DANMARKS NATIONALBANK

DANMARKS NATIONALBANK
WORKING PAPERS
2002 • 1

Heino Bohn Nielsen

Institute of Economics

University of Copenhagen

and

Dan Knudsen

Economics,

Danmarks Nationalbank

Robust estimation of the expected inflation

March 2002

The Working Papers of Danmarks Nationalbank describe research and development, often still ongoing, as a contribution to the professional debate.

The viewpoints and conclusions stated are the responsibility of the individual contributors, and do not necessarily reflect the views of Danmarks Nationalbank.

As a general rule, Working Papers are not translated, but are available in the original language used by the contributor.

Danmarks Nationalbank's Working Papers are published in PDF format at www.nationalbanken.dk. A free electronic subscription is also available at this Web site.

The subscriber receives an e-mail notification whenever a new Working Paper is published.

Please direct any enquiries to
Danmarks Nationalbank, Information Desk, Havnegade 5, DK-1093 Copenhagen K Denmark
Tel.: +45 33 63 70 00 (direct) or +45 33 63 63 63
Fax : +45 33 63 71 03
E-mail:info@nationalbanken.dk

Nationalbankens Working Papers beskriver forsknings- og udviklingsarbejde, ofte af foreløbig karakter, med henblik på at bidrage til en faglig debat.

Synspunkter og konklusioner står for forfatternes regning og er derfor ikke nødvendigvis udtryk for Nationalbankens holdninger.

Working Papers vil som regel ikke blive oversat, men vil kun foreligge på det sprog, forfatterne har brugt.

Danmarks Nationalbanks Working Papers er tilgængelige på Internettet www.nationalbanken.dk i pdf-format. På webstedet er det muligt at oprette et gratis elektronisk abonnement, der leverer en e-mail notifikation ved enhver udgivelse af et Working Paper.

Henvendelser kan rettes til :
Danmarks Nationalbank, Informationssektionen, Havnegade 5, 1093 København K.
Telefon: 33 63 70 00 (direkte) eller 33 63 63 63
E-mail: info@nationalbanken.dk

ISSN (Print/tryk) 1602-1185

ISSN (online) 1602-1193

Robust estimation of the expected inflation

Heino Bohn Nielsen
Institute of Economics,
University of Copenhagen
Studiestræde 6
DK-1455 Copenhagen K
heino.bohn.nielsen@econ.ku.dk

Dan Knudsen
Economics Department,
Danmarks Nationalbank
Havnegade 5
DK-1093 Copenhagen K
dkn@nationalbanken.dk

July 2, 2001

Abstract

Ignoring items with large price changes may enhance the informational content of a price index. As an application of the metrically trimmed mean (Kim, 1992) we suggest to discard the individual price changes that deviate the most from the median. Focusing on outliers increases the efficiency compared to always trimming equally in both tails and the implied bias problem seems small. The distribution of price changes is often skewed strongly to the right or to the left in a specific month but is much closer to symmetric for a longer period as a whole. This is also the case with Danish data analyzed in this paper. The suggested metrically trimmed mean gives a measure of expected inflation, which may help representing inflation in economic analyses.

Resume

Det kan forbedre informationen i et prisindeks at se bort fra poster med store prisændringer. Som en anvendelse af det metriske trim (Kim,1992) foreslår vi at udelade de prisændringer, der afviger mest fra medianen. Et sådant fokus på de ekstreme værdier forbedrer efficiensen i forhold til altid at trimme i begge haler af fordelingen, og det medfølgende bias-problem synes lille. Fordelingen af prisændringer er ofte stærkt højre- eller venstreskæv i en enkelt måned, men meget tættere på at være symmetrisk set over en længere periode. Det gælder også de danske data, der analyseres i nærværende papir. Det foreslåede metriske trim giver et mål for forventet inflation, som kan være med til at repræsentere inflationen i økonomiske analyser.

1 Introduction

Monthly inflation data typically show a considerable short-term variability, which reduces their usefulness for economic analyses. The volatility distorts the autocorrelation of the price series and the cross correlation to other economic time series. A simple reaction is to use a moving average of monthly figures or a lower frequency. It is, however, not necessarily all elements in the monthly price signal, which need to be downweighted. A big month-on-month change in a price index is typically created by a few subindices changing a lot rather than by all prices moving in parallel. Energy and unprocessed foods are typical candidates, but sharp price changes are recorded for most products.

To address this problem, we view changes in the subindices as observations of stochastic variables and proceed to estimate an expected increase in the consumer price index. If the distribution of price changes were Gaussian the best estimate would be the actual increase in the consumer price index. However, the outliers make the distribution of price changes non-Gaussian and it improves the robustness of the estimation to reduce the weight of extreme observations. A robust estimate of the price increase may serve as a measure of the underlying inflation trend and it may be helpful for representing inflation or expected inflation in economic analyses.

One frequently used robust estimator of the location of the central tendency of inflation is the so-called trimmed mean, which discards the largest and smallest observations in a given month and calculates the weighted mean of the remaining observations, see *inter alia* Bryan, Cecchetti and Wiggins II (1997). The standard trimmed mean is a symmetric estimator, which removes the same number of observations in both tails of the cross-sectional distribution every month. Some authors have noted the scope for asymmetric estimators. Roger (1997) reports that the distribution of individual inflation rates for New Zealand is on average skewed to the right and he suggests trimming relative to the average mean percentile to avoid a systematic difference between trimmed and actual inflation.

In this paper we are not particularly interested in the bias issue and suggest instead to minimise the variance and to analyse what this stabilised measure of price increases can be used for. Inspection of the data reveals that the skewness of price changes across consumer goods varies between left and right. This shifting from month to month in the position of outliers suggests that it may be efficient to allow trimming in only one tail of the distribution instead of trimming both tails every month. Specifically, we propose to apply the *metrically trimmed mean from the median*, see Kim (1992), and remove the observations that deviate the most from the median. The paper further argues that it could be preferable to remove whole subindices in the calculation rather than removing a precise proportion of the weights. It is illustrated that the proposed estimator seems efficient under a wide range of assumptions and the potential bias is small.

The rest of paper is organised as follows. Section 2 outlines the official Laspeyres price index and characterises the distribution of prices changes, which looks highly non-Gaussian. In section 3 some alternative robust measures of the inflation rate are discussed and in section 4 they are evaluated. The chosen measure is presented in section 5 and some applications are given. Finally, section 6 concludes.

2 The official index

The Danish consumer price index, P_t, is calculated as a Laspeyres index, i.e. as a weighted arithmetic mean of 121 subindices

$$P_t = \frac{\sum_{i=1}^{121} P_{it}Q_{i0}}{\sum_{i=1}^{121} P_{i0}Q_{i0}} = \sum_{i=1}^{121} \alpha_i \frac{P_{it}}{P_{i0}}, \tag{1}$$

where P_{it} and Q_{it} are price and quantity of good $i = 1, 2, ..., 121$ at time t, and $t = 0$ is the base period for the calculation. The weight $\alpha_i = P_{i0}Q_{i0}/\sum_{j=1}^{121} P_{j0}Q_{j0}$ indicates the share of good i in the budget of the average consumer and $\sum_{i=1}^{121} \alpha_i = 1$. The weights concern a certain year, but the base year and hence the weights have on average changed every 5 years. The subindices and the most recent weights are listed in table 3 and the modifications of the raw data applied in the current analysis are mentioned in the appendix.

In this paper we analyse the developments in the Danish consumer prices for the period $1981:1-2000:11$. The focus is on month-to-month changes, and the price indices are seasonally adjusted using X-11 to avoid that the variance is dominated by seasonal variations. It is appropriate to rank relative changes in logs, so the Laspeyres index (1) is approximated by cumulating the weighted average of the monthly inflation rates, $\pi_{it} = \log(P_{it}/P_{it-1})$, $i = 1, ..., 121$, i.e. cumulating

$$\overline{\pi}_t = \sum_{i=1}^{121} \alpha_i \pi_{it}. \tag{2}$$

Using this approximation, the overall price increase in a given month is the average of individual price increases. To mimic a Laspeyres index the weights in (2) should in principle be corrected every month to allow for relative price changes. However, the approximation with fixed weights is rather good as illustrated in figure 1 (A) and (B), which show the change month-to-month and year-to-year respectively in the CPI seasonally adjusted and in the approximation. In the rest of the paper, we use the approximation (2) as the baseline inflation measure.

The big picture of the Danish inflation from the beginning of the eighties till 2000 reflects the impact of the hard currency policy adopted in 1982. One of the accompanying measures was to abolish the price indexation of wages. The drop in inflation in the nineties to around two per cent came while the economy was weak, but it confirms a fundamental change over the sample that wage and price increases have responded only moderately to the economic upturn in the last half of the nineties.

The distribution of price changes. It is apparent from figure 1 (A) that the monthly inflation rate fluctuates a lot, and a single observation seems to contain little information on the tendency of the price development. To gain some insight on the variability we calculate for each of the 238 months in the sample the weighted mean, $\overline{\pi}_t$, of the 121 price changes $\{\pi_{it}\}_{i=1}^{121}$, the weighted standard deviation, σ_t, and the weighted skewness and kurtosis. The

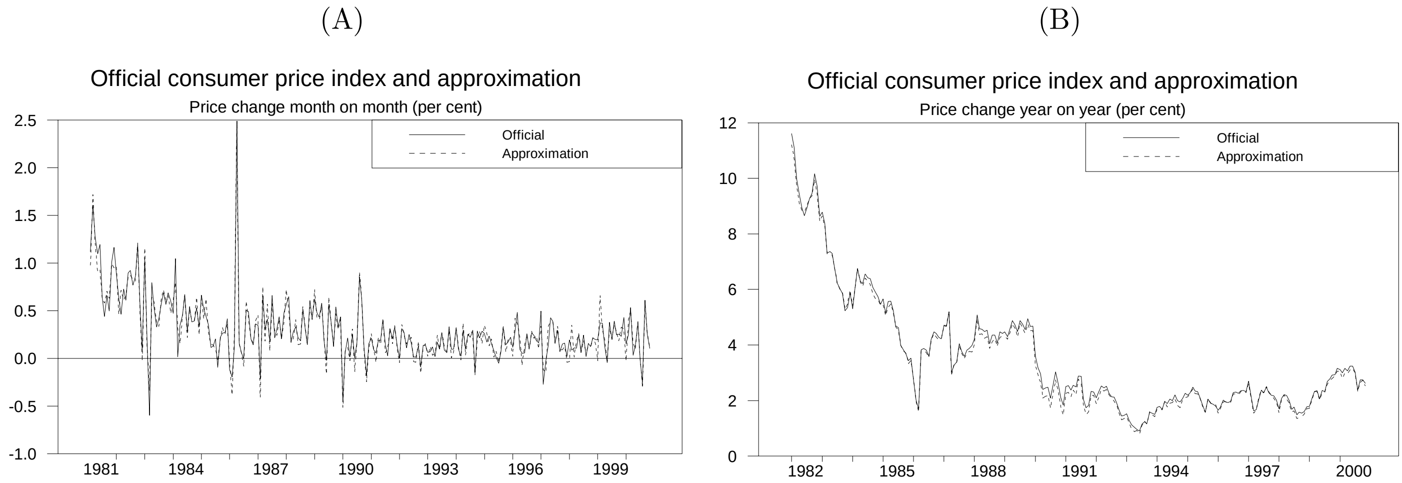

Figure 1. (A) and (B) compare the official consumer price index and the logarithmic approximation.

latter are the normalized third and fourth moment respectively, i.e.

$$m_t^r = \sum_{i=1}^{121} \alpha_i \left(\frac{\pi_{it} - \overline{\pi}_t}{\sigma_t} \right)^r, \quad r = 3, 4,$$

which relate the empirical distribution to the Gaussian that has a skewness of zero and a kurtosis of 3. Table 1 characterises the 238 distributions in the sample. The kurtosis is usually far greater than 3 indicating that most distributions are leptokurtic, i.e. have fat tails compared to the Gaussian distribution. The skewness exhibits large variation from month to month in a range of -20 to 20, but on average the calculated skewness is close to zero[1]. To illustrate we also calculate for each month the percentile of the ordered observations equivalent to the weighted mean. This mean percentile fluctuates in a wide range of 0.14 to 0.88 but has mean and median just above and below 0.50. Bryan et al. (1997) note that samples drawn from a symmetric but leptokurtic distribution are often skewed.

There are several possible factors behind the outliers, which give the price distribution its leptokurtic form. Price cartels may destabilise the price setting. Changing harvest outcomes, new technology and new products, sales to reduce stocks, animal disease etc., all create price swings. A broad explanation of price jumps is found in the theory of menu costs (Ball and Mankiw, 1995). If it is costly to change prices, they are presumably changed less frequently but in larger steps. A practical point is that the recorded price change in a subindex may be particularly large when the statistical agency substitutes a representative good by a new. In general, it may increase the outliers if a subindex is only based on a few representative goods. Moreover, the impact of indirect taxes and subsidies normally comes stepwise and also politically administered prices may move abruptly for instance at the start of a budget year.

[1] The measures used so far to characterize the distributions are themselves non-robust because the standardization is based on the sample mean and standard deviation. Several robust characterizations are possible, but that is not the aim of this section.

Table 1: Characteristics of the distribution of price changes over the sample.

Statistic	*Mean*	*Median*	*St. dev.*	*Minimum*	*Maximum*
Mean	0.303	0.242	0.333	-0.512	2.486
Standard deviation	1.189	1.059	0.595	0.434	5.882
Kurtosis	26.515	16.397	34.527	2.595	382.059
Skewness	0.412	0.375	5.223	-22.122	17.723
Mean percentile	0.507	0.494	0.170	0.136	0.884

Note: Mean, median, standard deviation, minimum and maximum over the sample of 238 observations for some statistics of the cross sectional distributions $\{\pi_{it}\}_{i=1}^{121}$.

3 Robust estimators

A simple way to deal with the noise in inflation data is to consider price developments over several months, which downweights all information in the current inflation figures. An alternative way to stabilise the inflation measure is to reduce the weight of the largest of the 121 monthly price changes. This is the essence of robust estimation. We consider the 121 monthly price changes as drawn from a distribution with all the non-Gaussian characteristics described in the previous section. The arithmetic mean is the least squares estimator, which is consistent and would be optimal if the distribution were Gaussian. The fat tails and high probability of outliers, however, destroys the efficiency of the mean since one extreme observation can remove the mean of the observations significantly from the expected price increase. Robust estimators focus on the central part of the distribution, which makes them less vulnerable.

A classical robust estimator is the median, which implies minimisation of absolute deviations. The median is extremely robust but is potentially inefficient by only using quantitative information from one observation. A class of estimators often used to calculate robust inflation measures is the trimmed means, of which the median is a special case. Instead of giving a weight of zero to all observations but the median one, the trimmed mean assigns the weight zero to a smaller proportion of the subindices and calculates the weighted mean of the remaining subindices.

Symmetrical trim. Usually in the case of price data the trimming is done symmetrically by removing observations corresponding to a weight of $100 \cdot \mu/2$ per cent from each tail of the sample distribution each month. This is called the $100 \cdot \mu$ per cent symmetrically trimmed mean and is calculated as

$$\widetilde{\pi}_t^{\mu} = \frac{1}{1-\mu} \sum_{i=1}^{121} \alpha_i \pi_{it} 1 \left\{ F_j^{\mu/2} [\pi_{jt}] < \pi_{it} < F_j^{1-\mu/2} [\pi_{jt}] \right\},$$

where e.g. $F_j^{\mu/2} [\cdot]$ is the $\mu/2-$percentile (over j) of the cross section in square brackets and $1\{\cdot\}$ is the indicator function. For $\mu = 0$ the estimator $\widetilde{\pi}_t^{\mu}$ is equal to the arithmetic mean and for $\mu \rightarrow 1$ the result converges to the median. This symmetrically trimmed mean is applied to price data in several studies, *inter alia* Bryan et al. (1997), Mio and Higo (1999) and Bryan and Cecchetti (1999b). Bakhshi and Yates (1999) discuss some drawbacks of the approach.

Several authors have noted the possible advantage of asymmetric estimators. Roger (1997) reports that the distribution of price changes for New Zealand on average is skewed to the right and he suggests to trim relative to the average mean percentile, in the specific example to trim in both tails of the distribution relative to the 57^{th} percentile, i.e. to $100 \cdot \mu/2 - 7$ per cent in the right hand tail and $100 \cdot \mu/2 + 7$ per cent in the left hand tail in each month. This is done to avoid an indicator with a full sample mean below the actual inflation, and the efforts entail a deliberate loss in efficiency. A mean at say the 57^{th} percentile normally implies that most outliers appear in the right hand tail, but most observations are trimmed in the left hand tail of the distribution.

Metrical trim. We prefer to put emphasis on variance reduction when trimming. The preceding section illustrated that the dominant position of outliers changes between the left and right tail. Thus, to focus on eliminating outliers one should not trim both tails with preset percentages but trim the $100 \cdot \mu$ per cent of the observations that deviate the most from the median, i.e.

$$\widehat{\pi}_t^{\mu} = \frac{1}{1-\mu} \sum_{i=1}^{121} \alpha_i \pi_{it} 1 \left\{ \left| \pi_{it} - F_j^{0.5} \left[\pi_{jt} \right] \right| < F_j^{1-\mu} \left[\pi_{jt} - F_k^{0.5} \left[\pi_{kt} \right] \right] \right\}.$$

This is an application of the *metrically trimmed mean from the median*, see Kim (1992). This trimming is flexible in the sense that when the distribution is skewed to the right, most observations are trimmed in the right hand tail and vice versa. Kim (1992) notes that for symmetric distributions, the metrically trimmed mean is consistent and asymptotically Gaussian. The breakdown point of the estimator is μ against $\mu/2$ for the symmetrically trimmed mean, which implies that a certain level of robustness can be achieved at the loss of fewer observations and less information.

It may reduce the volatility further to remove whole subindices in the calculation of the metrically trimmed mean instead of down-weighting the marginal index. Intuitively, it seems strange to stop trimming in the middle of an index, and any non-negligible remain of a marginal index may be more like an outlier than a part of the central distribution. The counterpart to removing whole indices is that the *de facto* trimming percentage varies over time and $100 \cdot \mu$ per cent is a minimum[2].

4 Test of estimators

Three estimators are tested in this section: The metrical trim with removal of whole subindices, the metrical trim with precise trim percentage and the symmetrical trim with precise trim percentage. First, a Monte Carlo simulation is performed to illustrate the relative virtues of the estimators. Second, bootstrap drawings from the price data are ap-

[2] Removing whole indices is not relevant for symmetric trimming as it destroys the symmetry.

plied to confirm the Monte Carlo set-up and to determine the optimal estimator and trim percentage for the present data set.

Monte Carlo simulation. The Monte Carlo set-up is designed to imitate some important features of the empirical distribution of price changes. Most observations are drawn from a standard Gaussian distribution, but 10 per cent of the observations are contaminated with noise, in the sense that they are drawn from a more dispersed or displaced Gaussian distribution, i.e.

$$\pi_{it} = (1 - s_{it}) \cdot \epsilon_{it} + s_{it} \cdot f_t \cdot \eta_{it},$$

for $i = 1, ..., 121$, $t = 1, ..., T$, where

$$\epsilon_{it} \sim N(0,1) \quad and \quad \eta_{it} \sim N(a,b).$$

Here s_{it} is binomial distributed with probability $P(s_{it} = 1) = 0.1$ for contamination, and f_t is a binomial distributed sign shift with probability $P(f_t = 1) = d$. The sign shift implies that the noisy observations are drawn from either $N(a,b)$ or $N(-a,b)$, i.e. the location of potential outliers changes between left and right and the distribution is on average symmetric if $a = 0$ or $\gamma = 0.5$. In the simulation we consider the values $a = 0, 3$; $b = 1, 3, 6$, $d = 0.5, 0.6$ and for each combination (a, b, d) we draw $T = 1000$ cross sections and trim the data using the three methods and applying the most recent set of weights.

In the choice of estimator, π_t^μ, there is a traditional trade-off between bias and efficiency. We measure the bias as the average deviation from the expected value, i.e.

$$Bias(\pi_t^\mu) = \frac{1}{T} \sum_{t=1}^{T} (\pi_t^\mu - E[\pi_t]),$$

while the efficiency is measured as the variance of the trimmed series

$$Var(\pi_t^\mu) = \frac{1}{T-1} \sum_{t=1}^{T} \left(\pi_t^\mu - \frac{1}{T} \sum_{i=1}^{T} \pi_i^\mu \right)^2.$$

A standard weighting of the bias an efficiency is the Mean Squared Error given by $MSE = \frac{T-1}{T} Var(\pi_t^\mu) + (Bias(\pi_t^\mu))^2$.

The results are reported in figure 2 for the bias, variance and MSE. In the first column $(0, b, 0.5)$, the distribution is fully symmetric but is leptokurtic for $b > 1$. In the second column $(3, b, 0.5)$, the contamination is drawn from a displaced distribution and the skewness change from month to month. On average, however, the distribution is symmetric and $E[\pi_t] = 0$. In the last column $(3, b, 0.6)$, the combined distribution is skewed and on average more outliers are located in the right hand tail.

For the standard Gaussian distribution, $a = 0, b = 1$, all estimators are consistent and the bias is in all cases negligible. The simple arithmetic mean is as expected preferable to the trimmed means in terms of variance. This result is mirrored in cases with minor deviations from the Gaussian distribution. As either the dispersion or the numerical mean of the contaminating distribution increases, implying fat tails and/or skewness in the sample distributions, the scope for the trimmed means increases. If the distribution is on average skewed, we observe a trade-off between bias and efficiency, but the bias problem looks small compared to the variance reduction, and the MSE is in all cases practically indistinguisable from the variance. The metrically trimmed mean is as expected marginally more biased than the symmetrically but is strongly preferable in terms of efficiency.[3]

One important difference between the results for the metric and the symmetric estimator is that the lowest variance of the metric estimator is typically obtained with a trim percentage close to the contamination probability $\delta = 0.1$. With the symmetric estimator a much larger proportion of the observations have to be discarded to obtain efficient estimates, and often the most efficient candidate is close to the median.

It also appears to marginally reduce the variance of the metrical trim to remove entire subindices rather than to use a precise trim percentage. The difference is clearest in the case of modest trim percentages.

Bootstrap. To analyse how the above Monte Carlo results translate to the data set, we use a bootstrap strategy to approximate the underlying price distribution, see *inter alia* Bryan and Cecchetti (1999b). We assume to have 121 distributions characterising the subindices, and 238 observations on each distribution. We now construct $t = 1, ..., 1000$ bootstrap cross sections by drawing (with replacement) one observation from each distribution $i = 1, 2, ..., 121$ in each month. In order to eliminate the effect of the change over time in the overall inflation rate and to focus on the short-term variation, the trend in all subindices has been removed with an HP filter[4]. The results regarding variance are illustrated in first part of table 2[5].

The outcome is similar to the Monte Carlo simulations with high kurtosis. More specifically, cases $(0, 6, 0.5)$, $(3, 6, 0.5)$ and $(3, 6, 0.6)$ in figure 2 resemble the empirical price distribution. The overall lowest variance is obtained by trimming 25 per cent using the metrically trimmed mean and removing whole indices. This estimator reduces the variance by over 75 per cent, and is chosen as the preferred in the following.

The bootstrap results on relative variability are repeated when applying the three trim methods to the historical data $1981 - 2000$. Second part of table 2 illustrates the variance of the trimmed historical data calculated around the full sample means of the series. Again the

[3] The result on relative efficiency also holds for e.g. a uniform contaminating distribution instead of a Gaussian, and the conclusions are robust to other measures of efficiency such as Mean Absolute Deviation.

[4] A smoothing of $\lambda = 1000$ was chosen, but that is not crucial for the results. Using an HP trend to represent the mean of the distribution obviously implies a measurement error. It is assumed that these errors are too small to matter in a characterization of outliers and their impact.

[5] The applied ordinary bootstrap algorithm eliminates the correlation structure of the data set. To assess the robustness of the results to the autocorrelation of the price changes, the stationary bootstrap algorithm of Politis and Romano (1994) was also applied. The idea is to draw several consecutive price changes at a time, where the block length is geometrically distributed around a fixed mean. This preserves the autocorrelation *within* the drawn blocks. We also tried to draw contemporaneous blocks for all subindices, thus preserving also the cross correlation between the subindices. The outcome was in all cases very similar to the results reported for the ordinary bootstrap.

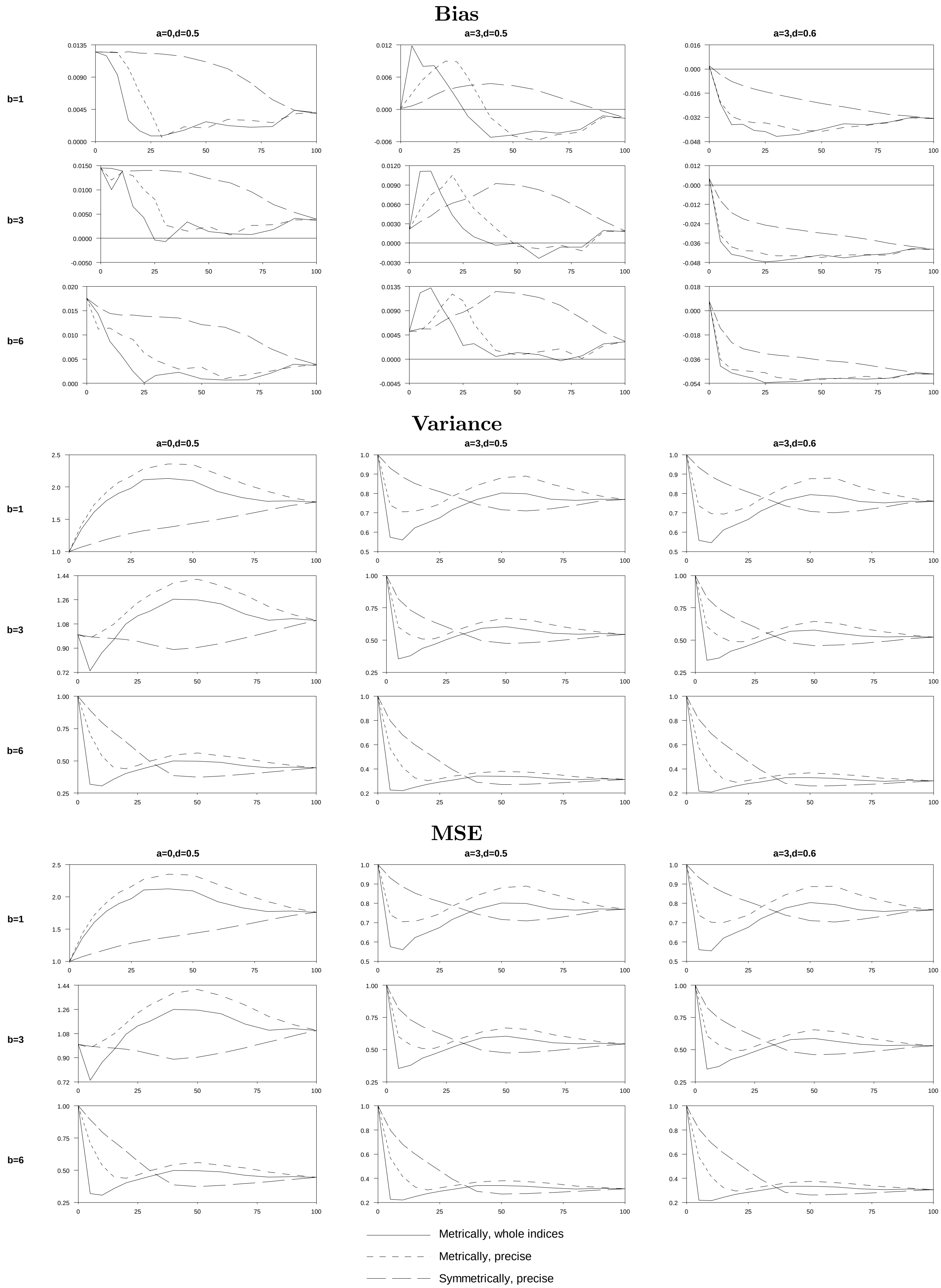

Figure 2: Monte Carlo Results. Bias, variance and MSE as a function of trim percentage.

Table 2: Variance of the trimmed series relative to the variance of arithmetic mean.

Trim percentage	0	5	10	15	20	25	30	40	50	60	70	80	90	100
							Bootstrapped cross sections							
Metrically, whole indices	1.000	.352	.254	.227	.218	.212	.224	.236	.253	.266	.260	.249	.261	.263
Metrically, precise	1.000	.398	.331	.282	.252	.242	.243	.257	.273	.291	.293	.281	.270	.263
Symmetrically, precise	1.000	.549	.441	.400	.372	.347	.324	.275	.240	.227	.227	.238	.252	.263
							Historical data							
Metrically, whole indices	1.000	.456	.341	.277	.252	.239	.244	.260	.263	.273	.276	.278	.279	.276
Metrically, precise	1.000	.515	.363	.311	.278	.253	.247	.256	.266	.272	.269	.271	.279	.276
Symmetrically, precise	1.000	.738	.581	.488	.429	.392	.369	.333	.310	.300	.290	.280	.271	.276

Note.: The bootstrap is based on 1000 simulations and the most recent set of weigths.

most stable estimate with a variance around $\frac{1}{4}$ of that in the untrimmed mean is obtained with the metrically trimmed mean and a trim percentage of 25. For the symmetrically trimmed mean, the most stable estimate is close to the median.

The true expected inflation is a theoretical magnitude, and we can only address the bias issue by making some assumptions. If for instance we want the same mean as the actual inflation over the full sample $1981 - 2000$, we should add 0.37 per cent p.a. to the trimmed inflation rate. Taking this as a bias measure implies that the squared bias amounts to 1 per cent of the difference in variance between trimmed and actual inflation. That is a relative magnitude as in the Monte Carlo experiment where for $(0, 6, 0.6)$ the squared bias amounts to $\frac{1}{2}$ per cent of the variance reduction at a trim percentage of 25. Instead of adding 0.37 per cent p.a. the sample mean of the actual inflation could also be reproduced by a metrical trim from the 60 percentile, but this would increase the simple variance in the measure by approximately 25 per cent compared to the preferred metrical trim from the median.

5 Results and applications

The changes in the arithmetic mean and in the preferred 25−per cent metrically trimmed mean from month to month and from year to year respectively are illustrated in figure 3 (A) and (B). The trimmed mean suggests a more stable picture where the underlying inflation is unchanged over most of the nineties. It appears that the sample difference in the year-on-year increases mostly relates to the first years where the overall inflation was still high. Thereafter the difference between trimmed and untrimmed mean becomes more unsystematic. The mentioned difference to actual inflation of 0.37 per cent p.a. for $1981 - 2000$ is reduced to a very insignificant 0.14 per cent p.a. for $1983 - 2000$.

Due to the strategy of excluding whole subindices, the *de facto* trim percentage changes over time. For the chosen sample and 25 per cent nominal trim the actual trimming percentage varies between precisely 25.00 and 42.71 per cent with an average of 26.56 per cent. The high trimming percentage that particular month reflects that subindex 54 of housing is the marginal index in the calculation.

Table 3 given on the last page illustrates how often the different subindices are removed in the preferred 25 per cent trimmed mean. As expected the prices of fuels and gasoline and of some unprocessed food articles disappear relatively often. A good part of the sales effect on garments is removed by the seasonal filter, but the timing of the winter and summer sales varies and leaves a lot of irregular price movements in not least women's garments.

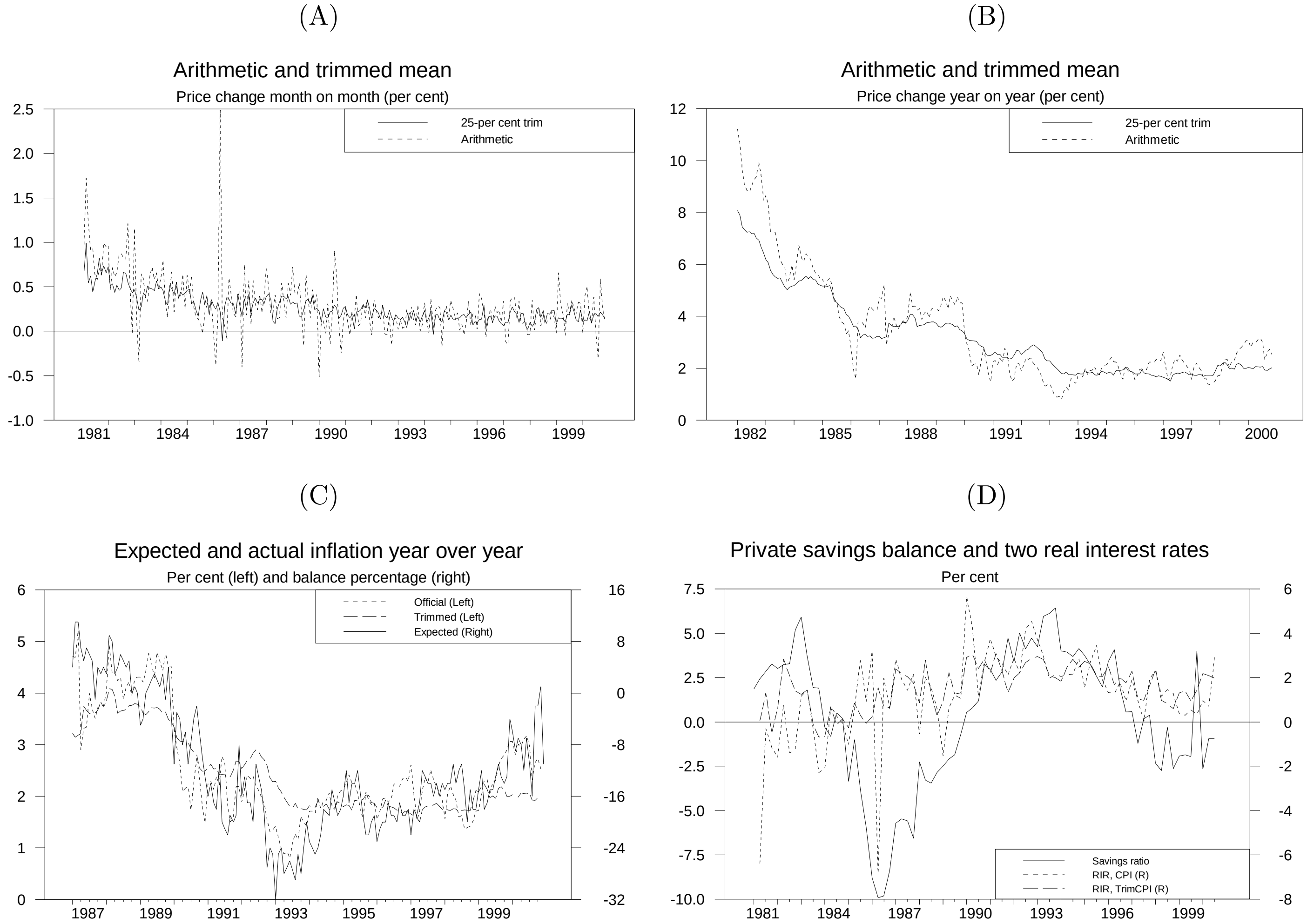

Figure 3. (A) and (B) compare the arithmetic mean with the preferred metrically trimmed mean. (C) Illustrates households expected inflation from the consumer surveys and the CPI and trimmed inflation. (D) Illustrates the private savings balance and real interest rates (RIR) based on the average bond yield and the increase in CPI and trimmed inflation respectively.

Prediction of inflation. The volatility of monthly CPI movements disturbs the autocorrelation of monthly price increases, and it is better to predict the monthly change in the full consumer price index by the less volatile monthly change in the trimmed index. A simple estimation yields

$$\log\left(\frac{P_t}{P_{t-1}}\right) = \underset{(0.1)}{0.00830} \log\left(\frac{P_{t-1}}{P_{t-2}}\right) + \underset{(7.5)}{1.08165} \log\left(\frac{\widehat{P}_{t-1}}{\widehat{P}_{t-2}}\right) + \underset{(0.0)}{0.00003}$$

for $t = 1981:3 - 2000:11$, where P_t is the price index accumulated from the arithmetic mean, $\overline{\pi}_t$, and $\widehat{P}_t$ is the index based on the 25 per cent metrically trimmed mean, $\widehat{\pi}_t^{25}$. Figures in parentheses are $t-$values.

Removing the first insignificant term increases the coefficient for the change in the trimmed index slightly to 1.092. A coefficient above one and a small positive constant lifts the forecast mean a little to make it equal the untrimmed inflation mean in the estimation sample. This automatic mean correction is not significant. Restricting the coefficient to one and the constant to zero can be done with an insignificant $F(2,235)$ statistic of 1.52.

Instead of the monthly price increase one often focuses on the relatively less volatile year-on-year change in the price index. Two consecutive year-on-year index changes have 11 monthly changes in common, and the ability of $\widehat{\pi}_t^{25}$ to predict π_t one month ahead can be brought to use by replacing the actual monthly price change 12 months back by the latest trimmed monthly change

$$\log\left(\frac{P_t}{P_{t-12}}\right) = -\underset{(0.1)}{0.0036} \log\left(\frac{P_{t-1}}{P_{t-13}}\right) + \underset{(16.3)}{1.0110} \left(\log\left(\frac{P_{t-1}}{P_{t-12}}\right) + \log\left(\frac{\widehat{P}_{t-1}}{\widehat{P}_{t-2}}\right)\right) - \underset{(0.2)}{0.0001}$$

The predictive gain reflects that a large price increase 11 months ago can make the present year-on-year increase both large and likely to moderate in the following month. The ability of the trimmed price change to predict the next actual price increase confirms that the trimmed change indicates a central tendency. This may be one point of departure for forecasting, but it goes without saying that other information can make the forecasted price increase differ from the latest observation on the trimmed increase.

Besides, the large price changes trimmed away are sometimes relevant to the forecasting. A jump in energy prices may first leave the trimmed inflation unchanged, but if energy prices stay up and trigger second round effects it will turn into a general price increase.

Household price expectations. The price expectations of Danish households have been measured by the monthly consumer survey since 1987. The answers to the survey's standard question on price increases are transformed to an index, which should correlate positively with the price increase expected over the coming 12 months. It turns out that this survey index correlates nicely with the year-on-year increase in the actual price index, better than with the year-on-year increase in the trimmed price index, cf. figure 3 (C).

Thus, households seem to expect the present headline inflation to repeat itself in the next 12 months. Specifically, households do not seem to weigh down the largest price movements. This may reflect that households get their information from the press mentioning the present headline inflation. Moreover, the press and the public debate are likely to focus on large movements in say energy, meat or coffee prices. However, the trimmed price increase may still be the most relevant for the investment and consumption decisions of economic agents.

Real interest rate. Another test could be to compare the use of trimmed inflation versus actual inflation in explaining consumption and investment. This amounts to testing whether trimmed price increases works better in the real rate of interest.

To simplify we concentrate on the private savings balance, i.e. savings minus investments in the total private sector comprising households and companies. This macro approach is facilitated by the Danish currency peg, which implies that the interest rate is given from abroad and not determined by the savings balance. The national account figures applied are quarterly and the price increases used are also quarterly. The nominal interest rate is the average bond yield after tax.

None of the two real interest rates (RIR) correlate that well with the savings balance, cf. figure 3 (D). Specifically, it seems difficult to explain the private savings balance during the eighties using only real interest rates. However, the real interest based on the trimmed price increase is less erratic and seems to be doing a little better as also indicated by a simple estimation

$$\begin{aligned} \textit{Savings balance} = {} & \underset{(4.0)}{2.16784} \ \textit{Interest rate} - \underset{(0.7)}{0.17588} \ \log\left(\frac{P_t}{P_{t-1}}\right) \\ & - \underset{(2.3)}{1.20755} \ \log\left(\frac{\widehat{P}_t}{\widehat{P}_{t-1}}\right) - \underset{(3.2)}{0.05552} \end{aligned}$$

estimated for quarterly data for $t = 1981:2 - 2000:3$.

6 Conclusion

This paper has focused on the trimmed mean as a robust estimator of expected inflation. The pattern of price changes over the last 20 years led us to prefer a metrical trim, which removes the largest outliers regardless of their sign. This seems to be the most efficient trim also taking into account the enhanced risk of bias compared to a symmetric trim. A bootstrap simulation suggested a trim percentage of 25 for the proposed metrical trim, and it indicated a minor advantage in removing whole subindices at a time instead of sticking to the 25 per cent sharp.

The trimmed inflation rate suggests a more stable picture of the actual inflation. This stabilised measure may be used to represent inflation or expected inflation in economic analyses, e.g. in applications involving a real interest rate. Trimmed inflation responds when the price development is general rather than concentrated on a few products. Furthermore the noise reduction enhances the correlation structure, which improves the ability of the trimmed rate to predict inflation. We are not suggesting to make this the only input in an inflation forecast, but the trimmed inflation could be a starting point.

In the economic policy framework, it may be advantageous to include a focus on the inflation trend represented by the trimmed inflation rate.

Appendix: The data

The price data used are the 121 subindices in the Danish CPI for the period January 1981 till November 2000, cf. table 3. All price data including the weighting scheme are from Statistics Denmark.

Statistics Denmark's revisions of weights are included in the calculation of our log indicator for the CPI. In the sample period the weight basis for the CPI was revised in April 1984, January 1991, September 1996 and December 1999. For instance the 121 monthly price changes March 1984 to April 1984 are weighted by the first set of weights, price changes April 1984 to May 1984 by the second. These revisions are intended to keep the content of the CPI up to date.

For one small item, services not included elsewhere, we did not have the official index but applied the price index for the main category miscellaneous goods up to 1991. Similarly, the price of the transport category is used for a small transport item, which only appears 1991-1996. These deviations to the official CPI calculation are too small to affect the results. Another minor data issue is that all subindices in the past were published without decimals although decimals were used for the calculation of the CPI. The increases the data noise but there is no remedy and, anyway, our log approximation based on a weighted average of changes in subindices comes close to the published CPI.

Consumers' price expectations are taken from Statistics Denmark's monthly consumer surveys. The quarterly private savings balance and the bond yield are from the data bank of the Mona model, Danmarks Nationalbank.

References

[1] Bakhshi, H. and T. Yates (1999) To trim or not to trim ? An application of the trimmed mean inflation estimator to the United Kingdom, Bank of England, Working Paper No. 97.

[2] Ball, L. and N.G. Mankiw (1995) Relative price changes as aggregate supply shocks, *The Quarterly Journal of Economics*, 110, 1, 161-193.

[3] Bryan, M.F. and S.G. Cecchetti (1999a) Inflation and the distribution of price changes, *Review of Economics and Statistics*, 81, 188-196.

[4] ——— and ——— (1999b) The Monthly Measurement of Core Inflation in Japan, Bank of Japan, Institute for Monetary and Economic Studies, Discussion Paper No. 99-E-4.

[5] ———, ——— and R.L. Wiggins II (1997) Efficient Inflation Estimation, NBER Working Paper No. 6183.

[6] Cecchetti, S.G. (1997) Measuring Short-run Inflation for Central Bankers, *Economic Review of the Federal Reserve Bank of St. Louis*, 79, 143-156.

[7] Kim, S.-J. (1992) The Metrically Trimmed Mean as a Robust Estimator of Location, *The Annals of Statistics*, 20, 3, 1534-1547.

[8] Mio, H. and M. Higo (1999) Underlying Inflation and the Distribution of Price Changes - Evidence from the Japanese Trimmed Mean CPI, Bank of Japan, Institute for Monetary and Economic Studies, Discussion Paper No. 99-E-5.

[9] Politis, D.N. and J.P. Romano (1994) The Stationary Bootstrap, *Journal of the American Statistical Association*, 89, 428, 1303-1313.

[10] Roger, S. (1997) A robust Measure of Inflation in New Zealand 1949-96, Reserve Bank of New Zealand, Discussion Paper No. G97/7.

Table 3: Proportion of the months a given subindex is trimmed with the metrical estimator and $\mu = .25$.

No.	Weigth	Index	Pct.	No.	Weigth	Index	Pct.
1	0.05	Rice	37.4	62	1.83	Furniture, etc.	8.0
2	0.09	Flour and cereals	30.3	63	0.26	Floor coverings	16.4
3	1.05	Bread, etc.	17.6	64	0.05	Rep. to furn. and floor cov.	10.1
4	0.51	Pastries, cakes, biscuits	19.3	65	0.77	Household textiles, etc.	29.4
5	0.21	Other cereal products, etc.	47.9	66	0.06	Repairs to other furnishings	17.6
6	0.70	Beef	49.6	67	0.82	Household appliances	25.6
7	0.10	Veal	39.5	68	0.09	Rep. to household appliances	12.6
8	0.61	Pigmeat	59.7	69	1.11	Glassware, tableware etc.	10.1
9	0.04	Lamb	71.4	70	0.31	Detergents	33.2
10	0.32	Poultry	43.3	71	0.66	Other non-dur. household goods	31.9
11	0.87	Cooked meat and sausages	40.8	72	0.09	Laundry, dry cleaning	13.4
12	0.71	Processed meat	18.5	73	0.13	Other household services	17.2
13	0.02	Other meat and offal	60.1	74	0.32	Domestic services	12.2
14	0.19	Fresh or frozen fish	60.1	75	0.77	Medicin and vitamins	32.4
15	0.08	Dried or smoked fish	56.3	76	0.61	Pharmaceut. and therapeutic eq.	14.3
16	0.23	Processed fish	26.5	77	0.05	Services of physicians	8.0
17	0.06	Other sea foods	62.2	78	0.75	Services of dentists	7.1
18	0.61	Fresh milk	35.3	79	0.46	Other health services	18.9
19	0.39	Other milk products	40.3	80	5.82	Purchase of car	16.4
20	0.65	Cheese	36.1	81	0.57	Purchase of other vehicle	15.1
21	0.19	Eggs	57.6	82	0.89	Tyres, spares, accessories	13.4
22	0.27	Butter	35.3	83	1.83	Repairs to vehicles	24.4
23	0.09	Margarines	46.2	84	2.81	Gasoline, oils and greases	76.9
24	0.05	Other oils and fats	37.0	85	0.67	Other expenditure	18.1
25	0.50	Fresh fruits	82.8	86	0.33	Taxis, removal, etc.	21.4
26	0.09	Dried fruits	52.5	87	0.46	Local bus transport	20.6
27	0.28	Frozen fruits, juices	44.5	88	0.81	Train and coach transport	20.2
28	0.65	Fresh vegetables	83.2	89	0.13	Sea transport	51.3
29	0.18	Frozen vegetables	39.9	90	0.21	Air transport	46.2
30	0.18	Processed vegetables	38.7	91	0.11	Postage	16.0
31	0.37	Potatoes	81.1	92	1.55	Telephone	12.6
32	0.09	Sugar	31.1	93	1.02	Radio and television sets	29.8
33	0.47	Coffee	67.2	94	0.05	Photographic equipment	36.6
34	0.06	Tea	45.0	95	1.17	Other major durable goods	40.8
35	0.01	Cocoa	33.2	96	0.38	Records, tapes, etc.	21.4
36	0.13	Preserves, jams, etc.	35.3	97	0.25	Sports and camping equipment	26.9
37	0.58	Chocolate	15.5	98	0.56	Games and toys	16.8
38	0.69	Sweets	20.2	99	0.06	Films, photo materials, etc.	18.9
39	0.38	Ice-cream	52.1	100	0.77	Flowers and plants	11.8
40	0.46	Condiments, spices, etc.	29.8	101	0.41	Pets, pet food and equipment	46.6
41	0.88	Non-alcoholic beverages	36.1	102	0.03	Other mat. for hobbies etc.	42.0
42	0.32	Alcohols (spirits)	13.9	103	0.19	Parts to recreat. equipment	18.1
43	1.03	Wines	17.2	104	0.35	Cinema, theatre, concerts	16.8
44	1.38	Beers	20.6	105	0.18	Other entertainment etc.	32.4
45	1.86	Cigarettes	16.8	106	0.51	Sporting	23.9
46	0.09	Cigars, incl. small cigars	18.9	107	1.06	Radio and TV licences	13.4
47	0.53	Other tobacco, etc.	31.1	108	1.20	Other recreational services	27.3
48	1.44	Men's garments	38.7	109	0.48	Books	21.0
49	2.21	Women's garments	60.5	110	1.28	Newspapers, magazines	22.3
50	0.70	Other cloth. acc., repairs	37.0	111	0.83	Education	11.3
51	0.37	Men's footwear	36.6	112	1.46	Children's day care instit.	13.4
52	0.65	Women's footwear	37.8	113	0.82	Hairdressers, beauty shops	19.3
53	0.03	Repairs, other footwear	17.6	114	0.14	Durable articles for pers.care	40.8
54	18.16	Accommodation: all-year dwel	8.4	115	0.97	Non-durable art. for pers.care	34.5
55	0.63	Accommodation: weekend dwel	10.5	116	0.25	Jewellery, watches etc.	14.7
56	1.44	Repairs and maintenance	16.4	117	0.40	Other personal goods	37.4
57	1.60	Water charges, refuse	57.1	118	0.21	Writing equipment, etc.	29.0
58	2.41	Electricity	24.4	119	4.00	Restaurants and cafes	9.2
59	0.70	Gas	65.1	120	0.57	Hotels	12.6
60	1.14	Liquid fuels	81.9	121	2.17	Serv. not elsewhere classified	10.1
61	2.12	Other fuels, etc.	25.2		100.00		

Note: Weigths are the most recent from 1999:12.