A Framework for Forecasting the Components of the Consumer Price Index: application to South Africa

JANINE ARON
Centre for the Study of African Economies,
Department of Economics, University of Oxford

JOHN N. J. MUELLBAUER
Nuffield College, University of Oxford

&

COEN PRETORIUS
South African Reserve Bank, Pretoria, South Africa

28 March 2004

Abstract: Inflation is a far from homogeneous phenomenon, but this fact is ignored in most work on consumer price inflation. Using a novel methodology grounded in theory, the ten sub-components of the consumer price index (excluding mortgage interest rates, or CPIX) for South Africa are modeled separately and forecast, four quarters ahead. The method combines equilibrium correction models in a rich multivariate form with the use of stochastic trends estimated by the Kalman filter to capture structural breaks and institutional change. This research is of considerable practical use for monetary policy, allowing sectoral sources of inflation to be identified. Aggregating the forecasts of the components with appropriate weights from the overall index, potentially indicates the gains to be made in forecasting the idiosyncratic sectoral behaviour of prices, over forecasting the overall consumer price index.

JEL codes: C22, C32, C51, C52, C53, E31, E52

* Paper presented at the CSAE Conference "Growth, Poverty Reduction and Human Development in Africa", St. Catherine's College, Oxford University, 21-22 March, 2004.

Earlier versions were presented at the South African Reserve Bank, April 2002 and December, 2003; at the Universities of Cape Town, Witwatersrand (joint with the Economic Association of South Africa) and Stellenbosch, August 2003; at Oxford University's Department of Economics Staff Seminar, November 2003; and at the University of Pretoria, December, 2003. We are very grateful for comments from these fora; and also for advice from C. Bowdler, G. Cameron, P. Collier and D. Hendry, Oxford University; B. Kahn, J. Van den Heever, O. Van der Merwe and P. Weideman, South African Reserve Bank; C. Morden, National Treasury; A. Black, University of Cape Town; G. Keeton, Anglo-American; J. du Toit, ABSA; and B. Smit, Bureau of Economic Research, South Africa. John Muellbauer acknowledges support from ESRC Grant No. 000 23-0244. We acknowledge funding from the Department for International Development (DFID), U.K., for the project "*Governance and Inflation Targeting in South Africa",* grant number ESCOR 7911. DFID supports policies, programmes and projects to promote international development. DFID provided funds for this study as part of that objective but the views and opinions expressed are those of the authors alone.

1. Introduction

South Africa adopted inflation targeting in 2000, departing from the previously applied "eclectic" monetary policy approach in which intermediate objectives, such as the growth in the money supply, played a prominent role. The move to inflation targeting aims to enhance policy transparency and accountability and thereby to decrease inflationary expectations, as well as private sector uncertainty. Currently, the inflation target has been specified as achieving a rate of increase in the overall consumer price index, excluding mortgage interest cost (the so-called CPIX), of between 3 and 6 percent per year.[1]

Inflation targeting is a forward-looking approach, with monetary policy based on the likely path of inflation. The fact that a definite time horizon (two years) is specified in inflation targeting makes it important that the central bank has a reliable forecasting framework. Amongst other requirements, the shift to inflation targeting demands good forecasting models of inflation and clarity on the mechanisms of monetary transmission (Leiderman and Svensson, 1995). The emphasis of the modelling activities in the South African Reserve Bank has shifted away from the maintenance of a single large-scale macroeconomic model towards a more compact or core model, supplemented by various other models. This is also in line with the international trend of using a "suite of models" approach.

In practice, forecasters employ a range of different approaches to forecast inflation. Most inflation models tend to forecast the total price index, e.g. the consumer price index (CPI or CPIX). Another less formal approach examines trends in different components of the price index, such as price indices for food, fuel, durable goods, financial and other services, housing and others. These trends are then projected ahead, often using fairly crude methods. The latter approach is under-researched and there are few formal econometric models in the literature which use information on the sub-components of the consumer price index to help forecast the overall CPI index. The two approaches are rarely combined, though Bryan and Cecchetti (1999) have examined the relationship between overall inflation and changes in the components of the consumer price index. Until quite recently, the role of relative price movements in the modelling of inflation has been neglected.

[1] A tightening in monetary policy to counter inflation pressures would cause interest rates to rise and be reflected in the interest cost component of measured inflation. This, in turn, could provoke a further tightening of monetary policy resulting in excessive movements in the inflation rate.

In a Governor's speech, the Bank of Norway refers to the use of a sub-index model for forecasting inflation, which model is described in a note by Akram and Bache (2001, English abstract). Single equation equilibrium correction models are apparently reported for nine consumer price sub-indices from 1986q1-1999q4.[2] All equations (save for housing rent and for fuels) employ a simple mark-up model: in the long-run prices are a weighted average of unit labour costs and foreign prices (converted to local currency). However, tax changes and domestic demand pressures are also allowed for in the dynamics. Fuel prices are modelled as a function of local currency oil prices and petrol taxes. Housing rent is based on the implicit user cost of housing capital owned by households.

In Espasa, Poncela and Senra (2002), U.S. CPI is broken down into four sub-components, corresponding to four groups of markets: energy, food, rest of commodities, and the rest of services. The different trending behaviour in the market prices suggests this disaggregation could help improve forecast accuracy. The individual components were forecast a year ahead for non-energy CPI, and it was found that aggregating the forecasts improved the overall accuracy for non-energy CPI by 23 percent, based on the root mean squared error.[3] Univariate ARIMA models in second differences are reported both for the one to twelve-month ahead forecasts of non-energy CPI and for its components.

In this paper, we span the two approaches, and use more sophisticated methods to forecast the sub-components of the price index. Our own approach is considerably more general, incorporating richly specified equilibrium correction models and stochastic trends We suggest modelling these components separately since international competition, changes in technology, the exchange rate, foreign price influences, domestic interest rate changes and housing market developments, will affect the various components differently. For instance, "clothing and footwear", and "furniture and equipment" categories have declined in relative price since 1970, while the relative price of "vehicles" rose from 1985 to 1994, probably related to import tariffs.

Four-quarter ahead forecasting models for the main sub-indices of CPIX are developed, combining equilibrium correction and stochastic trends, estimated using the Kalman filter. This type of model focuses on the long-run properties of the model, with

---

[2] The nine components are: agricultural and fishery products; other non-traded consumer goods (e.g. electricity); traded consumer goods (excluding fuels and lubricants); imported consumer goods without local competition; competing imported goods; housing rent; other services (with wages dominant); other services; and fuels and lubricants.

[3] This disaggregated approach draws on work by Espasa et al (1987) for Spanish inflation (see also Espasa and Matea, 1991). This work has been extended to EMU inflation by countries and sectors (Espasa et al, 2002,

inflation viewed as part of the process of relative price adjustment. The long-run solutions of the models include relative prices of sub-components of the index to the index as a whole. Economic interpretations of these long-run relative price adjustments are provided, and explanatory variables such as import liberalisation, the real exchange rate, world prices and changing technology are used in the modeling process.

The main aims of the paper are to develop a framework for more accurate overall inflation forecasting, and to develop a better understanding of inflationary pressures for particular components of the basket of consumer spending. Hypotheses about sectoral transmission of policy and shocks are often more specific than hypotheses about overall transmission. By exploiting idiosyncratic movements and separately forecasting the main components, such as price indices for food, housing, furniture, transport and others, possibly a more accurate aggregate price index forecast can then be derived from the individual components using the appropriate weights in the index. These weights shift at discrete intervals and structural shifts in the determination of CPIX can arise from such shifts. In a follow-up paper, we explore whether forecasts from this novel disaggregated approach adds anything of value to four-quarter-ahead forecasts for the aggregate index alone.

An improved understanding of inflationary pressures for particular components of the basket of consumer spending should help targeting micro-economic policy interventions, perhaps involving deregulation or the competition authorities. British experience in recent years suggests such policies can be useful complements to the objectives of monetary policy. This would complement methods already used at the SARB for forecasting CPI components on a monthly basis.

It should be noted that there are few research articles describing the modelling and forecasting of the inflation process in South Africa. Pretorius and Smal (1994) describes the price formation process in South Africa as a process in which output prices are mainly determined as a fixed mark-up over costs. The main conclusions are that changes in labour costs are at the core of the inflation process and that wage changes are largely driven by inflation expectations. Inflation expectations react very slowly to conventional monetary policy. Fedderke and Schaling (2000) use an augmented Phillips curve framework to investigate the link between inflation, unit labour costs, the output gap, the real exchange rate and inflation expectations. They also found evidence for mark-up behaviour of output prices over unit labour costs, which in turn are driven by inflation expectations. Aron, Muellbauer

2004). For annual GDP, disaggregation has also been applied by Zellner and Tobias (2000) and Zellner and Chen (2000).

and Smit (2003) estimate a seven equation model of the determinants of the inflation process in South Africa.[4] Information is provided on exchange rate pass-through and the various channels through which monetary policy influences inflation. The decline in inflation from 1990 is mainly ascribed to increased openness to international competition, lower world inflation and tight monetary policy.

The next section defines measures of consumer price inflation; and section 3 discusses the trends in the relative price components. This is followed by a description of the theoretical forecasting model framework in section 4; and the estimated equations of the price components are presented in section 5. In the conclusion, some policy implications are discussed.

2. Definitions of consumer price indices in South Africa[5]

The "Survey of Consumer Prices" is a monthly survey covering a sample of retailers operating in the South African economy. It is combined with price data obtained directly from insurance companies, electricity companies etc. to obtain prices for the Consumer Price Index (CPI). The weighting system for the CPI is calculated from the "Survey of Income and Expenditure of Households", last conducted in October 2000. The information obtained through this survey was reweighted according to the 1996 Population Census figures in order to represent all households in South Africa. Statistics SA conducts a "Survey of Income and Expenditure of Households" every five years, covering a sample of 30,000 households. In the year 2000 the survey collected information on approximately 1,000 different goods and services groups. Statistics SA made a further breakdown of these groups using supplementary sources. This process led to a list of approximately 1,500 groups on which the current calculation of the CPI is based.

The "Survey of Retail Prices" is a retail trade and service outlets sample survey covering prices of selected consumer goods and services sold to consumers in the 14 metropolitan and 39 other urban areas in the nine provinces[6]. Currently, an average of

---

[4] Including the exchange rate, import prices, unit labour costs, wholesale prices, food prices, house prices, and consumer prices.

[5] The consumer price deflator was used for forecasting inflation, prior to the publication of CPIX, in Aron and Muellbauer (2000c).

[6] Before 1996, the survey was carried out only in metropolitan areas. Before 2000, the CPI referred only to metropolitan areas, and then switched to the metropolitan and urban basis, computed retrospectively back to 1997.

110,000 price quotations are collected each month from approximately 2,200 outlets by means of 6,700 questionnaires. The indices are based on retail trade and service prices. Price information refers to the first seven days of the relevant month. The collection of prices depends on the frequency at which these prices tend to change[7].

2.1 The headline CPI

The CPI is a chained Laspèyres index with weights derived from consumer expenditure surveys in 2000, 1995, 1990, 1985 and earlier. Given processing delays, the 1990 weights were applied from August 1991 to December 1996, the 1995 weights from January 1997 to December 2001, and the 2000 weights from January 2002.[8] For each period of roughly 5 years, the index takes the form

$$\text{CPI} = \frac{\sum_i p_{it} q_i}{\sum_i p_{i0} q_{i0}} = \frac{\sum_i \frac{p_{it}}{p_{i0}} p_{i0} q_{i0}}{\sum_i p_{i0} q_{i0}} = \sum_i w_{i0} (p_{it} / p_{i0}) \tag{1}$$

where, for example, p_{i0} is the average price for the base period of the i$^{\text{th}}$ type of good, i=1, n, and q_{i0} is the quantity weight derived from the base period consumer expenditure survey. Note that $w_{i0} = p_{i0} q_{i0} / \sum_{j=1}^{n} p_{j0} q_{i0}$ is the survey expenditure share. Thus, the index currently applies the 2000 survey expenditure shares to the relative prices (p_{it} / p_{i0}), where t begins in January 2002, when t=0.

[7] Prices of items collected *monthly* include: bread, milk, meat, vegetables and fruit, other groceries, alcoholic beverages, sweets, non-alcoholic beverages, ice-cream and tobacco products, clothing and footwear; repairs of clothing, footwear and furniture; interest rates on mortgage bonds; coal and wood; new vehicles, repairs and services; motor spare parts and accessories; petrol and diesel. Prices of items collected *quarterly* are, in January, April, July and October: garden tools; washing, ironing and dry-cleaning; sport equipment; reading matter and stationery; tariffs of hairdressing services; in February, May, August and November: ironware and crockery and new and retread tyres; and in March, June, September and December: furniture and equipment; household textiles; electrical appliances and equipment; medical, toilet and photographic requisites and services and motor vehicle insurance. Prices of items collected *annually* are doctor's and dentist's fees; motor vehicle license and registration fees; toll-fees at toll-gates; school funds; university boarding and class fees; parking fees; telephone and postal tariffs; public transport tariffs; property taxes; refuse removal; sanitary fees; newspapers and magazines; entrance fees – drive-inns and cinemas; television licenses; maintenance of graves; and rent of dwellings. Prices of items collected *at other times* of the year are winter clothing; medicine; contribution to medical aid; property insurance; hospital fees; water; electricity; air transport fees and dog licenses.

[8] In earlier years, the weights were held constant for 1960 to 1977, January 1978 to October 1987, and November 1987 to July 1991.

2.2 CPI excluding interest rates on mortgage bonds (CPIX)

CPIX is defined as overall CPI excluding interest rates on mortgage bonds (the mortgage cost component of homebuyer's cost of housing). The Reserve Bank uses this measure for the inflation target, covering both metropolitan and urban areas.

2.3 Core inflation

For completeness, we include a definition of core inflation, which captures the underlying inflation pressures in the economy. The core index is derived by excluding items from the overall CPI basket (metropolitan and other urban areas) on the basis that changes in their prices are highly volatile, subject to temporary influences, or affected by government intervention and policy. Fresh and frozen meat and fish are excluded as prices may be highly volatile, particularly during and following periods of drought. Fresh and frozen vegetables and fresh fruit and nuts are excluded as prices may be highly volatile from month to month due to their sensitivity to climatic conditions. Interest rates on mortgage bonds are excluded due to their "perverse" effect on the CPI. Changes in VAT (Value Added Tax) are predominantly determined by government, and are also excluded. Assessment rates are predominantly determined by local government and are excluded.

3. Trends in relative price components

We have tracked the relative prices, p_i/CPI, of the 10 main components, i, of the CPI since 1970 (Figure 1). We also present the relative prices, p_i/CPIXC, using our constructed measure of CPIX ("metropolitan"), see Aron and Muellbauer (2004a). Figure 1 contains some surprising results.

3.1 Goods components

3.1.1 Food

The weight of food in the total CPI index for metropolitan areas for the base year 2000 is 22.1. However the weights differ very much over the expenditure groups, from 51.39 for the very low-income groups (up to R8,070 per annum), to 15.82 for the very high-income groups (R55,160 and more per annum). The food component is one of the most volatile components in the CPI index and the items contributing most to this volatility are vegetables, meat and fruits and nuts. The food components with the highest weights include meat (5.66), grain products (3.81), vegetables (2.00) and other food products (3.45).

The graph indicates that food prices have tended to increase faster than the overall consumer price index since 1970, though between 1995 and 2001 the relative price was stable. Although a fairly small percentage of food products are imported directly from abroad, the depreciating currency may have contributed to the increasing trend via the effect of rising transport cost and import tariffs. Erratic weather conditions have an important influence on the supply and the prices of food products. The sharp increases in the early 1980s and early 1990s can be associated with the drought conditions in South Africa over that period. It is interesting to note that meat prices, which have a relatively large weight in food prices (27 percent), tend to have a dampening impact on food price increases during periods of droughts. Farmers are usually forced to step up the marketing of livestock, and the increase in the supply of meat can dampen price increases in meat. However, once rainfall returns to normal, farmers usually replenish their livestock and the resultant lowering of supply could cause meat prices to increase faster. During droughts, the prices of items in the food price index such as milk, milk products and grain products also tend to accelerate.

Since the middle of 2001, prices in the main components of the food price index have risen sharply. More specifically, the price of both white and yellow maize more than doubled from June, 2001 to January, 2002. This price increase had a significant impact on the economy, since white maize is a staple food for many South Africans, while yellow maize is used as feedstock in the meat, dairy, poultry and egg industries.

The agricultural marketing boards were abandoned by 1996 (some earlier) and thereafter agricultural pricing has been determined by export and import parity pricing, as well as supply and demand factors affecting crop outputs. After the sharp depreciation of the Rand towards the end of 2001, maize prices increased to record Rand price levels. They moved away from trading at export parity prices to higher import parity prices because of lower domestic production and high regional demand as a result of low crops in neighbouring countries (see *Monetary Policy Review*, SARB, March, 2002). A similar pattern was also applicable to other grain products such as wheat and sunflower seeds.

3.1.2 Furniture and equipment

The weight of furniture and equipment in the total CPI index for metropolitan areas for the base year 2000 is 2.5 percent. This component comprises of three sub-components, namely furniture, appliances, and other household equipment and textiles. Price information is normally collected on a quarterly basis.

Furniture, with a weight of 0.95 percent, includes all household furniture, kitchen units, floor coverings and repair of furniture. Appliances, with a weight of 0.80 percent, include all electrical household appliances such as refrigerators, stoves and washing machines as well as non-electrical appliances. Other household equipment and textiles, with a weight of 0.78 percent includes such items as glassware, curtains, blankets and gardening equipment.

The strong declining trend in the relative price of the furniture component in Figure 1(6) clearly illustrates that prices of furniture and equipment items have increased at a much slower rate than the overall consumer price index. This is likely to reflect more rapid productivity growth in this sector and the opening of the economy to import competition.

3.1.3 Clothing and footwear

The weight of clothing and footwear in the total CPI index for metropolitan areas for the base year 2000 is 3.2 percent. The clothing component has a weight of 2.04 percent and includes all items of clothing as well as material, knitting wool and the cost of designing and repairing clothes. The footwear component has a weight of 1.21 percent. The information on the prices of the clothing and footwear category is collected on a monthly basis.

The graph indicates that prices of clothing and footwear show a strong declining trend relative to overall consumer prices, probably because of greater productivity growth. The more rapid decline since 1990 can be ascribed to greater international competition and the phasing out of import tariffs on clothing and footwear products.

3.1.4 Vehicles

The weight of vehicles in the total CPI index for metropolitan areas for the base year 2000 is 6.0 percent. This component includes prices of both new and used vehicles and price information is collected on a monthly basis.

The relative price ratio of vehicles shows a strong increase from 1985 to 1995, indicating that the prices of vehicles have increased much faster than the overall consumer price index. A major factor has been the weaker Rand after 1984, contributing to the increased cost of imported vehicles and imported vehicle components. Prices were steady from 1993, and from 1995, the relative prices of vehicles have been on a declining trend before stabilising from about 2000.

Trade policy has also influenced prices. Tariffs on built up vehicles were prohibitive (over 100 percent) until the early 1990s, with only exotic cars being imported. From 1984, the slightly higher local content may have influenced the price rise, but in practice, local content measured by value may well have declined (the local content requirement was set in volume terms). In 1989, local content requirements were effectively reduced; and tariffs were reduced slightly in 1993/94, and then phased down from 1995 (currently they are 38 percent). Local content requirements were abolished in 1995. Thus, the increased international competition and greater openness from 1995 probably contributed towards a slower rate of increase in vehicle prices.

Import duties on vehicles and components can be rebated by exporting, so in practice there is much more import competition than previously, and imports account for around 25 percent of the new vehicle market. There is currently an excise duty on new vehicles, the rate of which varies according to the value of the vehicle.

While this background helps explain the pattern of relative price change, the extent of the relative price rise between 1984 and 1993, when other transport goods, clothes and footwear and furniture and equipment all declined in relative terms, remains surprising, see further discussion in section 4.

3.1.5 Transport goods

The weight of transport goods in the total CPI index for metropolitan areas for the base year 2000 is 5.5 percent. This component includes prices of petrol and diesel (4.92 percent) as well as other running cost items such as oil, tyres, batteries and spare parts. Most of the price information is collected on a monthly basis. The price of petrol and diesel are classified as administrative controlled prices.

The running costs, which contain a substantial fuel element, respond to the second oil price shock (but not the first) and decline relative to the CPI from 1985-87 with the fall in oil prices, and thereafter were relatively stable until the oil price rises from 1999. Thus, the fairly volatile relative price ratio could be ascribed to fairly frequent changes in motor fuel prices, responding to world crude oil prices, and the impact of changes in the exchange rate.

3.1.6 Beverages and tobacco

The weight of beverage and tobacco in the total CPI index for metropolitan areas for the base year 2000 is 2.5 percent. This component includes alcoholic beverages and tobacco products. Non-alcoholic beverages are included in the food component. The weight of alcoholic beverages in the total CPI is 1.40 percent, and includes products such as wine, beer and spirits. The weight of the tobacco products is 1.14 percent and includes cigarettes, cigars and tobacco. Price information is collected on a monthly basis.

The relative price of "beverages and tobacco" has risen since 1988. The sharp upward movement in the relative price index since 1994 can be attributed to the high annual increases in excise duties on alcoholic and tobacco products, which far outpaced inflation.

3.1.7 *Other goods*

The weight of other goods in the total CPI index for metropolitan areas for the base year 2000, is 15.3 percent. This component includes prices for a wide variety of products.[9] The relative prices of the

[9] "Fuel and power", with a weight of 3.49, covers electricity, gas, spirits, wood, coal and other petroleum products. Fuel and power prices are classified as an administered controlled price and price information is collected in January, July and August. "Household consumables" (1.25) include household and swimming pool cleaning materials, fertilizer and other household consumables. "Reading matter" (0.39) includes books, reading matter and magazines. "Recreation and entertainment equipment" (2.44) includes music instruments, TV sets, sport and camping equipment, pet food, plants and computer and telecommunication equipment. "Water" (1.37), where the price is an administered controlled price, and price information is collected on an annual basis. "Personal care products" (3.06) include personal care products such as soap, skin care products, perfume and

other goods category show a strong upward movement, especially from 1985 onwards. The components showing the biggest increase include household consumables, medical products, water tariffs and personal care products.

Prices of household consumables increased by 22 percent in 1985 and by another 40 percent in 1986. This was considerably faster than price increases in other products and may be attributed to the sharp depreciation of the Rand over the same period. This component includes cleaning and chemical products with a high import content. Increases in prices of household consumables continued to outpace the average inflation for most of the remainder of the sample period.

Prices of medical products also showed a fairly substantial increase, especially from 1990 onwards. On average, they increased by 4.5 percentage points faster than the overall cpi-inflation rate over this period. The increased cost of imported medicine could be part of the explanation.

Water tariffs also accelerated from the early 1990s, outpacing the average rate of inflation. The shortage of water during periods of drought and a rapidly increasing population was at the base of a decision to invest in water resources. After the 1994 election a concerted effort was also made to make access to drinking water available to more people, resulting in an increase in general water tariffs.

The relative price of personal care products such as perfume, deodorant, hair preparations also showed a noticeable increase during the 1980s, but has since then, on average, increased in line with prices of other products.

3.2 Services components

3.2.1 Housing

The weight of housing in the total CPI index for metropolitan areas for the base year 2000 is 24.3 percent. This component includes prices for rent (4.56), homeowner's costs (15.21), domestic workers (3.48) and boarding costs (1.0). The rental component comprises house rent, flat rent and town house rent. The interest costs (11.43) is the main element of homeowner's cost. The rest of the homeowner's costs consists of assessment rates, sanitary services, refuse removal, insurance, and maintenance costs.

---

other. "Medical care and health products" (2.77) includes medical and pharmaceutical products. "Other goods" (0.54) includes watches, jewelry, luggage bags and other.

This treatment of home-owners' costs can be criticized from several points of view. It measures only the cost of borrowing a given sum of money, not any increase in the price of housing that the given sum of money can buy, as noted by Haglund (2000). Note that the average level of nominal interest rates in the recent past is no higher than it was in the 1980s, while the price of housing has risen along with that of other goods. Hence this treatment of home-owners' costs in the CPI will result in the housing component of the CPI and hence the total CPI increasing *less* in the long-run than CPIX. And it neglects the fact that, in the context of an increasingly liberal mortgage market, an increase in nominal interest rates caused by a rise in general inflation, may not have the same cash flow implications as was once the case. Households with significant net equity can refinance and stabilize the real cash flow burden of their mortgage debt.

Figure 1 shows the ratio to CPI of total housing and of housing excluding the mortgage interest costs. The latter is also shown relative to CPIX. Note that the econometric work presented in this paper is for forecasting CPIX (as defined for metropolitan areas – see section 5.1), not the headline CPI. Therefore, the housing CPI component we model below, excludes the mortgage bond rate. The relative price ratio of the total housing component increased noticeably during the debt crisis period from 1984 to 1985, reflecting the substantial increase in interest rates. The relative price ratio thereafter declined sharply as prime interest rates declined from 25 percent at the end of 1985 to 12.5 percent at the end of 1987. More recently, the price ratio rose when interest rates increased in 1998 by more than 5 percentage points.

For housing, rents and mortgage bond rates move differently and are driven by different factors. We expect rents to adjust slowly to house prices and interest rates. The relative price of total "housing" in the CPI declines over time as noted above, despite the sharp increases in the mortgage bond rate in the 1997-1999 period.

The measurement of rents has gone badly wrong in recent years after the rent survey was discontinued in 1999, as a cost-cutting measure. In April 2003, Statistics South Africa were obliged to revise the data back to January, 2002, bringing down the overall inflation rate by almost 2 percent from its peak rate since January, 2002. In earlier years, the rent data were based on an annual survey carried out in October, and analysed by January, so that, for seasonally adjusted data, other months were interpolated.

3.2.2 Transport services

The weight of transport services in the total CPI index for metropolitan areas for the base year 2000 is 3.38 percent. The transport services component includes prices of running costs such as repairs, servicing, retreading, washing and related services; licences and registration fees; insurance; parking fees and other running costs as well as public and hired transport. The biggest weight is assigned to public and hired transport (1.84 percent). Public and hired transport is subdivided into bus (0.3), train (0.14) aircraft (0.22) and taxi and hired transport (1.16) and other (0.02). Some of the price information is collected on a monthly basis (vehicle servicing costs), while others are collected on a quarterly basis (insurance) or annually (public transport). Certain components of public transport, namely bus and train tickets, as well as licence and registration fees are categorised as administrative controlled prices. Their combined weight is 0.52 percent.

The relative price of "transport services" declines from 1985 to 1989, paradoxically at the same time as the relative price of vehicles is rising at its most rapid rate. A possible explanation for the renewed decline since 1995 could be the change in the government. An effort was made to keep the regulated ticket prices of buses and trains affordable for workers and therefore it increased at a lower rate than total CPI. The average annual inflation rate for transport services was less than 3 percent per annum from 1997 onwards.

3.2.3 Other services

The weight of other services in the total CPI index for metropolitan areas for the base year 2000 is 15.2 percent, and the component includes prices of a wide variety of products.[10]

Prices of the other services components have increased considerably faster than the total CPI index, especially from 1990 onwards. The main components contributing to this trend include educational costs, the cost of medical services and other services, which encompass banking costs.

The cost of education increased by 58.6 percent in 1993, compared to an average inflation rate of 9.7 percent. This can be attributed to the decision to extend and improve

[10] "Household services" (0.09) includes laundry and dry-cleaning and other services. "Communication" (2.98) includes internet, telephone and postage services. Prices of this component are classified as part of the administered price component, and are collected annually in April. "Education" (3.48) includes tuition and attendance fees. Educational costs are classified as an administered prices component and price information is normally collected annually in March. "Medical services" (4.38) includes doctors and hospital related fees, contributions to aid funds and insurance. Medical service costs are administered prices, and price information is collected on a quarterly basis. "Personal care services" (0.61) includes beauty care and hair dressing services. "Recreation and entertainment services" include television licenses, membership fees for clubs and libraries.

education to all population groups. Subsidies to tertiary institutions and the so-called Model C schools were reduced substantially, increasing the cost of education. Compared to the average inflation rate, the cost of education increased by almost 6 percentage points per annum faster over the entire sample period.

The prices of medical services also outstripped the average price increases in the total CPI index constantly from 1989 onwards. The increased cost of medical services can be linked to the privatisation of some hospitals and clinics as well as the increased contribution to medical aid contributions. A concerted effort was made to make medical services accessible to all South Africans.

The other services component also increased at a much faster rate since 1990, reflecting the relative increased costs associated with bank charges and interest on loans. The annual average increase in the price increases of the other services component was almost 3 percentage points higher than the total inflation rate since 1990.

4. Methodology: using sub-indices to help forecast the consumer price index

4.1 Multi-step forecasting and structural breaks

In our approach, the dependent variable is the four-period-ahead rate of inflation, in single equation equilibrium correction models. The multi-step forecasting single equation models developed here have the advantage of simplicity over a full Vector Autoregressive Model (VAR). There are substantial difficulties in interpreting and using VARs for policy and forecasting, arising from omitted variables (a restriction in another form), omitted structural breaks and relevant lags, omitted non-linearities, and the use of sometimes doubtful identifying restrictions to give economic interpretations to shocks. Multi-step models for inflation forecasting have been popularised by Stock and Watson (1999, 2001).

Methodologically, multi-step models can be regarded as single equation, reduced-forms of the related VAR system. Recent research suggests that where VAR models suffer from specification errors such as omitted moving average error components or certain kinds of structural breaks – both important in South Africa - single-equation, multi-step models can provide more robust forecasts (Weiss, 1991; Clements and Hendry, 1996, 1998).

Television licenses are classified as an administered price and measured annually in October. “Other services”

The impact of institutional changes or structural breaks is usually difficult to model. We pay careful attention to testing for structural breaks, and, where necessary and possible, model them. Otherwise, we use a smooth non-linear stochastic trend to help capture such shifts - effectively the Kalman filter applied to a time-varying intercept - while VAR models generally do not.[11] We follow Harvey (1993) and Harvey and Jaeger (1993) in defining the stochastic trend μ_t as follows:

$$\begin{aligned} \mu_t &= \mu_{t-1} + \gamma_t + \eta_{1t} \\ \gamma_t &= \gamma_{t-1} + \eta_{2t} \end{aligned} \tag{2}$$

where η_{it} are white noise errors. When var $\eta_{2t} = 0$, μ_t is an I(1) trend with drift. When var $\eta_{1t} = 0$, μ_t is a smooth I(2) trend, and this is the type we use to capture the evolution of the supply side. These non-linear trends can be estimated, via the Kalman filter, in the STAMP package (Koopman et al, 2000).

We begin with richly parameterised equations. Economic principles imply a strong set of sign priors on the effects of variables in the long-run, and often also in the dynamics. Using a general-to-specific model selection strategy, these equations are reduced to parsimonious forms on a single equation basis, guided by the sign priors. In Aron, Muellbauer and Smit (2003) we describe the process of reducing the dynamic variables, where the coefficients may suggest moving averages, or longer changes than one quarter, or lagging the level of the price terms. We do not carry out formal tests for cointegration since most of our equations include stochastic trends, and we cannot then compute the size of the relevant test statistics.

We are able to test for long lags without loss of parsimony by restricting the nature of longer lags. In most VARs, lag lengths are restricted to one or two quarters and rarely beyond four. The longer the lag, the less likely is it that the precise timing can be estimated, since one expects such effects to die away as the lag length increases e.g., distinguishing between effects at t-5, t-6 and t-7. This supports the case for allowing the possibility of longer lags, but restricting the effects to $\Delta_4 X_{t-4}$ and $\Delta_4 X_{t-8}$ or 4-quarter moving averages at such lags. For a variable such as the terms of trade, subject to erratic movements, annual or biannual changes or moving averages tend to smooth out erratic jumps in the data.

(2.78) include interest on bank and loan charges, gambling and funeral services.

[11] In some VAR studies, the Hodrick-Prescott filter is used to de-trend output and other variables in advance of estimation. While this is different, and in our view inferior, to estimating the trend within the model (see Harvey and Jaeger, 1993), in some contexts the Hodrick-Prescott filter may give broadly similar results.

We also check for asymmetries, for example, that in the short-run, oil price increases are passed on more quickly than price decreases. Asymmetric effects can be tested for as follows. Suppose the short-run inflationary effect of a 10 percent rise in the oil price is more than the dis-inflationary effect of a 10 percent price fall. If ΔX is the change in the log oil price, we can test for this through including the terms: $\varsigma_1 \Delta X + \varsigma_2 |\Delta X|$. If $\varsigma_2 = \varsigma_1$, the effect of a fall in the oil price is zero, while the effect of a rise is $2\varsigma_1$. If $\varsigma_2 = 0$, there is no asymmetry. If $0 < \varsigma_2 < \varsigma_1$, we have an intermediary case, in which the effect of an oil price fall is less than that of an increase.

4.2 A forecasting model framework

We now propose a new modelling framework for forecasting 4-quarter-ahead inflation rates for each of the ten components of the CPI (or other aggregate price measures).

We propose a model of the following general type:

$$
\begin{aligned}
\Delta_4 \log p_{i,t+4} = {} & \alpha_i + \gamma_i (\Delta_4 \log \hat{p}_{t+4}) \\
& + \delta_i (\log p_t - \log p_{i,t}) + \sum_{j=0}^{k} \delta_{i,j} \Delta \log p_{i,t-j} \\
& + \eta_i (\log wpi_{i,t} - \log p_{i,t}) + \sum_{j=0}^{k} \eta_{i,j} \Delta \log wpi_{i,t-j} \\
& + \theta_i (\log mulc_t - \log p_{i,t}) + \sum_{j=0}^{k} \theta_{i,j} \Delta \log mulc_{t-j} \\
& + \lambda_i (\log wpimp_t - \log p_{i,t}) + \sum_{j=0}^{k} \lambda_{i,j} \Delta \log wpimp_{t-j} \\
& + \phi_i (\log otherp_t - \log p_{i,t}) + \sum_{j=0}^{k} \phi_{i,j} \Delta \log otherp_{t-j} \\
& + \sum_{l=1}^{n} \beta_{i,l} X_{l,t} + \sum_{l=1}^{n} \sum_{j=0}^{k} \beta_{i,l,j} \Delta X_{l,t-j} \\
& + \sum_{j=0}^{k} \omega_i \Delta \log p_{i,t-j} + \mu_{i,t} + \varepsilon_{i,t} \qquad (3)
\end{aligned}
$$

where ε_t is white noise plus, possibly, a moving average error component, and μ_t is a smooth stochastic trend reflecting institutional changes. All the variables are defined in Table 2, together with statistics and stationarity characteristics for the data.

Consider the elements of equation (3). The term on the first line after the constant, $\gamma_i(\Delta_4 \log \hat{p}_{t+4})$, captures the influence of the four-quarter-ahead forecast of the annual CPIX inflation.[12] This term has a dual role. First, it can be considered as a proxy for overall inflationary expectations, which are likely to influence pricing behaviour in the different sectors, as well as feeding into wage costs - which in turn may feed into sectoral prices. Secondly, it offers a helpful way of comparing the forecast performance of the multi-sectoral method of forecasting inflation with whatever pre-existing method is being used.[13]

The next term in equation (3), $\delta_i(\log p_t - \log p_{i,t})$, with $\delta_i>0$, allows for the possibility that there is some tendency for the i[th] component of the CPI index to track total CPI, so that if it rises relatively, there will be some pressure to revert to the CPI trend.[14] The dynamics are captured by the corresponding changes in log *p*.

In the following two lines, we think of $wpi_{i,t}$ as the most relevant wholesale price index component that might influence sector i, though it could also be the unit labour costs for the manufacturing sector ($mulc_t$), or for the wider economy. In service sectors, unit labour costs[15] or renumeration per worker (allowing the stochastic trend to pick up productivity trends) may be particularly relevant, since there is no obvious related WPI component. The dynamics are captured by the corresponding changes in the logs of these prices.

The following line captures both the dynamics and "equilibrium correction" mechanisms for foreign prices, in a similar way to the wholesale price components or unit labour cost measures above. We use the log of import prices from the wholesale price index, $wpimp_t$. Alternative measures of international prices include the U.S. wholesale price, and world commodity price indices, both converted into rands.

The last error correction term captures other prices relevant to particular components, $otherp_{i,t}$,again relative to the price sub-component. For instance, the raw food component

---

[12] In principle, such a forecast could come from any source, including the SA Reserve Bank's own forecasts, or forecasts from a single equation of the type Aron and Muellbauer (2000c) have estimated for the consumer expenditure deflator, or for CPIX (Aron and Muellbauer, 2004b).

[13] This can be seen as follows. Since $\Delta_4 \log p_{t+4} \approx \sum_{i=1}^{10} w_{i,0} \Delta_4 \log p_{i,t+4}$, finding that $\sum_{i=1}^{10} w_{i,0}\gamma_i \approx 1$ would suggest that gains from the multi-sectoral approach are likely to be small, since then the weighted combination of the other terms in equation (3) would be likely to be small also. In other words, the fitted value of $\sum_{i=1}^{10} w_{i,0} \Delta_4 \log p_{i,t+4}$ would forecast no better than $\Delta_4 \log \hat{p}_{t+4}$, the pre-existing aggregate forecast.

[14] One instance of this reversion might be found with administered prices indexed to the CPI; but more generally, if wages are loosely indexed to the CPI and there is some wage matching across industries, the rise in wage costs in the i[th] sector will be linked to the CPI.

extracted from the agricultural wholesale price index, and U.S. maize prices converted into rands, are used in the food equation. Most of the equations initially employ an international oil price converted into rands, expressed relative to the i^{th} component of the CPI index. Corresponding rates of change in the rand oil price are included.

Relative price trends in other countries may also have forecasting implications (e.g., the U.S.). The stochastic trends in each equation allow for a deviation between log p_t and log $p_{i,t}$, due to changes in technology. Given U.S. technological leadership the long-run solutions for (log p_t – log $p_{i,t}$) in equation (3) may be forecastable from long-run trends in the U.S. This suggests (in future work) including the following term using U.S. CPI data as another regressor, where the i^{th} components approximately correspond:

$$+\varsigma_i(\log p_t^{US} - \log p_{i,t}^{US}) + \sum_{j=0}^{k} \varsigma_{i,j}\Delta(\log p^{US} - \log p_i^{US})_{t-j} \qquad (4)$$

In principle, there may be other information that could be relevant, for example, interest rates, indirect tax rates, “sin taxes” on alcohol and tobacco, exchange rates, terms of trade, measures of excess demand, or information derived from the Bureau of Economic Research (BER) business, consumer or inflation surveys. Institutional changes are also likely to be very important, for instance, changing trade policy has a key effect on certain price components through increased import competition or openness (φ_i). These potential determinants are collected as the X_l variables on the penultimate line. Institutional changes we are unaware of, or find hard to quantify, will be captured in the stochastic trend, $\mu_{i,t}$.

The set of *X* variables employed in the sub-component equations is as follows:

$$\begin{aligned} X_l = \{&\log(REERN); \log(TOT); OUTGAP; RCASUR; VAT; \\ &LREXISEDUTY; RPRIME; DUMOPEN\} \end{aligned} \qquad (5)$$

With the sign priors given in parentheses, the terms are: the log of our constructed[16] real exchange rate index (-), where more expensive imports through real exchange rate depreciation feeds into living costs (pass-through); the log of the terms of trade index (+), as

[15] The “normalized” unit labour cost measures were calculated by subtracting a measured (stochastic) trend in log productivity from the log wage in the manufacturing sector – see a full description in Aron, Muellbauer and Smit (2003).
[16] The real exchange rate is, in effect, another relative price term, and by definition increases with appreciation. Details of the construction of the real exchange rate index are given in Aron, Muellbauer and Smit (2003).

as a terms of trade boom drives up non-traded prices, and also manufactured prices, depending on the extent to which they are shielded from world prices[17]; the output gap (+) - either an economy-wide or a sector-specific output gap[18] effect - since with higher excess demand, consumer prices may increase relative to wage costs and wholesale prices[19]; the ratio of the current account surplus to GDP (-), partly as an excess demand indicator, and partly as a predictor of exchange rate movements; the rate of sales tax or indirect taxes relative to consumer expenditure (+) are expected to increase consumer prices (ideally, taxes specific to the commodity in question e.g. excise duties for particular goods, such as beverages and tobacco).[20]

The real prime interest rate is included to capture the effect of business costs increasing when interest rates rise, or the disinflationary effect via lower demand pressure, exchange rate appreciation, or inflation expectations. Rates of change of all nominal variables appearing in ratio form in equation (3) also appear in the general model.

The 'cost channel' of monetary policy via the real prime rate is likely to be controversial for central bankers and has been the subject of a significant literature, see Barth and Ramey (2001). A related mechanism by which higher interest rates may raise subsequent inflation is via the effect of high interest rates on investment and bankruptcies. These reduce capacity and so may increase inflation in subsequent upturns. In principle, an output gap measure incorporating the well-measured capital stock should capture this effect. However, official capital stock estimates are typically based on fixed service life assumptions for equipment, and will not fully reflect scrapping of capital arising through business closure. A third related mechanism by which higher interest rates could raise subsequent inflation can arise when businesses rebuild profit margins or balance sheets after suffering losses or reduced profits during a period of high interest rates. Chevalier and Scharfstein (1996) have

---

[17] However, our prior on this sign is weak, since an improvement in the terms of trade may cause the exchange rate to appreciate and so have a disinflationary effect.

[18] The output gap measure is derived by regressing log (GDP) on a stochastic trend, lagged log (GDP), and distributed lags in capacity utilisation and changes in the log terms of trade, to proxy cyclical effects. The output gap is then defined as the deviation of log output from the stochastic trend, scaled appropriately.

[19] Regarding the output gap, we have to be aware that, at the sectoral level, a negative coefficient is a possibility. Suppose the ith good's price does not respond at all to the output gap. We know that CPIX responds positively to the output gap. Then, *relative to CPIX*, the ith good's price will fall when the output gap rises. Suppose the four-quarter-ahead equation for the ith good's price includes an important ECM term involving CPIX, and the four-quarter-ahead forecast for CPIX is also important, making the relative price to CPIX important for the dynamics as well as the long-run. Then one would expect a negative output gap effect in the price equation for the ith price.

[20] Note that at a sectoral level, indirect tax increases are likely rapidly to feed through into consumer prices, and hence we would not expect to find significant effects of tax rates in a four-quarter-ahead approach. However, there is scope for separately developing forecasting models for indirect tax rates, which could take into account

provided evidence for counter-cyclical mark-ups for U.S. supermarkets. It seems plausible that some service sector companies, such as insurance companies, are also subject to such behaviour.

An important institutional variable included in the long-run equilibrium term is changing openness to international trade: higher foreign trade taxes are consistent with raised producer prices from more expensive imported inputs, which may feed into consumer prices. A comprehensive measure of openness for South Africa, *DUMOPEN*, was derived by Aron and Muellbauer (2002a), and uses data on the share of manufactured imports in domestic demand for manufactured goods (Figure 2). Our openness measure includes both reported tariffs and import surcharges, and captures unmeasured quotas and the effect of international trade sanctions on South Africa – the latter two especially important in the 1970s and 1980s. We expect a negative sign for *DUMOPEN*, which increases with increasing liberalisation for those CPI components where import competition is more important. (For less tradable goods and services, the price relative to the CPI may increase with increases in trade liberalization, implying a positive effect from *DUMOPEN*). Potentially, interaction effects could also be important for other variables. For example, if the cost channel discussed above works through the capital stock, the effect may have been overshadowed in more recent years by the opening of the economy to international trade and capital flows.

4.3 Weighting components for an aggregate forecast

Table 1 shows the current weights of the components in the CPI index. In what follows we separate out from the housing sub-component the contribution due to the mortgage interest rate (see section 5.3.1), using the appropriate weights, which change over time, given that we are modelling CPIX and not CPI. The individual forecasts from section 5 can be weighted and reaggregated to produce an overall forecast for consumer prices. In a follow-up paper we compare such aggregate forecasts with our forecast for aggregate CPIX (see section 5.1).

5. Forecasting equations

fiscal deficits. Our framework otherwise implicitly incorporates the effects of fiscal deficits through the inclusion of the output gap, interest rates and inflationary expectations.

In this section we present four-quarter ahead "forecasting" equations for the ten separate sub-components of the CPI index (adjusting the housing sub-component to remove the mortgage interest rate). Section 5.4 explains the further steps needed to obtain practical forecasts from these equations. We use the equilibrium correction forecasting framework shown in equation (3), which includes our CPIX forecast; various relative price terms in the price sub-component with the corresponding dynamics; a range of other variables expected to influence prices, denoted X_l , with the corresponding dynamics; a stochastic trend and a measure of trade openness; and the lagged dependent variable.

The parsimonious equations following a general-to-specific methodology are shown in Tables 3-12. In the equations reported, all long-run explanatory variables are I(1). Standard Dickey-Fuller tests suggest that over 1979-2002, each log $p_{i,t}$ is I(1), implying that $\Delta_4 \log p_{i,t}$ is a stationary variable (Table 2). This would imply that the stochastic trend, μ, the X_l variables, and the relative prices in log p_i are cointegrated. The fact that μ is by construction an I(2) variable, is, at first sight, problematic. However, a low variance stochastic trend closely resembles a segmented linear trend so that the linear combination can easily be I(0) over the relevant samples.

For each sub-component, an equation is estimated from the second quarter of 1979, when the exchange rate began to float (in one form or another), to 2002Q2 (i.e. forecasting one year further ahead). Estimates for two shorter samples, 1979:2 to 1990:1 and 1988:1 to 2002:2, were taken for the same equation to examine robustness of the estimates as STAMP does not allow CHOW tests for parameter stability. These estimates are not reported in this version of the paper, but are discussed for each component. In the equations reported, all long-run explanatory variables are I(1), as expected, except for the output gap which is (borderline) I (0) – see Table 2. However, the output gap is persistent and we treat it as part of the long-run solution. Note that in general we did not find evidence for asymmetric effects (see section 4.1) for the four-quarter-ahead equations.

5.1 CPIX Forecasting Equation

As noted in section 2, the official CPIX for metropolitan and urban areas has been published only from 1997. We have therefore constructed our own measure of CPIX for metropolitan areas only back to 1970, which we denote CPIXC (constructed) in this paper.[21] Our methodology for forecasting four quarters ahead broadly follows the basic structure of equation (3): with γ_i set to zero, and with log CPIX (log p_t) in place of the i^{th} component, log p_{it} , and is explained in detail in Aron and Muellbauer (2004b).

Key elements in the long-run solution for log CPIX are equilibrium correction terms in the log ratio to CPIX of unit labour costs and wholesale prices, and logs of relative prices, especially the real exchange rate and the terms of trade. Other elements are the output gap, the real prime rate of interest and our index of openness. The last has a positive long-run effect on CPIX relative to wholesale prices and unit labour costs in manufacturing, since CPIX has been less subject to the competitive pressures stemming from increased international competition as trade barriers have fallen. Indeed, we find that increased openness has had further effects in altering the size of impact on the economy of the terms of trade and interest rates. When the economy was relatively closed, a rise in the terms of trade (e.g. due to gold and platinum prices rising internationally) had net inflationary effects on CPIX relative to wholesale prices and unit labour costs. Currently, however, with much greater openness, presumably not only on the trade account but also on the capital account, the currency appreciating effects of such terms of trade movements, impart a net dis-inflationary effect on CPIX. The greater priority in monetary policy placed on controlling inflation may also have played a part in this shifting role for the terms of trade. In the early 1980s, the Reserve Bank may have been more concerned with preventing 'Dutch disease' consequences of the gold price boom, through active exchange rate intervention. Furthermore, it appears that the cost channel of a rise in real interest rates e.g., via the cost of capital or via the effect on capacity, has weakened substantially with the opening of the economy. This can be interpreted in terms of the weaker constraint of domestic interest rates on the cost of capital and so on investment.

The short-term dynamics of the model also include negative effects on inflation from the change in the prime rate and from the two-year change in the trade surplus to GDP ratio. Together with the terms of trade effects, these are consistent with an exchange rate channel operating on CPIX four quarters ahead - supporting the results in Aron, Muellbauer and Smit (2003) on the determination of the real exchange rate.

5.2 Goods Forecasting Equations

[21] Our method, unlike that of Statistics South Africa, removes the mortgage bond component consistently, see Aron and Muellbauer (2004a) on the bias in the official CPIX.

5.2.1 Food

South Africa has a substantial agricultural sector, and production is frequently subject to major weather shocks (as it is throughout the neighbouring region). These shocks have significant inflationary consequences. With a large fraction of the population on low incomes, they also have important welfare consequences.

The food price inflation process has been subject to several structural changes. Maize is the major food staple. In 1987, the maize marketing board was abolished, while other marketing boards were abolished by the early 1990s. With the opening of the economy to imports in the 1990s and the removal of trade sanctions, it seems likely that food prices in South Africa became more subject to international competition. In 1996, SAFEX, the South African Futures Exchange began operation, providing a further factor in increasing the influence of foreign prices and the exchange rate.

In Figure 1, the log ratio of the food price component deflated by the CPI is shown, see discussion in section 3.1.1.[22]

Two additional error correction price terms are introduced into equation (5): one for the log ratio of the raw food price index to the CPI food component (+); and the other for the log ratio of the index of U.S. maize prices (in Rands) to the CPI food component (+), as more expensive imported maize, especially in times of domestic shortages, translates into higher food prices. Other price terms include our constructed real effective exchange rate index (-), since appreciation reduces the price of agricultural outputs as imported raw foodstuffs are then cheaper[23]; and the terms of trade (+), where enhanced demand from trade booms elevates the prices received for agricultural output. Since South Africa is a major food exporter for certain cash crops, the terms of trade will reflect these export prices, which seem likely to have a positive impact on food prices within South Africa. The terms of trade excluding gold will measure these effects better than the terms of trade including gold, and we test for both definitions.

[22] All the CPI sub-component data suffer from rounding errors in the earlier years due to the unfortunate practice of *Statistics SA* throwing away accuracy when rebasing. To illustrate, on the current base of 100 in the year 2000, in August and September, 1979, the food price index was 7.6, and in October and November, 1979, it was 7.7. The lack of a second decimal place reduces the accuracy of the data compared with current data by the order of 12-fold.

[23] Furthermore, with South Africa a net exporter of many agricultural commodities, appreciation makes exporting less profitable. The diversion of supply to the domestic market then reduces domestic food prices.

We allow the stochastic trend to reflect structural changes in the food inflation process caused by the removal of the maize marketing board, and the associated decline in prices with market competition. We also allow it to pick up various drought episodes when supply shortages drove up prices. Note that the subtleties concerning import parity pricing post 1996 are not specifically addressed in our model.

The parsimonious equation following a general-to-specific methodology, imposing sign priors, are shown in Table 3b. The diagnostics are satisfactory and the parameter estimates are similar over the different samples. Two equilibrium correction terms are significant: one links food prices with CPIX, and the other with the import price index. The other terms in the long-run solution are the output gap, suggesting food prices respond to excess demand pressure, and the terms of trade together with the stochastic trend. The real exchange rate has a negative coefficient, and there is some evidence that this had a larger negative impact with the opening up of the economy, as one would expect, but the effects are not statistically significant, and hence omitted in the reported estimates.

The long run solution is as follows:

$$\begin{aligned}\log FD = 0.58\log CPIXC + 0.42\log WPIMP + 0.21\log TOTXG \\ + 0.35 OUTGAP + \mu\end{aligned} \tag{6}$$

The stochastic trend trends upwards, particularly from 1989, by about 0.2 in the long run solution from 1989 to 2001, see Figure 3. The upward movement reflects the upward trend in the relative price of food, which, as Figure 1 showed, had begun already in 1970. At first sight, this seems surprising, particularly as the raw food price index has declined since 1980, not just relative to CPIX, but even against the wholesale price index for manufactures, which has itself risen less than CPIX. The model controls for import prices, which take the exchange rate into account, so that it is hard to argue that the stochastic trend is a proxy for a missing exchange rate effect.

By contrast, the relative price of the food component to the overall CPI index in the U.S. has shown a general downward trend since 1980. One reason for the different behaviour in South Africa is likely to be the fact that, relative wages of unskilled workers, such as those employed in retailing, have risen in South Africa, but declined in the U.S. Increased openness of the economy in the 1990s leading to rapid productivity gains in the manufacturing sector and competitive pressures on prices of tradeables, plausibly had a smaller effect on food prices, so raising their relative price. Another possibility for South Africa is that the proportion of processed food and meals taken outside the home has increased over time. These items are more expensive but presumably reflect a quality improvement e.g., in time saved. Thus, a more detailed comparison with experiences in other countries and a check on the procedures used by Statistics S.A. to take account of new goods, and quality change in the food sector, needs to be undertaken. A study of retail margins and the evolution of concentration ratios in retailing might also yield useful clues to help understand these trends. Further, it has also been argued that the proliferation of new shopping centres in South Africa has encouraged retailers to open new stores with existing sales simply spread across an increasing number of outlets, thus raising costs.[24]

On the face of it, the rise in the price of food relative to the CPIX has unfavourable implications for the distribution of real living standards, since food accounts for a larger part of total expenditure for poor households. However, if the price of staples such as maize has actually declined, in the long run, relative to CPIX, the opposite conclusion follows. In practice, the raw food component (from the agricultural wholesale price) and the factory food wholesale price component have declined relative to the CPI food component (Figure 4).

The dynamics in the equation also includes lagged rates of change in the food CPI component (indicating a strong tendency to correcting previous overshoots); the rate of change over three quarters in the Rand US maize price; and the CPIX inflation forecast over the next four quarters. Robustness checks with alternative CPIX forecasting models give very similar results. Note that implicit in the parsimonious long-run terms are implications for the dynamics. For example, the import price error correction term is a four-quarter-moving average, the CPIX error correction term enters with a one quarter lag and the terms of trade are smoothed over two quarters, and lagged one quarter.

5.2.2 Furniture and equipment

In Figure 1, the log ratio of the furniture price component deflated by the CPI is shown. It exhibits a downward trend, sharper from 1990, attributable to productivity improvements in the sector, as well as the opening of the South African economy to competing imports, which reduced the pricing power of producers in this sector. We allow the stochastic trend to reflect productivity changes in the sector and changing openness, though we also test for a separate effect from our openness dummy.

[24] We are grateful to Gavin Keeton for this point.

The detailed estimates are shown in Table 3c. The model for this sector is a relatively simple one, with equilibrium correction of the price index for furniture and equipment to both the wholesale price index for all goods and to CPIX. The other element in the long-run solution is a strong real exchange rate effect, as one might expect from such a tradable set of goods. The long run solution is as follows:

$$\log FR = 0.55 \log CPIXC + 0.45 \log WPMTOT - 0.092 \log REERN + \mu \qquad (7)$$

The short-term dynamics incorporate the annual change in the trade balance to GDP ratio, which we know helps to forecast the exchange rate, and can also be an excess demand or supply indicator, the forecast CPIX rate of inflation, and the rate of change in unit labour costs. The equation standard error is comparatively low and the stochastic trend the smoothest of any in the ten equations. Parameter stability is excellent.

The long-run effect of the stochastic trend from 1979 to 2002 on the relative price of furniture and equipment is minus 0.65, indicating a halving of the relative price over this period. This is most obviously due to a combination of technological cost reductions and quality improvements, and of greater openness to international competition: the steeper decline after 1990 is characteristic of the latter.

5.2.3 Clothing and footwear

The relative price of clothing and footwear exhibits a similar downward trending pattern to that of furniture and equipment, and for similar reasons: productivity growth, and especially greater openness to competitive trade from 1990 onwards (Figure 1). We allow the stochastic trend to reflect productivity changes in the sector and changing openness, though we also test for a separate effect from our openness dummy.

A very simple equation describes the price behaviour, with long run equilibrium correction terms linking clothing prices with the wholesale price for clothing a feedback to CPIX. In each case, the parsimonious parameterisation is a three-quarter moving average (which embodies two lagged terms in each of CPIX and in the wholesale price index for clothing). However, the second ECM term with a feedback to CPIX is fairly collinear with the ECM term in wholesale prices. When both are included the model proves not to be stable over split samples, and the omission of the second ECM term hardly worsens the overall fit

(Table 3b). We therefore prefer the simpler model. There is no effect from the forecast aggregate CPIX. The long run solution is as follows:

$$\log CL = \log WPCL + \mu \tag{8}$$

The dynamics in the equation includes lagged rates of change in the clothing and footwear CPI component (indicating a strong tendency to correcting previous overshoots); and the annual change of the current account surplus.

The stochastic trend shows a decline from 1990, which looks very like an openness effect, see Figure 3. If the stochastic trend is omitted, the openness dummy has a significant negative effect and the fit of the equation deteriorates only marginally, consistent with the openness interpretation of the stochastic trend.

5.2.4 Vehicles

Figure 1 shows a very sharp increase in relative vehicle prices from the mid-1980s to the mid-1990s. Prices stabilised from 2000, after following a declining trend for about five years.

The vehicles equation reported in Table 3b incorporates three significant equilibrium correction terms, suggesting that, in the long run, vehicle prices depend on the CPIX, on wholesale import prices, and on unit labour costs in manufacturing. The other factors in the long run solution are the output gap, indicating a greater response to excess demand than present in the CPIX, and the terms of trade. The long run solution is as follows:

$$\begin{aligned} \log VH &= 0.85\log CPIXC + 0.08\log WPIMP + 0.07\log MANULC \\ &\quad + 0.17\log TOTIG + 0.32 OUTGAP + \mu \end{aligned} \tag{9}$$

The stochastic trend rises from 1982 to 1994 by around 0.8 (0.5 in terms of the long-run effect) and then declines by around 0.2 (0.13 in the long-run) to 2002 see Figure 3. A rise of 0.5 in a log relative price corresponds to a 1.65 fold increase, i.e. a rise from an index of 100 to 165. In other words, between 1982 and 1994, *after* the effects of general rises in the CPIX, import prices, etc are accounted for, vehicle prices rose by almost two thirds. As we noted in Section 3.1.4 above, tariffs on built up vehicles were over 100 percent from the early 1980s to the early 1990s, and began to be cut in 1993/4 and then more sharply thereafter, and have now fallen to 38 percent. Local content requirements were tightened in 1984 probably raising prices. These institutional facts are consistent with the stochastic trend as pictured in Figure 3. Still, as we noted in Section 3.1.4, the extent of the relative price rise between the mid 1980s and 1993 does raise questions about whether adequate allowance was made for quality improvements in the vehicles component of the CPI. A further consideration is that the wholesale price index for imports may not fully compensate for the Rand's weakness against the Yen and Deutsche Mark, given that both these currency experienced a structural strengthening against the U.S. dollar for 1986-94 (most vehicle imports come from Japan and Germany).[25]

The dynamic terms in the model include lags of rates of change of vehicle prices with negative coefficients, suggesting a correction for recent price overshoots [26]; and the forecast rate of change of CPIX for the next four quarters. Checks for parameter stability were carried out by estimating the model for different samples and proved satisfactory.

5.2.5 Transport goods

The transport goods equation reported in Table 3d incorporates three significant equilibrium correction terms, suggesting that, in the long run, prices of transport goods are determined in the long-run by CPIX, unit labour costs and the Rand price of oil. The latter is unsurprising, given the inclusion of petrol and diesel in the price index (Figure 1). The other term in the long-run solution is the output gap, suggesting some impact from excess demand pressures on prices. The long run solution is as follows:

$$\begin{aligned} \log TG = 0.73\log CPIXC + 0.23\log MANULC + 0.04\log POIL \\ +1.52 OUTGAP + \mu \end{aligned} \qquad (10)$$

In the short-run dynamics, there is an important effect from the forecast rate of inflation of CPIX, and an effect from the annual change in the trade surplus to GDP ratio, which helps forecast the exchange rate and is an excess demand indicator.

[25] We are grateful to Gavin Keeton for this point.

[26] Another way of looking at these negative feedback terms is as compensation for the four-quarter moving average in the ECM term in log CPIX minus log vehicle price. Reformulating this as the four-quarter moving average of CPIX minus the current log vehicle price, effectively eliminates rates of change in the vehicle price lagged one and two quarters, and sharply reduces the coefficient on the current rate of change. Similar considerations arise in other equations.

The stochastic trend broadly declines to 1998 and then rises from 1998-2000 for reasons we do not understand, perhaps an increase in fuel duties. The very different pattern from the "vehicles" sub-component is interesting. Presumably there would have been low tariffs on parts to encourage local assembly and high tariffs on finished vehicles. The local contents requirements in the 1980s and 1990s may have increased investment in parts manufacture, bringing economies of scale. Perhaps the substitution of cheaper local parts for expensive imported parts brought down prices. However, the index is dominated by petrol and diesel and a puzzling feature of our results is that the Rand oil price effect does not have a larger influence.

5.2.6 Beverages and tobacco

In Figure 1, the relative log ratio of beverages and tobacco prices moves sharply upwards from about 1987, having declined during the previous fifteen years. Cheaper tobacco and alcohol products through the expansion of local manufacturing capacity in the 1970s was rapidly reversed with higher taxation, especially following election of the ANC government in 1994. We allow the stochastic trend to capture these tax changes.

The parsimonious equations are shown in Table 3c. One equilibrium correction term is significant, linking beverage and tobacco prices with CPIX. The other terms in the long-run solution are the terms of trade; the real prime rate of interest with a positive coefficient, representing the business cost channel; and the stochastic trend. However, as in the aggregate CPIX forecasting equation discussed in Section 5.1, there is some evidence of a weakening terms of trade effect as openness increased, represented by an interaction effect of the terms of trade with our openness measure. There is also slight evidence of a weakening in the real interest rate effect with openness, as in the aggregate CPIX equation, but as its statistical significance was weak, the effect is omitted in the reported estimates.

The long run solution is as follows:

$$\begin{aligned}\log BEVT = \log CPIXC + (0.42 - 0.26RDUMOPEN)\log TOTIG \\ + 0.65RPRIME + \mu\end{aligned} \tag{11}$$

The dynamics in the equation include the usual negative feedback effects in the lagged dependent variable (indicating a strong tendency to correcting previous overshoots); and current and lagged changes in the current account surplus (expressed as an annual change). Further dynamics are implicit in the long-run parameterisations.

The stochastic trend is consistent with indirect tax effects, examining excise duties data from the National Treasury: a strong rise in duties on cigarettes, tobacco and alcoholic beverages from 1995 coincides with a steep rise in the stochastic trend (see Figure 3), by far the biggest rise over the period. In terms of the long run solution, the rise since 1995 is around 0.8, implying a more than doubling of the relative price of beverages and tobacco. In terms of forecasting, these results suggest that judgemental or other forecasts of rises in duties on beverages and tobacco, would be an important complement to the model derived here, since they would be likely to eliminate much of the role the rather volatile stochastic trend is playing at present.

5.2.7 Other goods

The relative price of "other goods" shows trend rises from 1975-80, 1985-86 and from about 1997 onwards (Figure 1). As with "other services", this category covers a very wide group of goods, but imported goods, or goods using foreign imports are well represented, suggesting the likely importance of the exchange rate. `

One equilibrium correction term appears (Table 3d), linking "other goods" prices prices with CPIX. The other terms in the long-run solution are the output gap, the real exchange rate and the terms of trade. The long run solution is as follows:

$$\begin{aligned}\log OG = \log CPIXC + 0.25 OUTGAP + 0.09 \log TOTXG \\ - 0.10 \log REERN + \mu\end{aligned} \quad (12)$$

The stochastic trend has a range of less than 0.1 in the long-run solution, far smaller than is the case for the other nine CPIX components. It increases by around 0.1 between 1983 and 1985 and decreases by a similar amount between 1991 and 1996.

The dynamics in the equation include the usual negative feedback effects in the lagged dependent variable (indicating a strong tendency to correcting previous overshoots); the forecast rate of CPIX inflation and the lagged three-quarter change in unit labour costs. Further dynamics are implicit in the long-run parameterisations, in which all the relative price terms enter as four-quarter moving averages.

5.3 Services Forecasting Equations

5.3.1 Housing

The housing CPI component we model below excludes the mortgage bond rate (Figure 1).[27] Housing costs excluding mortgage interest are an amalgam of rents, costs of repairs and maintenance and domestic staff costs. Some of these are likely to follow CPIX, and the forecast rate of inflation in CPIX plays an important role. However, the ECM is with respect to the wholesale price index, other candidates for ECMs all proving insignificant. The long-run solution also has significant real interest rate effect, suggesting a cost channel, probably feeding into rents.

The long run solution takes the form:

$$\log HX = \log WPDTOT + 0.20RPRIME - 0.21FLIB + \mu \qquad (13)$$

In the dynamics, the rates of growth of house prices, also likely to feed into rents, of CPIX, of unit labour costs in manufacturing, and exchange rate depreciation all play a role, in addition to that of the inflation forecast. Interestingly, we find a negative role for the index of financial liberalisation of consumer credit markets in South Africa developed in Aron and Muellbauer (2000a,b). With liberalisation, access to mortgage credit became easier for potential owner-occupiers, suggesting a fall in demand for rental accommodation and so a slower rate of increase in rents. The stability of the equation looks satisfactory, with a slight tendency for exchange rate depreciation to become more important in more recent years.

5.3.2 Transport services

The relative price of transport services shows a distinct downward trend from the mid-1980s, probably attributable to attempts to lower the costs of publicly-owned bus and other transport in the difficult climate after 1984, and then after the elections of 1994 (Figure 1). The

[27] The method we use to remove the mortgage interest component from the housing component follows the method to remove the mortgage interest component from total CPI ("metropolitan areas"), explained in Aron and Muellbauer (2004a).

stochastic trend would be expected to capture changes in the administered prices of transport, for which we do not have figures.

The transport services equation reported in Table 3a incorporates three significant equilibrium correction terms, suggesting that, in the long run, transport services prices depend on the CPIX, on wholesale import prices, and on unit labour costs in manufacturing. The other factors in the long run solution include the output gap and the stochastic trend.

The long run solution is as follows:

$$\log TS = 0.41\log CPIXC + 0.45\log MANULC + 0.14\log WPIMP \\ + 0.9OUTGAP + \mu \qquad (14)$$

The stochastic trend declines, especially during 1984-88, perhaps a symptom of price controls in a period of general inflation, and at a slower rate after 1990, perhaps connected with deregulation of bus and taxi services.

In the dynamics we find the usual negative feedback to the last four quarters' price changes in transport services and a lagged effect from the rate of change of Rand oil prices. Other dynamic terms are implicit in the parameterisations of long-run terms.

5.3.3 Other services

The relative price of other services shows a gently declining trend from1970-85 and then a steeper rise from 1985 (Figure 1). As with "other goods", this sector encompasses a broad range of goods, with medical, communication and education services having the highest weights, and interest on bank and loan charges included in the category. The trend is likely to be important to achieving an economic interpretation of this equation, but will itself not be easily interpretable.

One equilibrium correction term appears (Table 3a), linking "other services" prices with CPIX. The other term in the long-run solution is the output gap. The long run solution is as follows:

$$\log OS = \log CPIXC + 0.17OUTGAP + \mu \qquad (15)$$

The stochastic trend shows a rise from 1985 and especially from 1990. This suggests that the relative price of other services has risen with increased openness in the economy, consistent with the relatively non-tradable aspect of other services. The rise from 1990 to 2000 corresponds to a long–run effect of 0.23, a rise of around 25 percent in the price relative to the CPIX.

The dynamics in the equation include the usual negative feedback effects in the lagged dependent variable (indicating a strong tendency to correcting previous overshoots); the forecast CPIX inflation rate and the annual inflation rate in wholesale prices. Further dynamics are implicit in the long-run parameterisations, specifically the four-quarter moving average in the ECM.

5.4 Modelling versus Forecasting

The models we have developed for the components of CPIX have incorporated important stochastic trends estimated in STAMP. In some cases, these stochastic trends are quite volatile. It is unlikely that, for out-of-sample forecasting, these models will be optimal. To explain why, note that the models can be written in schematic form:

$$\Delta_4 \log(p_{it+4}) = \mu_t + \beta_t X_t + \varepsilon_t \qquad (16)$$

Estimating up to T-4, we have the recursively-estimated parameter vector β_{T-4} and an estimate of μ_{T-4}. To forecast for $\Delta_4 \log(p_{i,T+4})$, at date T, we assume β_{T-4} is unchanged and project μ_{T-4} forwards by four quarters. Given equation (2) with η_1 zero, it is clear that the best projection is

$$\hat{\mu}_T = \mu_{T-4} + 4 * \gamma_{T-4}$$

This arises from the fact that η_2 has a zero expected value for T-4 to T, so that $\mu_t = \mu_{t-1} + \gamma_{T-4}$ over this interval.

If the stochastic trend μ_t is quite volatile, then this four-quarter-ahead linear projection could be quite poor. Thus, a good fit over the historical sample does not guarantee a good out-of-sample forecasting performance. Models in which the stochastic trend is less volatile and where the economic variables have somewhat greater relative explanatory power, may well forecast better out-of-sample, even if the within-sample fit is worse.

Within STAMP, it is possible to constrain the variance of η_2 *a priori*, so forcing greater smoothness on the stochastic trend. It is then possible that a general-to-specific model selection strategy could, in some cases, select a different model for some price index components than the ones presented above. The task that remains is to develop such models for a small set of alternative restrictions on the smoothness parameter. Then, the individual CPIX component forecasts can be aggregated into an overall CPIX forecast and the performance of models under different assumptions about the smoothness parameter can be compared. This task is devolved into another paper.

6. Conclusion

In this paper, using more richly specified equilibrium correction models than hitherto and employing stochastic trends, we have modelled separately the inflation rates for the ten major groups of goods and services that make up the consumer price index for metropolitan areas, excluding the mortgage cost component, i.e. CPIX ("metropolitan"). The models have been designed to explain the separate inflation rates four quarters ahead, bringing to bear relevant economic and institutional knowledge. As well as serving as a prelude to designing practical forecasting models for overall inflation, these models cast important light on the complex forces acting on the relative prices and inflation rates of the different goods and services, explaining higher inflation rates in some sectors and the persistence of the residual disturbances. An improved understanding of inflationary pressures for particular components of the basket of consumer spending could potentially help targeting micro-economic policy interventions, perhaps involving deregulation or the competition authorities, and the phasing of taxation.

We have found important evidence for the influence of the increasing degree of openness, which has brought down relative prices and inflation rates of the more tradable goods and of those goods where the pressure of international competition has contributed to higher productivity growth. Conversely, the relative prices of goods and services more sheltered from these international competitive pressures have tended to rise.

Other important factors with differential impacts on prices of different goods and services include the exchange rate, oil prices and other terms of trade shocks. However, evidence both from our overall CPIX forecasting equation and from some of the detailed inflation equations suggest that terms of trade shocks now have less inflationary effects, or even disinflationary effects, via exchange rate appreciation, than was true in the early 1980s. Changes in institutions, such as marketing boards and tariffs and quotas, including local content restrictions on foreign vehicle manufacturers, have also had their impact on prices of particular goods and services. The evidence from some goods is for important fiscal effects on prices, as indirect tax rates on, for example, alcohol and tobacco, have risen. As we have no time series data on concentration ratios in particular sectors, we have unfortunately not been able to model the impact of changing concentration in South Africa industry on the mark-up of prices over costs.[28]

We have considerable evidence from our aggregate CPIX forecasting equation that the inflation process in South Africa has not been stable over the last quarter of a century. In particular, increased openness has not merely affected the overall inflation rate but seems to have altered the role of both the terms of trade and interest rates in explaining inflation. The inflationary role of terms of trade shocks and of the cost channel of monetary policy appear to have declined, altering the monetary policy transmission.

We have thus established good framework for modelling at the sectoral level. There may be scope for testing for further interaction effects and asymmetries. In some cases the availability of data on sectoral output gaps and sectoral wages may throw further light on the sectoral inflationary processes. In a follow-up paper, we explore whether forecasts from this novel disaggregated approach produce improved predictions over the four-quarter-ahead forecasts for the aggregate index alone.

Our research also highlights some important measurement issues. The most obvious is how important it is that the statistical service follows transparent, systematic and internationally-recognised measurement procedures[29] and is able to access historical data and provide it to the central bank, as well as to the private sector. We suggest in a parallel paper that the lack of a clear handbook of CPI methodology has led *Statistics South Africa* to measure CPIX inconsistently with their methods for measuring CPI (Aron and Muellbauer, 2004a)[30]. The combination of the lack of a clear handbook[31] and *Statistics South Africa's* inability to access and provide detailed pre-1997 price data may have contributed to the errors

[28] Fedderke et al (2000) find evidence that industry concentration exerts a positive influence on the mark-up in a study of 3-digit manufacturing sectors in South Africa.
[29] The issuance of index-linked bonds in principle legally requires such transparency. We are grateful to Paul Collier for this point.
[30] In each case, for both the "metropolitan" and the "metropolitan and urban" definitions.
[31] Specifically, it is not possible to check whether weighted components add up to the total CPI without a clear understanding of the methodology followed by Statistics SA in constructing CPI.

in data on the CPI components published up to June 2003, which we uncovered in the course of this research, see Appendix 1. Even now, as our Appendix shows, new data for the period 1998 to January 2002 provided by *Statistics South Africa* in September 2003 appear to violate the basic adding-up criterion according to which the CPI is defined. The aggregate discrepancy is not large, a maximum of around 2 percent, but this could correspond to substantially larger errors in individual components. The discrepancies are large enough to cause difficulties for the modelling effort described in this paper and in the extension of the method for practical forecasting.

At a deeper level, issues of quality correction and the treatment of new goods, discount outlets and substitutes, as highlighted by the Boskin Commission in the U.S. (Boskin et al, 1996), appear to have been somewhat neglected in South Africa. Some of our findings suggest the need for further research into these issues, for example, for vehicles and food. Issues of the statistical quality of price quotations and around the introduction of new retail outlets, are ever-present in the measurement of price indices and need to be kept under periodic review. We would argue that the brief examination of the issues conducted by Haglund (2000), though very useful, is insufficient. By contrast, at the Bureau of Labour Statistics (U.S.), at the European Central Bank and Eurostat, and at the Bank of England and the Office of National Statistics, amongst very many other reputable institutions, these issues are of ongoing long-term concern, and the subject of intense internal scrutiny and research programmes, together with interaction with academia.

Given the inflation targeting regime, and the effect of perceptions of economic performance on long term capital inflows, it would be unfortunate if inflation was being overestimated and growth underestimated because of lack of investment and lack of capacity in the over-stretched statistical agencies.

REFERENCES

Akram, Q. Farooq and Ida Wolden Bache.2001. "Empirical modelling of sub-indices of the Norwegian CPI" Research Department, Norges Bank, March.

Aron, J. and J. Muellbauer. 2000a. "Personal and Corporate Saving in South Africa", *World Bank Economic Review* 14 (3): 509-544.

Aron, J. and J. Muellbauer. 2000b. "Financial Liberalisation, Consumption and Debt in South Africa." Working Paper Series 2000: 22, Centre for the Study of African Economics, Oxford University.

Aron, J. and J. Muellbauer. 2000c. "Inflation and Output Forecasts for South Africa: Monetary Transmission Implications." Working Paper Series 2000: 23, Centre for the Study of African Economics, Oxford University.

Aron, J. and J. Muellbauer. 2002a. “Interest rate effects on output: evidence from a GDP forecasting model for South Africa.” *IMF Staff Papers* 49 (November, *IMF Annual Research Conference*): 185-213.

Aron, J. and J. Muellbauer. 2002b. "Estimating Monetary Policy Rules for South Africa", in Norman Loayza and Klaus Schmidt-Hebbel (eds) "Monetary Policy: Rules and Transmission Mechanisms", *Series on Central Banking, Analysis and Economic Policies*, Volume 4, Central Bank of Chile, pages 427-475.

Aron, J., J. Muellbauer and B. Smit. 2003. “Understanding the Inflation Process in South Africa.” Keynote Address, *Eight Annual Conference on Econometric Modelling for Africa*, Stellenbosch University, South Africa, 1-4th July, 2003.

Aron, J. and J. Muellbauer. 2004a. “Construction of CPIX Data for Forecasting and Modelling in South Africa.” CSAE Working Paper Series, Centre for the Study of African Economics, Oxford University (forthcoming, April).

Aron, J. and J. Muellbauer. 2004b. “Multi-Step Inflation Forecasting Models: an Application to South Africa.” CSAE Working Paper Series, Centre for the Study of African Economics, Oxford University (forthcoming).

Barth, M. and V. Ramey. 2001. “The Cost Channel of Monetary Transmission.” *NBER Macroeconomics Annual 16*: 199-240.

Bryan, M. and S. Cecchetti. 1999. “Inflation and the distribution of price changes.” *Review of Economics and Statistics* 81 (2): 188-196.

Boskin, M.J., E. Dulberger, R. Gordon, Z.Griliches and D. Jorgenson. 1996. “Toward a more accurate measure of the cost of living.” Final Report to the Senate Finance Committee, December 4.

Chevalier, J. and D. Scharfstein. 1996. “Capital-Market Imperfections and Countercyclical Markups: Theory and Evidence.” *American Economic Review* 82(2): 346-353.

Clements, Michael P. and David F. Hendry. 1996. “Multi-step estimation for forecasting.” *Oxford Bulletin of Economics and Statistics* 58: 657-684.

Clements, Michael P. and David F. Hendry. 1998. *Forecasting Economic Time Series.* Cambridge: Cambridge University Press.

Espasa, A. and R. Albacete. 2004.“ Econometric modelling for short-term inflation forecasting in the EMU.” Mimeo, Univ. Carlo III de Madrid, Madrid Spain.

Espasa, A., P.Poncela and E. Senra. 2002. “Forecasting monthly US consumer price indexes through a disaggregated I(2) analysis.” Mimeo, Univ. Carlo III de Madrid, Madrid Spain.

Espasa, A., E. Senra and R. Albacete. 2000.“Forecasting EMU inflation: A disaggregated approach by countries and by sectors.” *International J. of Services and Technology and Management* 59-70.

Espasa, A. and M.L. Matea. 1991. “Underlying inflation in the Spanish economy: estimation and methodology.” Working paper 9130, Banco de España.

Espasa, A., M.L. Matea, M.C. Manzano and V. Catasús. 1987. “La inflación subyacente en la economía española: estimación y metodología.” *Boletín Económico*, Banco de España, March.

Fedderke, J., C. Kularatne and M. Mariotti. 2003. “Mark-up Pricing in South Africa Industry.” ERSA, Unversity of the Witwatersrand. Mimeograph.

Fedderke, J. and E. Schaling. 2000. “Modeling Inflation in South Africa: a Multivariate Cointegration Analysis.” Department of Economics Research Papers 0010, RAU, South Africa.

Haglund, Mats. 2000. "The South African CPI - the reliability of present practices and some potential for improvements." Mission Report RSASTAT 2000:2, Statistics Sweden, 2000.

Harvey, A.C. 1993. *Time Series Models,* 2nd Edition *(*Deddington: Philip Alan).

Harvey, A.C. and A. Jaeger. 1993. “Detrending, Stylized Facts and the Business Cycle.” *Journal of Applied Econometrics* 8: 231-47.

Johansen, S. 1988. “Statistical analysis of cointegration vectors.” *Journal of Economic Dynamics and Control* 12: 231-54.

Johansen, S. and K. Juselius. 1990. "Maximum Likelihood Estimation and Inference on Cointegration - with Applications to the Demand for Money." *Oxford Bulletin of Economics and Statistics* 52 (2): 169-210.

Koopman S.J., A.C. Harvey, J.A. Doornik, and N. Shephard, 2000, *STAMP: Structural Time Series Analyser, Modeller and Predictor (*London: Timberlake Consultants Press).

Leiderman, L. and L. E. O. Svensson. 1995. *Inflation Targets.* London: Centre for Economic Policy Research.

Pretorius, C.J. and M.M. Smal. 1994. “A Macro-economic Examination of the Price-formation Process in the South African Economy.” SARB Quarterly Bulletin, March, 25-36.

Sims, C. 1980. “Macroeconomics and reality.” *Econometrica* 48: 1-48.

Sims, C. 1996. “Macroeconomics and methodology.” *Journal of Economic Perspectives* 10: 105-20.

South African Reserve Bank. 2000. A New Monetary Policy Framework. Appendix, Statement of the Monetary Policy Committee, 6 April, 2000.

Stock, J. and M. Watson. 1999. “Forecasting inflation.” Working Paper, No. 7023, National Bureau of Economic Research, Cambridge.

Stock, J., and M. Watson. 2001. “Forecasting Output and Inflation: The Role of Asset Prices.” Working Paper, No. 8180, National Bureau of Economic Research, Cambridge.

Weiss, Andrew A. 1991. “Multi-step estimation and forecasting in dynamic models.” *Journal of Econometrics* 48: 135-149.

Zellner, A. and B. Chen. 2000. “Bayesian modelling of economies and data requirements.” Working paper, H.G.B. Alexander Research Foundation GSB, The University of Chicago.

Zellner, A. and J. Tobias. 2000 “A note on aggregation, disaggregation and forecasting performance.” *Journal of Forecasting.* 19 (5): 457-469.

Table 1: Weights for the consumer price index for metropolitan areas

	Components	Apr 70 - Dec 77	Jan 78 - Oct 87	Nov 87 – Jul 91	Aug 91 - Dec 96	Jan 97- Dec 01	Jan 02 -
Services	Housing	21.6	19.5	22.5	20.5	26.0	24.3
	Mortgage interest cost	3.61	3.4	9.47	11.51	12.91	11.43
	Transport	4.9	3.7	5.9	4.3	4.3	3.4
	Other	7.1	9.7	11.1	17.3	14.7	15.2
	Total	33.6	32.9	39.5	42.1	45.0	42.9
Goods	Food	23.9	25.5	23.2	19.3	18.8	22.1
	Furniture and equipment	7.8	6.0	4.7	5.5	3.9	2.5
	Clothing and footwear	9.6	8.8	6.0	7.0	4.8	3.2
	Vehicles	6.7	5.6	5.5	5.5	5.3	6.0
	Transport goods: running costs	5.0	5.6	5.9	4.6	5.2	5.5
	Beverages and tobacco	4.1	3.8	2.3	2.2	2.1	2.5
	Other	9.3	11.8	12.9	13.8	14.9	15.3
	Total	66.4	67.1	60.5	57.9	55.0	57.1
Total		100.0	100	100.0	100.0	100.0	100.0

SOURCE: SARB and Statistics SA

Notes on the composition of the deflators:
Vehicles: Include personal transport equipment and motorcars tyres, parts and accessories.
Transport running costs: Include Household fuel and power and petroleum products.
Other goods: Include Recreational and entertainment goods, other durable goods, household textiles, furnishings and glassware, personal goods and writing and drawing equipment, household consumer goods, medical and pharmaceutical products.
Housing: Include rent, incorporating rent for owner-occupied dwellings and
Transport: Include transport and communication services household services.
Other services: medical services, recreational, entertainment and educational services and other miscellaneous services.

Table 2: Statistics and Variable Definitions: 1979:2-2003:1

Sub-component Equations	*Definition of Variable*	*Mean*	*Standard Dev.*	*I(1)*[a,b]	*I(2)*[a,b]
Housing services					
log (HX)	Log of the "housing" sub-component of CPI (metropolitan), with the mortgage interest component subtracted, using the correct weights for each period of weighting	3.4	0.83	-3.33*	-7.85**
FITCPIXC	Our forecast of aggregate CPIX (details in (Aron and Muellbauer, 2004b), where CPIX is constructed by the authors (see Aron and Muellbauer, 2004a)	0.11	0.04	-3.29	-3.72**
ECM2HX	ECM2HX=log(WPDTOT)-log(HX), where WPDTOT is the wholesale price for domestic goods	0.35	0.16	-1.62	-6.32**
Log (CPIXC)	Log of CPIX, the consumer price index excluding mortgage interest costs – constructed	-1.12	0.81	-2.46	-3.57*
Log (MANULC)	Log of manufacturing unit labour costs, adjusted to remove a productivity trend, see section 4.2. [d]	3.8	0.75	-4.02**	-8.78**
Log (HPRICE)	House price, total SA: New & Old - Medium - Purchase Price - Smoothed (Source: ABSA HOUSE PRICE DATABASE)	11.6	0.62	-5.30*	-4.15**
Log (NEERN)	Log of the constructed nominal effective exchange rate (see text)	2.8	0.68	-2.47	-4.48**
FLIB	Financial liberalisation measure – see text	-	-	-	-
Transport services					
Log (TS)	Log of the "transport services" sub-component of CPI (metropolitan)	3.9	0.59	-1.78	-3.70*
ECM1TS	ECM1TS=log(CPIXC)-log(TS)	-5.1	0.24	-2.80	-6.39**
ECM3TS	ECM3TS=log(WPIMP)-log(TS) , where WPIMP is the wholesale price for imported goods	-0.13	0.13	-1.03	-3.28*
ECMULCTS	ECMULCTS=log(MANULC)-log(TS) , using manufacturing unit labour costs, adjusted to remove a productivity trend (section 4.2)	-0.19	0.18	-2.80	-9.47**
OUTGAP	Measure of the output gap derived as explained in section 4.2.	3.96	0.025	-3.57*	-5.72**
log (OIL)	Log of (Rand) Brent oil price	4.1	0.58	-2.38	-4.96**
Other services					
log (OS)	Log of the "other services" sub-component of CPI (metropolitan)	3.4	0.91	-2.87	-7.97**
FITCPIXC	Our forecast of aggregate CPIX (details in (Aron and Muellbauer, 2004b), where CPIX is constructed by the authors (see Aron and Muellbauer, 2004a)	0.11	0.04	-3.29	-3.72**
ECM1OS	ECM1OS=log(CPIXC)-log(OS)	-4.5	0.10	-3.28	-9.94**
log (WPDTOT)	Log of the wholesale price for domestic goods	3.8	0.70	-1.54	-4.41**
OUTGAP	Measure of the output gap derived as explained in section 4.2	3.96	0.025	-3.57*	-5.72**
Food					
log (FD)	Log of the "food" sub-component of CPI (metropolitan)	3.6	0.85	-2.42	-4.87**
FITCPIXC	Our forecast of aggregate CPIX (details in (Aron and Muellbauer, 2004b), where CPIX is constructed by the authors (see Aron and Muellbauer, 2004a)	0.11	0.04	-3.29	-3.72**

Sub-component Equations	***Definition of Variable***	***Mean***	***Standard Dev.***	***I(1)***[a,b]	***I(2)***[a,b]
ECM1FD	ECM1FD=log(CPIXC)-log(FD)	-4.7	0.049	-3.34	-6.94**
ECM3FD	ECM3FD=log(WPIMP)-log(FD), where WPIMP is the wholesale price for imported goods	0.25	0.23	-1.38	-4.52**
Log (USMAIZER)	The rand price of US maize (Chicago)using the bilateral rand/$ exchange rate	5.5	0.65	-3.99*	-7.51**
OUTGAP	Measure of the output gap derived as explained in section 4.2.	3.96	0.025	-3.57*	-5.72**
Log (TOTXG)	Log terms-of-trade (excluding gold)	4.6	0.051	-3.04*	-14.5**
Vehicles					
Log (VH)	Log of the "vehicles" sub-component of CPI (metropolitan)	3.5	0.98	-2.20	-3.89**
FITCPIXC	Our forecast of aggregate CPIX (details in (Aron and Muellbauer, 2004b), where CPIX is constructed by the authors (see Aron and Muellbauer, 2004a)	0.11	0.04	-3.29	-3.72**
ECM1VH	ECM1VH=log(CPIXC)-log(VH)	-4.6	0.18	-1.32	-5.13**
ECM3VH	ECM3VH=log(WPIMP)-log(VH), where WPIMP is the wholesale price for imported goods	0.31	0.35	-1.71	3.16*
ECMULCVH	ECMULC=log(MANULC)-log(VH) , using manufacturing unit labour costs, adjusted to remove a productivity trend (section 4.2)	0.25	0.24	-2.12	-7.60**
OUTGAP	Measure of the output gap derived as explained in section 4.2.	3.96	0.025	-3.57*	-5.72**
Log (LTOTIG)	Log terms-of-trade (including gold)	4.6	0.066	-3.36*	-12.6**
Clothing					
log (CL)	Log of the "clothing" sub-component of CPI (metropolitan)	4.0	0.64	-2.64	-3.94*
ECM2CL	ECM2CL=log(WPCL)-log(CL) , where WPCL is the wholesale price for clothing	-0.17	0.11	0.882	-3.82**
RCASUR	Current account surplus to GDP ratio, both seasonally adjusted	0.0040	0.033	-4.59**	-14.1**
Furniture					
log (FR)	Log of the "furniture" sub-component of CPI (metropolitan)	4.0	0.61	-3.20*	-2.96*
FITCPIXC	Our forecast of aggregate CPIXC (details in (Aron and Muellbauer, 2004b), where CPIXC is constructed by the authors (see Aron and Muellbauer, 2004a)	0.11	0.04	-3.29	-3.72**
ECM1FR	ECM1FR=log(CPIXC)-log(FR)	-5.1	0.22	-2.07	-8.29**
ECM2FRTOT	ECM2FRTOT=log(WPMTOT)-log(FR), where WPMTOT is the wholesale price for manufactured goods	-0.22	0.13	-0.543	-8.13**
log(MANULC)	Log of manufacturing unit labour costs, adjusted to remove a productivity trend, see section 4.2. [d]	3.7	0.75	-4.02**	-8.78**
RCASUR	Current account surplus to GDP ratio, both seasonally adjusted	0.0040	0.033	-4.59**	-14.1**
Log (REERN)	Weighted average of the log of the SARB's real effective exchange rate & bilateral US$/R rate – see text.	2.2	0.17	-1.96	-7.82**
Beverages & Tobacco					
Log (BT)	Log of the "beverages and tobacco" sub-component of CPI (metropolitan)	3.3	0.93	-0.593	-6.88**
ECM1BT	ECM1BT=log(CPIXC)-log(BT)	-4.4	0.15	-1.93	-7.16**

Sub-component Equations	***Definition of Variable***	***Mean***	***Standard Dev.***	***I(1)***[a,b]	***I(2)***[a,b]
RCASUR	Current account surplus to GDP ratio, both seasonally adjusted	0.0040	0.033	-4.59**	-14.1**
RPRIME	Prime rate/100 less the annual change in the log of CPIXC	0.056	0.054	-4.06*	-6.77*
log (LTOTIG)	Log terms-of-trade (including gold)	4.6	0.066	-3.36*	-12.6**
DUMOPEN* Log (LTOTIG)	The saved stochastic trend from the share of import demand equation in Table 3, column 1, minus (4.30 x RTARIF(-1)), Aron and Muellbauer (2002a)	-	-	-	-
Transport goods					
log (TG)	Log of the "transport goods" sub-component of CPI (metropolitan)	3.6	0.68	-3.34	-5.05**
FITCPIXC	Our forecast of aggregate CPIXC (details in (Aron and Muellbauer, 2004b), where CPIXC is constructed by the authors (see Aron and Muellbauer, 2004a)	0.11	0.04	-3.29	-3.72**
ECM1TG	ECM1TG=log(CPIXC)-log(TG)	-4.7	0.18	-2.46	-4.78**
ECMULCTG	ECMULCTG=log(MANULC)-log(TG), using manufacturing unit labour costs, adjusted to remove a productivity trend (section 4.2)	0.15	0.18	-2.31	-4.66**
OUTGAP	Measure of the output gap derived as explained in section 4.2	3.96	0.025	-3.57*	-5.72**
RCASUR	Current account surplus to GDP ratio, both seasonally adjusted	0.0040	0.033	-4.59**	-14.1**
log (ROILTG)	Log(ROIL)=log(OIL)-log(TG)	0.46	0.31	-2.13	-9.18**
Other goods					
Log (OG)	Log of the "other goods" sub-component of CPI (metropolitan)	3.6	0.83	-1.79	-2.97*
FITCPIXC	Our forecast of aggregate CPIXC (details in (Aron and Muellbauer, 2004b), where CPIX is constructed by the authors (see Aron and Muellbauer, 2004a)	0.11	0.04	-3.29	-3.72**
ECM1OG	ECM1OG=log(CPIXC)-log(OG)	-4.7	0.039	-2.30	-8.21**
Log(MANULC)	Log of manufacturing unit labour costs, adjusted to remove a productivity trend, see section 4.2.[d]	3.7	0.75	-4.02**	-8.78**
OUTGAP	Measure of the output gap derived as explained in section 4.2	3.96	0.025	-3.57*	-5.72**
Log (REERN)	Weighted average of the log of the SARB's real effective exchange rate & bilateral US$/R rate – see text	2.2	0.17	-1.96	-7.82**
Log (TOTIG)	Log terms-of-trade (including gold)	4.6	0.066	-3.36*	-12.6**

a. The variables are those included in sections (5.2)-(5.3). Tables 3A-3D present parsimonious equations where some variables are transformed into moving averages and delta terms, during general to specific reduction.

b. For a variable X, the augmented Dickey-Fuller (1981) statistic is the t ratio on π from the regression: $\Delta X_t = \pi X_{t-1} + \Sigma_{i=1,k} \theta_i \Delta X_{t-i} + \psi_0 + \psi_1 t + \varepsilon_t$, where k is the number of lags on the dependent variable, ψ_0 is a constant term, and t is a trend. The kth-order augmented Dickey-Fuller statistic is reported, where k is the last significant lag of the 3 lags employed. The trend is included only if significant. For null order I(2), ΔX replaces X in the equation above. Critical values are obtained from MacKinnon (1991). Asterisks * and ** denote rejection at 5% and 1% critical values. Stationarity tests are performed for the variables in levels before time-transformation i.e. before taking moving averages and changes.

c. Variables from SARB except house prices (ABSA), prime rate (IMF), US wholesale prices (IMF), maize prices (IMF) and Brent oil prices (IMF).

d. The apparent stationarity of log(manulc) is spurious: the error correction coefficient is less than 1 percent.

Table 3a. Services sub-component price equations

Housing services	***Housing services***	***Transport services***	***Transport services***	***Other services***	***Other services***
Dependent variable: $\Delta_4$log (HX)(+4)	Full sample 1979:2-2002:2	Dependent variable: $\Delta_4$log (TS)(+4)	Full sample 1979:2-2002:2	Dependent variable: $\Delta_4$log (OS)(+4)	Full sample 1979:2-2002:2
Regressors	***STAMP results***	***Regressors***	***STAMP results***	***Regressors***	***STAMP results***
Δlog (HX)	-1.03 [11.6]	Δlog (TS)	-1.14 [12.6]	Δlog (OS)	-0.67 [8.0]
FITCPIXC	0.76 [4.8]	Δlog (TS(-1)	-1.06 [9.8]	Δlog (OS(-1)	-0.44 [5.7]
ECM2HX(-1)	0.90 [8.9]	Δlog (TS(-2)	-0.44 [3.5]	Δlog (OS(-2)	-0.13 [1.7]
RPRIME(-3)	0.18 [2.1]	Δlog (TS(-3)	-0.21 [2.2]	FITCPIXC	0.70 [5.0]
$\Delta_4$log (MANULC)	0.084 [2.1]	ECM1TS(-2)	0.46 [3.2]	ECM1OSMA4	1.2 [5.3]
Δlog (HPRICE)	0.30 [3.1]	ECM3TS(-3)	0.17 [2.6]	$\Delta_4$log (WPDTOT)	0.25 [2.3]
Δlog (NEERN)	-0.053 [2.7]	ECMULCTSMA4	0.50 [3.2]	OUTGAP(-1)	0.28 [1.9]
FLIB	-0.19 [1.6]	OUTGAP	1.01 [5.0]	-	-
DLCPIXC	0.84 [4.9]	Δlog (OIL) (-1)	0.017 [2.5]	-	-
Diagnostics		***Diagnostics***		***Diagnostics***	
Equation Std error	0.0144	Equation Std error	0.0136	Equation Std error	0.0142
Rmse of slope	0.0084	Rmse of slope	0.0097	Rmse of slope	0.0057
Durbin-Watson	1.77	Durbin-Watson	1.92	Durbin-Watson	1.68
R^2	0.588	R^2	0.689	Adjusted R^2	0.581

(Absolute values of asymptotic *t*-ratios in parentheses)

1. See Table 1 for definitions of the variables (before transformation).
2. All samples begin in 1979:2 (floating exchange rate regime)

Table 3b. Goods sub-component price equations

Food	***Food***	Vehicles	***Vehicles***	***Clothing***	***Clothing***
Dependent variable: $\Delta_4$log (FD)(+4)	Full sample 1979:2-2002:2	Dependent variable: $\Delta_4$log (VH)(+4)	Full sample 1979:2-2002:2	Dependent variable: $\Delta_4$log (CL)(+4)	Full sample 1979:2-2002:2
Regressors	***STAMP results***	***Regressors***	***STAMP results***	***Regressors***	***STAMP results***
Δlog (FD)	-0.96 [11.1]	Δlog (VH)	-0.80 [9.4]	Δlog (CL)	-0.97 [9.9]
$\Delta_4$log (FD)(-1)	-0.21 [2.5]	Δlog (VH) (-1)	-0.55 [5.5]	Δlog (CL) (-1)	-0.56 [4.8]
FITCPIXC	0.39 [2.5]	Δlog (VH) (-2)	-0.27 [3.3]	Δlog (CL) (-2)	-0.21 [1.9]
ECM1FD(-1)	0.84 [4.5]	FITCPIXC	0.33 [3.1]	ECM2CLMA3	0.62 [3.6]
ECM3FDMA4	0.61 [4.2]	ECM1VHMA4	1.34 [5.6]	Δ_4RCASUR	-0.11 [3.0]
$\Delta_3$log (USMAIZER)	0.024 [2.8]	ECM3VH	0.13 [2.4]	-	-
OUTGAP	0.51 [2.4]	ECMULCVH(-1)	0.11 [2.1]	-	-
(LogTOTXGMA2) (-1)	0.31 [3.5]	OUTGAP	0.49 [2.8]	-	-
-	-	Log (LTOTIGMA3)(-2)	0.26 [3.6]	-	-
Diagnostics		***Diagnostics***		***Diagnostics***	
Equation Std error	0.0151	Equation Std error	0.0120	Equation Std error	0.0117
Rmse of slope	0.0114	Rmse of slope	0.0111	Rmse of slope	0.0093
Durbin-Watson	1.88	Durbin-Watson	1.93	Durbin-Watson	1.86
R^2	0.739	R^2	0.738	R^2	0.511

(Absolute values of asymptotic *t*-ratios in parentheses)

1. See Table 1 for definitions of the variables (before transformation).
2. All samples begin in 1979:2 (floating exchange rate regime)

Table 3c. Goods sub-component price equations

Furniture	***Furniture***	***Beverages and Tobacco***	***Beverages and Tobacco***
Dependent variable: $\Delta_4$log (FR)(+4)	Full sample 1979:2-2002:2	Dependent variable: $\Delta_4$log (BT)(+4)	Full sample 1979:2-2002:2
Regressors	***STAMP results***	***Regressors***	***STAMP results***
Δlog (FR)	-1.03 [8.5]	Δlog (BT)	-1.07 [12.7]
Δlog (FR) (-1)	-0.81 [7.0]	Δlog (BT) (-1)	-0.55 [6.4]
Δlog (FR) (-2)	-0.55 [5.3]	Δlog (BT) (-2)	-0.47 [4.4]
FITCPIXC	0.37 [4.0]	Δlog (BT) (-3)	-0.21 [2.5]
ECM1FRMA4	0.76 [5.1]	ECM1BTMA2(-1)	0.51 [2.5
ECM2FRTOTMA4(-1)	0.62 [4.1]	Δ_4RCASUR	-0.14 [3.2]
$\Delta_2$log(MANULC)	0.088 [2.8]	Δ_4RCASUR(-1)	-0.10 [2.2]
Δ_4RCASUR	-0.086 [2.7]	RPRIME (-2)	0.33 [3.4]
Log (REERNMA4)	-0.13 [4.2]	(log TOTIGma2)(-3)	0.21 [3.5]
-	-	DUMOPEN* (Log LTOTIGma2)(-3)	-0.13 [1.6]
Diagnostics		***Diagnostics***	
Equation Std error	0.00995	Equation Std error	0.01338
Rmse of slope	0.0036	Rmse of slope	0.0112
Durbin-Watson	1.55	Durbin-Watson	1.91
R^2	0.59	R^2	0.698

(Absolute values of asymptotic *t*-ratios in parentheses)

1. See Table 1 for definitions of the variables (before transformation).
2. All samples begin in 1979:2 (floating exchange rate regime)

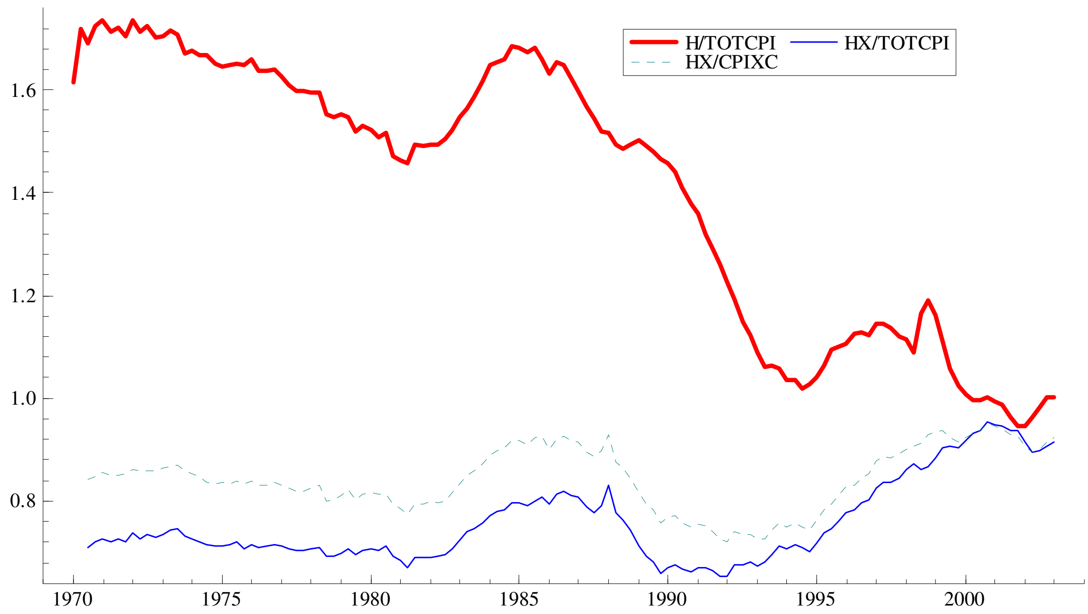
Table 3d. Goods sub-component price equations

Transport goods	***Transport goods***	***Other goods***	***Other goods***
Dependent variable: $\Delta_4$log (TG)(+4)	Full sample 1979:2-2002:2	Dependent variable: $\Delta_4$log (OT)(+4)	Full sample 1979:2-2002:2
Regressors	***STAMP results***	***Regressors***	***STAMP results***
Δlog (TG)	-1.12 [13.3]	Δlog (OG)	-0.53 [5.6]
Δlog (TG) (-1)	-0.90 [4.8]	Δlog (OG) (-1)	-0.46 [4.9]
Δlog (TG) (-2)	-0.87 [4.5]	Δlog (OG) (-2)	-0.22 [2.3]
FITCPIXC	0.85 [2.5]	FITCPIXC	0.68 [5.2]
ECM1TG(-3)	1.02 [4.6]	ECM1OGMA4	1.70 [11.5]
ECMULCTG(-1)	0.33 [2.2]	D3LMANULC-1	0.085 [2.2]
log (ROILTG) (-3)	0.055 [2.0]	OUTGAP	0.42 [2.8]
OUTGAPMA3	2.13 [3.7]	Log (REERNMA4)	-0.17 [3.7]
Δ_4RCASUR(-1)	-0.24 [2.1]	Log (LTOTIGMA4)(-1)	0.15 [2.1]
Diagnostics		***Diagnostics***	
Equation Std error	0.0342	Equation Std error	0.0132
Rmse of slope	0.0200	Rmse of slope	0.0060
Durbin-Watson	1.78	Durbin-Watson	1.71
R^2	0.717	R^2	0.633

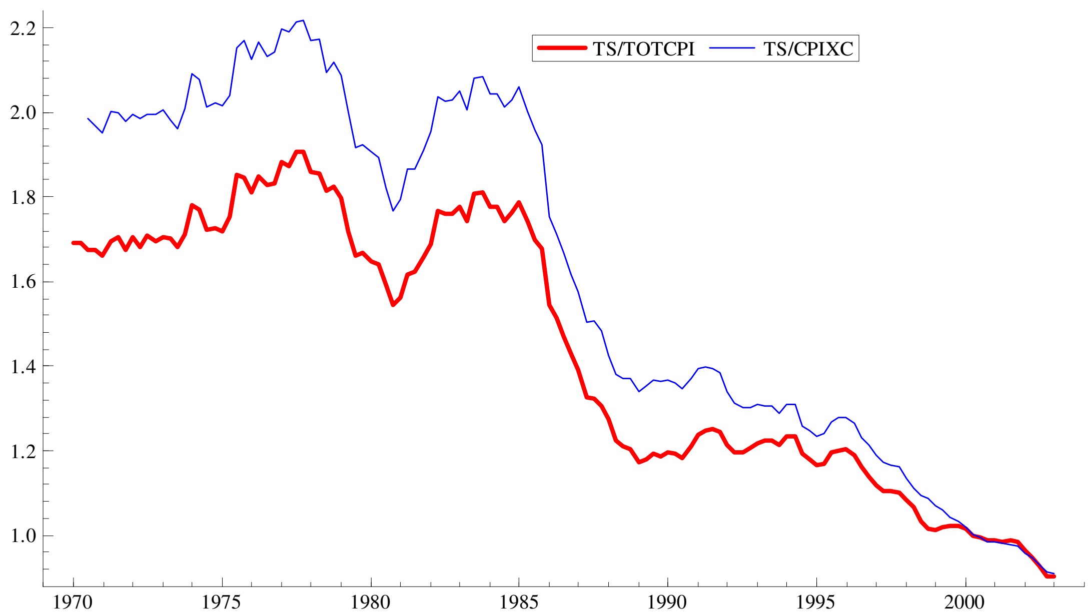
(Absolute values of asymptotic *t*-ratios in parentheses)
1. See Table 2 for definitions of the variables (before transformation).
2. All samples begin in 1979:2 (floating exchange rate regime)

FIGURE 1: Relative prices of CPI components to headline CPI (metropolitan) index (data from September 2003 Quarterly Bulletin data)

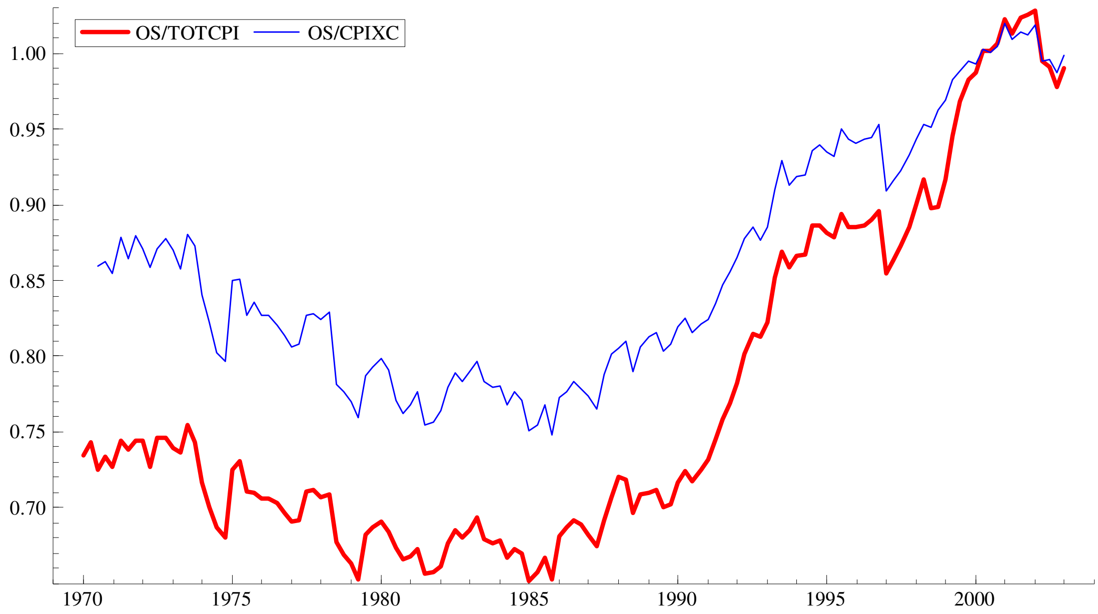
(i) Relative price of the housing component



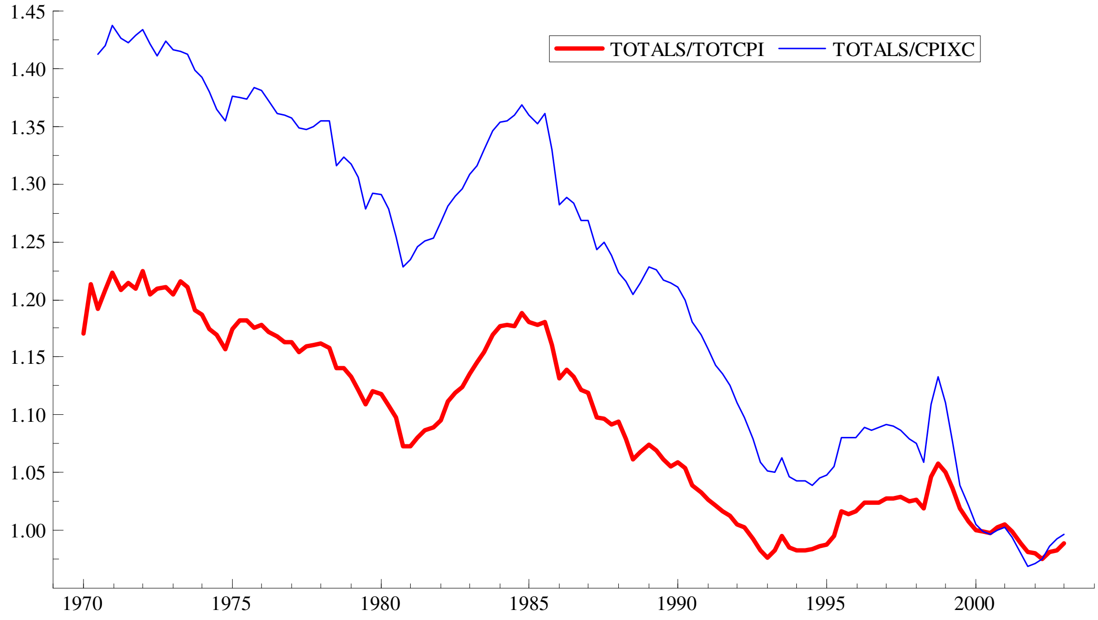
(ii) Relative price of the transport services component



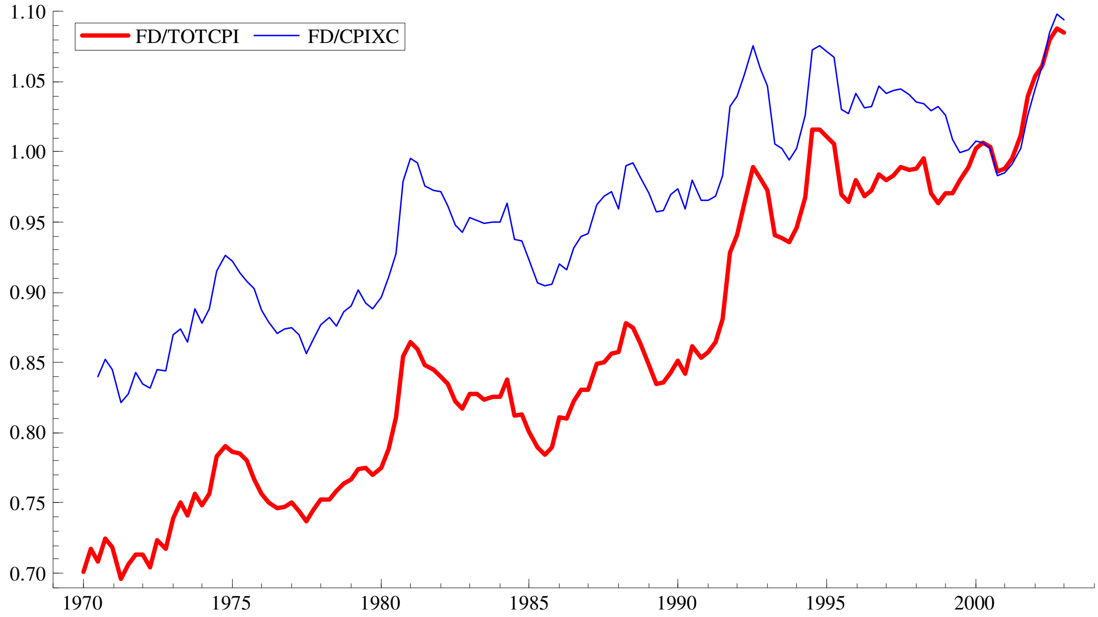
(iii) Relative price of the other services component



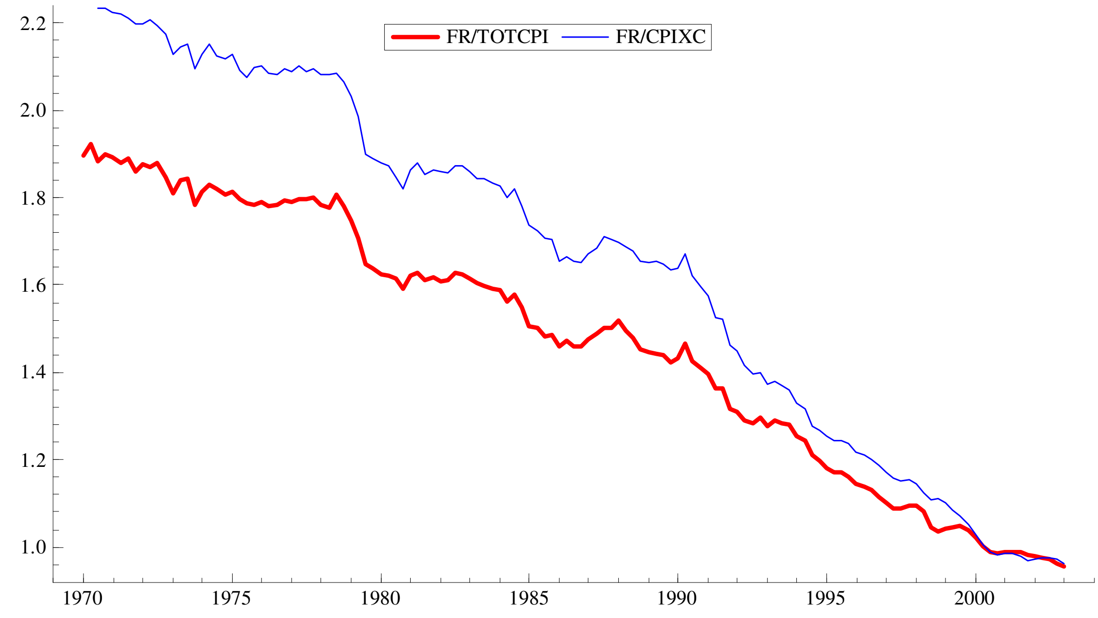
(iv) Relative price of <u>total services</u> component (includes mortgage interest cost)



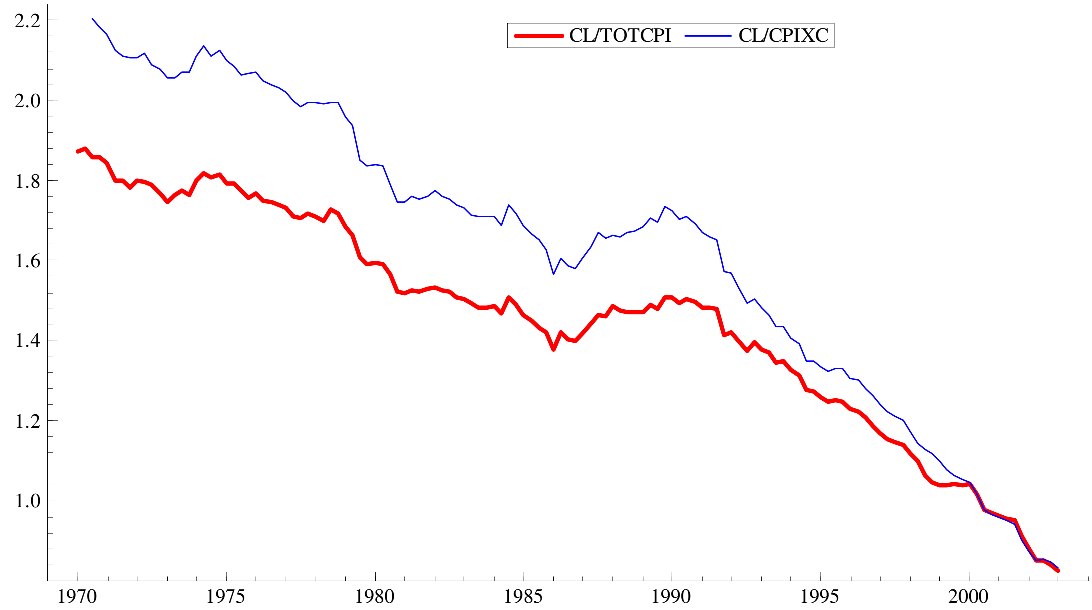
(v) Relative price of the food component



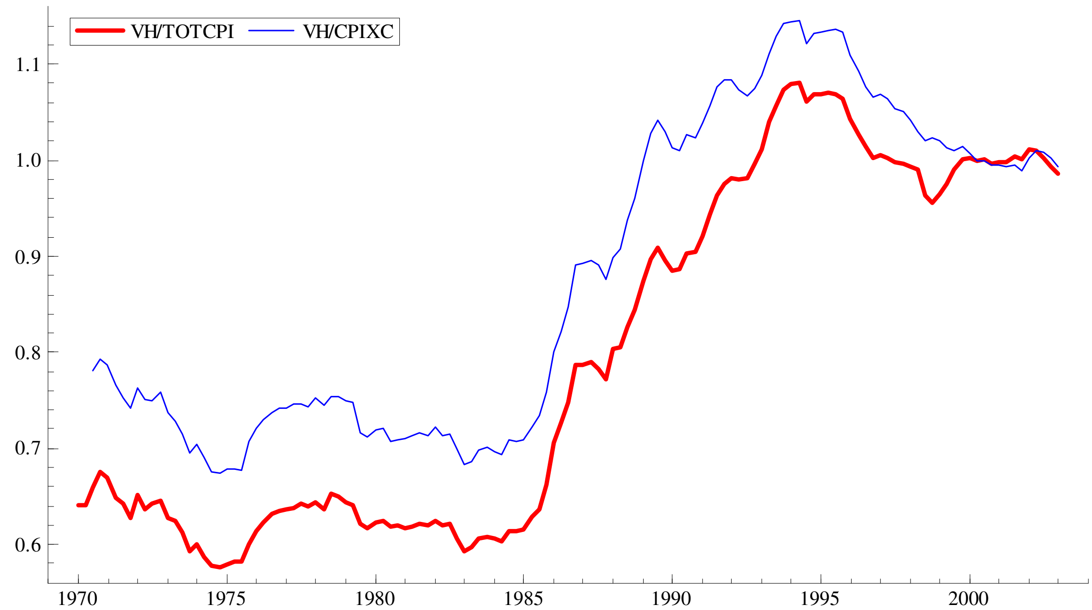
(vi) Relative price of the furniture and equipment component



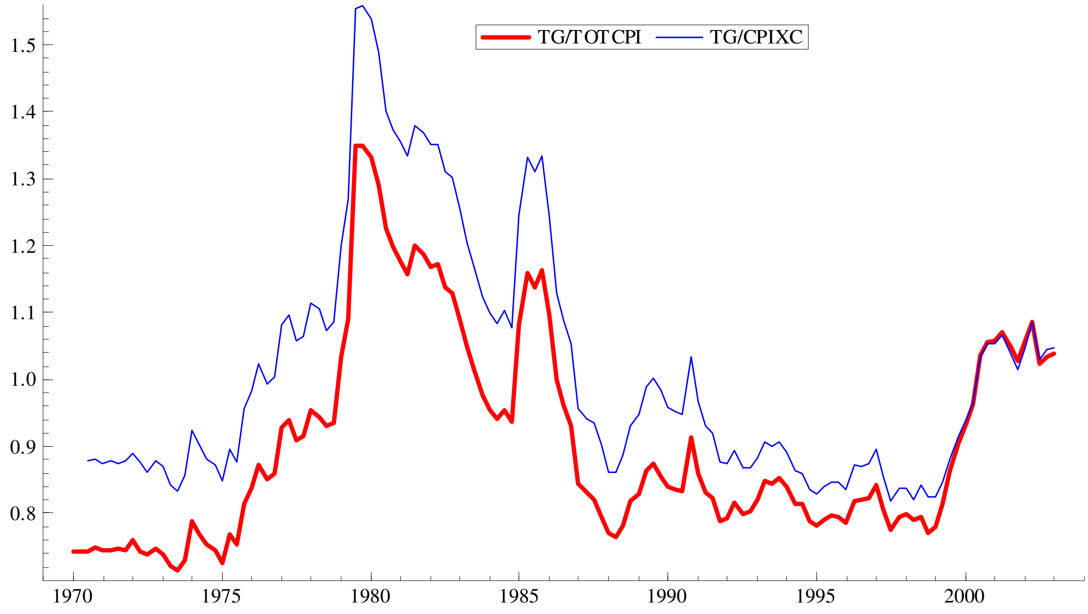
(vii) Relative price of the clothing and footwear component



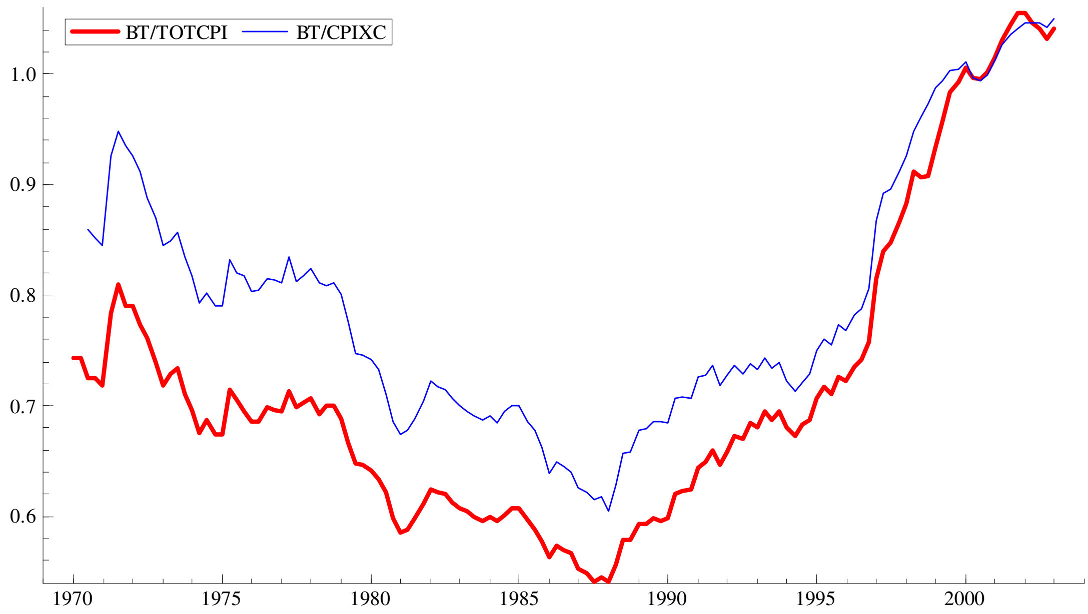
(viii) Relative price of the vehicles component



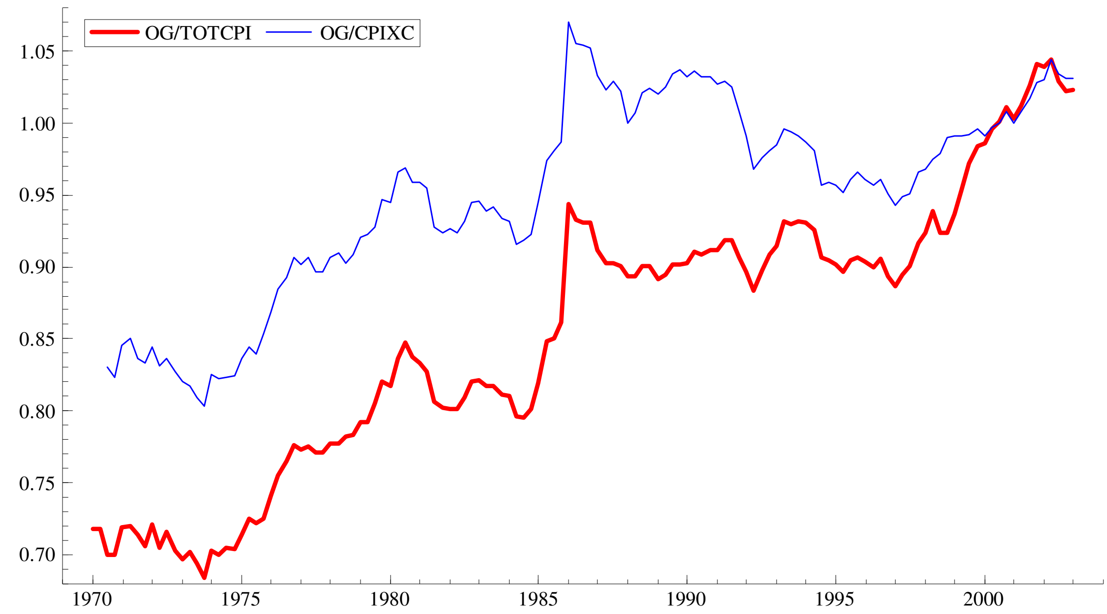
(ix) Relative price of the transport goods component



(x) Relative price of the beverages and tobacco component



(xi) Relative price of the other goods component



(xii) Relative price of total goods component

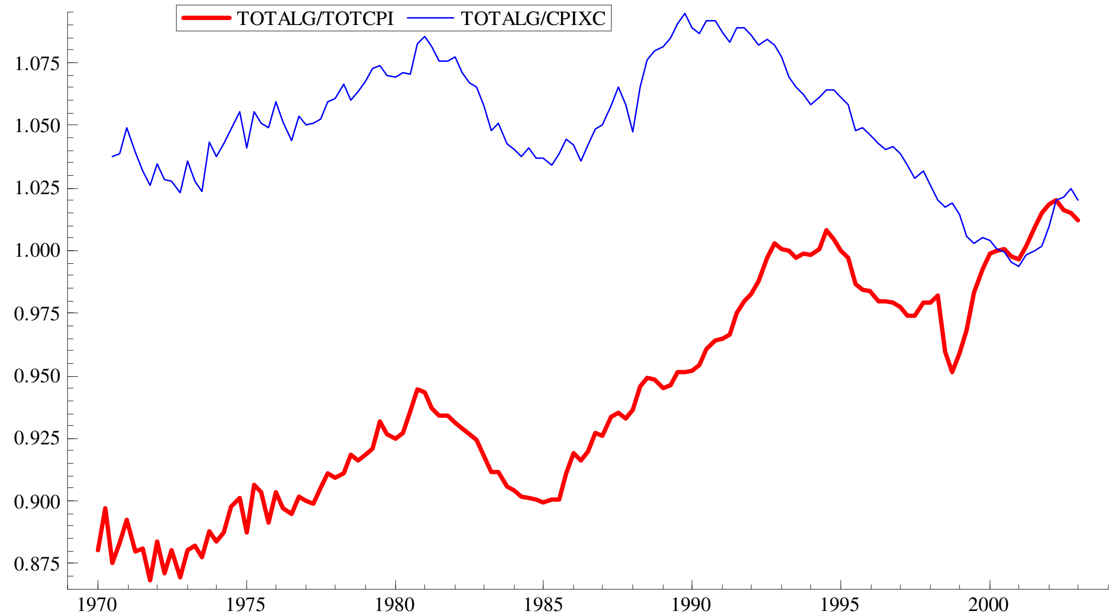


Figure 2. The Openness Indicator for South Africa.

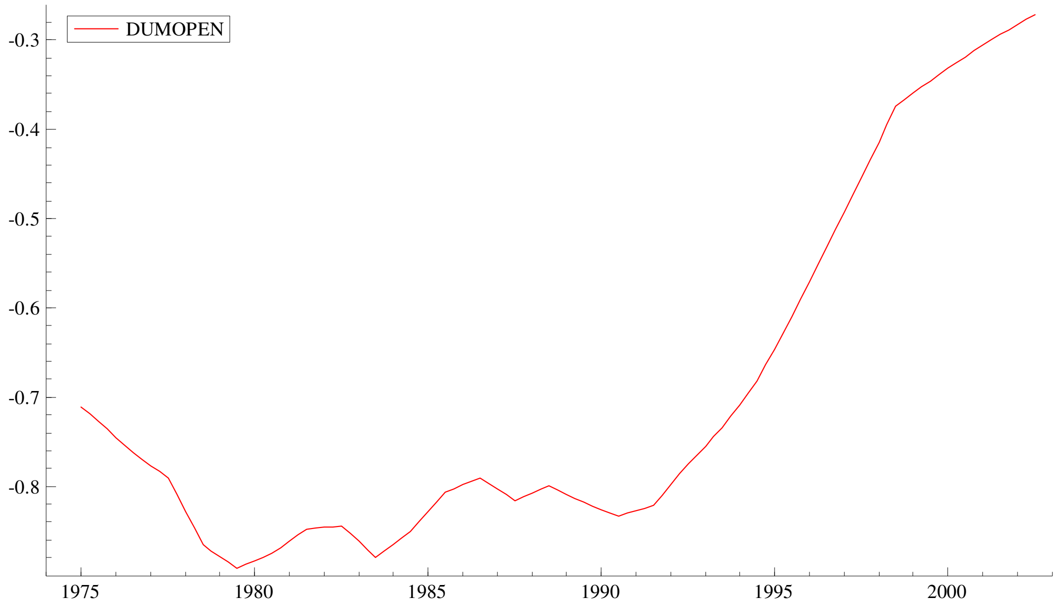
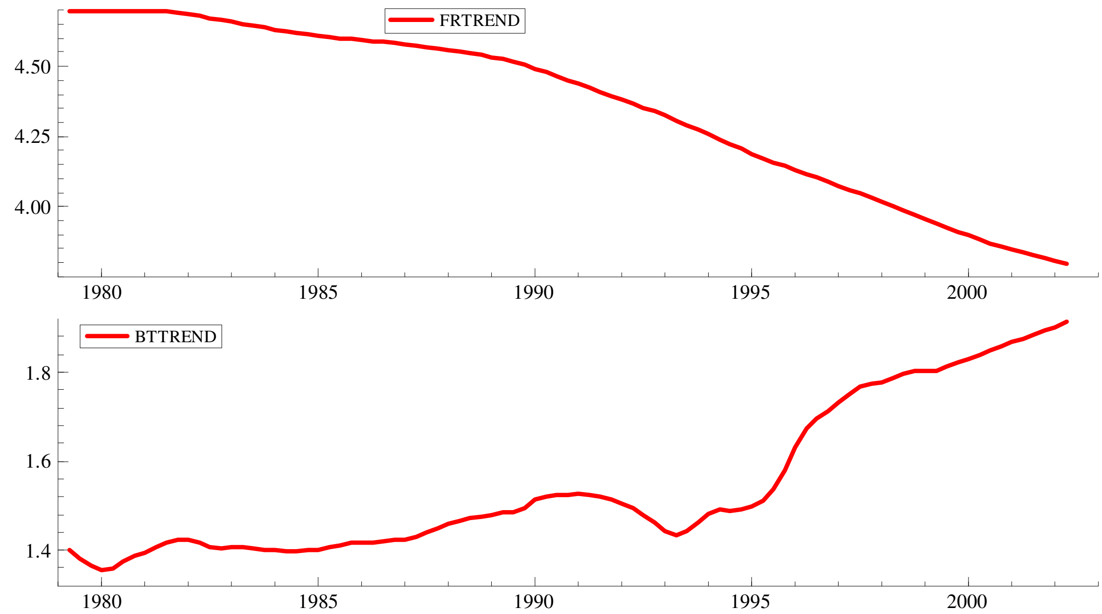


FIGURE 3: Stochastic trends from the forecasting equations of the ten components of the CPI index


FRTREND
4.50
4.25
4.00
1980
1985
1990
1995
2000
BTTREND
1.8
1.6
1.4
1980
1985
1990
1995
2000

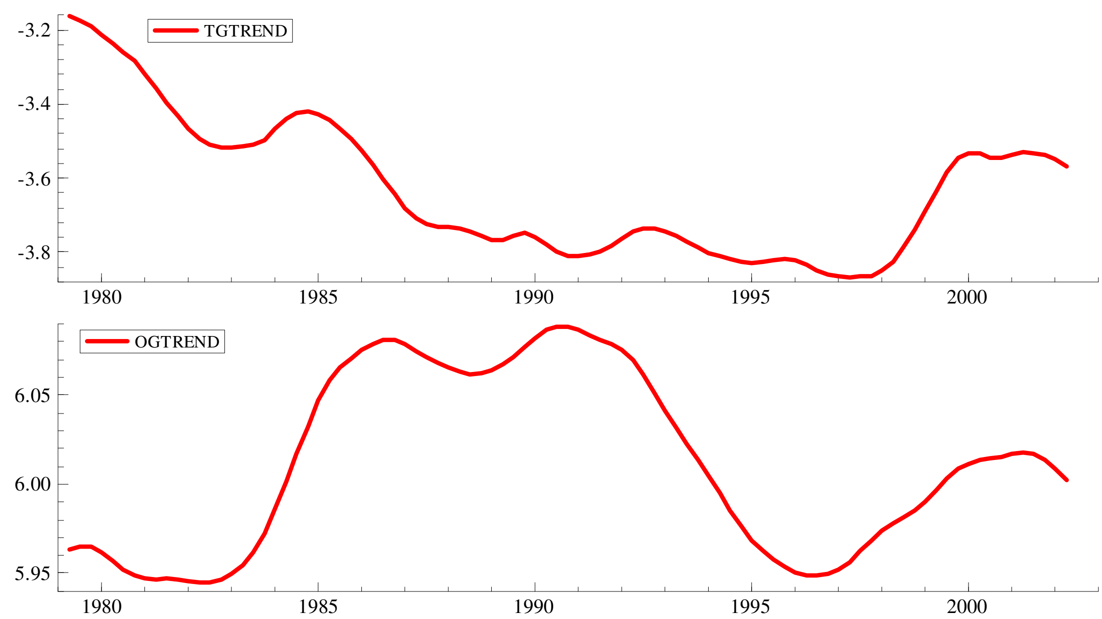
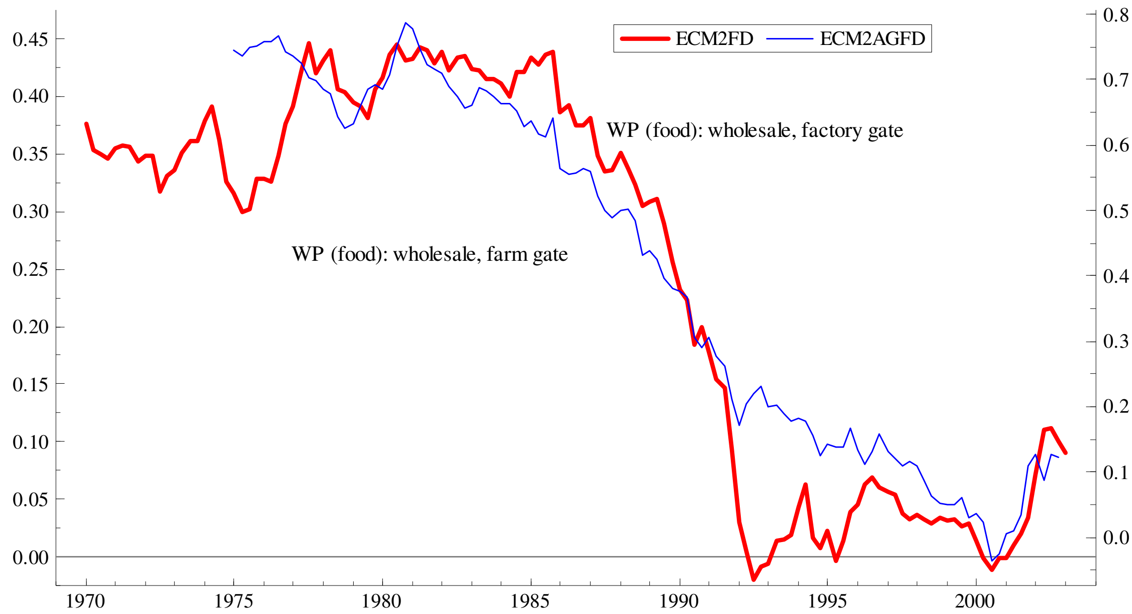

TGTREND
-3.2
-3.4
-3.6
-3.8
1980
1985
1990
1995
2000
OGTREND
6.05
6.00
5.95
1980
1985
1990
1995
2000

FIGURE 4: Relative prices of wholesale food components to the food component of the CPI index



APPENDIX 1: Do the components add up to the total CPI when weighted?

Table A1: Periods where discrepancies exceed 1 percent:

JUNE 1999 SARB Website, 1995=100	JUNE 2003 SARB Website, 2000=100 has Stats SA data (seas adj) only from 1994	SEPT. 03 SARB Website, 2000=100 has Stats SA data (seas adj) from 1986 (1st time on website)
1978:9-1984:4	1979:1-1984:1	
1986:1-1987:11	1986:1-1987:1	1986:1-1986:12
		1999:12-2002:1

Summary

We test whether the components of the CPI, when weighted correctly, add up to the total CPI. Taking the SARB *Quarterly Bulletin* June 1999 (which contained the new presentation of the National Accounts data), as a benchmark, we find discrepancies appear during 1978:9-1984:4 and 1986:1-1987:11 (column 1 above, details available from authors). These discrepancies were sustained until June 2003. When we alerted the SARB to data errors for several CPI sub-components (notably in "other goods" and "other services" and "housing") these data were replaced on the SARB Web from September 2003 with seasonally-adjusted data from Statistics SA back to 1986; while errors in earlier data (seasonally-adjusted by the SARB) were largely eliminated. However, while this unfortunately retained the problems of not adding up in 1986, it also introduced new problems from 1998 to 2003, which were not there in June 2003.

Note that *Statistics SA* seasonally adjust totals such as CPI ("metropolitan") independently of seasonal adjustment applied to the components. Thus some difference is expected between the weighted sum of the seasonally-adjusted components and the separately seasonally-adjusted total. However, discrepancies of this magnitude above (e.g. 2 percent in January, 2002) could not have arisen from such methods of seasonal adjustment.

Repeated attempts to find out the source of these new errors from *Statistics SA* have not borne fruit, and hence we are using these data subject to the caveat in this Appendix.

Method

We have carried out a consistency check on the seasonally-adjusted data for the 10 sub-components, beginning in 1978, using the weights for the expenditure survey years 1975, 1985, 1990, 1995, 2000, applied over the samples shown below, when these weights were effective. We constructed a weighted sum of the components, and divided this by the value of the weighted sum in the first month of the sample. We also divided total CPI by its value in the same month. We did this for each sub-period. This should show whether the overall index, constructed as a weighted average, matches the overall CPI (met) index already on the SARB web site. The following TSP code illustrates the consistency check with monthly data when using weights for 1975 expenditure patterns for the ten components and the total CPI (met).

```
FREQ M; SMPL 1978:1 1978:1;
?list the components and total CPI
LIST CPICOMP H TS OS FD FR CL VH TG BT OG totcpi;
?label the base year variables in 1978:1
DOT CPICOMP; msd .; set .b78 =@mean; enddot;
?weighted average, 1975 weights, for the period of their application
FREQ M; SMPL 1978:1 1987:11;
WCOMB78=(19.5/100)*H/Hb78+(3.7/100)*TS/TSb78+(9.7/100)*OS/OSb78
```

```
+(25.5/100)*FD/FDb78+(6/100)*FR/FRb78+(8.8/100)*CL/CLb78+(5.6/100)*VH/VHb78
+(5.6/100)*TG/TGb78+(3.8/100)*BT/BTb78+(11.8/100)*OG/OGb78;
TOTCPI78=TOTCPI/TOTCPIb78;
?ratio of weighted average and total CPI each relative to their respective
base years
r78=WCOMB78/TOTCPI78; print r78;
```

Below follow the ratios for the five periods, with discrepancies greater than 0.75 percent indicated in bold.

SARB SEP 2003
CONSISTENCY CHECK

	R78	R87	R91	R97	R00
1978:1	1.00000	.	.	.	.
1978:2	0.99653	.	.	.	.
1978:3	0.99782	.	.	.	.
1978:4	0.99509	.	.	.	.
1978:5	0.99852	.	.	.	.
1978:6	0.99350	.	.	.	.
1978:7	0.99655	.	.	.	.
1978:8	**0.98684**	.	.	.	.
1978:9	0.99154	.	.	.	.
1978:10	**0.98852**	.	.	.	.
1978:11	**0.98866**	.	.	.	.
1978:12	0.99541	.	.	.	.
1979:1	0.99249	.	.	.	.
1979:2	0.99240	.	.	.	.
1979:3	0.99729	.	.	.	.
1979:4	0.99471	.	.	.	.
1979:5	0.99315	.	.	.	.
1979:6	**0.98964**	.	.	.	.
1979:7	1.00922	.	.	.	.
1979:8	0.99849	.	.	.	.
1979:9	0.99720	.	.	.	.
1979:10	1.00576	.	.	.	.
1979:11	1.00361	.	.	.	.
1979:12	0.99825	.	.	.	.
1980:1	0.99896	.	.	.	.
1980:2	1.00600	.	.	.	.
1980:3	1.00094	.	.	.	.
1980:4	1.00610	.	.	.	.
1980:5	0.99647	.	.	.	.
1980:6	1.00883	.	.	.	.
1980:7	1.00439	.	.	.	.
1980:8	1.00589	.	.	.	.
1980:9	1.00533	.	.	.	.
1980:10	1.00619	.	.	.	.
1980:11	1.00173	.	.	.	.
1980:12	1.00777	.	.	.	.
1981:1	1.00706	.	.	.	.
1981:2	1.00285	.	.	.	.
1981:3	1.00913	.	.	.	.
1981:4	1.00530	.	.	.	.
1981:5	1.00561	.	.	.	.
1981:6	1.00229	.	.	.	.
1981:7	1.00510	.	.	.	.
1981:8	1.00287	.	.	.	.
1981:9	0.99846	.	.	.	.
1981:10	1.00225	.	.	.	.
1981:11	1.00386	.	.	.	.
1981:12	0.99797	.	.	.	.
1982:1	1.00040	.	.	.	.
1982:2	0.99993	.	.	.	.
1982:3	1.00236	.	.	.	.
1982:4	1.00131	.	.	.	.

1982:5	1.00280	.	.	.	.
1982:6	1.00261	.	.	.	.
1982:7	1.00194	.	.	.	.
1982:8	0.99881	.	.	.	.
1982:9	1.00034	.	.	.	.
1982:10	0.99723	.	.	.	.
1982:11	0.99901	.	.	.	.
1982:12	0.99978	.	.	.	.
1983:1	0.99904	.	.	.	.
1983:2	1.00063	.	.	.	.
1983:3	1.00634	.	.	.	.
1983:4	1.00157	.	.	.	.
1983:5	0.99956	.	.	.	.
1983:6	1.00058	.	.	.	.
1983:7	0.99912	.	.	.	.
1983:8	0.99767	.	.	.	.
1983:9	1.00118	.	.	.	.
1983:10	0.99790	.	.	.	.
1983:11	0.99960	.	.	.	.
1983:12	1.00231	.	.	.	.
1984:1	1.00314	.	.	.	.
1984:2	1.00092	.	.	.	.
1984:3	1.00136	.	.	.	.
1984:4	1.00122	.	.	.	.
1984:5	0.99403	.	.	.	.
1984:6	1.00467	.	.	.	.
1984:7	0.99765	.	.	.	.
1984:8	0.99759	.	.	.	.
1984:9	0.99416	.	.	.	.
1984:10	0.99468	.	.	.	.
1984:11	0.99835	.	.	.	.
1984:12	1.00234	.	.	.	.
1985:1	0.99910	.	.	.	.
1985:2	0.99901	.	.	.	.
1985:3	1.00337	.	.	.	.
1985:4	1.00481	.	.	.	.
1985:5	1.00483	.	.	.	.
1985:6	1.00321	.	.	.	.
1985:7	1.00173	.	.	.	.
1985:8	1.00079	.	.	.	.
1985:9	1.00166	.	.	.	.
1985:10	1.00186	.	.	.	.
1985:11	1.00358	.	.	.	.
1985:12	1.00214	.	.	.	.
1986:1	**1.01463**	.	.	.	.
1986:2	**1.01774**	.	.	.	.
1986:3	**1.01667**	.	.	.	.
1986:4	**1.01567**	.	.	.	.
1986:5	**1.01806**	.	.	.	.
1986:6	**1.01524**	.	.	.	.
1986:7	**1.01959**	.	.	.	.
1986:8	**1.01499**	.	.	.	.
1986:9	**1.01714**	.	.	.	.
1986:10	**1.01600**	.	.	.	.
1986:11	**1.01662**	.	.	.	.
1986:12	**1.01876**	.	.	.	.
1987:1	1.00732	.	.	.	.
1987:2	1.00253	.	.	.	.
1987:3	1.00569	.	.	.	.
1987:4	1.00567	.	.	.	.
1987:5	1.00470	.	.	.	.
1987:6	1.00524	.	.	.	.
1987:7	1.00357	.	.	.	.
1987:8	1.00778	.	.	.	.
1987:9	1.00430	.	.	.	.
1987:10	1.00214	.	.	.	.
1987:11	1.00286	1.00000	.	.	.
1987:12	.	0.99988	.	.	.

1988:1	.	1.00053	.	.	.
1988:2	.	1.00097	.	.	.
1988:3	.	1.00177	.	.	.
1988:4	.	0.99880	.	.	.
1988:5	.	1.00044	.	.	.
1988:6	.	1.00083	.	.	.
1988:7	.	0.99968	.	.	.
1988:8	.	0.99740	.	.	.
1988:9	.	0.99702	.	.	.
1988:10	.	1.00350	.	.	.
1988:11	.	0.99867	.	.	.
1988:12	.	1.00078	.	.	.
1989:1	.	0.99813	.	.	.
1989:2	.	0.99912	.	.	.
1989:3	.	0.99886	.	.	.
1989:4	.	0.99893	.	.	.
1989:5	.	0.99966	.	.	.
1989:6	.	0.99913	.	.	.
1989:7	.	0.99884	.	.	.
1989:8	.	0.99998	.	.	.
1989:9	.	0.99888	.	.	.
1989:10	.	0.99696	.	.	.
1989:11	.	0.99966	.	.	.
1989:12	.	0.99470	.	.	.
1990:1	.	0.99810	.	.	.
1990:2	.	0.99880	.	.	.
1990:3	.	1.00069	.	.	.
1990:4	.	0.99895	.	.	.
1990:5	.	0.99782	.	.	.
1990:6	.	0.99803	.	.	.
1990:7	.	0.99736	.	.	.
1990:8	.	0.99817	.	.	.
1990:9	.	0.99598	.	.	.
1990:10	.	0.99640	.	.	.
1990:11	.	0.99928	.	.	.
1990:12	.	0.99948	.	.	.
1991:1	.	0.99562	.	.	.
1991:2	.	0.99715	.	.	.
1991:3	.	0.99453	.	.	.
1991:4	.	0.99301	.	.	.
1991:5	.	0.99369	.	.	.
1991:6	.	0.99558	.	.	.
1991:7	.	0.99733	.	.	.
1991:8	.	0.99715	1.00000	.	.
1991:9	.	.	1.00139	.	.
1991:10	.	.	0.99988	.	.
1991:11	.	.	1.00017	.	.
1991:12	.	.	1.00096	.	.
1992:1	.	.	0.99884	.	.
1992:2	.	.	0.99949	.	.
1992:3	.	.	0.99891	.	.
1992:4	.	.	0.99929	.	.
1992:5	.	.	1.00090	.	.
1992:6	.	.	1.00043	.	.
1992:7	.	.	1.00098	.	.
1992:8	.	.	1.00212	.	.
1992:9	.	.	1.00099	.	.
1992:10	.	.	1.00212	.	.
1992:11	.	.	1.00010	.	.
1992:12	.	.	0.99947	.	.
1993:1	.	.	0.99623	.	.
1993:2	.	.	0.99673	.	.
1993:3	.	.	0.99682	.	.
1993:4	.	.	0.99696	.	.
1993:5	.	.	0.99921	.	.
1993:6	.	.	0.99886	.	.
1993:7	.	.	1.00037	.	.
1993:8	.	.	1.00134	.	.

1993:9	.	.	1.00287	.	.
1993:10	.	.	0.99969	.	.
1993:11	.	.	0.99813	.	.
1993:12	.	.	0.99891	.	.
1994:1	.	.	0.99743	.	.
1994:2	.	.	0.99727	.	.
1994:3	.	.	0.99708	.	.
1994:4	.	.	0.99647	.	.
1994:5	.	.	0.99921	.	.
1994:6	.	.	0.99891	.	.
1994:7	.	.	1.00077	.	.
1994:8	.	.	1.00381	.	.
1994:9	.	.	1.00275	.	.
1994:10	.	.	1.00259	.	.
1994:11	.	.	1.00208	.	.
1994:12	.	.	1.00031	.	.
1995:1	.	.	1.00086	.	.
1995:2	.	.	0.99916	.	.
1995:3	.	.	0.99899	.	.
1995:4	.	.	0.99994	.	.
1995:5	.	.	1.00088	.	.
1995:6	.	.	1.00175	.	.
1995:7	.	.	1.00338	.	.
1995:8	.	.	1.00514	.	.
1995:9	.	.	1.00342	.	.
1995:10	.	.	1.00074	.	.
1995:11	.	.	1.00049	.	.
1995:12	.	.	1.00318	.	.
1996:1	.	.	1.00192	.	.
1996:2	.	.	1.00168	.	.
1996:3	.	.	1.00236	.	.
1996:4	.	.	1.00154	.	.
1996:5	.	.	1.00288	.	.
1996:6	.	.	1.00281	.	.
1996:7	.	.	1.00291	.	.
1996:8	.	.	1.00459	.	.
1996:9	.	.	1.00223	.	.
1996:10	.	.	1.00207	.	.
1996:11	.	.	1.00089	.	.
1996:12	.	.	1.00294	.	.
1997:1	.	.	0.99566	1.00000	.
1997:2	.	.	.	1.00175	.
1997:3	.	.	.	0.99928	.
1997:4	.	.	.	0.99967	.
1997:5	.	.	.	0.99936	.
1997:6	.	.	.	0.99978	.
1997:7	.	.	.	0.99844	.
1997:8	.	.	.	1.00031	.
1997:9	.	.	.	0.99988	.
1997:10	.	.	.	1.00022	.
1997:11	.	.	.	1.00146	.
1997:12	.	.	.	1.00380	.
1998:1	.	.	.	1.00368	.
1998:2	.	.	.	1.00328	.
1998:3	.	.	.	1.00320	.
1998:4	.	.	.	1.00297	.
1998:5	.	.	.	1.00321	.
1998:6	.	.	.	1.00249	.
1998:7	.	.	.	1.00364	.
1998:8	.	.	.	1.00413	.
1998:9	.	.	.	1.00381	.
1998:10	.	.	.	1.00380	.
1998:11	.	.	.	1.00393	.
1998:12	.	.	.	1.00637	.
1999:1	.	.	.	1.00679	.
1999:2	.	.	.	1.00652	.
1999:3	.	.	.	1.00639	.
1999:4	.	.	.	1.00578	.

1999:5	.	.	.	1.00611	.
1999:6	.	.	.	1.00806	.
1999:7	.	.	.	1.00858	.
1999:8	.	.	.	1.00745	.
1999:9	.	.	.	1.00814	.
1999:10	.	.	.	1.00829	.
1999:11	.	.	.	1.00930	.
1999:12	.	.	.	**1.01135**	.
2000:1	.	.	.	**1.01100**	.
2000:2	.	.	.	**1.01003**	.
2000:3	.	.	.	**1.01022**	.
2000:4	.	.	.	**1.01142**	.
2000:5	.	.	.	**1.01203**	.
2000:6	.	.	.	**1.01349**	.
2000:7	.	.	.	**1.01398**	.
2000:8	.	.	.	**1.01401**	.
2000:9	.	.	.	**1.01568**	.
2000:10	.	.	.	**1.01463**	.
2000:11	.	.	.	**1.01558**	.
2000:12	.	.	.	**1.01620**	.
2001:1	.	.	.	**1.01607**	.
2001:2	.	.	.	**1.01594**	.
2001:3	.	.	.	**1.01613**	.
2001:4	.	.	.	**1.01588**	.
2001:5	.	.	.	**1.01728**	.
2001:6	.	.	.	**1.01763**	.
2001:7	.	.	.	**1.01747**	.
2001:8	.	.	.	**1.01594**	.
2001:9	.	.	.	**1.01807**	.
2001:10	.	.	.	**1.01829**	.
2001:11	.	.	.	**1.01960**	.
2001:12	.	.	.	**1.01899**	.
2002:1	.	.	.	**1.02129**	**1.00000**
2002:2	.	.	.	.	1.00193
2002:3	.	.	.	.	1.00180
2002:4	.	.	.	.	1.00407
2002:5	.	.	.	.	1.00285
2002:6	.	.	.	.	1.00159
2002:7	.	.	.	.	1.00367
2002:8	.	.	.	.	1.00402
2002:9	.	.	.	.	1.00468
2002:10	.	.	.	.	1.00606
2002:11	.	.	.	.	1.00671
2002:12	.	.	.	.	1.00623
2003:1	.	.	.	.	1.00704
2003:2	.	.	.	.	1.00721
2003:3	.	.	.	.	1.00736
2003:4	.	.	.	.	1.00797
2003:5	.	.	.	.	1.00586
2003:6	.	.	.	.	1.00362