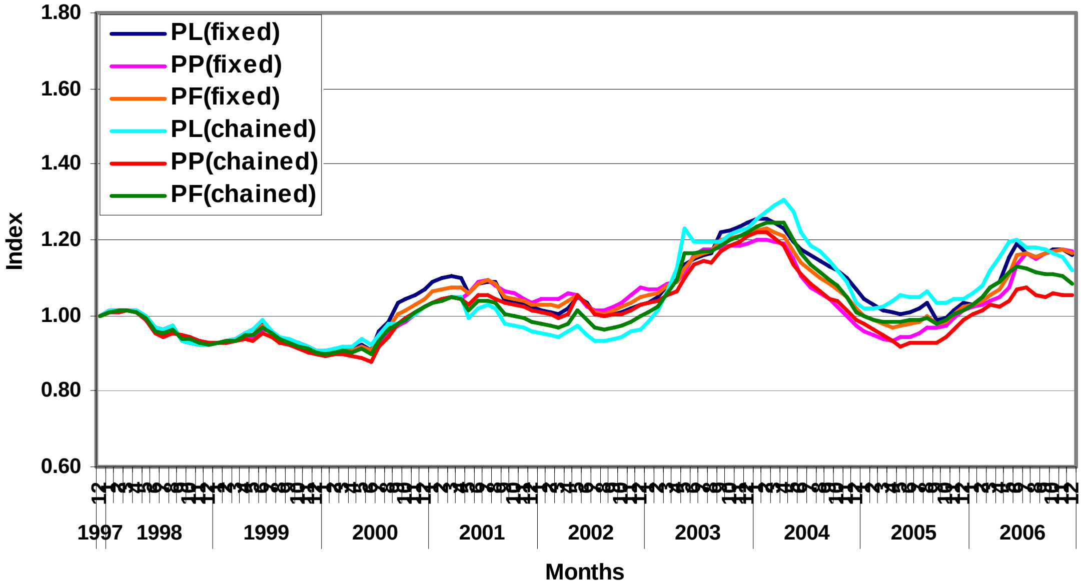


Figure 4: Rolling Year Approximate Laspeyres, Paasche, and Fisher Price Indices

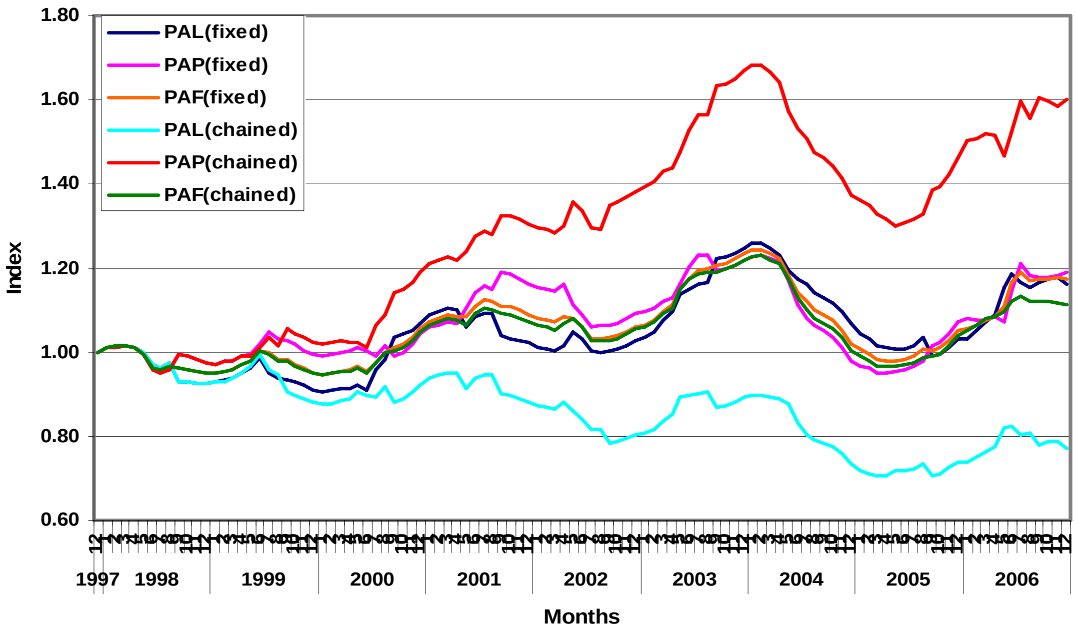
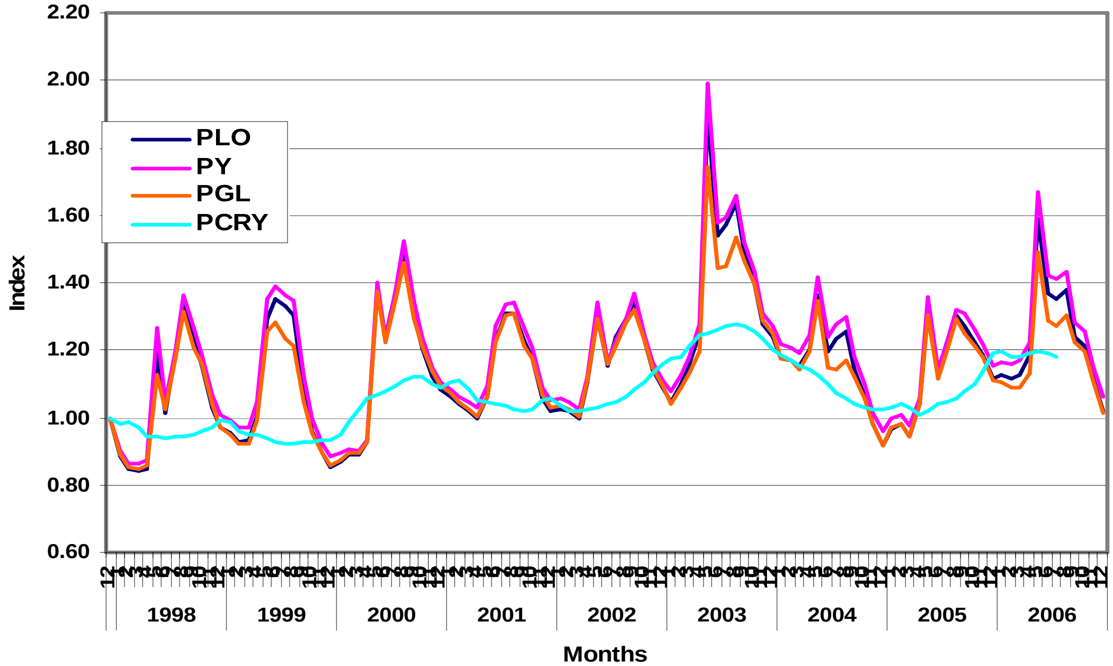
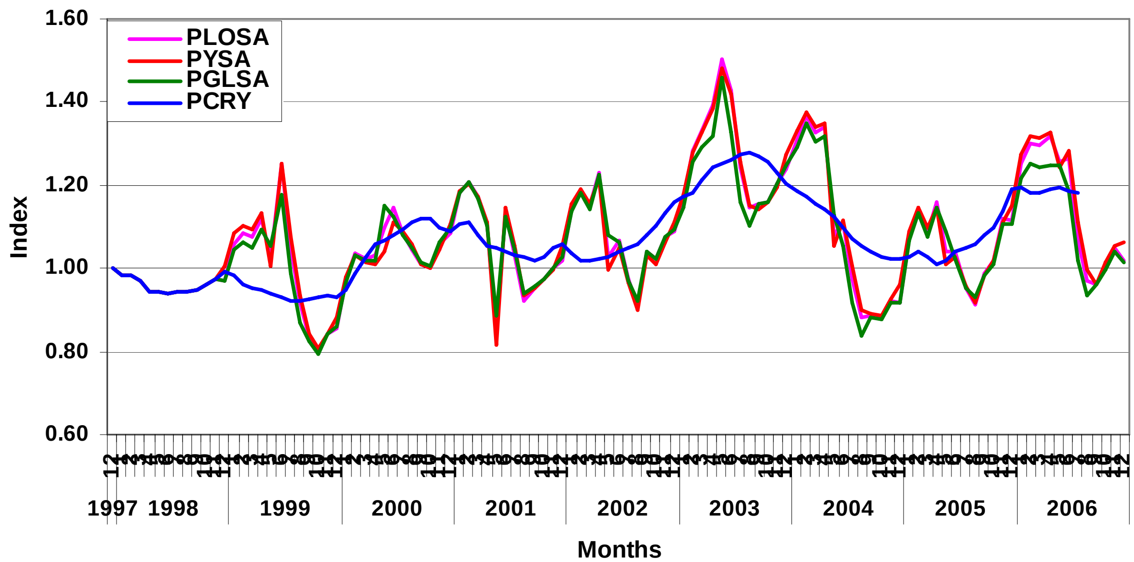


Figure 5: Lowe, Young, Geometric Laspeyres, and Centered Rolling Year Indices with Carry Forward Prices



Extreme seasonal fluctuations of annual basket indices (Lowe, Young and Geometric Laspeyres), which use monthly prices but annual base quantities / expenditure shares, make them unsuitable predictors for their seasonally adjusted rolling year counterparts (Figure 5).

Figure 6: Seasonally Adjusted Lowe, Young, and Geometric Laspeyres Indices with Carry Forward Prices and the Centered Rolling Year Index

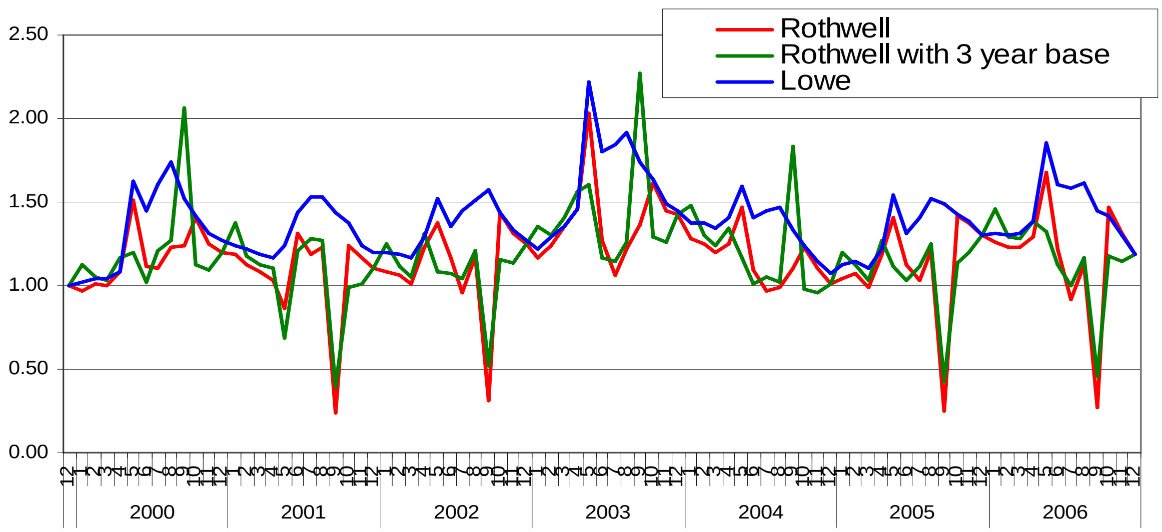


The artificial dataset produces regular seasonal movements, which are then exploited to construct seasonally adjusted indices. Peaks and troughs of the indices constructed on the basis of the true dataset occur at different months and hold for different periods of time. In this case, seasonally adjusted annual basket indices

(Figure 6) may not be so close to the Centered Rolling Year Index, as is the case with the artificial dataset. In this case, using an average of several years seasonal adjustment factor can produce better, "smoother" results.

Until 1987 the price indices of fresh fruits and vegetables in the Israeli CPI were computed according to the Rothwell formula that expresses changes in quantities as well as in prices, whereas the base year quantities were a three-year average, in order to overcome the sample errors in the Household Expenditure Survey. As of January 1988 these indices are computed according to the annual basket month-to-month Lowe formula. Missing prices are inserted using the imputed prices method (as opposed to the carry forward or maximum overlap methods). The CPI Manual studies these two indices and concludes that the Rothwell index exhibits smaller seasonal movements and is less volatile in general. In contradiction to the findings of the Manual, in our study the Rothwell index is much more volatile than the Lowe index. We also constructed a three-year quantities' based Rothwell index, in order to imitate the Israeli CPI method before 1988. Three indices are presented in Figure 7.

Figure 7: The Lowe with Carry Forward Prices, Normalized Rothwell and Normalized Rothwell with 3-year base quantities



Rothwell Indices have lower means than Lowe indices, but are much more volatile, as can be seen in the table below. It is striking that the three-year based Rothwell has a higher mean and standard deviation, than the regular one.

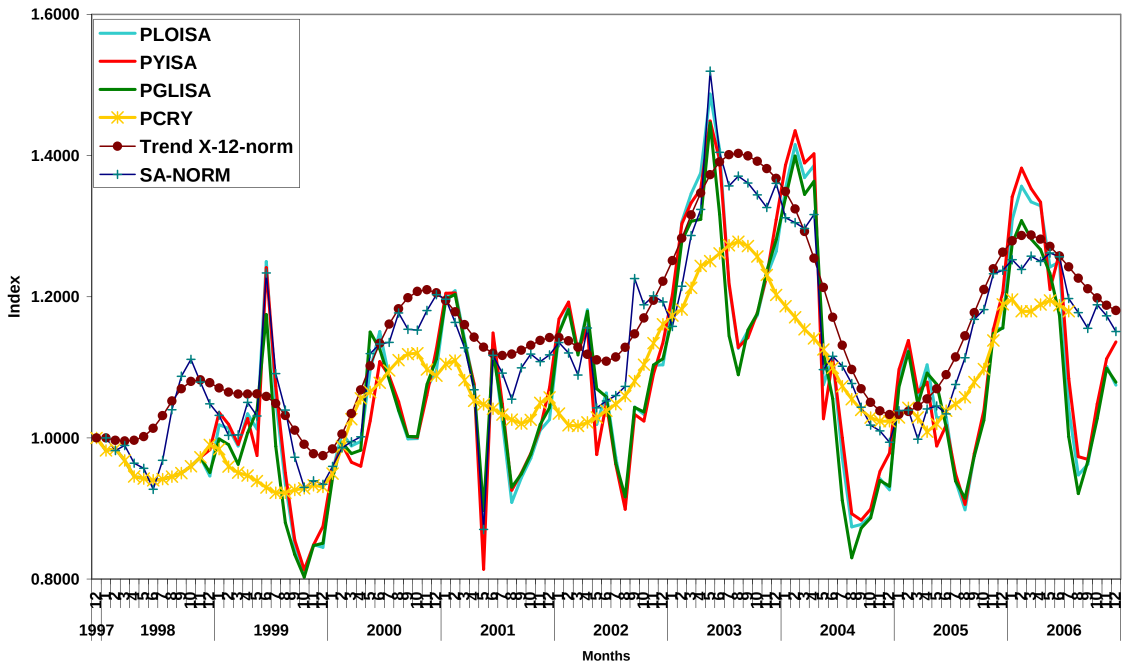
Table 4: Mean and Standard Deviation of Normalized Lowe and Rothwell Indices, and Rothwell Index with 3-year base

	Lowe 1998-2006	Rothwell 1998-2006	Lowe 2000-2006	Rothwell 2000-2006	Rothwell with 3 year-base 2000-2006
Mean	1.1612	0.9921	1.3911	1.1702	1.1785
St. Dev.	0.1888	0.2443	0.2165	0.2713	0.2728

In Figure 7 one also notices the almost- zero downfalls of the Rothwell indices, occurring in September in the years 2001, 2002, 2005 and 2006. It also happens in the years 1998 and 1999 which are not presented here. The reasons for this, as well as the failure of the Rothwell index to overcome this problem are explained in section IV of this paper.

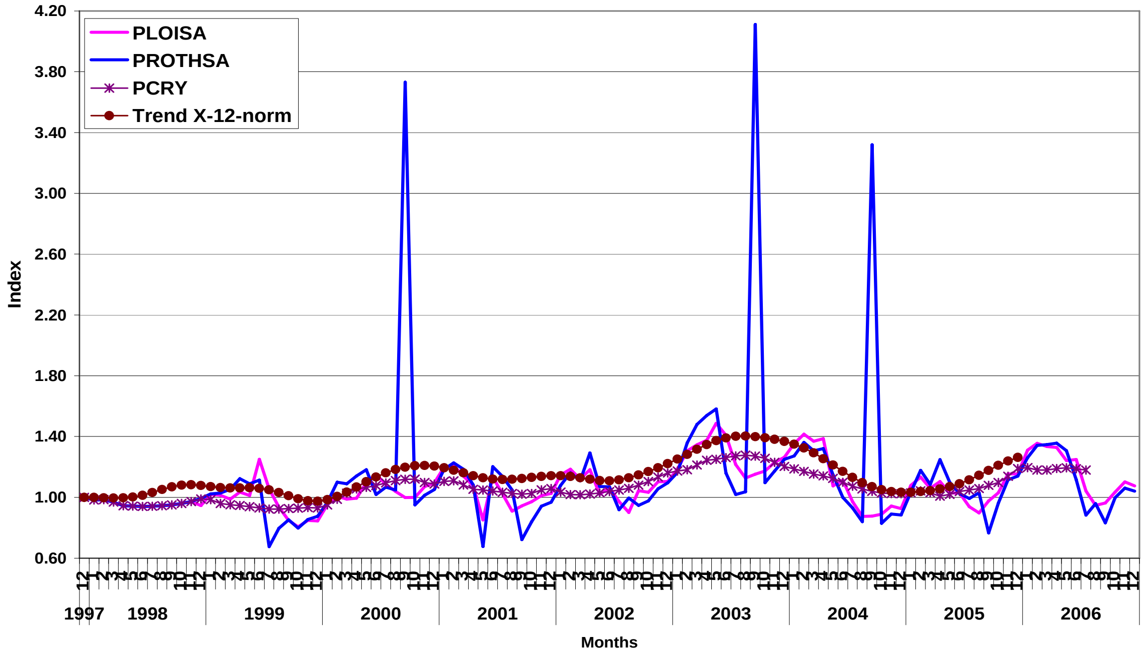
As offered in Chapter 22 of the Manual, we made an attempt to forecast Rolling Year Indices using month-to-month annual basket indices. We also added the 7-item group Seasonally Adjusted Lowe Index, obtained by the ARIMA-12, and its trend. In order to make them comparable, all the series were normalized by dividing the original indices by the first observation. Figure 8 shows that the predicted values of these "Seasonally adjusted" indices have still very large seasonal movements. Although they are closer to the target Centered Rolling Year Index, they are still very volatile. Lowe indices adjusted using the ARIMA-12 methods are closer to the Centered Rolling Year Index. It is remarkable that even after all the adjustments, the trend looks like a roller-coaster. The findings in Chapter 22 of the Manual show very smooth seasonal adjusted indices.

Figure 8: Seasonally Adjusted Lowe, Young, and Geometric Laspeyres Indices with Imputed Prices, Centered Rolling Year Indices, Seasonally and Trend adjusted ARIMA 12 Index and Seasonally Adjusted ARIMA-12



The most confusing result of all seasonally adjusted indices was displayed by the seasonally adjusted Rothwell index, presented in Figure 9.

Figure 9: Seasonally Adjusted Lowe, Rothwell and Centered Rolling Year Indices, with ARIMA-12 seasonally and trend adjusted index



Three extreme peaks during 9 years exhibited by the seasonally adjusted Rothwell index, are absent in the annual basket indices (Lowe, Young and Geometric Laspeyres – only the Lowe index is shown in Figure 9 for the sake of better presentation). This strengthens the argument in favor of using the Lowe index in the Israeli CPI. The reasons for these peaks in the Rothwell index will be explained in the next section.

IV. The "Watermelon mine": dealing with the sensitivity of the price index to entries and exits of commodities

The 2004 paper discussed several differences between the theoretical approach to dealing with seasonal products and its practice. These differences might be the reason for the contradictions between the finding of the paper and those of the CPI Manual.

Of the four problems mentioned, the unsynchronized seasonal cycles seem to raise the highest challenge to most of the index formulas discussed in Chapter 22. Accordingly, the dataset was modified in order to fit the "theoretical approach". Commodities were forced to appear and disappear during the same month, each year, by dropping observations or carrying forward last prices. 16 observations were modified for 6 years 1997-2002 (in the 2004 paper). In this manner, 27 more were changed for the four years 2003-2006 (16 of them due to the fact that in these years grapefruits ceased to be measured in the months July-September, making them strongly seasonal commodities, after they had had only weak seasonality before).

This "perfectly-modified" dataset permits, for example, the month-to-month yearly indices to perform much better, because it overcomes the main drawback of the formula: the sum of price relatives multiplied by expenditure shares might be very low if some of the products were not in the market, compared to the same month in the previous year, making the dramatic decrease of the price index due to poor commodities overlap, and not to any real price change. The fall of the index is not so extreme in the perfectly-modified dataset. Toying with the market data in order to achieve nicer, smoother results may be problematic in reality, however.

What exactly are the properties of the Israeli dataset that undermine the performance of the seasonal price index methods so extremely? Is it possible in practice, on an everyday basis, to overcome the problems without changing the whole database? Every statistical agency dealing with seasonal prices has learnt the pitfalls and the sensitivity of its data, and developed methods to deal with them.

In the database used in this paper, the high-weighted mine that "spoils" the smooth and theoretically acceptable performance of the price indices is the watermelon. This commodity, which appears for three to five months during the year, entering the market most of the time in May, sometimes in April or June, and exiting in August or September, has not only literally, but also in expenditure terms, the highest weight alongside with the lowest price level. This high expenditure weight exceeds at times 80% of total household expenditure for fresh fruits (out of 7 commodities listed here). Low price levels cause the slightest monetary change to be translated into high percentage change. Finally, unstable seasonal cycle exacerbates the problem.

The prices of the watermelon during the years 1997-2006 are listed in Table 5.

Table 5: Watermelon Prices, 1997-2006

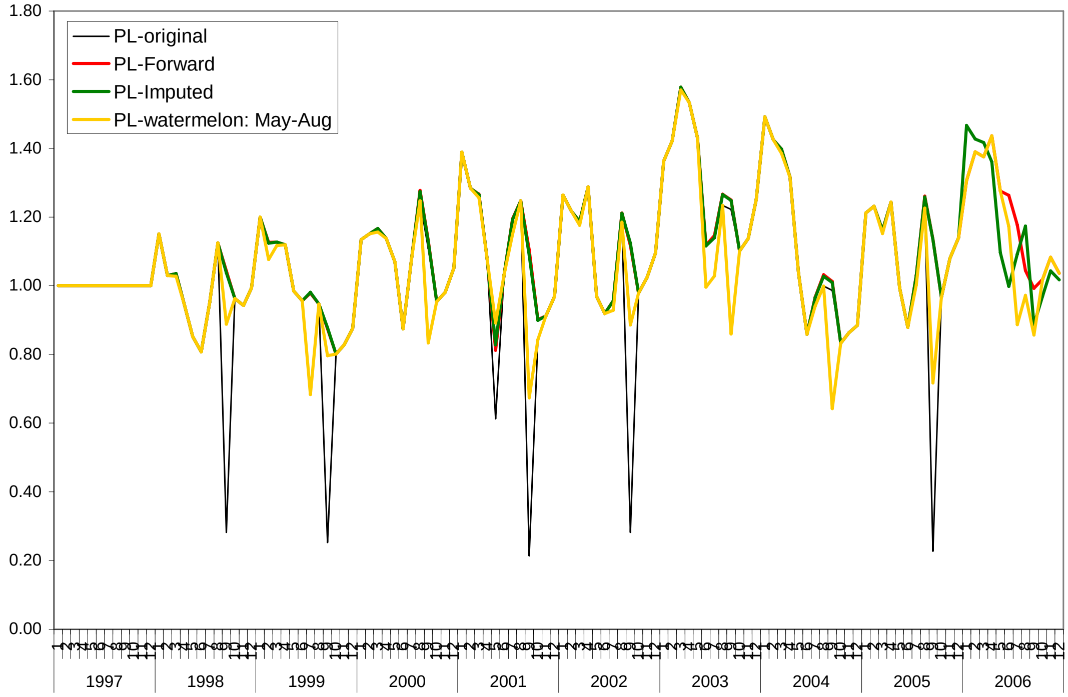
Month \ Year	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	4.06*	4.05*	4.4*	4.73*	4.78*	0
5	3.65	3.34	2.47	2.74	2.69*	2.89	4.4	2.4	2.92	3.38
6	2.03	1.67	1.7	1.65	2.21	1.99	2.08	1.67	1.87	2.21
7	1.56	1.57	1.4	1.76	1.97	1.6	1.7	1.57	1.79	1.9
8	1.46	1.74	1.42	1.93	1.96	1.91	2	1.62	2.01	1.83
9	1.56	0	0	1.93	0	0	2.17	1.79	0	0
10	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0

*The price did not participate in the calculation of the price index in the original CPI.

Instead of building a perfectly-modified dataset where all items appear and disappear on the same months every year, we have only made changes to the watermelon item: we made it enter the market in May and to exit the market in August (instead of September). One missing observation in May 2001 was inserted, and four September observations were removed. These are the same modifications that were made in the perfectly-modified dataset with regard to the watermelon commodity, except for the fact that in May 2001 the price used was that of June.

Figure 10 presents the Laspeyres price index, which looked so problematic in Figure 1, with only the watermelon observations modified. A simple imputation method when whole-group inflation is assumed for the missing prices or carrying forward the last prices method[1] could have "stabilized" the price index in the same way as the watermelon modification did.

Figure 10: Year Over Year Monthly Fixed Base Regular Laspeyres, Carry Forward and Imputed Price Indices, and Regular Laspeyres for the Watermelon-modified dataset



It can be seen, that the problem of extreme downfalls in the fixed base Laspeyres Index (which uses base year expenditure shares) in Figure 10 arises solely because in the base year, 1997, watermelon prices participated in the calculation of the price index, while in 1998, and in every other year in which the index fell (1998, 1999, 2001, 2002, 2005, 2006), they were not. Therefore, one has to be very cautious when deciding about the base year shares.

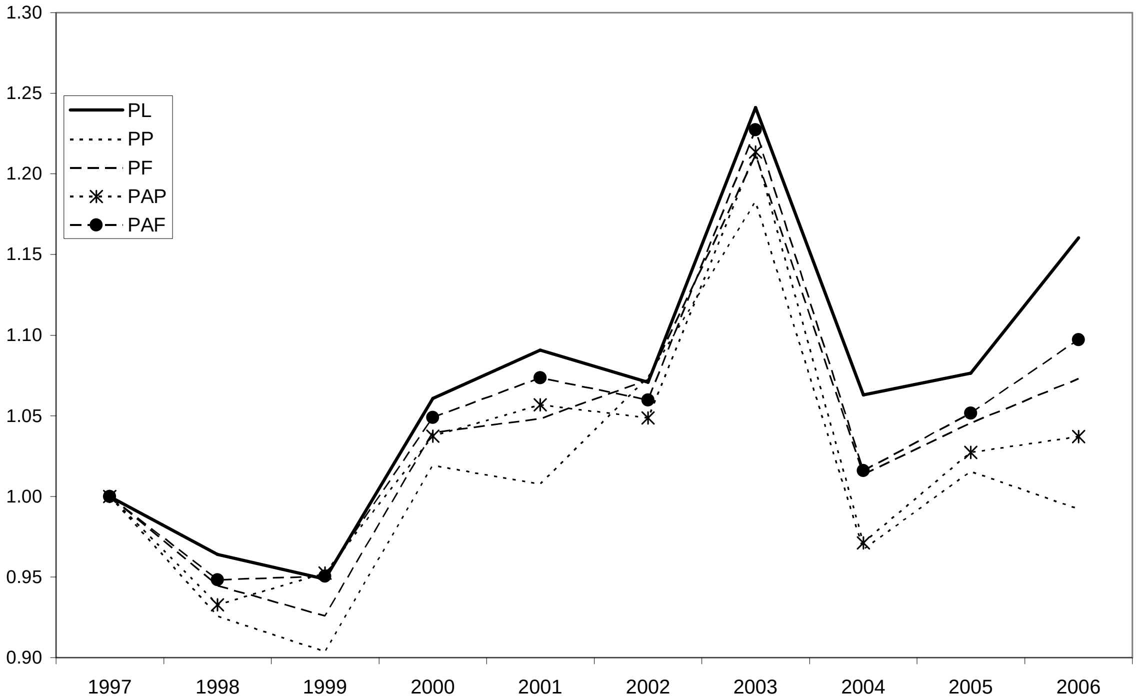
Annual Indices, presented in Figure 2, with their poor approximation, are also "spoiled" by the watermelon inconsistency in market cycle: Figure 11 presents the "corrected" version, where only comparable months were used (i.e. June to August only). Excluding 1999 and 2001, when the Paasche approximate index moves up and away from its true counterpart, all approximated indices follow close the true ones, with the Laspeyres being the highest of all.

Recall Tables 1-3, which compare true yearly month-to-month indices with their approximations. "Smoothing" the data in order to increase the comparability of prices performed in the 2004 paper produces better approximation of month-to-month indices (in accordance to the Manual), both in fixed base and chained indices for the years 1998-2002. However, extending the dataset reveals that this is true only for the

[1] Yearly month-to-month indices with imputed or carried-forward prices were not analyzed in Chapter 22 of the Manual, neither in the 2004 paper. We introduce this possibility here as a method to deal with problematic data.

fixed-base Fisher index, with chained indices still being approximated poorly[2]. It is interesting to note that removing watermelon observations from September instead of modifying the whole dataset improves slightly the approximation of chained indices.

Figure 11: Annual Fixed base Laspeyres, Paasche and Fisher Indices, and their Approximate counterparts: watermelon measured from June to August



Recall that for the original dataset, the spread between Laspeyres and Paasche month-to-month indices (Table 3) is small to negative. This spread is much higher for chained indices in the modified dataset[3].

One of the interesting findings in Section III were Rolling Year Indices, presented in Figures 3-4: approximate chained indices of Laspeyres and Paasche showed huge spreads that started around 2000 and widened further. For the perfectly modified dataset, as well as in the watermelon-modified dataset, this spread is minor for approximate chained indices (Figure 13), but is significant specifically for the true chained indices (Figure 12), with Laspeyres index being the highest, and Paasche – the lowest. This watermelon example reveals the high sensitivity of chained Laspeyres and Pasche indices and suggests that the Fisher formula is preferable.

[2] In the perfectly-modified dataset, for the fixed-base indices, only 6 observations out of 108 (for the years 1998-2006) showed a difference of more than 5% between the true and approximate Fisher indices. For chained indices, 45 observations had a spread of 5% or more. For the watermelon-modified dataset, 12 observations were significantly different for the fixed indices, and 33 observations for the chained indices.

[3] The ratio between Laspeyres and Paasche Indices being 1.033 for Fixed-base indices and 1.175 for chained indices in perfectly-modified dataset. The results for watermelon-modified dataset are 1.015 and 1.106 respectively.

Figure 12: Rolling Year Fixed Base and Chained Laspeyres, Paasche, and Fisher Indices with watermelon-modified dataset

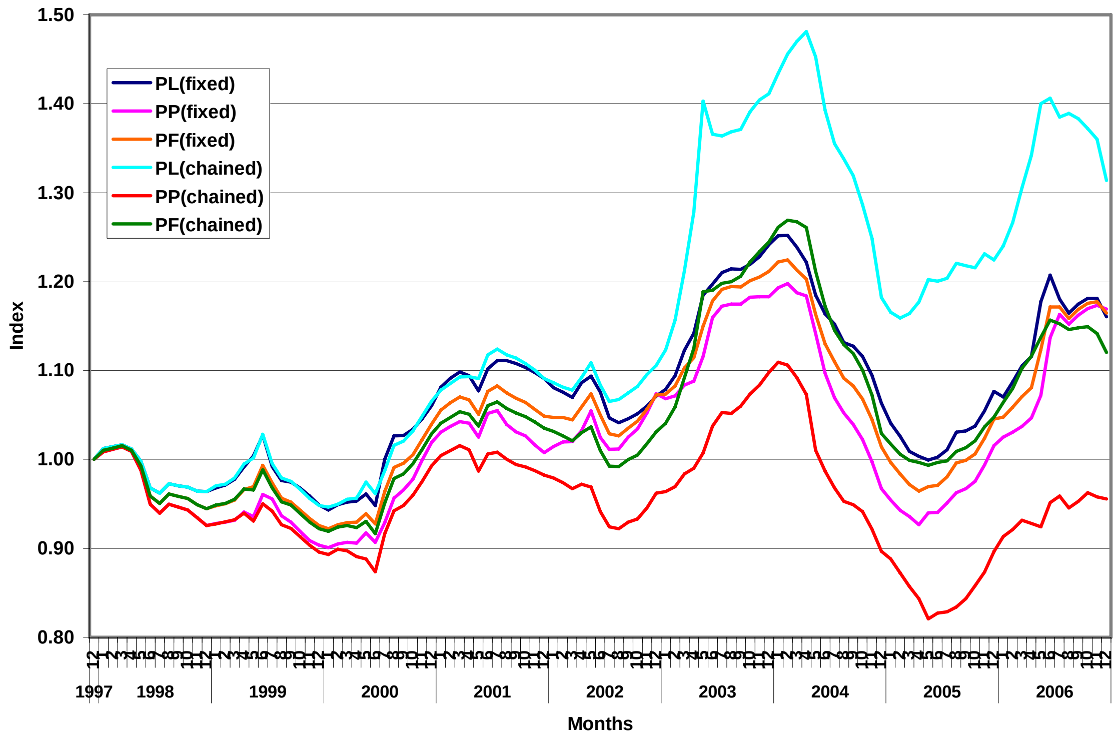
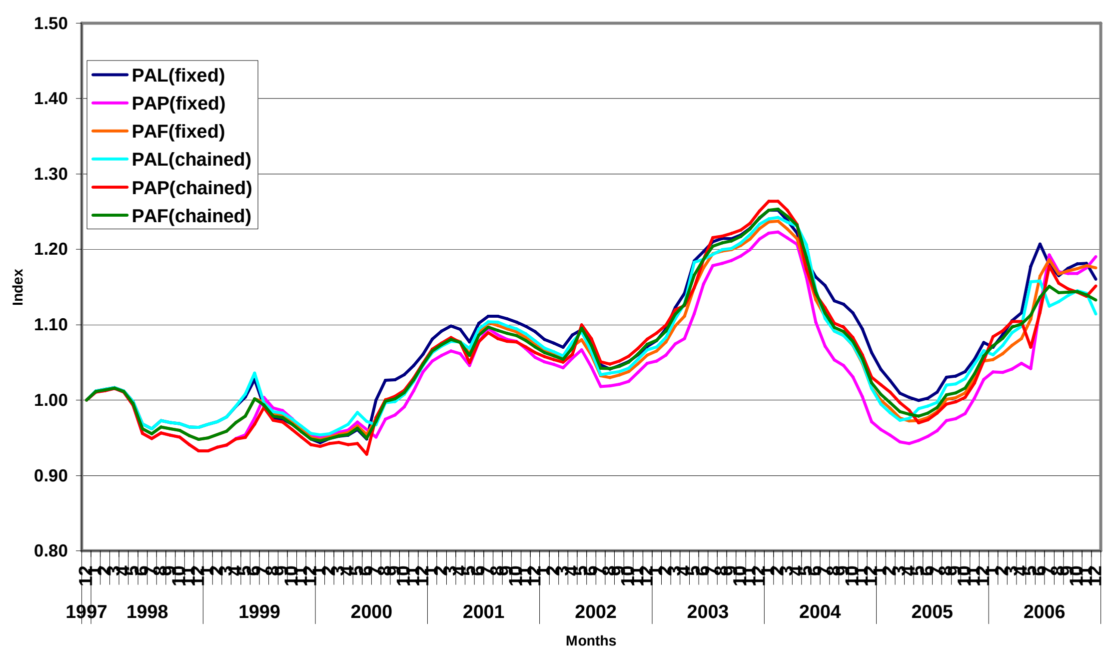


Figure 13: Rolling Year Approximate Laspeyres, Paasche, and Fisher Price Indices

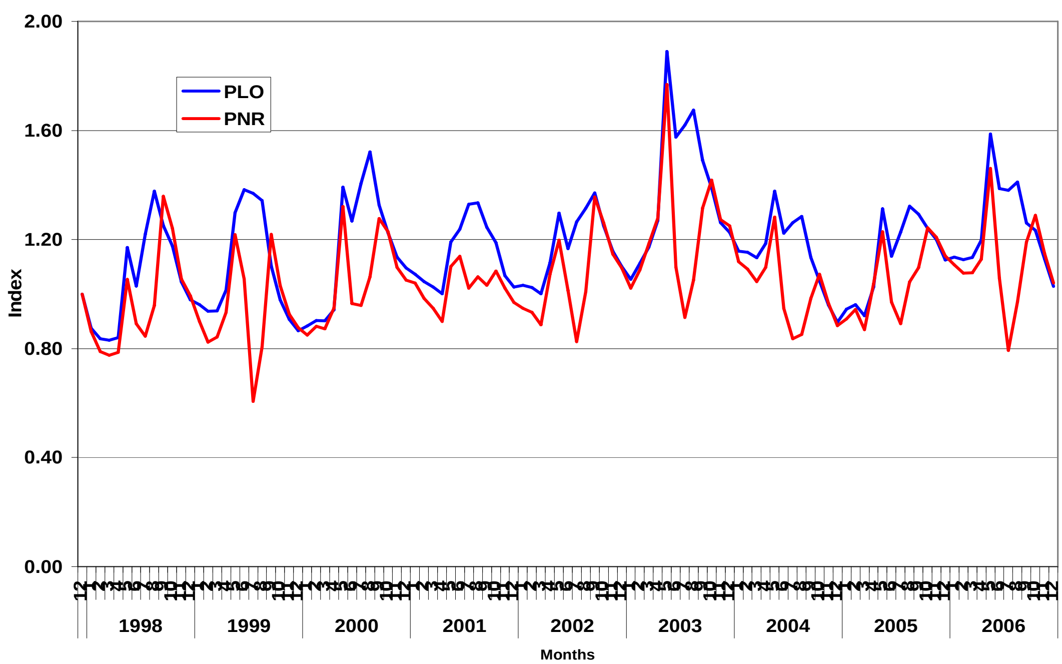


The comparison between Lowe and Rothwell Indices in the perfectly-modified dataset yields results that differ to those presented in Table 4: now the Rothwell index has the lower mean and standard deviation. This strengthens the findings of the 2004 paper. In fact, opposite results could be obtained merely by adjusting the period of watermelon presence in the market.

Table 6: Mean and Standard Deviation of Normalized Lowe and Rothwell Indices, for the perfectly-modified and watermelon-modified dataset

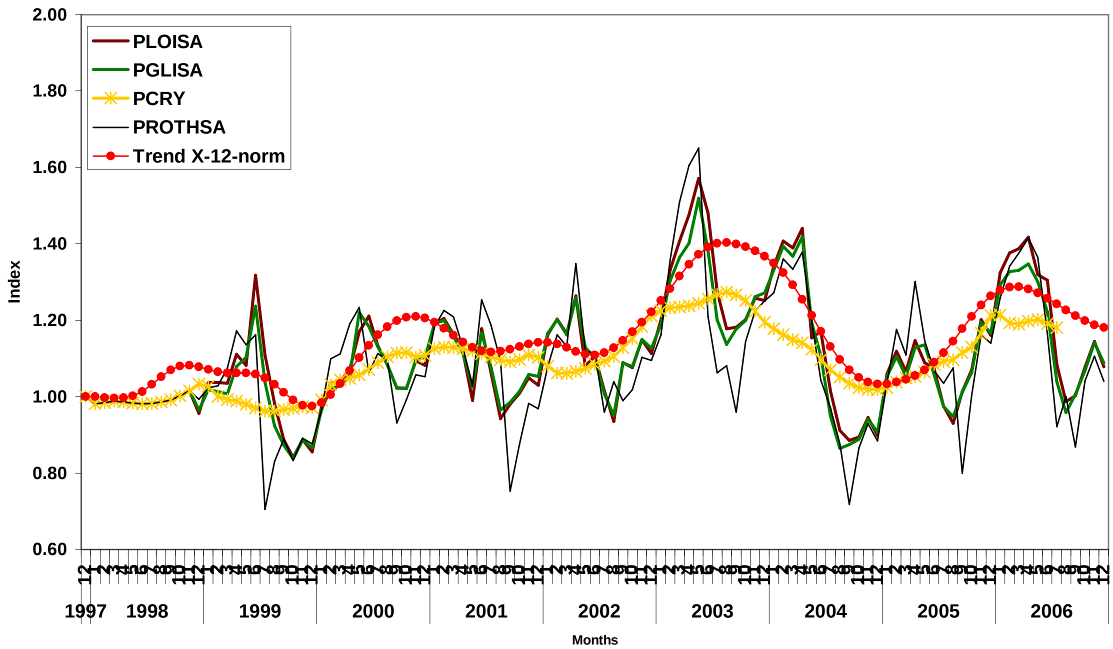
	Lowe 1998-2006	Rothwell 1998-2006
Perfectly-modified dataset		
Mean	1.1559	1.0613
St. Dev.	0.1894	0.1735
Watermelon-modified dataset		
Mean	1.1731	1.0476
St. Dev.	0.1953	0.1724

Figure 14: The Lowe with Carry Forward Prices and Normalized Rothwell in watermelon-modified database



The seasonally adjusted annual basket indices with imputed prices are less successful in repeating the trend of the target Centered Rolling Year Index, although the Rothwell Index does not show peaks as extreme as with the original dataset.

Figure 15: Seasonally Adjusted Lowe and Geometric Laspeyres Indices with Imputed Prices, Centered Rolling Year Indices, Rothwell Seasonally Adjusted Index for the watermelon-modified dataset and Seasonally and Trend adjusted ARIMA 12 Index for the original dataset



The troughs of the Rothwell index in Figure 7 that happen in September are also explained by the watermelon problem. This is the index that reflects changes in prices as well as quantities, when the base year quantities and (unit value) prices are used. Zero prices and quantities in the years 1998, 1999, 2002, 2002, 2005 and 2006, as in the case with yearly month-to-month indices, caused the fall of the price index. As a consequence, the seasonally adjusted Rothwell index, introduced in Figure 9, which is the original Rothwell multiplied by the Seasonal Adjustment Factor (SAF), reflects extreme peaks in the years when the watermelon did appear in the market: 2001, 2003 and 2004. The SAF, based on 1997, contains the price of the commodity actually sold in September of 1997. Each time the commodity was sold again, its already high price index was multiplied again by the high SAF.

The sensitivity of the price index to the timing of the appearance of a certain item raises the question – how should one decide whether to include a particular commodity in one month or another? Should we use it in the index calculation from the moment of first observation, no matter its price or the quantity sold?

In Israel, during some years, the watermelon appeared in April and remained in the market as late as October. Every time the commodity reappears, considerations are made whether to include it in the calculations or not according to the price level at which it enters the market, the relative price of the commodity in comparison to the previous months / years in the beginning of the season and at its end, etc… The most important factor in the decision is the relation between the current price and the price at the last month of calculation. This is because the strong seasonally products can shock the price index, artificially, without actually reflecting the "true" price changes.

Apart from price related considerations, there is a formal approach to defining the minimum number of observations that would be representative enough.

We start by defining the months in which each observation can be measured. For this purpose, the three-year average of traded quantities of each commodity is retrieved for each month from the agricultural databases (Table 7).

Table 7: Watermelons sold on the Israeli market in the years 2001-2003 (kilograms)

Month	2	3	4	5	6	7	8	9	Average
Average	1,372	6,631	18,251	34,239	43,813	36,635	13,680	4,891	19,939
2001	478	2,722	17,618	36,198	37,592	42,862	16,214	7,222	20,113
2002	3,311	14,821	28,149	41,952	47,534	21,846	2,644	2,528	20,348
2003	328	2,351	8,985	24,567	46,312	45,196	22,182	4,924	19,356

If the three-year average quantity is at least 30% of the total monthly average, the month is set to be a "reference month". Hence, the reference period for including the watermelon items in the calculations is March to September.

In the second stage, we calculate the basic average number of observations in the price index survey for each of the reference months. The decision whether to include the calculation of the specific item in the price index is based on whether the number of price observations exceeds 70% of the basic average.

The number of observations that appeared in the CPI on “border” months are listed in Table 8 (only the data for the years 2001-2006 are available).

Table 8: Number of observations for the watermelon commodity in the Israeli CPI:

	April	May	Measured in the CPI	September	Measured in the CPI
2001	37	41	NO	9	NO
2002	39	41	YES	13	NO
2003	32	35	YES	19	YES
2004	37	40	YES	21	YES
2005	35	43	YES	21	NO
2006	38	43	YES	18	NO

Tables 7 and 8 show that September is the most controversial month, and most of the time it was decided not to enter the watermelon item into the index calculation. April was excluded because of the extremely high entering price, leading to a conclusion that the actual consumption of the commodity was too low to consider. However, the exclusion of May 2001 seems confusing.

We tested the influence on the price index of fresh fruits of entering the watermelon into the calculation on some month other than May, such as April or June. The true April prices are available for the years 2001-2005. We used the average April prices in these years in order to impute April prices for the rest of the years, taking into account that the watermelon prices for other months were on similar levels. The results for the yearly month-over-month Laspeyres, Approximate Fisher, seasonal adjusted Rothwell, Lowe and Centered Rolling Year Indices are presented in Figures A-E of Appendix 3. Means, standard deviations and pearson coefficients between the results of different datasets are presented in Table 9.

The most striking result is that if we include the watermelon as early as April, in most of the cases we actually obtain a lower index level than if it were entered in May or June. Looking at the prices data (Appendix 1) may offer an explanation: in May or June apricots enter the seasonal cycle, and in February, persimmons exit it. Stretching the participation of watermelons simply smoothes the transfer from one seasonal commodity to another.

Table 9: Means, Standard deviations and Pearson coefficients for yearly month-over-month Laspeyres, Approximate Fisher, seasonal adjusted Rothwell, Lowe and Centered Rolling Year Indices

	PL	PAF	PLOISA	PROTHSA	PCRY
mean	1.0517	1.0739	1.0735	1.1321	1.0676
st. dev	0.2543	0.1793	0.1476	0.4716	0.0961
April-to-August watermelon included					
mean	1.0678	1.0632	1.0998	1.0834	1.0919
st. dev	0.1934	0.1812	0.1422	0.1650	0.0853
Pearson with the original data	0.8624	0.9664	0.9598	0.1440	0.9693
Pearson with May-to-August	0.9989	0.9987	0.9611	0.9975	1.0000
May-to-August watermelon included					
mean	1.0701	1.0655	1.1092	1.0865	1.0929
st. dev	0.1965	0.1841	0.1524	0.1702	0.0863
Pearson with the original data	0.8627	0.9681	0.9841	0.1473	0.9683
June-to-August watermelon included					
mean	1.0824	1.0779	1.1212	1.1043	1.0992
st. dev	0.2054	0.1899	0.1643	0.1971	0.0855
Pearson with the original data	0.8360	0.9338	0.9133	0.1336	0.9664
Pearson with May-to-August	0.9709	0.9655	0.9515	0.9476	0.9980

V. Summary and Conclusion

Chapter 22 of the CPI Manual introduces various methods how to deal with seasonal products, using an artificial dataset where seasonal cycles of each product repeat every year. Using real market datas on strong seasonality products shows that if seasonal cycles are not stable, some of the methods fail to produce results similar to those of Chapter 22 and contradict its conclusions. Fixing base-year quantity or expenditure basket may undermine the calculation of the index since for the real market data, especially for seasonal products, one year is never like the other. A single problematic item (in our case, the watermelon) may alter the results.

Each dataset should be studied alone in order to retrieve the specific drops, problematic commodities, price and expenditure patterns, etc. Only then can one conclude what methods would be most useful for it. Hence, more empirical research on seasonal products is necessary.

Bibliography

Alterman, W.F., W.E. Diewert and R.C. Feenstra, (1999), *International Trade Price Indexes and Seasonal Commodities*, Bureau of Labor Statistics, Washington D.C.

Baldwin, A. (1990), “Seasonal Baskets in Consumer Price Indexes”, *Journal of Official Statistics* 6:3, 251-273.

Balk, B.M. (1980a), Seasonal Products in Agriculture and Horticulture and Methods for Computing Price Indices, *Statistical Studies* no. 24, The Hague: Netherlands Central Bureau of Statistics.

Balk, B.M. (1980b), “Seasonal Commodities and the Construction of Annual and Monthly Price Indexes”, *Statistische Hefte* 21:2, 110-116.

Balk, B.M. (1980c), “A Method for Constructing Price Indices for Seasonal Commodities”, *Journal of the Royal Statistical Society A* 143, 68-75.

Balk, B.M. (1981), “A Simple Method for Constructing Price Indices for Seasonal Commodities”, *Statistische Hefte* 22, 72-78.

Bean, L. H. and O. C. Stine (1924), “Four Types of Index Numbers of Farm Prices”, *Journal of the American Statistical Association* 19, 30-35.

Burck L. and E. Salama (2003). “Seasonal and Prior Adjustment Factors for 2003”, *Current Statistics* no. 7, 2003, Central Bureau of Statistics, Israel

Dalén, J. (1992), “Computing Elementary Aggregates in the Swedish Consumer Price Index”, *Journal of Official Statistics* 8, 129-147.

Diewert, W.E. (1983c), “The Treatment of Seasonality in a Cost of Living Index”, pp. 1019-1045 in *Price Level Measurement*, W.E. Diewert and C. Montmarquette (eds.), Ottawa: Statistics Canada.

Diewert, W.E. (1995a), “Axiomatic and Economic Approaches to Elementary Price Indexes”, Discussion Paper No. 95-01, Department of Economics, University of British Columbia, Vancouver, Canada.

Diewert, W.E. (1996b), “Seasonal Commodities, High Inflation and Index Number Theory”, Discussion Paper 96-06, Department of Economics, University of British Columbia, Vancouver, Canada, V6T 1Z1.

Diewert, W.E. (1998b), “High Inflation, Seasonal Commodities and Annual Index Numbers”, *Macroeconomic Dynamics* 2, 456-471.

Diewert, W.E. (1999a), “Index Number Approaches to Seasonal Adjustment”, *Macroeconomic Dynamics* 3, 1-21.

Diewert, W.E. (2002c), “Harmonized Indexes of Consumer Prices: Their Conceptual Foundations”, *Swiss Journal of Economics and Statistics* 138:4, 547-637.

Diewert, W.E. and P. Armknecht (2003), “The Treatment of Seasonal Products”, Producer Price Index (PPI) Manual, Ch. 22.

Diewert, W.E., Y. Finkel and Y. Artsev (2004), “On the Treatment of Seasonal Commodities in CPI: The Israeli Experience”, paper presented at the SSHRC International Conference on Index Number Theory and the Measurement of Prices and Productivity, Vancouver, June 30-July 3, 2004.

Finkel, Y. and Y. Artsev (2005), "“On the Treatment of Seasonal Commodities in CPI: The Israeli Experience”, *Statistical Journal of the United Nations* ECE 21, (2), pp. 95-113.

Fisher, I. (1922), The Making of Index Numbers, Boston: Houghton Mifflin Co.

Karshai, R. (1992), The Consumer Price Index*, Technical Publication Series* No. 60, Central Bureau of Statistics, Israel.

Lowe, J. (1823), The Present State of England in Regard to Agriculture, Trade and Finance, second edition, London: Longman, Hurst, Rees, Orme and Brown.

Rothwell, D.P. (1958), “Use of Varying Seasonal Weights in Price Index Construction”, *Journal of the American Statistical Association* 53, 66-77.

Sabag, M. and Y. Finkel (1994), “ The Israeli CPI: Alternative Computing Methods for an Item Price Index”, *Statistical Journal of the United Nations ECE* no. 11 1994, 95-118

Turvey, R. (1979), “The Treatment of Seasonal Items in Consumer Price Indices”, *Bulletin of Labour Statistics*, Fourth Quarter, International Labour Office, Geneva, 13-33.

Young, A. (1812), *An Inquiry into the Progressive Value of Money in England as Marked by the Price of Agricultural Products*, London.

Zarnowitz, V. (1961), “Index Numbers and the Seasonality of Quantities and Prices”, pp. 233-304 in *The Price Statistics of the Federal Government*, G.J. Stigler (Chairman), New York: National Bureau of Economic Research.

Appendix 1: Prices and Expenditures of fresh fruits, 1997-2006

Table A: Prices, p_n^m

Year, n	Month, m	Lemons	Apricots	Avocado	Watermelon	Persimmon	Grapefruit	Bananas
1997	1	3.42	0	3.42	0	5.81	2.81	3.79
	2	3.34	0	3.71	0	5.81	2.74	3.88
	3	3.43	0	3.78	0	6.67	2.78	3.76
	4	3.89	0	4.03	0	0	2.9	4.24
	5	4.35	0	5.07	3.65	0	2.81	5.39
	6	6.76	8.81	6.44	2.03	0	3.01	6.77
	7	7.7	8.01	7.25	1.56	0	3.41	9.73
	8	9.15	0	0	1.46	0	3.63	9.43
	9	8.36	0	7.65	1.56	0	4.48	7.57
	10	6.47	0	5.65	0	6.7	4.31	7
	11	4.79	0	4.35	0	5.34	3.61	6.74
	12	3.9	0	3.95	0	5.44	2.9	5.86
1998	1	3.51	0	3.82	0	5.75	2.69	4.49
	2	3.45	0	3.72	0	5.88	2.42	4.09
	3	3.42	0	3.78	0	0	2.46	4
	4	3.68	0	3.98	0	0	2.57	3.98
	5	4.19	0	4.6	3.34	0	2.95	3.97
	6	5.9	6.11	5.18	1.67	0	3.39	5.01
	7	6.38	6.64	5.81	1.57	0	3.77	7.12
	8	7.39	0	9.16	1.74	0	3.7	7.52
	9	7.58	0	8.79	0	0	4.74	5.88
	10	7.06	0	6.34	0	7.85	4.36	5.71
	11	5.88	0	5.44	0	6.71	4.16	4.76
	12	4.94	0	5.71	0	7.1	3.37	4.19
1999	1	4.55	0	6.28	0	7.61	3.2	3.89
	2	4.22	0	6.14	0	0	3	3.75
	3	4.17	0	6.5	0	0	3.05	3.67
	4	4.62	0	7.56	0	0	3.22	4.16
	5	5.47	0	10.58	2.47	0	3.45	5.07
	6	7	8.82	13.66	1.7	0	3.88	6.14
	7	7.88	0	0	1.4	0	0	6.36
	8	7.96	0	0	1.42	0	4.13	5.94
	9	7.19	0	8.42	0	0	4.4	4.69
	10	5.68	0	5.84	0	7.5	4.33	4.3
	11	4.85	0	4.95	0	6.23	3.73	4.06
	12	4.32	0	4.64	0	6.41	3.4	3.81
2000	1	4.06	0	4.56	0	7.14	3.29	4.07
	2	3.83	0	4.35	0	7.66	3.19	4.4
	3	3.69	0	3.85	0	0	3.1	4.58
	4	3.49	0	3.67	0	0	3.27	5.13
	5	4.24	0	4.48	2.74	0	3.44	7.58
	6	5.7	7.32	5.56	1.65	0	3.87	7.58
	7	8.15	7.61	6.55	1.76	0	4.24	8.12

Year, n	Month, m	Lemons	Apricots	Avocado	Watermelon	Persimmon	Grapefruit	Bananas
	8	10.92	0	9.04	1.93	0	0	7.85
	9	7.84	0	9.26	1.93	0	0	6.12
	10	6.18	0	6.55	0	7.41	5.65	5.83
	11	5.3	0	5.09	0	6.11	4.26	5.71
	12	4.65	0	4.93	0	6.02	3.73	5.49
2001	1	4.15	0	5.03	0	6.35	3.41	5.33
	2	3.86	0	4.86	0	7.01	3.19	5.11
	3	3.7	0	5.04	0	0	3.17	4.84
	4	3.91	0	5.14	0	0	3.32	4.45
	5	4.4	0	6.73	0	0	3.59	4.66
	6	5.78	8.45	8.33	2.21	0	3.75	5.31
	7	6.46	8.86	0	1.97	0	4.66	6.56
	8	6.69	0	0	1.96	0	5.69	6.42
	9	5.62	0	8.88	0	0	0	5.42
	10	5.21	0	6.69	0	7.77	0	5.4
	11	4.57	0	4.97	0	6.75	4.12	4.91
	12	4.31	0	4.75	0	6.82	3.9	4.56
2002	1	4.1	0	4.97	0	7.15	3.56	4.65
	2	3.91	0	4.62	0	7.76	3.48	4.67
	3	3.67	0	4.32	0	0	3.44	4.54
	4	3.94	0	4.7	0	0	3.64	5.72
	5	4.05	10.6	4.74	2.89	0	3.75	5.94
	6	4.21	6.46	5.07	1.99	0	4	6.2
	7	5.84	6.51	0	1.6	0	3.83	7.81
	8	6.58	0	0	1.91	0	0	7.64
	9	6.19	0	9.61	0	0	5.69	6.8
	10	5.48	0	6.32	0	7.93	5.11	6.52
	11	4.8	0	6.22	0	6.28	4.23	5.84
	12	4.22	0	6.33	0	5.91	3.76	5.36
2003	1	3.91	0	6.78	0	5.65	3.69	4.77
	2	3.71	0	7.71	0	6.32	3.44	5.25
	3	3.65	0	9	0	0	3.53	5.58
	4	3.79	0	11.85	0	0	3.77	5.88
	5	4.36	0	16.63	4.4	0	4.01	6.98
	6	5.41	9.92	0	2.08	0	4.75	7.1
	7	6.38	9.77	0	1.7	0	0	8.18
	8	6.1	0	0	2	0	0	8.33
	9	6.58	0	11.44	2.17	0	0	7.5
	10	6.7	0	9.21	0	8.18	5.2	6.69
	11	6.16	0	7.27	0	7.44	4.25	6.04
	12	5.25	0	7.23	0	7.57	3.82	5.87
2004	1	4.62	0	7.54	0	8.45	3.56	5.07
	2	4.38	0	7.2	0	8.74	3.45	5.17
	3	4.15	0	7.39	0	0	3.88	4.84

Year, n	Month, m	Lemons	Apricots	Avocado	Watermelon	Persimmon	Grapefruit	Bananas
	4	4.9	0	8.01	0	0	4.06	5.1
	5	5.65	11.22	10.38	2.4	0	4.38	5.65
	6	5.76	5.53	11.12	1.67	0	4.54	5.3
	7	5.72	5.65	13.61	1.57	0	0	5.24
	8	5.68	0	13.79	1.62	0	0	5.38
	9	5.4	0	8.95	1.79	0	0	5.02
	10	5.15	0	6.53	0	6.98	5.19	4.78
	11	4.95	0	5.23	0	5.93	4.79	4.27
	12	4.75	0	4.74	0	5.91	4.39	3.76
2005	1	4.34	0	5.01	0	6.17	4.2	4.36
	2	4.24	0	5.36	0	6.64	4.12	4.45
	3	4.01	0	5.18	0	0	4.11	4.07
	4	3.71	0	5.58	0	0	4.37	5.25
	5	3.82	9.75	6.03	2.92	0	4.73	5.73
	6	3.8	5.61	6.61	1.87	0	4.99	5.72
	7	4.42	5.84	7.64	1.79	0	0	6.49
	8	5.67	0	0	2.01	0	0	6.94
	9	5.73	0	9.86	0	0	0	5.93
	10	5.59	0	8.18	0	8.03	5.42	5.63
	11	5.13	0	7.41	0	6.84	5.07	5.62
	12	4.53	0	7.28	0	6.74	4.64	4.96
2006	1	4.36	0	8.59	0	7.17	4.31	4.77
	2	4.14	0	9.09	0	7.54	4.13	4.55
	3	4.01	0	9.46	0	0	4.3	4.55
	4	3.84	0	10.85	0	0	4.62	4.89
	5	4.06	12.01	14.15	3.38	0	5.19	5.46
	6	4.88	6.73	0	2.21	0	5.63	5.39
	7	6.97	0	0	1.9	0	0	5.44
	8	7.55	0	0	1.83	0	0	5.83
	9	7.36	0	8.65	0	0	0	5.65
	10	6.94	0	8.07	0	7.93	6.35	5.44
	11	6.25	0	6.56	0	6.82	5.46	4.98
	12	5.06	0	6.27	0	7.06	4.63	4.23

Table B: Expenditures, $(p_n^{t,m} * q_n^{t,m})$

Year, n	Month, m	Lemons	Apricots	Avocado	Watermelon	Persimmon	Grapefruit	Bananas
1997	1	1.7	0.1	4.3	0	1.1	0.3	17.7
	2	2.5	0	3.6	0.3	0.9	1.7	14.8
	3	1.9	0	3.7	0.7	0.2	1.4	15.3
	4	2.4	0.1	3	3.4	0	1.4	17.5
	5	2.1	0.2	2.7	11	0	1.5	11.8
	6	3.2	6.4	2.3	28.9	0	1.6	6.2
	7	2.2	7.4	1.6	27.8	0.1	0.8	1
	8	2.9	0.8	0.5	22.2	0	0.7	1.5
	9	2.8	0	0.5	13.3	0	0.5	2.4
	10	2.8	0	1.7	2.7	1.6	0.6	6.7
	11	2.4	0	3.5	0.4	3.6	1.2	13.3
	12	2.2	0.1	5.1	0.2	3.6	0.9	15
1998	1	1.7	0	3.8	0	3	0.7	14.4
	2	2.2	0	4.8	0	1.9	0.8	16.9
	3	2.6	0.1	3.8	0.6	0.7	1	17.4
	4	2.8	0.2	3.2	2.6	0.1	1.7	17.5
	5	2.6	1.1	2.8	22.2	0.1	0.7	12.5
	6	2.4	10.4	1.6	26	0	0.4	7.2
	7	3.7	6.9	1.4	23.6	0	0.7	3.2
	8	2.6	0.3	0.8	24.6	0.2	0.7	3.2
	9	2.9	0.1	1.1	11.7	0.2	1.1	4.5
	10	3.6	0.1	2.1	1.8	0.8	0.4	8.9
	11	3.2	0	4.3	0.3	3.4	1.2	13.8
	12	2.8	0	4	0.2	2.7	0.9	14.7
1999	1	2.1	0.1	4.3	0	1.5	1	16.1
	2	2.4	0.1	4.3	0.1	1.7	1.1	14.2
	3	2.1	0	4.4	0.4	0.3	1.1	15
	4	3	0.1	4	4.3	0.2	1.2	13.6
	5	3.2	2.2	2.2	21	0.3	1.7	11.5
	6	2.8	11	1.9	26.7	0	0.8	6.7
	7	3.1	6	0.4	25.7	0	0.8	4
	8	2.6	0.5	0.2	19.4	0.1	0.4	3.7
	9	2.8	0.2	1.1	9.4	0.4	0.6	6.1
	10	2.8	0	2.6	1.4	1.6	0.9	8.3
	11	2.5	0.3	5.2	0.2	3.7	1.4	12.7
	12	2.6	0	4.4	0	3.4	0.7	12.3
2000	1	2.2	0.2	3.7	0	2.9	1	11
	2	2.7	0	4.2	0	2.3	1	13.6
	3	3.1	0	3.6	0.1	0.6	1.4	12.7
	4	2.6	0	3.2	3.6	0.1	1	14.2
	5	3.1	1.2	3	18	0	1.1	8.5
	6	2.4	8.9	1.6	25.4	0.1	0.6	4.7
	7	3.2	7.1	1.6	25.7	0	0.2	2

Year, n	Month, m	Lemons	Apricots	Avocado	Watermelon	Persimmon	Grapefruit	Bananas
	8	3.8	0.4	1.1	21.3	0	0.5	2.5
	9	2.6	0.1	1.1	9.9	0.1	0.3	4.6
	10	2.9	0.1	2.5	1.8	1.5	0.6	9.4
	11	3.1	0	5	0.7	3.9	1	11.3
	12	2.6	0.3	4	0.2	3.5	0.9	13.7
2001	1	2.3	0	4.2	0	4.1	1.4	13.5
	2	2.9	0.2	3.7	0.2	2	0.9	14
	3	2.6	0.2	3.6	0.8	1.7	1.2	13.8
	4	2.9	0.1	3	5.9	0.1	0.8	13.7
	5	2.4	2.5	2.5	21.1	0	0.7	10.1
	6	2.8	10.1	2.3	23.3	0	0.9	5.8
	7	3	3.6	2.1	23.6	0	0.4	4
	8	3.3	0.1	1.1	17	0	0.3	3
	9	3.4	0.1	1.4	5.2	0.2	0.1	4.8
	10	3.7	0.2	4.1	1.8	2.4	0.6	9.4
	11	3.1	0	6.3	0.5	4.4	1	13.2
	12	2.5	0.1	5	0.5	3.4	0.8	13.5
2002	1	1.8	0.2	16	0	3.2	1.2	12.7
	2	3	0.1	15	0.5	1.9	1.3	16.4
	3	3.5	0.1	14.2	0.7	1.1	1.1	14.4
	4	3.7	0.1	14.2	9.9	0.4	1.1	14.5
	5	2.8	4	13.6	17.3	0.1	1.2	11
	6	2.8	10.6	13.1	21.6	0	0.7	6.9
	7	3.3	3.8	16	25.1	0	0.7	3.4
	8	3.9	0.3	20.5	18.4	0	0.1	3.2
	9	3.4	0.1	16.8	10.6	0.2	0.4	3.8
	10	2.8	0	17.3	1	2.1	0.4	7.3
	11	2.8	0	16.7	0.3	4.5	0.7	11.5
	12	2.8	0	21.4	0.4	5	0.9	14.7
2003	1	2	0.1	4.4	0	2.9	1.3	15
	2	2.2	0	5	0	2.9	3.1	14.4
	3	2.8	0	5.5	0.3	2	0.9	15.2
	4	2.4	0	3.5	6.9	0.2	0.9	14.8
	5	3.3	1.3	2.4	15.3	0.1	0.9	11.3
	6	3	9.6	1.1	24.6	0	0.6	6.8
	7	3.2	7.1	0.1	29	0	0.4	3.2
	8	4.4	1.2	0.2	20.6	0	0.2	2.4
	9	3.6	0.2	1.1	9.1	0.1	0.7	3.4
	10	3.4	0.1	2.2	1.4	1	0.4	7.1
	11	3.4	0	5.5	1.2	3.3	1	13.1
	12	3.7	0	6.5	0	3.3	0.8	14.5
2004	1	1.9	0	5.7	0	3.4	0.4	11.9
	2	3.1	0.1	5.9	0	2.2	0.5	15.9
	3	2.9	0	4.1	0.7	0.5	1.3	13.1

Year, n	Month, m	Lemons	Apricots	Avocado	Watermelon	Persimmon	Grapefruit	Bananas
	4	3.6	0	5.7	7.7	0.2	0.5	15.5
	5	2.8	3.2	3.7	17	0	0.5	10
	6	2.5	9.7	2.6	19.6	0.2	0.7	5.3
	7	3.3	5.2	1.6	21.3	0	0.2	3.9
	8	3.6	0.4	1.6	14.2	0	0.2	3.6
	9	3.4	0.2	1.5	8	0.2	0.1	5
	10	3.8	0.1	3.7	1.1	1.6	0.3	8.1
	11	3	0.1	5.2	1.4	3.9	0.5	11.4
	12	3.2	0	5.4	0.3	3.2	0.5	11.9
2005	1	2.2	0	3.8	0	1.4	1.9	10
	2	2.8	0	5.1	0	2.8	1.2	13.7
	3	3.9	0	4.9	0.5	1.2	1.1	13.2
	4	3.5	0	3.7	5.3	0.4	1.5	11.3
	5	2.8	3.9	3.9	12.8	0.1	1.5	10.7
	6	2.9	10.3	2.8	20	0.1	0.6	5.4
	7	2.8	4	2.2	22.5	0	0.3	3.2
	8	3.2	0.5	1.7	17.4	0.1	0	2.4
	9	4	0.2	1.8	8	0.1	0.1	5.1
	10	4.2	0.2	3	1.6	1.3	0.3	8.2
	11	3.8	0.2	4	1.3	3	0.5	10.7
	12	2.5	0	6.4	0.3	2.8	0.6	13.1
2006	1	2.70	0.21	7.38	0	3.37	1.64	14.16
	2	2.99	0.05	5.88	0.02	2.34	0.78	14.39
	3	2.44	0.08	4.55	0.49	1.13	0.78	13.66
	4	2.97	0.28	4.99	4.34	0.29	1.01	14.44
	5	2.95	2.81	3.87	15.53	0.10	1.71	11.50
	6	3.28	9.61	3.03	22.41	0.04	0.22	7.51
	7	2.40	3.89	0.97	20.96	0.00	0.32	3.02
	8	4.14	0.58	1.03	16.99	0.00	0.08	4.15
	9	4.72	0.09	2.72	8.38	0.36	0.02	4.08
	10	4.60	0.16	3.80	2.09	0.93	0.14	8.63
	11	3.76	0.05	5.40	0.83	2.95	0.83	12.91
	12	3.37	0.00	5.99	0.10	4.90	1.22	12.58

Appendix 2: Formulae for Methods of Treatment of Seasonal Products

1. Year over Year Monthly Indices

For each month m=1,2,…,12, let S(m) denote the set of products that are available for purchase in each year t=0,1,…,T. For t = 0,1,…,T and m = 1,2,…,12, let $p_n^{t,m}$ and $q_n^{t,m}$ denote the price and quantity of product n that is available in month m of year t for n belongs to S(m). Then *the year over year monthly Laspeyres, Paasche and Fisher indices* going from month m of year t to month m of year t+1, in price relative and monthly revenue share form, can be defined as follows:

$$P_L = \sum_{n\in S(m)} s_n^{t,m}\left(p_n^{t+1,m} / p_n^{t,m}\right); \qquad m=1,2,\ldots,12;$$

$$P_P = \left[\sum_{n\in S(m)} s_n^{t+1,m}\left(p_n^{t+1,m} / p_n^{t,m}\right)^{-1}\right]^{-1}; \qquad m=1,2,\ldots,12;$$

$$P_F = \sqrt{P_L P_P}\ .$$

where the monthly revenue share for product n∈S(m) for month m in year t is defined as:

$$s_n^{t,m} = \frac{p_n^{t,m} q_n^{t,m}}{\sum_{i\in S(m)} p_i^{t,m} q_i^{t,m}}; \qquad m=1,2,\ldots,12;\ n\in S(m);\ t = 0,1,\ldots,T$$

Approximate year over year monthly Laspeyres and Paasche indices are defined as follows:

$$P_{AL} = \sum_{n\in S(m)} s_n^{0,m}\left(p_n^{t+1,m} / p_n^{t,m}\right); \qquad m=1,2,\ldots,12;$$

$$P_{AP} = \left[\sum_{n\in S(m)} s_n^{0,m}\left(p_n^{t+1,m} / p_n^{t,m}\right)^{-1}\right]^{-1}; \qquad m=1,2,\ldots,12;$$

Where $s_n^{0,m}$ is the base period monthly revenue share.

2. Year over Year Annual Indices

Using the notation introduced above, *the Laspeyres and Paasche annual (chain link) indices* comparing the prices of year t with those of year t+1 can be defined as follows:

$$P_L = \sum_{m=1}^{12} \sum_{n\in S(m)} \sigma_m^t s_n^{t,m}\left(p_n^{t+1,m} / p_n^{t,m}\right);$$

$$P_P = \left[\sum_{m=1}^{12} \sum_{n\in S(m)} \sigma_m^{t+1} s_n^{t+1,m}\left(p_n^{t+1,m} / p_n^{t,m}\right)^{-1}\right]^{-1}.$$

where the revenue share for month m in year t is defined as:

$$\sigma_n^t = \frac{\sum_{n\in S(m)} p_n^{t,m} q_n^{t,m}}{\sum_{i=1}^{12} \sum_{j\in S(i)} p_j^{t,i} q_j^{t,i}}; \qquad m = 1,2,\ldots,12;\ t = 0,1,\ldots,T$$

The current year weights, $s_n^{t,m}$ and σ_m^t and $s_n^{t+1,m}$ and σ_m^{t+1} can be approximated by the corresponding base year weights, $s_n^{0,m}$ and σ_m^0.

There is no need to restrict attention to calendar year comparisons: any 12 consecutive months of price and quantity data could be compared to the price and quantity data of the base year, provided that the January data in the non-calendar year is compared to the January data of the base year, the February data of the non-calendar year is compared to the February data of the base year, …, and the December data of the non-calendar year is compared to the December data of the base year. Alterman, Diewert, and Feenstra (1999; 70) called the resulting indices *rolling year* or *moving year indices*.

3. Maximum Overlap Month to Month Price Indices

Let there be N products that are available in some month of some year and let $p_n^{t,m}$ and $q_n^{t,m}$ denote the price and quantity of product n that is in the marketplace in month m of year t (if the product is unavailable, define let $p_n^{t,m}$ and $q_n^{t,m}$ to be 0). Let $p^{t,m} \equiv [p_1^{t,m},p_2^{t,m},\ldots,p_N^{t,m}]$ and $q^{t,m} \equiv [q_1^{t,m},q_2^{t,m},\ldots,q_N^{t,m}]$ be the month m and year t price and quantity vectors respectively. Let S(t,m) be the set of products that is present in month m of year t *and the following month*.

Define the revenue shares of product n in month m and m+1 of year t, using the set of products that are present in month m of year t and the subsequent month, as follows:

$$\text{(a)}\ \ s_n^{t,m}(t,m) = \frac{p_n^{t,m} q_n^{t,m}}{\sum_{i\in S(t,m)} p_i^{t,m} q_i^{t,m}}; \qquad m=1,2,\ldots,11;\ n\in S(t,m);$$

$$\text{(b)}\ \ s_n^{t,m+1}(t,m) = \frac{p_n^{t,m+1} q_n^{t,m+1}}{\sum_{i\in S(t,m)} p_i^{t,m+1} q_i^{t,m+1}}; \qquad m=1,2,\ldots,11;\ n\in S(t,m);$$

$s_n^{t,m+1}(t,m)$ has to be distinguished from $s_n^{t,m+1}(t,m+1)$. The revenue share $s_n^{t,m+1}(t,m)$ is the share of product n in month m+1 of year t but where n is restricted to the set of products that are present in month m of year t and the subsequent month, whereas $s_n^{t,m+1}(t,m+1)$ is the share of pruduct n in month m+1 of year t but where n is restricted to the set of products that are present in month m+1 of year t and the subsequent month.

If product n is present in month m of year t and the following month, define $s_n^{t,m}(t,m)$ using (a); if this is not the case, define $s_n^{t,m}(t,m) = 0$. Similarly, if product n is present in month m of year t and the following month, define $s_n^{t,m+1}(t,m)$ using (b); is this is not the case, define $s_n^{t,m+1}(t,m) = 0$.

Using these share definitions, Laspeyres and Paasche formulae can be written in revenue share and price form as follows[4]:

[4] It is important that the revenue shares that are used in an index number formula add up to unity. The use of unadjusted expenditure shares would lead to a systematic bias in the index number formula.

$$P_L = \sum_{n \in S(t,m)} s_n^{t,m}(t,m)\left(p_n^{t+1,m} / p_n^{t,m}\right); \qquad \text{m=1,2,…,11;}$$

$$P_P = \left[\sum_{n \in S(t,m)} s_n^{t,m+1}(t,m)\left(p_n^{t+1,m} / p_n^{t,m}\right)^{-1}\right]^{-1}; \qquad \text{m=1,2,…,11;}$$

4. Annual Basket Indices

The *Lowe index* for month m is defined by the following formula:

$$P_{LO} = \frac{\sum_{n=1}^{N} p_n^m q_n}{\sum_{n=1}^{N} p_n^0 q_n}$$

where $p^0 \equiv [p_1^0,\ldots,p_N^0]$ is the price reference period price vector, $p^m \equiv [p_1^m,\ldots,p_N^m]$ is the current month m price vector and $q \equiv [q_1,\ldots,q_N]$ is the weight reference year quantity vector.

The *Young (1812) index* is defined as follows:

$$P_Y = \sum_{n=1}^{N} s_n (p_n^m / p_n^0)$$

where $s \equiv [s_1,\ldots,s_N]$ is the weight reference year vector of revenue shares.

The *geometric Laspeyres index* is defined as follows:

$$P_{GL} = \prod_{n=1}^{N} (p_n^m / p_n^0)^{s_n}$$

Thus the geometric Laspeyres index makes use of the same information as the Young index except that a geometric average of the price relatives is taken instead of arithmetic one.

It is of interest to compare the above three indices that use annual baskets to the fixed base Laspeyres rolling year indices. However, the rolling year index that ends in the current month is centered five and a half months backwards. Hence the above annual basket type indices may be compared with an arithmetic average of two rolling year indices that have their last month 5 and 6 months forward. This latter *centered rolling year index* is labeled P_{CRY} and is mentioned in Figures 5 and 7 in the paper.

5. Bean and Stine Type C or Rothwell Indices

The *Bean and Stine Type C* (1924; 31) or *Rothwell* (1958; 72*) index* makes use of seasonal baskets in the base year, denoted as the vectors $q^{0,m}$ for the months m = 1,2,…,12. The index also makes use of a vector of base year unit value prices, $p^0 \equiv [p_1^0,\ldots,p_5^0]$ where the nth price in this vector is defined as:

$$p_n^0 \equiv \frac{\sum_{m=1}^{12} p_n^{0,m} q_n^{0,m}}{\sum_{m=1}^{12} q_n^{0,m}};$$

The *Rothwell price index* for month m in year t can now be defined as follows:

$$P_R = \frac{\sum_{n=1}^{N} p_n^{t,m} q_n^{0,m}}{\sum_{n=1}^{N} p_n^{0} q_n^{0,m}}; \qquad\qquad m = 1,\ldots,12.$$

To make the different series more comparable, the *normalized Rothwell index* P_{NR} is introduced; this index is simply equal to the original Rothwell index divided by its first observation.

6. Forecasting Rolling Year Indices using Month to Month Annual Basket Indices

For each of the series, Lowe, Young and Geometric Laspeyres, a seasonal adjustment factor (SAF) is defined, as the centered rolling year index P_{CRY} divided by P_{LO}, P_Y and P_{GL}, respectively for the first 12 observations. Now for each of the three series, repeat these 12 seasonal adjustment factors for the remaining observations. These operations will create 3 SAF series for all the observations (label them SAF_{LO}, SAF_Y and SAF_{GL}, respectively).
Finally, define *seasonally adjusted Lowe, Young and Geometric Laspeyres indices* by multiplying each unadjusted index by the appropriate seasonal adjustment factor.

$$P_{LOSA} \equiv P_{LO} SAF_{LO}; P_{YSA} \equiv P_Y SAF_Y; P_{GLSA} \equiv P_{GL} SAF_{GL}.$$

A seasonally adjusted version of the Rothwell index presented in the paper may also be defined in the same way.

Appendix 3: Yearly month-over-month Laspeyres, Approximate Fisher, seasonal adjusted Rothwell, Lowe and Centered Rolling Year Indices in the original dataset, April-to-August watermelon participation, May-to-August and June-to-August

Figure A: Yearly month-over-month Laspeyres Index for the original dataset, Aprl-to-August, May-to-August and June-to-August watermelon participation

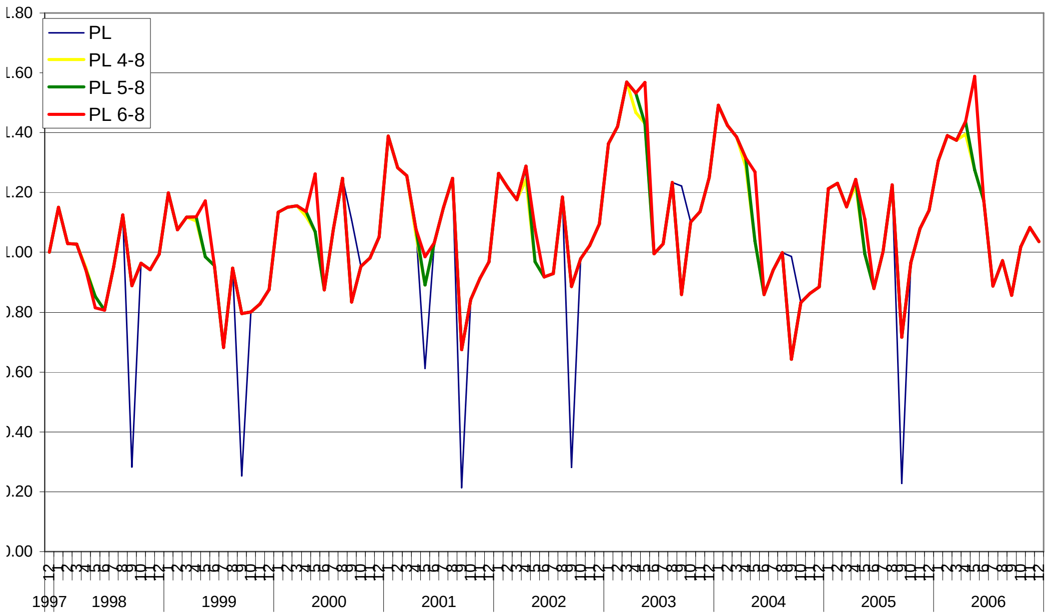


Figure B: Yearly month-over-month approsimate Fisher Index for the original dataset, Aprl-to-August, May-to-August and June-to-August watermelon participation

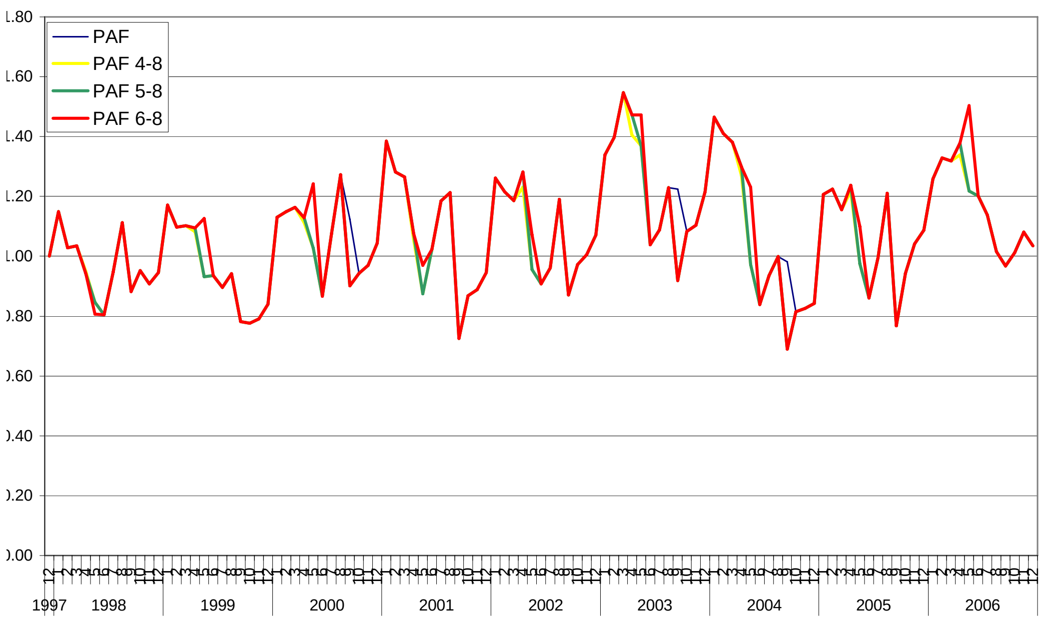


Figure C: Seasonally adjusted Lower Index for the original dataset, Aprl-to-August, May-to-August and June-to-August watermelon participation

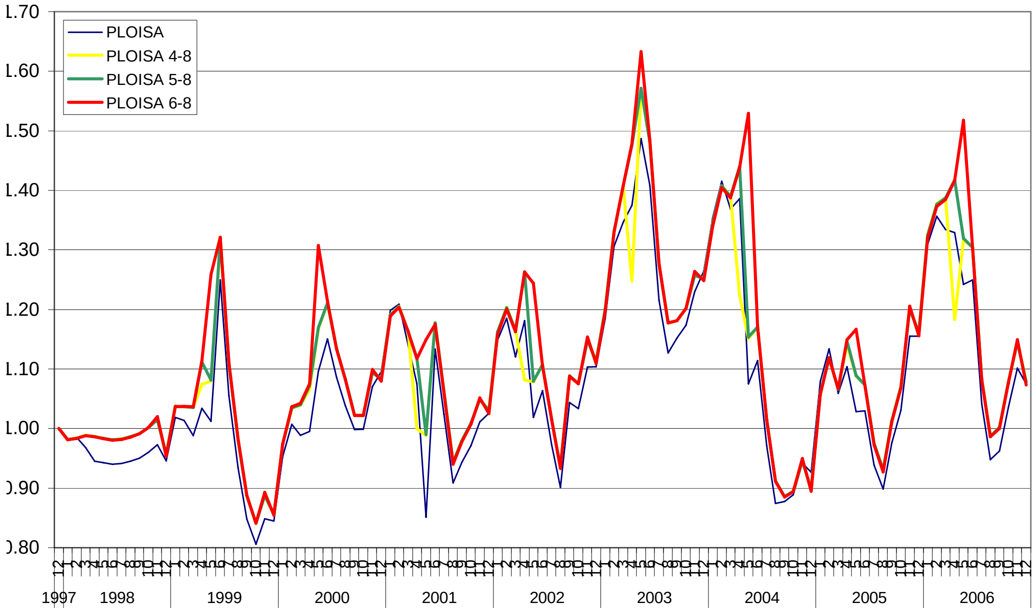


Figure D: Seasonally adjusted Rothwell Index for the original dataset, Aprl-to-August, May-to-August and June-to-August watermelon participation

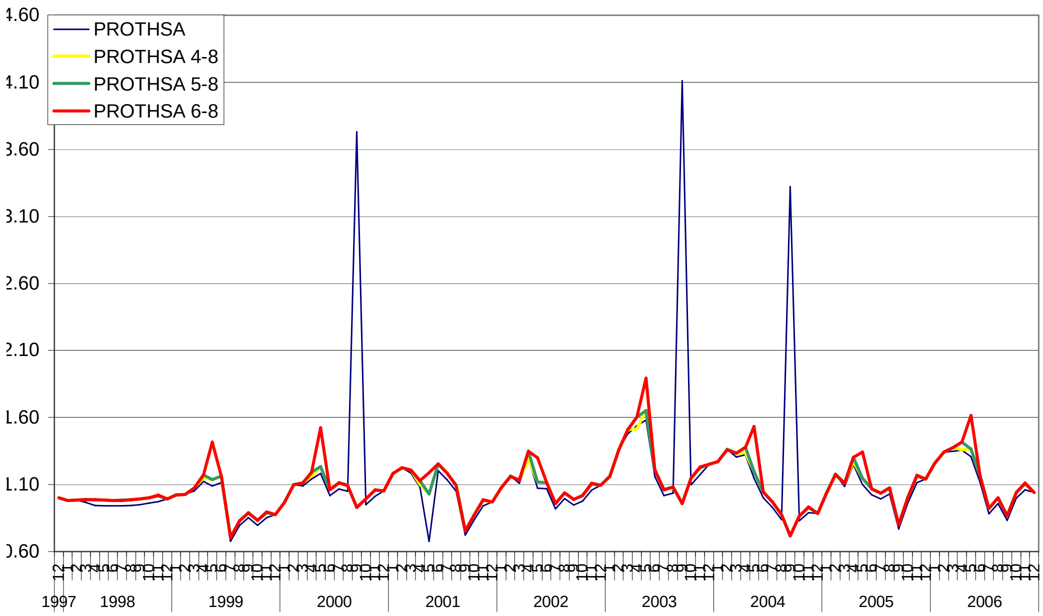


Figure E: Centered Rolling Year Index for the original dataset, Aprl-to-August, May-to-August and June-to-August watermelon participation