Consumer Price Index. Does the Price Collection Frequency Matter?

Some Monte Carlo Results.

By

Dikaios Tserkezos

Greek Econometric Institute

and
Department of Economics.
University of Crete.
Gallos, GR-74100, Rethymno, GREECE.
Phone: 28310 77415.
Fax: 28310 77415.
e-mail:dtsek@tellas.gr

Abstract.

In this paper, we examine the effects of data collection frequency on the computation of the Consumer Price Index (CPI). Using stochastic simulation techniques, we conclude that the frequency of data collection has a considerable effect on CPI values. Our findings confirm the need for high levels of data collection frequency for the purpose of computing the CPI and, in general, for the more effective monitoring of developments in the cost of living, as this is approached on the basis of Consumer Price Index values.

Keywords: *Frequency of Data Collection, Consumer Price Index, Stochastic Simulation.*

JEL classification: C81 - C82 - E30

1. Introduction.

This paper investigates the effectiveness of the methodology for calculating the Consumer Price Index (CPI) when the frequency of data collection differs. Figure 1 shows the stochastic simulation results for the relatively extreme – though not altogether unlikely – case in which the CPI is calculated using the same data on a daily and on a monthly basis[1].

It is clear that when the prices of the various products (goods) which contribute to the shaping of the CPI on a monthly basis are used, we obtain a completely different picture of price developments, at least relative to the respective index when calculated on a daily basis. Given that the prices of most – if not all – goods are shaped on a daily basis[2], for the purpose of calculating the values of a Consumer Price Index, the choice of frequency of data collection would appear to play an important role in the compiling of accurate indices.

This is the main purpose of this paper. We shall look at the effects on the Consumer Price Index when the data used for its compilation is collected at different frequencies. Because such types of de-aggregated data are not available, we examined the impact of varying data collection frequency on the values of the Consumer Price Index using stochastic simulation techniques. We created prices for 60 goods, for 10 different urban areas and for five data collection stages in each urban area, and experimented by applying the CPI methodology 5,000 times at 30 different levels of data collection frequency.

The results of our experimentation confirm that the effects of the frequency of data collection are determinative for the numerical approach to the inflationary pressures on consumers, at least in the context of the methodology presently being used, in which the prices of most goods are collected on a monthly or, at best, fortnightly basis. Our results point to a decrease of 11% in the percentage change of the Consumer Price Index for a period of 365 days when there is only one data collection in a period of 30 days. Our conclusions clearly show the necessity for Consumer Price Indices, in the immediate future, to be calculated – if not on a weekly basis – then at least a fortnightly one[3].

Although numerous studies have been conducted on the effects of time aggregation and systematic sampling in many methodological approaches, we have not come across such applications for the Consumer Price Index. This means we are unable to make relevant references and comparisons which would augment the conclusions of this paper. It should also be noted that the effects of the frequency of data collection on the values of the CPI are

[1] *The monthly data correspond to 30 days.*
[2] *There are also instances of – mainly public – goods, the prices of which remain unchanged for long periods irrespective of the data collection areas and stages.*

not combined with other impacts on the Consumer Price Index such as Consumer Price Index Bias[4].

Figure 1. Graphical comparison of the diachronic behavior of the Consumer Price Index estimated using data in Daily and Monthly Frequencies. Source: (Simulation Results).

This paper is structured as follows: The second part presents the methodology for calculating the Consumer Price Index and the stochastic experiment design. In the third part, we present our results and our conclusions are set out in the final part.

[3] *Naturally, the possibility should not be ruled out of data in the future being collected on a daily basis through the application of modern data collection methods.*

[4] *We are referring to Substitution Bias, Elementary Index Bias, Outlet Substitution Bias, Quality Adjustment Bias (or Linking Bias) and finally New Goods Bias. Diewert (2001).*

2. The Consumer Price Index and the Simulation Experiment.

According to the standard methodology[5] Consumer Price Indexes can be estimated using LASPEYRES[6] type indexes of the form:

$$CPI_t = \sum_{j=1}^{N} w_j \left[\sum\nolimits_{s=1}^{m} \left(\frac{1}{g} \sum\nolimits_{j=1}^{g} \frac{p^t{}_{jsg}}{p^o{}_{jsg}} \right) \Phi_s \right] \qquad (1)$$

where:

CPI: Consumer Price Index.

N: Number of Goods and Services

m: Total Number of Stages.

G : Total Number of Outlets

$\Phi_{s=1,2,..,m}$: Weights of the Stages

$p_{j(j=1,2,...,N),s(s=1,2,...,m),g(g=1,2,..,G)}$: Price of the *j* Good and Services at stage *s* in the *g* *outlet.*

$w_{j(j=1,2,...,N)}$: Weights of Goods and Services.

$j = 1,2,3,..., N$ Number of Goods and Services.

$s = 1,2,3,..., m$ Number of Stages.

$g = 1,2,..., G$ Price Collection Outlets.

According to the usual methodology, prices of goods are collected on a monthly or, at best, fortnightly basis. Of course there are also instances of – mainly public – goods, for which prices are collected at long time intervals. Examples of such goods are electricity, public transport fares, etc.

---

[5] National Statistical Service of Greece, NSSG, 1998.In the presentation of the CPI methodology we did not take into account the case of Groups of Goods and Services. This helps us to simplify the computations in the stochastic experiments without, of course, diminishing the significance of our results.

[6] Kiochos, P.,'Statistics'., Interbooks., Athens 1993, pp. 323-340.

The Experiments.

Examination of some actual price series suggested that the simulations of the $j = 1,2,3,....,(N = 60)$ goods for the $s = 1,2,3,...,(m = 10)$ Urban Areas for the $g = 1,2,...,(G = 5)$ Stages, can be well characterized by random processes of the following sort:

$$p_{jt} = p_{jt-1}(mg_j + gg_{jt} + e_{jt}) \qquad (2)$$

with p_{jt} the price of the $j = 1,2,3,....,(N = 60)$ good at period t, and $mg_{j=1,2,...,(N=60)}$, $gg_{j=1,2,...,(N=60)}$ the corresponding rates of change and g^i is one plus a rate of growth, e_t^i is a normally distributed error term specific to series i, u_t is a normally distributed error term common to all the series, and a^i is a constant. Thus (2) implies that prices on the average grow at a constant rate varying among prices, but this is modified by random shocks. Some of these random shocks affect individual series only, while some affect all the series to some degree (reflecting the role of the Business Cycle) of only peripheral interest[7].

In our experiments we specified the components on the right hand side of (2) as follows:

$$mg_j \in (1.0013 - 1.0075) \qquad (3)$$

$$gg_{jt} \approx NID(0, gg_j) \qquad (4)$$

μὲ

$$gg_j = 0.004\left(\frac{mg_j - 1}{mg^*_j - 1}\right)^2 \qquad (5)$$

καὶ

$$mg^* = 1.0045 \qquad (6)$$

$$e_{jt} \approx NID(0, \sigma^2_e = 0.023) \qquad (7)$$

$$p_{jo} = .20 \qquad (8)$$

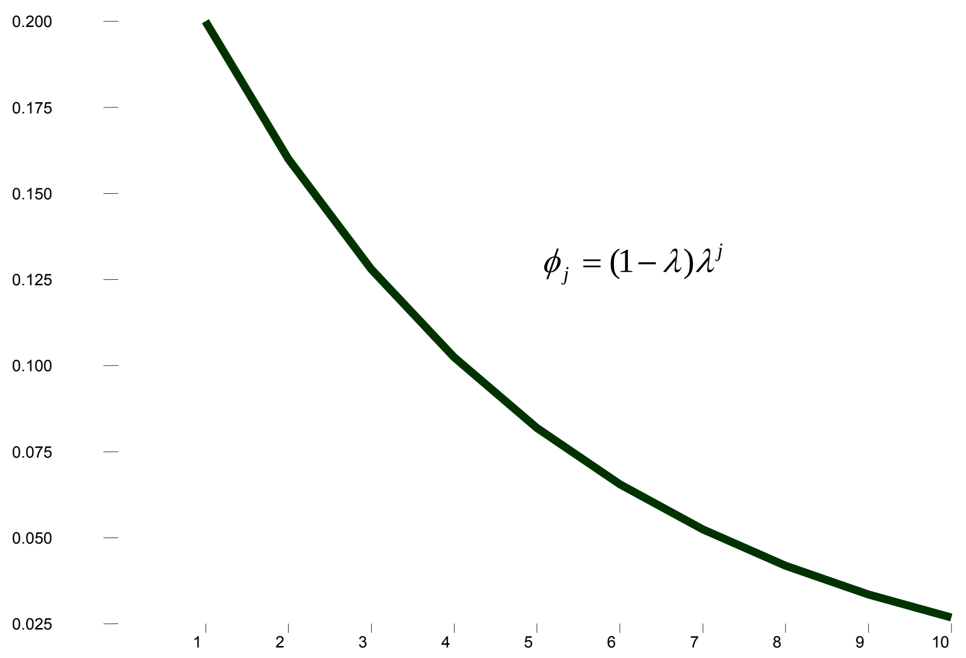
$$j = 1,2,3,....,60$$

We conducted our experiments with 5000 iterations for 30 different levels of data collection frequency using a data sample of 365 days. In each experiment we obtained some useful statistical measures and we used their frequency distributions in our conclusions. In order to calculate the Consumer Price Index in our experiments, we need the weights of price collection outlets and the weights of the 60 goods participating in the Consumer Price Index; the weights of the 10 different Stages and the Goods and Services were estimated as follows:

Price Collection Outlets.

In our experiments we assumed 10 Urban Areas for price data collection. The Urban Area weights $\phi_{j=1,2,...,10}$ were estimated using the relation[8] $\phi_j = (1-\lambda)\lambda^j$ with $\lambda = .65$ an estimated adjusted parameter based on the geometric lag distribution. A graphical presentation of the Urban Area weights is given in Figure 2.

URBAN AREA WEIGHTS.



Urban Areas.

Figure 2: Assumed Weights of the 10 Urban Areas - Distributions. (Source: Simulation Results)

---

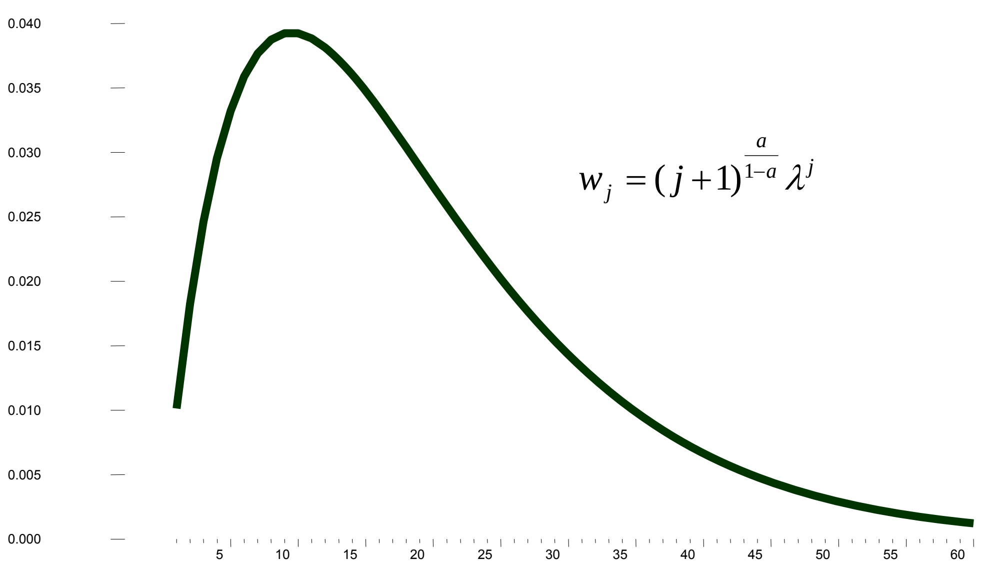
[7] More on sampling designs in constructing consumer price indices can be found in Boon, M., (1997),

[8] The parameter λ was estimated using the $\min_{\hat{\lambda}} S(\hat{\lambda}) = \min_{\hat{\lambda}} \sum_{j=o}^{10} ((1-\hat{\lambda})\hat{\lambda}^j - n_j)^2$ with $n_{j=1,2,..,10}$ the actual Urban weights according to data of the Greek National Statistical Service (NSSG 1998) and λ an adjusted parameter lying in the interval $0 \le \lambda \le 1$. More about this geometric lag specification can be found in Almon (1965).

Weighting coefficients of Goods.

The weights of goods $w_{j=1,2,...,60}$ are estimated using a Gamma distributed lag distribution, the relation[9] $w_j = (j+1)^{\frac{a}{1-a}} \lambda^j$ $0 \le \lambda \le 1$ and $0 \le \alpha \le 1$ with $\lambda = .65$ and $a = .67$[10] . A graphical presentation of the weights of goods is given in Figure 3.

WEIGHTING COEFFICIENTS OF GOODS AND SERVICES



Goods and Services.

Figure 3: Assumed Frequency Distribution weights of the Goods and Services participating in the CPI construction. (Source: Simulation Results)

We conducted our experiments with 5000 iterations for 30 different levels of data collection frequency using a data sample of 365 days. Lastly, it should be noted that with respect to the procedure for the stochastic simulation of prices, we made no assumption relating to the distribution of expenditures at the household level in different Urban Areas and Households.

---

[9] Schmidt Peter.,1974.

[10] The parameters λ and a were estimated using the $\min_{\hat{\lambda},\hat{\alpha}} S(\hat{\lambda},\hat{\alpha}) = \min_{\hat{\lambda},\hat{\alpha}} \sum_{j=1}^{60} ((j+1)^{\frac{\hat{a}}{1-\hat{a}}} \hat{\lambda}^j - \psi_j)^2$ with $0 \le \lambda \le 1$, $0 \le a \le 1$ and

4. Results of the Experiments.

The results of our experiments are presented in Figure 4 and Table 1. Figure 4 shows the frequency distributions of the 5000 experiments for the yearly CPI percentage changes at 30 different levels of price collection frequency. It is clear that as the number of price collections decreases, the estimated value of the cost of living also decreases, at least to the extent that this can be approached with the existing Consumer Price Index methodology. The magnitude of the effects of time aggregation is indicated by line A in Figure 4.

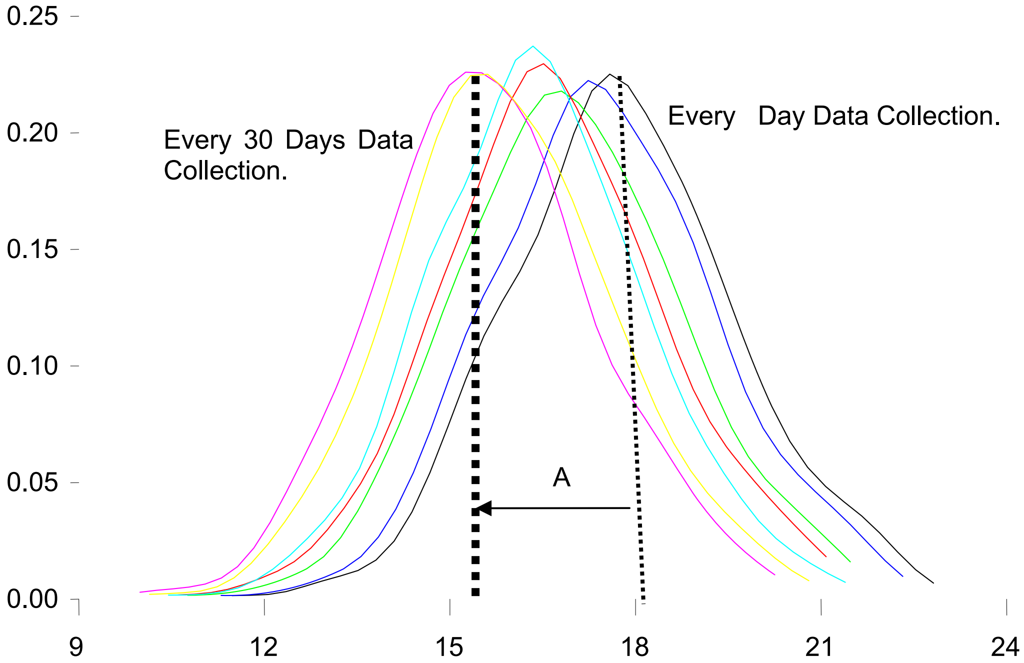


Figure 4. Frequency Distributions of the yearly CPI Percentage Changes at Different Price Collection Frequencies. (<u>Source</u>: Simulation Results)

$\psi_{j=1,2,\ldots,60}$ the actual weights of Goods and Services according to data of the National Statistical Service of Greece (NSSG 1998).

TABLE 1. Estimates of Yearly Price Changes and their statistics at Different Price Collection Frequencies.

Price Collection Frequency Levels	Average Change	Minimum Change	Maximum Change	Standard Deviation	Kurtosis	Symmetry	Jarque Berra	Significance
1 (Every 1 Day)	17,79	11,48	22,82	1,78	0,07	-0,04	0,90	0,64
5 (Every 5 Days)	17,50	11,30	22,32	1,76	0,07	-0,08	1,06	0,59
10 (Every 10 Days)	16,95	10,76	21,80	1,75	0,08	-0,05	1,29	0,53
15 (Every 15 Days)	16,67	10,60	21,52	1,74	0,09	0,01	1,49	0,48
20 (Every 20 Days)	16,43	10,45	21,40	1,69	0,11	0,04	1,95	0,38
25 (Every 25 Days)	15,61	9,99	20,96	1,67	0,11	0,08	2,38	0,30
30 (Every 30 Days)	15,89	10,14	20,80	1,69	0,12	0,00	2,52	0,28

Source: Our Estimates.

Table 1 presents a series of statistical criteria of a descriptive nature for the average[11] yearly changes of the Consumer Price Index at different price collection frequencies. For systematic sampling level 30, which means that we collect prices every 30 days, the average yearly CPI change appears to be underestimated by 11%, whilst if the maximum and minimum changes are taken into consideration, this underestimation is around 12% and 9% respectively. If one takes into account that the total sample of stochastic experiments is 365 observations, i.e. one year, which is also used as the main basis for reference and comparison of medium-term inflationary pressures, the size of the underestimation of inflationary pressures would appear to be relatively large.

Moreover, the data in Table 1 reveal an increase of positive symmetry in conjunction with a tendency for the distribution of CPI changes to become more platykurtic at a low number of price collections. These results complement the negative effects due to the reduction of the number of price collections, given that the chances increase for a greater possible spectrum of CPI price changes (kurtosis) with a simultaneous tendency for a concentration of yearly percentage changes at relatively higher price levels (positive asymmetry). Indirect confirmation of the above is provided by the values of the Jarque-Berra[12] criterion for testing the normality of the percentage changes in the Consumer Price Index.

[11] With respect to the number of iterations.
[12] Jarque-Berra 1987

4. ***Conclusions***.

The criticism leveled at the methodology for calculating the level of the cost of living on the basis of the Consumer Price Index is both heavy and diverse. To this one might add the criticism relating to frequency of the number of price collections. As emerges from the application of stochastic simulation techniques, a small number of price collections may lead to an underestimation of the real inflationary effects based on the assessment of usual Consumer Price Indices. In addition, this underestimation may be accompanied by an increase in the likelihood of obtaining estimates of CPI percentage changes in a greater spectrum of prices with a tendency for their concentration at high levels.

Given the importance of the Consumer Price Index in the implementation of economic policy, but above all for estimating the quality of life of consumers, the results of these experiments should be taken into serious consideration when formulating the methodology for calculating the Consumer Price Index.

References.

1. Almon S.,(1965), ‘The Distributed Lag Between Capital Appropriations and Expenditures’ Econometrica, 33:178-196.

2. Boon, M., (1997), Sampling designs in constructing consumer price indices: current practices at statistical offices. Interim report on the Eurostat project ‘Development of methodologies in consumer price indices and purchasing power parities’ (SUP-COM 1996, Lot 1).

3. Diewert E. W .,(2001)., ‘Sources of Bias in Consumer Price Indexes.’ Discussion Paper , Scholl of Economics, The University of New South Wales.

4. Drakatos K.Γ., Greek Economics Statistics, Papazisis Publications.,1982
NSSG, (1998), ‘Rivesed Consumer Price Index’. National Statistical Service for Greece. Methodological Studies 21, Athens.

5. Haan, J. De and E. Opperdoes, (1997α), Estimation of the coffee price index using scanner data: the choice of the micro index. Internal Report (Statistics Netherlands, Voorburg).

6. Haan, J. De, E. Opperdoes and C. Schut, (1997β), Item sampling in the consumer price index: a case study using scanner data. Internal Report (Statistics Netherlands, Voorburg).

7. Jarque C.M and Bera.,1978.,A test fir Normality of Observations and Regression Residuals.,International Statistical Review.,Vol.55.

8. Κιόχος Π.,Statistics’. Interbooks.,Αθήνα 1993.

9. Koyck L.M.(1954), Distributed Lags and Investment Analysis., North Holland Publishing Company, Amsterdam 1954.

10. Philips Louis, (1974), ‘Applied Consumption Analysis.’ North –Holland Publishing Company.

11. Sogel Thomas ,(2000), ‘Comments on Price Indices and Inflation.’ Research Department , Federal Reserve Bank, Minneapolis, Minnesota, 2001

12. Schmidt Peter.,1974., An Augment of the Usefulness of the Gamma Distributed Lags Model., International Economic Review, February 1974.

13. Tserkezos Dikaios, 2001, ‘Class Consumer Price Indices’. To “To Vima Sunday” (16/10/2001) http: //tovima.dolnet.gr

14. Tserkezos Dikaios, 2001, ‘The Class CPI’, Oikonomikos Tachydromos, http://tovima.dolnet.g