

**An Empirical Illustration of Index Construction
Using Israeli Data on Vegetables**

**Meeting of the Expert Group on
Consumer Price Indices
Geneva, 26-28 May 2014**

Erwin Diewert

**University of British Columbia and
University of New South Wales**

Introduction

- **This paper is a follow up to the work of Diewert, Finkel and Artsev (2009) on seasonal commodities in the Israeli CPI.**
- **Since we did our research on this topic, Rolling Year GEKS has come on to the scene as a possible solution to the chain drift problem.**
- **A major goal of this paper is to see how RYGEKS performs using seasonal data (and how badly chained Fishers perform).**
- **We will also compare the RYGEKS indexes to traditional annual basket Lowe and Young indexes using the Israeli data to see how much bias is in these traditional indexes.**
- **The Israeli data is unusual in that monthly expenditure information on seasonal commodities is collected on a continuous basis.**
- **Another major goal of this paper is to convince you to follow the example of Israel and collect expenditure information on a continuous basis.**

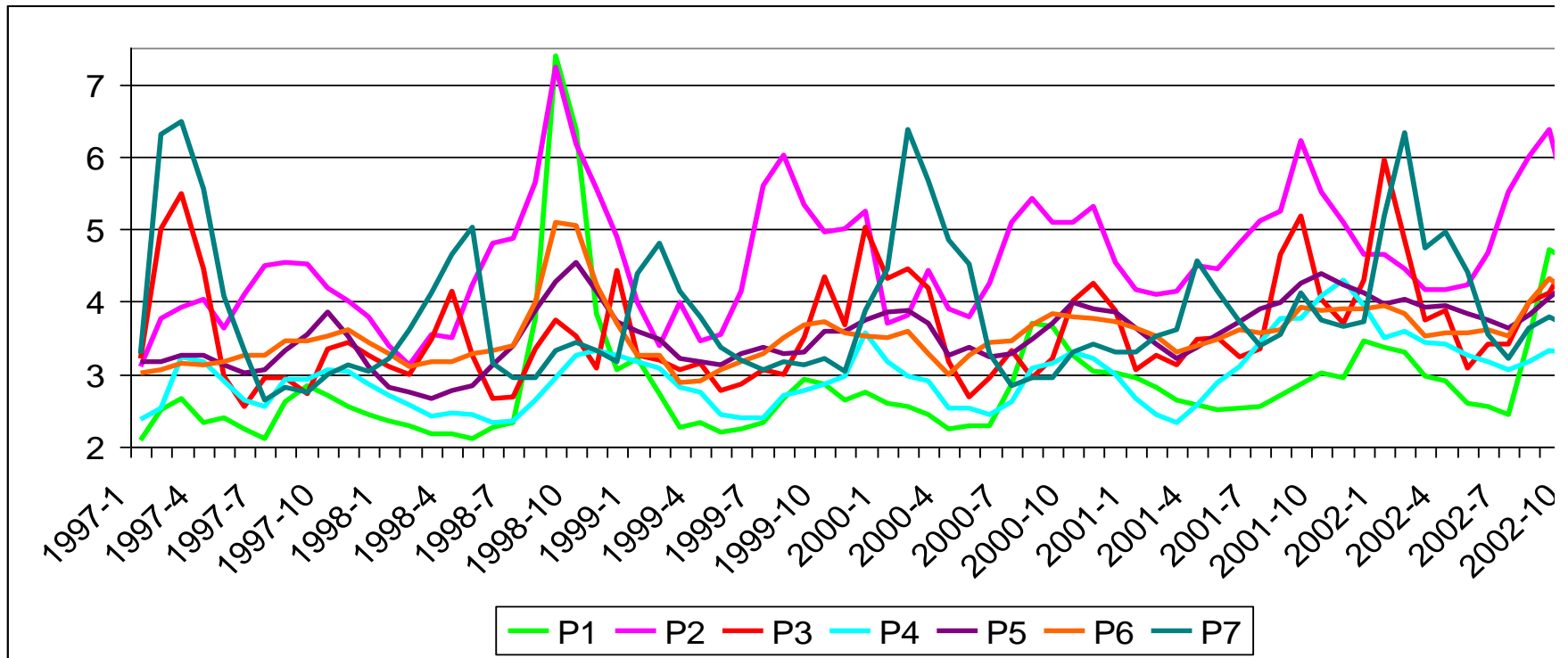
The Data

- **The price data consist of average monthly prices for 7 types of vegetable consumed by households in Israel and cover the 6 years 1997-2002.**
- **The Israeli CPI program has a continuous Household Expenditure Survey and so estimates of monthly household expenditure on the 7 vegetable groups are available.**
- **Group 1: Cabbages;**
- **Group 2: Cauliflower;**
- **Group 3: Cucumbers;**
- **Group 4: Potatoes;**
- **Group 5: Carrots;**
- **Group 6: Lettuce and**
- **Group 7: Eggplants.**

The Data (cont)

Chart 9: Monthly Prices for Seven Kinds of Vegetable, 1997:1-2002:12

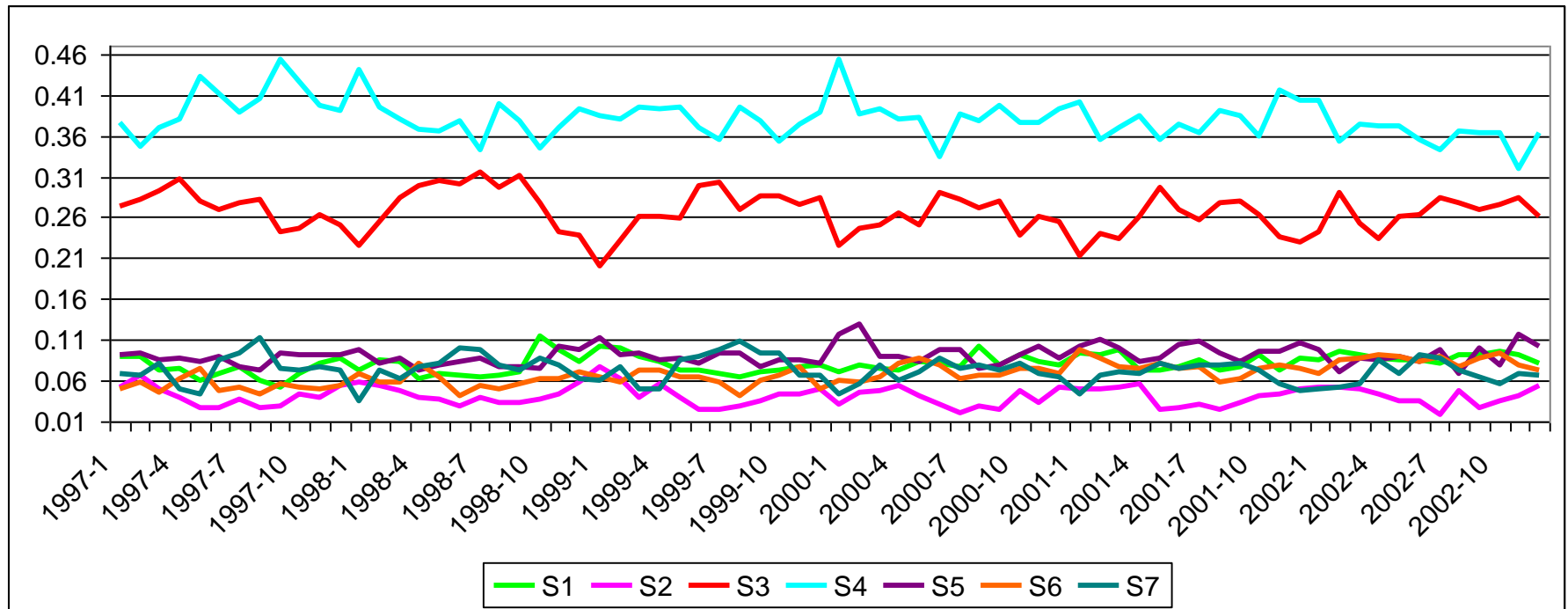
The tremendous seasonality in the price data is evident below.



The Data (cont)

Table 10: Monthly Expenditure Shares for Seven Kinds of Vegetable, 1997:1-2002:12

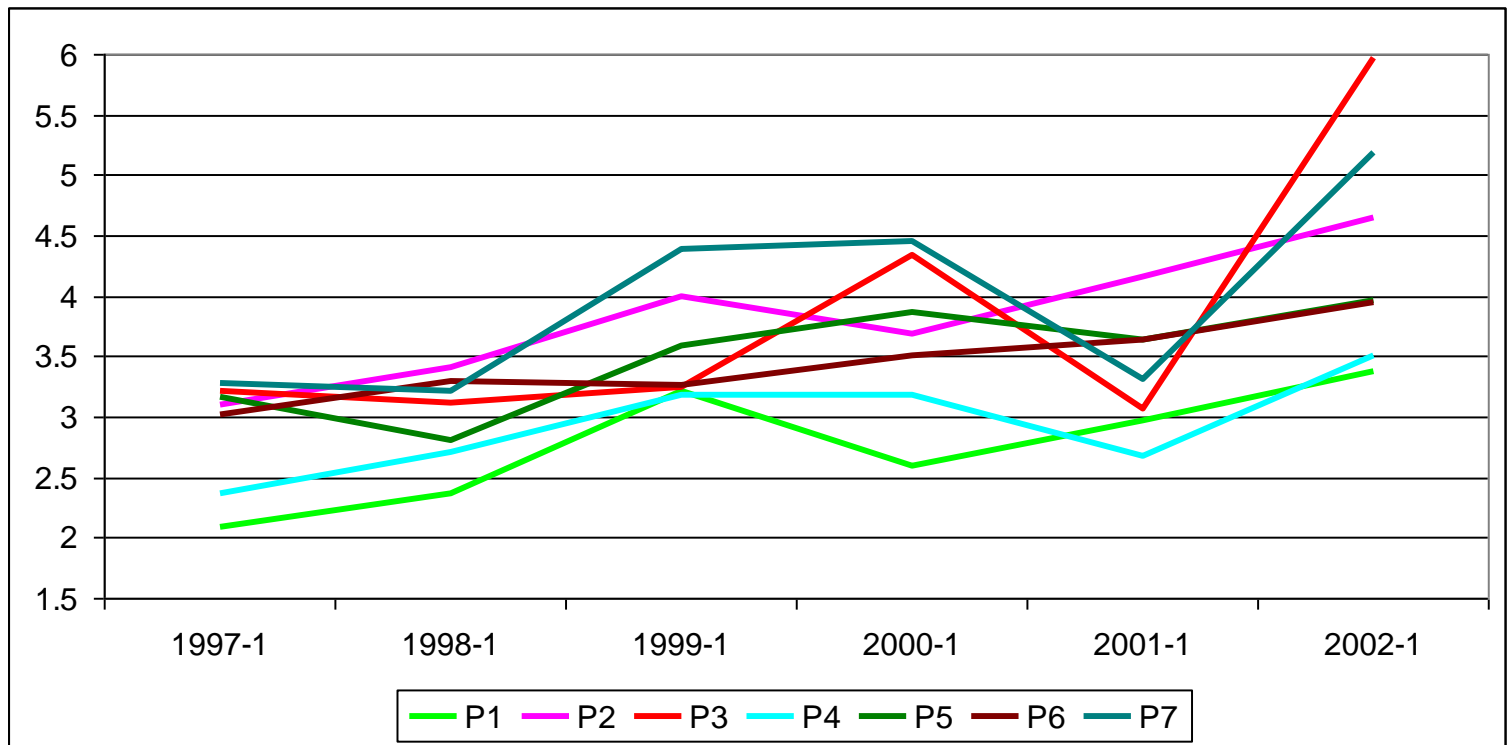
- The monthly expenditure shares are a lot smoother
- Cucumbers and potatoes are the big share items (3 and 4)



The Data (cont)

However, when we look at year over year data for the same month, the data look a lot smoother. Here are plots of the year over year January prices

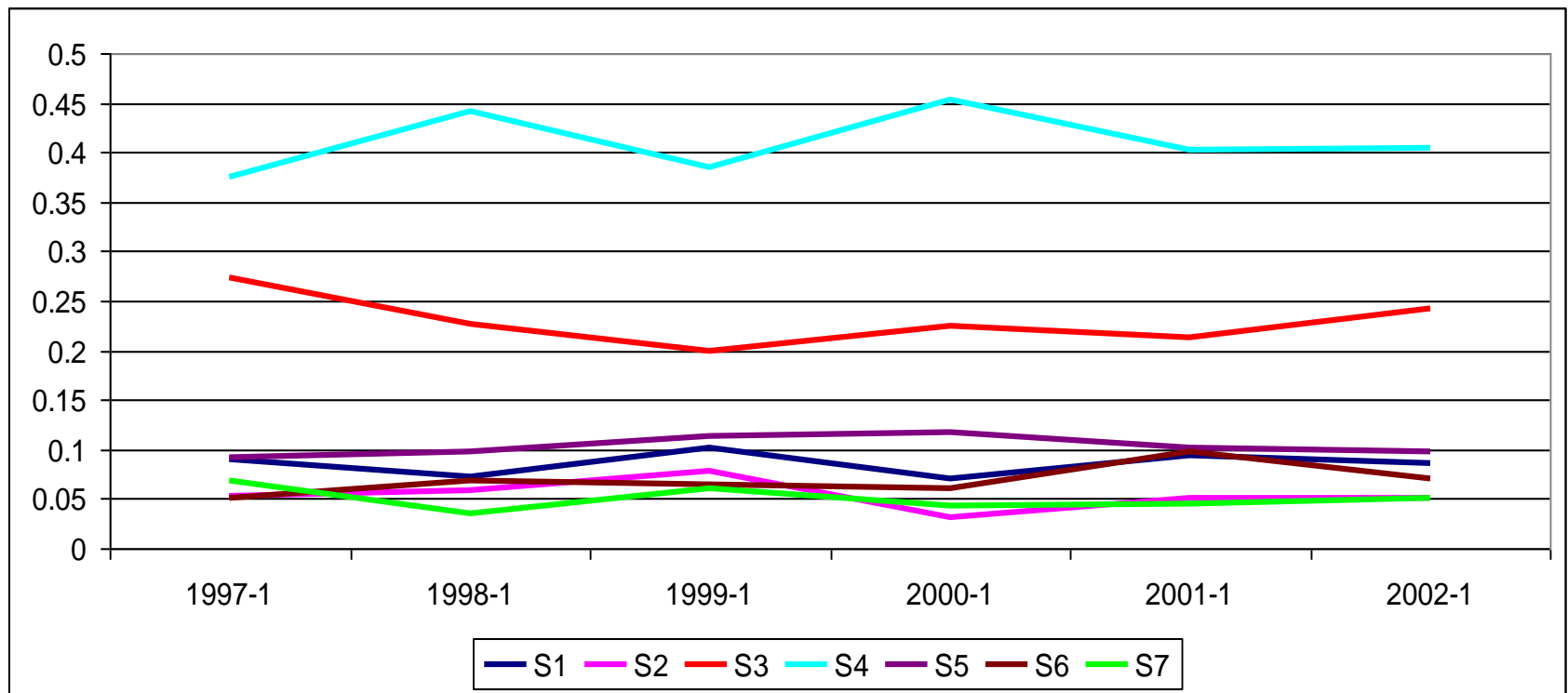
Chart A1: January Prices for Seven Kinds of Vegetable



The Data (cont)

Here are the corresponding January expenditure shares for each year and they are pretty smooth.

Chart A2: January Expenditure Shares for Seven Kinds of Vegetable



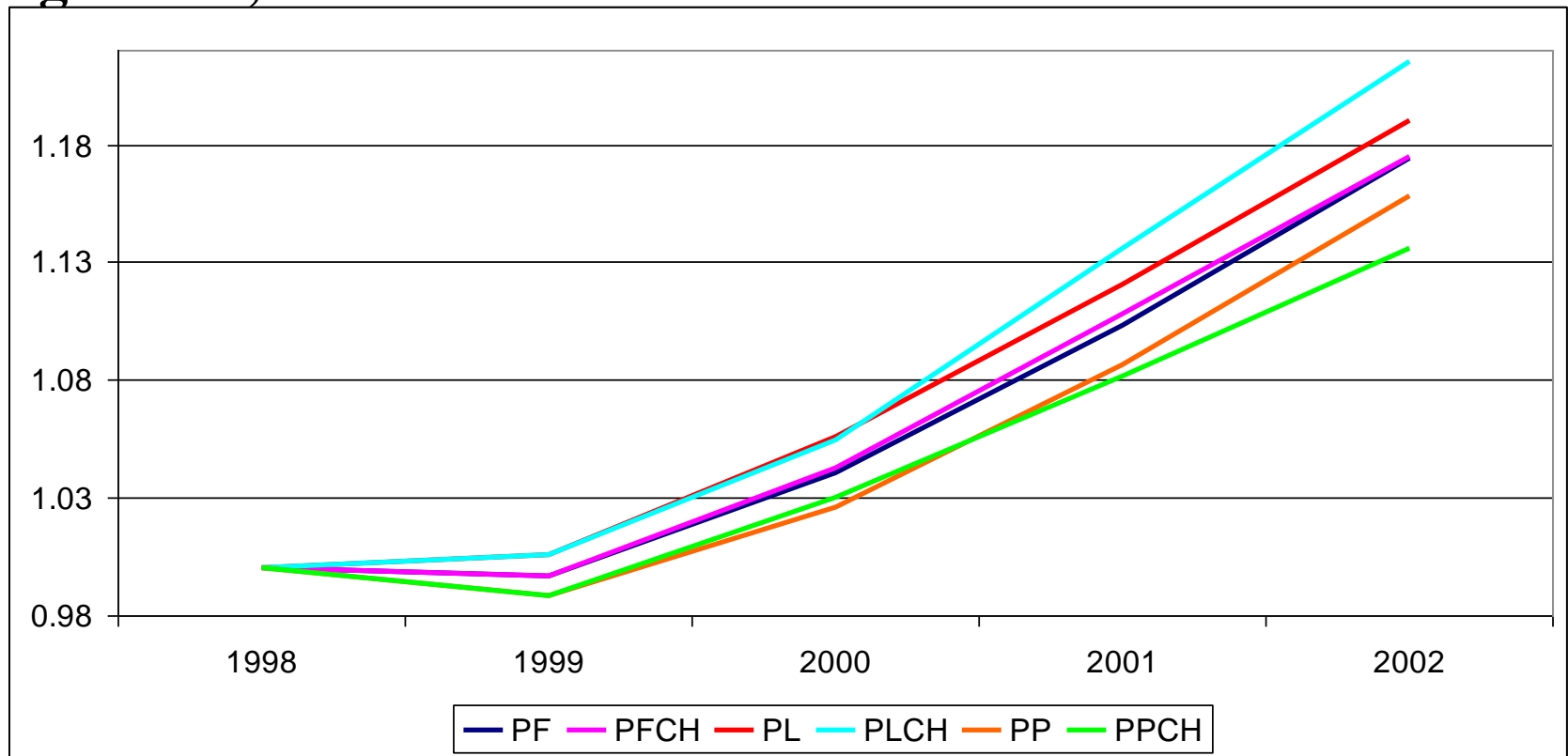
Mudgett Stone Annual Indexes

- Mudgett (1955) Stone (1956) indexes are my preferred **annual indexes**, particularly when dealing with seasonal commodities.
- However, one must have monthly expenditure shares along with monthly data in order to calculate these indexes.
- What is the methodology behind these indexes? We simply treat each commodity in each season as a separate commodity and apply normal index number theory.
- In the present case, we have 7 monthly commodities so this translates into $7 \times 12 = 84$ annual commodities that are inserted into your favourite index number formula.
- The resulting annual indexes (fixed base and chained) are plotted in the next slide for the Laspeyres, Paasche and Fisher formulae.

Mudgett Stone Annual Indexes (cont)

- The Fisher fixed base and chained indexes are close.
- There is **substitution bias** in the other indexes.

Chart 5: Annual Mudgett Stone Fixed Base and Chained Fisher, Laspeyres and Paasche Price Indexes for Seven Kinds of Vegetables, 1998-2002



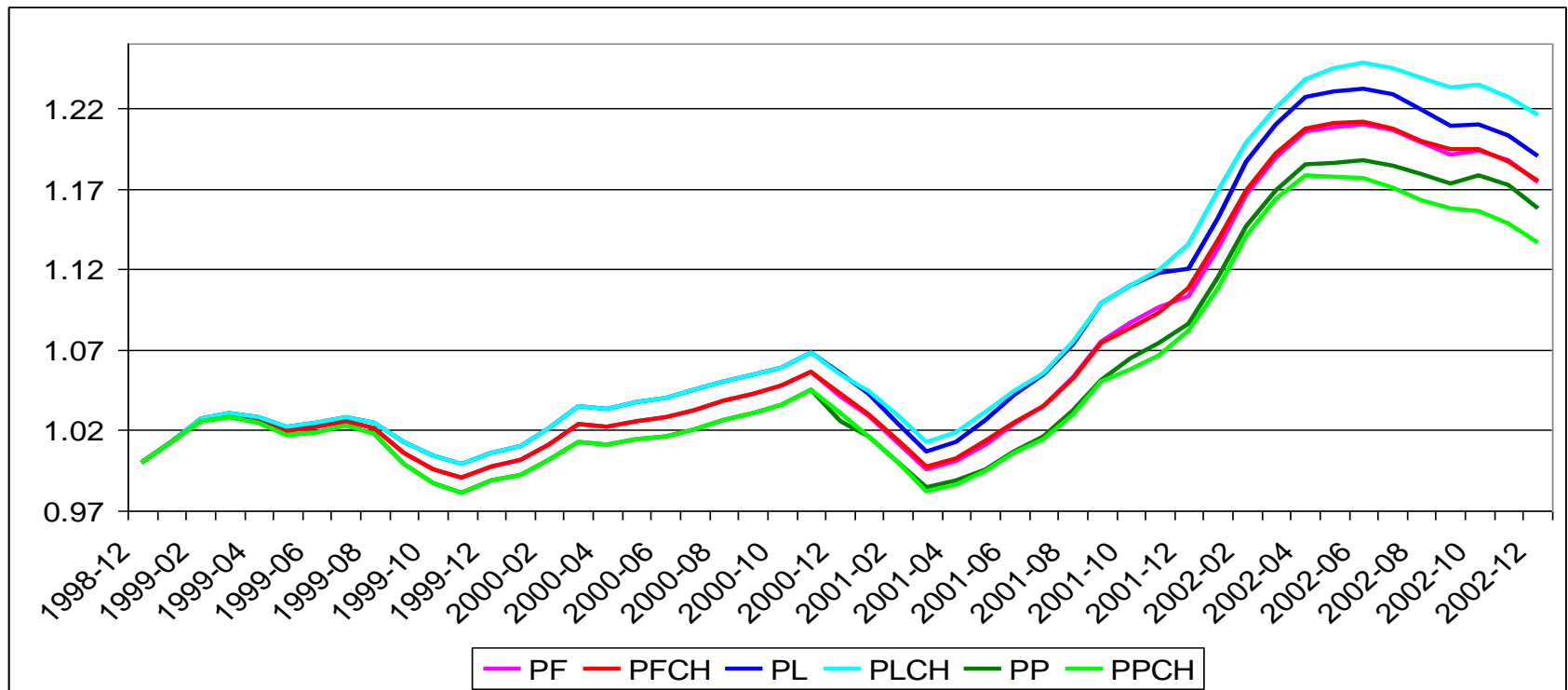
Mudgett Stone Annual Indexes (cont)

- **The Laspeyres fixed base and chained Mudgett Stone indexes grew on average 0.35 and 0.91 percentage points per year more rapidly than their fixed base Fisher index counterparts.**
- **The Paasche fixed base and chained indexes grew on average 0.35 and 0.85 percentage points less rapidly than their fixed base Fisher index counterparts.**
- **This indicates that indexes that do not allow for substitution behavior are likely to have substantial biases.**
- **The Mudgett Stone idea can be generalized. Basically, calendar years are somewhat arbitrary years.**
- **In order to get a new look at annual inflation on a monthly basis, we can compare a Rolling Year of 12 months of data with the price and quantity data of a base year.**
- **This leads to Rolling Year Mudgett Stone indexes.**

Rolling Year Mudgett Stone Indexes

- Here are the **R_Y Mudgett Stone indexes** for fixed base and chained Laspeyres, Paasche and Fisher indexes

Chart 8: Rolling Year Mudgett Stone Fixed Base and Chained Fisher, Laspeyres and Paasche Price Indexes for Seven Kinds of Vegetables, 1998:12-2002:12



Rolling Year Mudgett Stone Indexes (cont)

- There are no obvious seasonal fluctuations in the above Rolling Year Mudgett Stone indexes.
- The Rolling Year fixed base and chained Laspeyres indexes are well above our preferred target index, P_F .
- The Rolling Year fixed base and chained Paasche indexes are well below our preferred target index, P_F .
- Finally it can also be seen that there is little difference between P_F and its chained counterpart P_{FCh} .
- What are the RYMS indexes good for? They are good for providing ex post seasonally adjusted and smoothed estimates of inflation that is centered in the middle of the Rolling Year. **These indexes should be suitable target indexes for the inflation targets of central banks.**
- The only drawback associated with this use of the RYMS indexes is that the inflation measure is 6 months out of date.

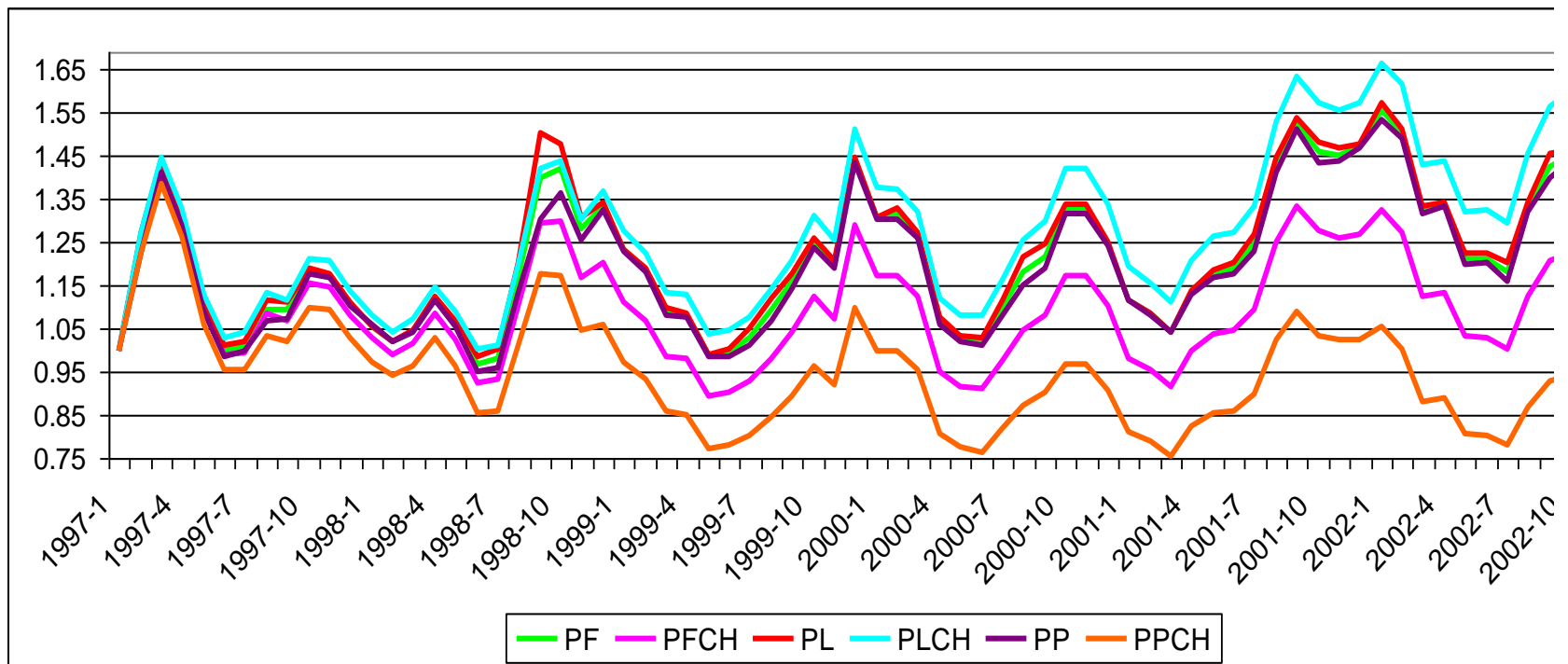
Month to Month Indexes

- **The Rolling Year Mudgett Stone indexes are useful analytic indexes that could usefully be produced by statistical agencies but it will still be necessary to have a month to month CPI of the usual type so that central banks and other interested parties can monitor sudden changes in general inflation. Thus we now turn our attention to these month to month indexes in the context of our seasonal data.**
- **The Consumer Price Index Manual basically recommended the use of chained superlative indexes as a suitable target index for statistical agencies to produce. However, recent research has shown that sales cause this methodology to fail; chained superlative indexes will usually have a substantial downward drift due to the sales problem.**
- **It is of some interest to see if the chain drift problem materializes in the context of seasonal data and so we will now check whether our Israeli data set on vegetables is subject to the chain drift problem.**

Month to Month Indexes (cont)

- We start off by computing **fixed base and chained Laspeyres, Paasche and Fisher** month to month indexes.

Chart 11: Monthly Fixed Base and Chained Fisher, Laspeyres and Paasche Price Indexes for Seven Kinds of Vegetables, 1997:1-2002:12



Month to Month Indexes (cont)

- **The fixed base Fisher, Laspeyres and Paasche indexes are all tightly clustered in a narrow band when we look at the previous Figure.**
- **However, in reality, there are substantial differences between the 3 fixed base indexes.**
- **The fixed base Laspeyres indexes are on average 1.38 percentage points above their fixed base Fisher counterparts.**
- **The fixed base Paasche indexes are 1.35 percentage points below their Fisher counterparts on average.**
- **However, the differences between the fixed base Fisher and the chained Fisher and chained Paasche are massive: the chained Fisher indexes are on average a huge 11.6 percentage points below their fixed base Fisher counterparts.**
- **Our conclusion is that month to month chained Fisher indexes over seasonal goods are subject to the same sort of chain drift bias that has shown up in scanner data studies.**

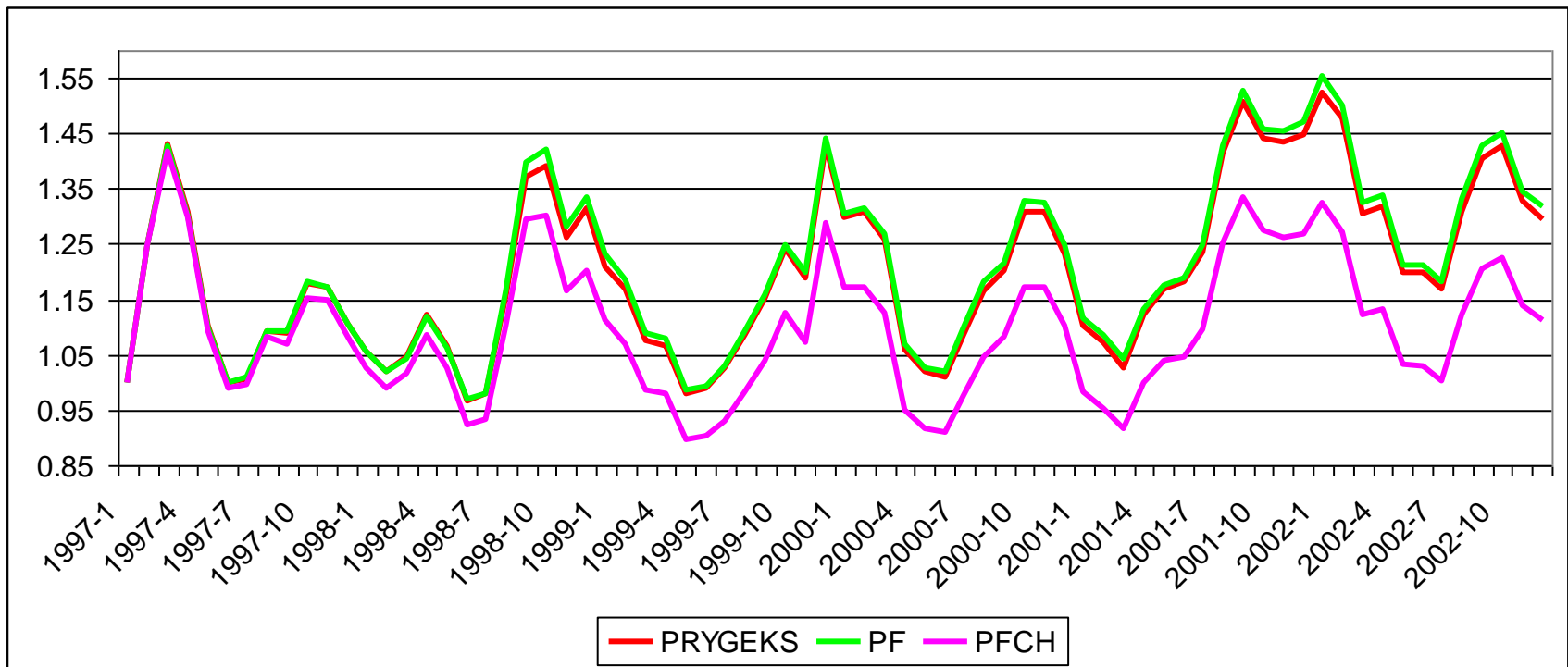
Rolling Year GEKS Indexes

- **Why not just use the fixed base Fisher indexes as the target index in our seasonal context?**
- **The problem is that the expenditure patterns and prices in the first month of 1997 may not be representative of what happens throughout the entire year. Thus the base month plays a very asymmetric role in the fixed base Fisher indexes.**
- **The solution is to use Rolling Year GEKS in the present context. We choose a 13 month window for the usual reasons and compute the resulting indexes.**
- **The following slide plots the fixed base and chained month to month Fisher indexes that appeared previously along with the corresponding Rolling Year GEKS index.**

Rolling Year GEKS Indexes (cont)

- It can be seen that for our Israeli data set, **the RYGEKS and fixed base Fisher do not differ substantially.**

Chart 12: Monthly Rolling Year GEKS, Fixed Base and Chained Fisher Price Indexes for Seven Kinds of Vegetables, 1997:1-2002:12



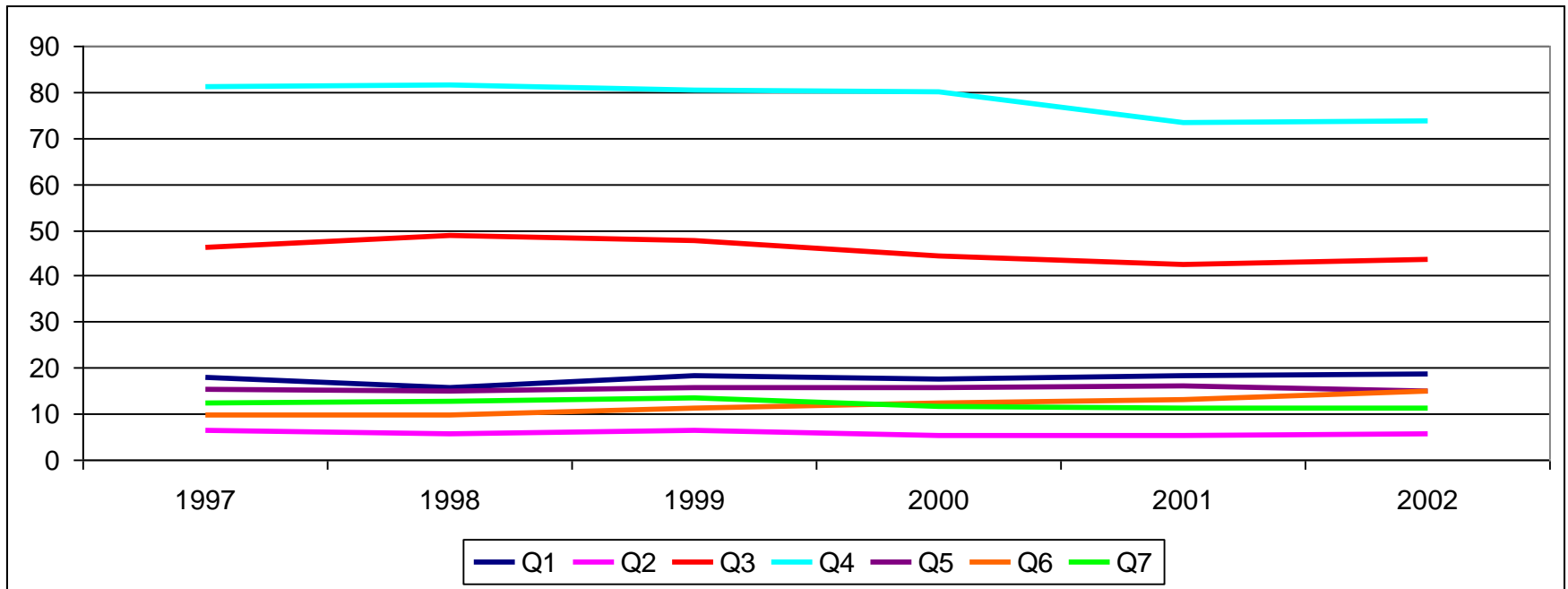
Rolling Year GEKS Indexes (cont)

- The peaks in the fixed base Fisher indexes seem to be consistently higher than the corresponding peaks in the RYGEKS indexes.**
- The downward drift in the chained Fisher indexes is again apparent.**
- It can be seen that the seasonal peaks do not remain constant across the years but there is a definite peak during the last 3 months of each year and there is a definite valley around June of each year but these peaks and valleys are not completely regular.**
- Finally, we will compare the above RYGEKS indexes with the various month to month “practical” indexes that are used by statistical agencies. These practical indexes rely on out of date annual quantity baskets or expenditure shares to weight the monthly prices.**

Comparison with “Practical” Indexes

- **“Practical” consumer price indexes rely on annual baskets.**
- **We plot the implied annual baskets for our Israeli data.**

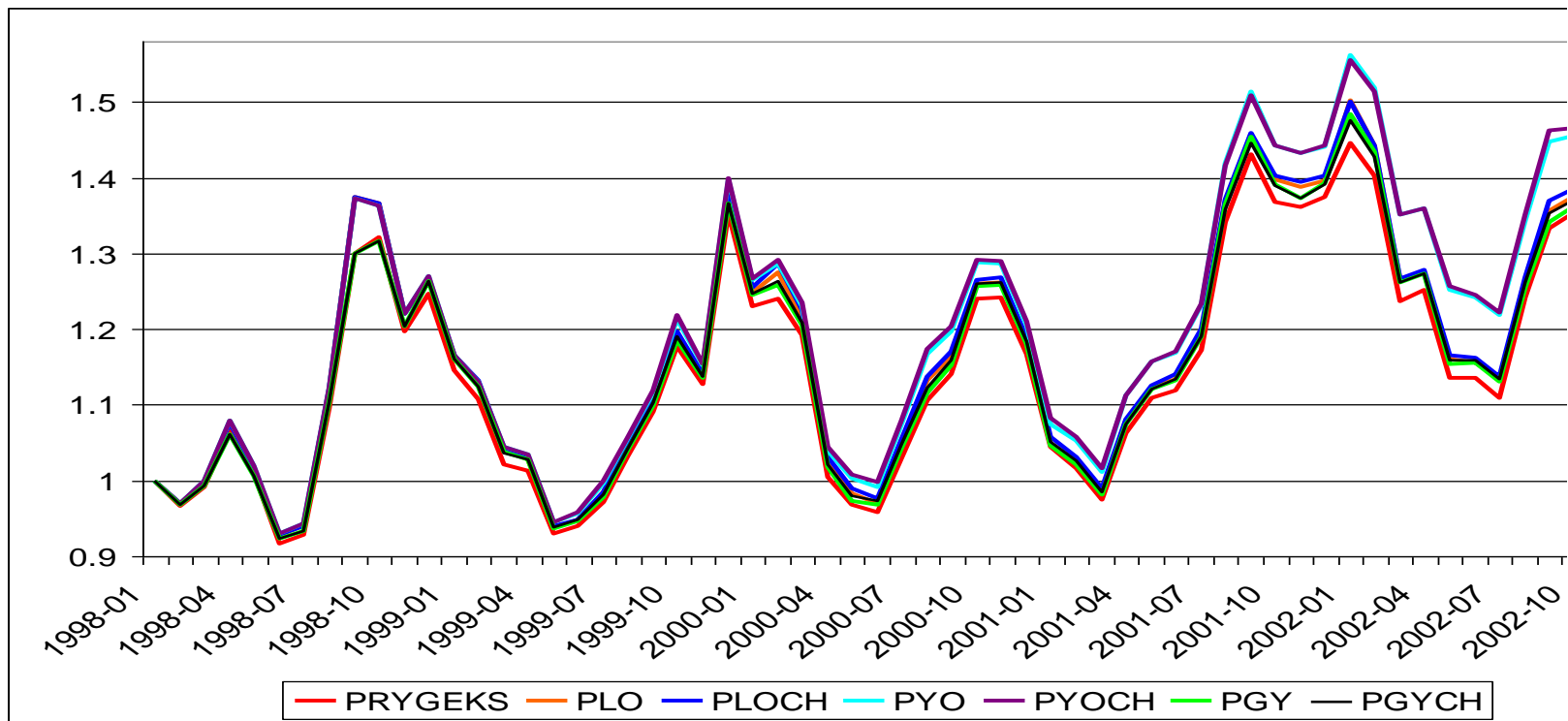
Chart 13: Annual Quantity Baskets for Seven Kinds of Vegetable, 1997-2002



Comparison with “Practical” Indexes (cont)

- Time does not permit me to define the various practical indexes here; please refer to the paper for definitions.

Chart 15: Month to Month **Rolling Year GEKS, Lowe, Chained Lowe, Young, Chained Young, Geometric Young and Chained Geometric Young Price Indexes, 1998:1-2002:12.**



Comparison with “Practical” Indexes (cont)

- The Young and Chained Young tend to be well above the other indexes.
- The Lowe indexes and the Geometric Young indexes are fairly close to each other but the Rolling Year GEKS index (the red line) tends to be markedly below the other indexes at peaks and valleys.
- **All 6 of the practical month to month indexes that use annual out of date expenditure or quantity information are above our preferred Rolling Year GEKS month to month index. (Substantially in the case of Young Indexes).**

Y/Month	P_{RYGEKS}	P_{Lo}	P_{LoCh}	P_{Yo}	P_{YoCh}	P_{GY}	P_{GYCh}
1998-01	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
2002-12	1.22932	1.24903	1.25464	1.32220	1.32619	1.24652	1.24902
Diff.	0	0.01987	0.02381	0.05056	0.05275	0.01198	0.01346

Conclusion and Recommendations

- **All of the indexes that rely on out of date annual quantity or expenditure information are subject to some substitution bias for the case of our Israeli data set, which can be quite considerable in the case of the Young indexes.**
- **Thus it is important for statistical agencies to collect household expenditure data on a continuous basis so that substitution bias in “practical” Consumer Price Indexes can be evaluated on an ex post basis.**
- **It appears that chained superlative indexes used in the context of seasonal commodities can be subject to a considerable amount of chain drift bias.**
- **Thus we recommend the use of Rolling Year GEKS indexes for the computation of month to month target indexes.**
- **Finally, if continuous household expenditure data are available to the statistical agency, then a useful analytical index to produce is the Rolling Year Mudgett Stone index.**