South African Reserve Bank

Working Paper Series

WP/25/03

Whose inflation expectations forecast best? Alternatives based on survey and financial data

Monique Reid and Pierre Siklos

Authorised for publication by Konstantin Makrelov

14 March 2025

SOUTH AFRICAN RESERVE BANK

South African Reserve Bank Working Papers are written by staff members of the South African Reserve Bank and, on occasion, by consultants under the auspices of the South African Reserve Bank. The papers deal with topical issues and describe preliminary research findings and develop new analytical or empirical approaches in their analyses. They are solely intended to elicit comments and stimulate debate.

South African Reserve Bank Working Papers are externally refereed.

Information on South African Reserve Bank Working Papers can be found at
https://www.resbank.co.za/en/home/publications/Papers/working-papers.

Enquiries relating to the Working Paper Series can be addressed to:
Head: Economic Research Department
South African Reserve Bank
P O Box 427
Pretoria 0001

Tel. +27 12 313 3911

Whose inflation expectations forecast best? Alternatives based on survey and financial data

Monique Reid* and Pierre Siklos†

Abstract

Lars Svensson (1997) argued that inflation targeting should be called inflation forecast targeting, capturing the fact that monetary policy is necessarily forward looking. Inflation forecast performance is therefore a critical element for good conduct in monetary policy. As more measures of inflation expectations have become available, it is worth asking whether some are better than others at forecasting inflation and whether the long-held belief that forecast averaging outperforms individual forecasts continues to hold. We consider five sources of inflation forecasts for South Africa, including three unique quarterly surveys of firms, financial analysts and trade unions. We find that a linear combination of forecasts obtained from a factor model can improve the accuracy of forecasting over alternative forms of aggregation.

JEL classification

C82, E31, E37

Keywords

Forecasting, inflation expectations, survey data

* Lead economist, South African Reserve Bank (SARB); Extraordinary Associate Professor, Stellenbosch University. Corresponding author: monique.reid@resbank.co.za.

† Professor of Economics, Wilfrid Laurier University, Canada; Balsillie School of International Affairs, Canada; Research fellow, SARB; Research fellow, Stellenbosch University.

1. Introduction[1]

In the *Handbook of forecasting* (2013), Faust and Wright, focusing on the United States (US) experience, reviewed developments in inflation forecasting and identified four key principles from their forecast comparison exercise. They argued that (1) subjective forecasts (forecasts which incorporate judgement) do best; (2) good forecasts must account for slow changes in average inflation; (3) inflation forecasts benefit from high-quality nowcasts; and (4) focusing only on the most important information (heavy shrinkage) improves inflation forecasts. In this paper, we apply principle 1 by using survey data that represents subjective data, and principle 4 by selecting a single measure of inflation expectations that captures the most important information from the available surveys to use in a forecasting model, such as that used by the South African Reserve Bank (SARB).

It is worth noting that when Faust and Wright (2013: 20) claimed that subjective forecasts were the only ones that ‘consistently significantly improve on our simple benchmark’, they were relying on subjective forecasts of financial specialists (Blue Chip, SPF) and Greenbook forecasts published by the US Federal Reserve. In fact, in reflecting on their findings, they argue that these private sector forecasters and policymakers are likely to have access to forecasting models to which they add their professional judgement. In the language of more recent literature, these are the more ‘attentive’ decision-makers in the economy who also are more informed about the topic (typically financial specialists of some kind).

In 2013, when the review article was written, most studies that used survey data to capture inflation expectations would have used the expectations of either professional forecasters (often financial analysts) or household consumers (such as the Michigan survey), which were the two groups for which these surveys were typically available. However, events since the global financial crisis (GFC) have increased the urgency to improve our understanding of the way expectations are formed and the implications for

[1] Previous versions of this paper were presented at a SARB webinar, the Bureau of Economic Research at Stellenbosch University, the University of Cape Town and the Central Bank Business Survey Conference in Rome, Italy.

modelling and the use of inflation expectations data. Partly in response to these events, the number and type of inflation forecasts has mushroomed over time.

Following the GFC, when mainstream macroeconomic models came under scrutiny, the widespread treatment of expectations as rational attracted renewed criticism. In the face of the zero lower bound, the increased use of central bank communication to manage inflation expectations also meant that survey data were keenly analysed. The missing inflation and missing disinflation (see Coibion and Gorodnichenko (2015) and others) that followed the crisis and then the high levels of inflation experienced after the pandemic, led many macroeconomists to reflect further on the behaviour of inflation expectations and how these may influence the sensitivity of inflation to changes in economic slack. For example, research focusing on the possibility that the Phillips curve is non-linear (Gagnon and Collins 2019; Forbes, Gagnon and Collins 2021) and the idea that there may be some threshold beyond which the public becomes more attentive to inflation have become highly relevant to policymakers.

Researchers acknowledge that macroeconomics needs to be based on more satisfactory (evidence-based) micro foundations. This reflection by macroeconomists includes the observation that what we really need is to understand the behaviour of the price setters (groups other than financial markets and financial analysts), which academics and policymakers alike began to link more frequently with consumer and firm-level decisions.

Over the last decade there has been a surge in macroeconomic literature that relies on microdata. This helps policymakers and researchers interpret many observations with more confidence and better understanding. The implications for modelling, however, are less obvious. New macroeconomic theories about how expectations are formed have emerged (learning, sticky information, rational inattention and behavioural macroeconomic models), but these all really fit into more theory-based macroeconomic models, such as dynamic stochastic general equilibrium (DSGE) models, which are trying to adapt to changing views about what drives expectations and economic fluctuations (e.g. see Coletti (2023)). They offer limited guidance to modellers that require a single measure of expected inflation to use in forecast models such the Quarterly Projection Model of the SARB.[2]

One simple question that modellers who use the more parsimonious models are likely to have is which single measure of inflation expectations is best for particular applications. The horizon over which models are viewed as producing useful forecasts will also matter. Generally, central banks will generate forecasts for two to three years ahead to observe the full impact of policy changes on inflation outcomes given lags in the transmission mechanism. However, policymakers will also be interested in longer-term forecasts because these may be more informative about the credibility of a central bank, as measured by the degree to which expectations are anchored.

In this paper, we will focus on which choice of inflation expectations data in South Africa performs best for forecasting inflation. The purpose is to decide which measure of inflation expectations (or which aggregation of these data) would perform best when used in a model such as the SARB's Quarterly Projection Model. The SARB currently uses the Bureau for Economic Research (BER) aggregate survey measure of inflation expectations. The BER surveys three groups – financial analysts, the business sector (firms) and trade unions – and aggregates them by taking the simple average of their forecasts to form a single number reflecting average inflation expectations of all the groups. We stress the uniqueness of this survey as it has been carried out since the start of the SARB's inflation-targeting policy regime in 2000. Few other firm-level surveys of its kind and time span exist elsewhere.[3] This raises one of the challenges faced by macroeconomic models, namely the need to decide how to aggregate the underlying data, notably inflation expectations, included in the model. Leading theorists are exploring ways to capture heterogeneity within macroeconomic models, but the degree of heterogeneity that can be dealt with is not yet very high and the models are

---

[2] For more information about the SARB's Quarterly Projection Model, see Botha et al. (2017) and Pirozhkova et al. (2023).

[3] While a number of countries around the world have started to conduct surveys of firms' inflation expectations, most of them have begun relatively recently. Where there is firm-level data from earlier years, they are often snapshot surveys (once-off, rather than time series) or do not include forecasts of other macroeconomic variables, as the BER survey does. The BER survey of firm-level inflation expectations, which offers a time series at a quarterly frequency for almost 25 years, is notable, especially for an emerging market economy.

complex. These developments hold particular promise for analysing policy-relevant questions that benefit from a general equilibrium setup; however, from a forecasting perspective, simple models that rely on data aggregated in some manner still perform comparatively well. In this paper, we will evaluate a set of choices about which microdata to use and how to aggregate it for the sake of forecasting inflation.

Addressing this issue matters to policymakers and the wider public because reliable forecasts are central to the implementation of inflation targeting.[4] This was emphasised by Svensson (1997), who preferred to call the framework 'inflation forecast targeting'. Central bank forecasts support both the monetary policy decisions made by the monetary policy committee and the central bank's communication of these decisions (Bernanke 2024). Policy choices made by SARB decision-makers may be influenced by the expectations used in their forecast models, and so it is prudent for researchers to regularly evaluate and aim to improve the quality of the data used in central bank models.

The remainder of the paper is organised as follows. The next section describes the data. We then consider ways in which the BER survey's inflation expectations of the various sub-groups can be aggregated. This is followed by an empirical examination of forecast errors obtained by comparing the inflation expectations of the various aggregates to actual inflation outcomes, and then statistical analysis and formal tests of the forecast performance. The paper concludes with a summary and discussion of policy implications.

Our results suggest that a linear combination of the survey's sub-groups obtained from a factor model can improve forecast performance over available alternative forms of aggregation. We also find differences in forecast performance of the various sub-groups depending on the underlying state of the economy and the forecast horizon, which reveal that simply relying on financial professionals who are typically found to be more

---

4 This paper also refers to the heterogeneity of the inflation expectations within and across the groups surveyed. Evaluating this heterogeneity (its character or drivers) is not within the scope of this paper; this topic has been investigated in other South African literature. Heterogeneity is only relevant to this paper in so far as it has an impact on forecasting performance.

'rational' could mean overlooking important information. The financial analysts tend to perform particularly well when economic conditions are calm. Trade unions and businesses may have something more to offer at the longer horizons and when economic conditions are poor, when forecasting models themselves perform less well. The factor model is a form of aggregation that allows us to retain this information from trade unions and businesses without weighting them equally, as is the case in the simple arithmetic mean.

2. Data

Since 2000, the BER has surveyed the inflation expectations of trade unions (or labour), businesses (or firms)[5] and financial analysts on a quarterly basis.[6] The firm-, trade union- and financial analyst-level surveys aim to ensure good representation of a cross-section of the South African economy.[7] Table 1 provides an overview of the number of observations in the three surveys for the full sample considered in this study, namely 2000Q2–2023Q4.[8] It is worth noting that the sample of firms is far larger than that of the financial analysts or trade unions. This set of surveys is rich by international standards[9]

---

5 In keeping with the terminology used in the existing literature that investigates expectations of businesses, we will also refer to this group interchangeably as the 'firms' in our statistical investigation.

6 Between 2000 and 2003, the quarterly surveys were conducted in February, May, August and October. Since that time, the February and October surveys were shifted to March and November. The timing of the remaining two surveys is unchanged.

7 Since each respondent is identified only by an ID number, we are also able to establish that there are only a very small number of duplicate individuals surveyed over time. More precisely, 7.45% of trade union respondents, 6.50% of businesses, and 5.08% of financial analysts are duplicates over the complete sample. There is therefore limited scope to exploit this panel dimension. Repeated surveying of the same individuals would be valuable in the sense that we can study how their views evolve over time. However, potential drawbacks include the fact that these individuals might learn through participation in the survey (especially if historical inflation figures are given to the respondents as part of the survey question) and they might become more attentive to inflation because it is regularly brought to their attention through participation in the survey, which would not be true for the rest of the population.

8 Information about the distribution of observations by year and quarter is reported in Reid and Siklos (2022).

9 Beyond the usual survey questions asking for (headline) inflation and economic growth (percentage change in real GDP) forecasts, the survey is notable in at least two respects. First, it also asks for forecasts for a wide range of key macrofinancial variables. Hence, the survey sent to trade unions and firms requests forecasts for the prime interest rate (interest rate charged by commercial banks for loans to their best customers), wage and salary growth, and the rand/US dollar exchange rate. Surveys of financial analysts add questions that elicit expectations about

and, unlike virtually all firm-level forecasts we are aware of (e.g. see Coibion et al. (2020) and Reid and Siklos (2022)), the BER data set generates a quarterly time series that is more than 20 years long.

The precise wording for the inflation question is: "What do you expect the average headline inflation rate to be during the year?" [10] For the longer-term inflation expectations question, respondents are asked: "What do you expect the average CPI inflation rate to be over the next five years?" Respondents are then asked to fill in boxes for the current calendar year and the next two. The phrasing of the question for the other series surveyed is comparable. There is some 'priming' because respondents are provided with average inflation rates (actual inflation outturns) for the previous calendar year as well the mean inflation rate for the last five years.[11] The horizon for the forecasts is the calendar year, like most fixed-event surveys. Respondents are anonymous to researchers.

Table 1 reveals that considerable diversity exists in the average inflation expectations across groups surveyed, which the present study seeks to exploit to better understand what the public and markets think about the future evolution of inflation. A few salient features emerge. There is less disagreement among the financial analysts than the other two groups surveyed, as captured by the standard deviations of their inflation forecasts across the various forecast horizons, but none of the categories shown contains any notably large deviations in views about expected inflation rates.

---

growth rates in the M3 money stock (broad money measure), the yield on long-term government bonds and capacity utilisation in the manufacturing sector (i.e. percentage utilisation of production capacity). Second, in addition to current year forecasts, one-year-ahead forecasts are recorded for all the variables except inflation, which includes the two- and five-year-ahead (since 2011Q2) horizons.

10 As measured by the annualised percentage change in the consumer price index (CPI). Between 2000 and 2008, both the CPI and the CPI excluding mortgages (CPIX) were surveyed. Thereafter, only the CPI data have been collected. CPIX includes the cost of shelter but not the investment portion of housing investment. Instead, a measure of the imputed rent is included. In the empirical portion of this study, only CPI data are considered throughout.

11 The impact of priming remains an unresolved issue. At the very least, priming is thought to alter the distribution of inflation forecast responses and may be a function of the level of inflation. See, for example, Reid, Siklos and Du Plessis (2021), Niu and Harvey (2023) and references therein.

Finally, aggregating data at the level of each one of the surveys as a whole, we can also generate trimmed measures of inflation. While central banks have long resorted to evaluating core measures of inflation to exclude the most volatile prices in the CPI (i.e. food and energy products), there has more recently been a shift towards measures of inflation that instead exclude the tails of the distribution of prices. In this manner, some biases inherent in traditional core measures of inflation can be overcome somewhat by leaving out the most volatile prices regardless of their origin. Table 1 provides some trimmed estimates of inflation and these reveal a modest decline in long-term inflation expectations relative to the short horizon forecasts. This may indicate the credibility of the inflation-targeting regime.[12] The trimmed forecasts also reveal that removing the tails of inflation forecasts has a small impact on average inflation expectations.[13]

Table 1: Summary statistics: inflation expectations under different levels of aggregation
(a) By firm size: all surveys

Firm size (no. of employees)	**Horizon**	**Business**		**Financial analysts**		**Trade union**	
		Inflation expectations – mean (S.D.)	# OBS	Inflation expectations – mean (S.D.)	# OBS	Inflation expectations – mean (S.D.)	# OBS
Micro (< 21)		Max no. obs.	8 465	Max no. obs.	243	Max no. obs.	691
	T0	6.44 (2.14)	8 114	5.55 (1.76)	232	6.26 (2.04)	663
	T1	6.58 (2.12)	7 026	5.50 (1.32)	230	6.35 (2.03)	607
	T2	6.67 (2.29)	6 768	5.28 (1.17)	215	6.39 (2.09)	593
	5Y	6.15 (1.54)	2 918	4.98 (0.51)	123	5.65 (1.27)	267
Small (≥21 & <50)		Max no. obs.	6 007	Max no. obs.	138	Max no. obs.	109
	T0	6.45 (1.99)	5 752	5.54 (1.67)	135	6.76 (2.21)	107
	T1	6.55 (1.93)	5 234	5.48 (1.10)	134	6.54 (1.96)	102
	T2	6.57 (2.05)	5 041	4.94 (1.09)	131	6.34 (1.92)	98
	5Y	6.26 (1.36)	2 102	4.42 (0.79)	64	5.85 (0.65)	11
Medium (≥51 & ≤200)		Max no. obs.	8 798	Max no. obs.	348	Max no. obs.	295
	T0	6.41 (2.00)	8 506	5.89 (2.15)	335	5.62 (1.76)	281
	T1	6.47 (1.87)	7 846	5.41 (1.20)	335	5.62 (1.64)	261
	T2	6.46 (1.93)	7 698	5.13 (0.86)	303	5.67 (1.75)	254
	5Y	6.13 (1.40)	2 569	5.33 (0.51)	111	5.50 (1.30)	151
Large (>200)		Max no. obs.	6 919	Max no. obs.	912	Max no. obs.	282
	T0	6.35 (2.00)	6 733	5.67 (1.78)	891	5.95 (1.84)	271
	T1	6.31 (1.74)	6 328	5.36 (0.98)	891	5.81 (1.45)	237
	T2	6.23 (1.72)	6 250	5.22 (0.72)	808	5.76 (1.39)	232
	5Y	6.02 (1.10)	1 977	5.24 (0.59)	442	5.31 (1.25)	125

12 This interpretation should be read with some caution as the five-year-ahead inflation forecasts are substantively different from the others. More on this below.

13 We also conducted more formal tests for the presence of outliers in inflation forecasts (results not shown). When they are found, observed inflation is high, as during the first few years of inflation targeting and the recent post-pandemic surge in inflation, which has been a global phenomenon.

(b) By trimming

Variable	Trimmed CPI inflation forecast mean (S.D.)			Not trimmed inflation forecast mean (S.D.)		
Survey source	Business	Fin. analysts	Trade union	Business	Fin. analysts	Trade union
T0	6.18 (1.53)	5.60 (1.76)	5.96 (1.56)	6.19 (1.53)	5.96 (1.56)	5.96 (1.56)
T1	6.20 (1.21)	5.37 (0.82)	6.02 (1.44)	6.29 (1.25)	5.37 (0.82)	6.05 (1.42)
T2	6.18 (1.06)	5.19 (0.47)	6.04 (1.33)	6.31 (1.09)	5.18 (0.48)	6.08 (1.32)
5Y	5.88 (0.55)	5.18 (0.48)	5.49 (0.71)	5.95 (0.55)	5.15 (0.44)	5.54 (0.72)
Observed inflation						5.32 (2.56)

Notes: Sample is 2000Q2–2023Q4, except for 5Y data which begins 2011Q3. All figures are in %. 'Max no. obs.' refers to the number of observations if all respondents provide a forecast. All values rounded to two decimal places. Trimming cuts off the top and bottom 20% of the inflation or observed inflation distributions. T0 are current calendar year forecasts, T1 are one-year-ahead calendar year forecasts, T2 are two-year-ahead calendar year forecasts and 5Y are five-year-ahead average inflation forecasts.

We can also learn from the information about the distribution of the expectations of the different groups and how the distribution of forecasts evolves over time. Greater dispersion is likely to reduce forecast accuracy, and it also has implications for the best way to evaluate forecast accuracy (we will argue below that the mean scaled forecast error statistics may be more appropriate than the commonly used root mean squared errors metric). As an illustration of how the expectations differ across groups as well as time, Figure 1 shows the distribution of one-year-ahead inflation expectations for select years from the business and financial analysts surveys.[14] The years shown consider three periods: the first years of inflation targeting in South Africa (2002 and 2004), the peak and end of the GFC (2008 and 2010), and the COVID-19 pandemic and subsequent surge in inflation (2020 and 2022). To conserve space and simplify calculations, the data for entire calendar years are used.

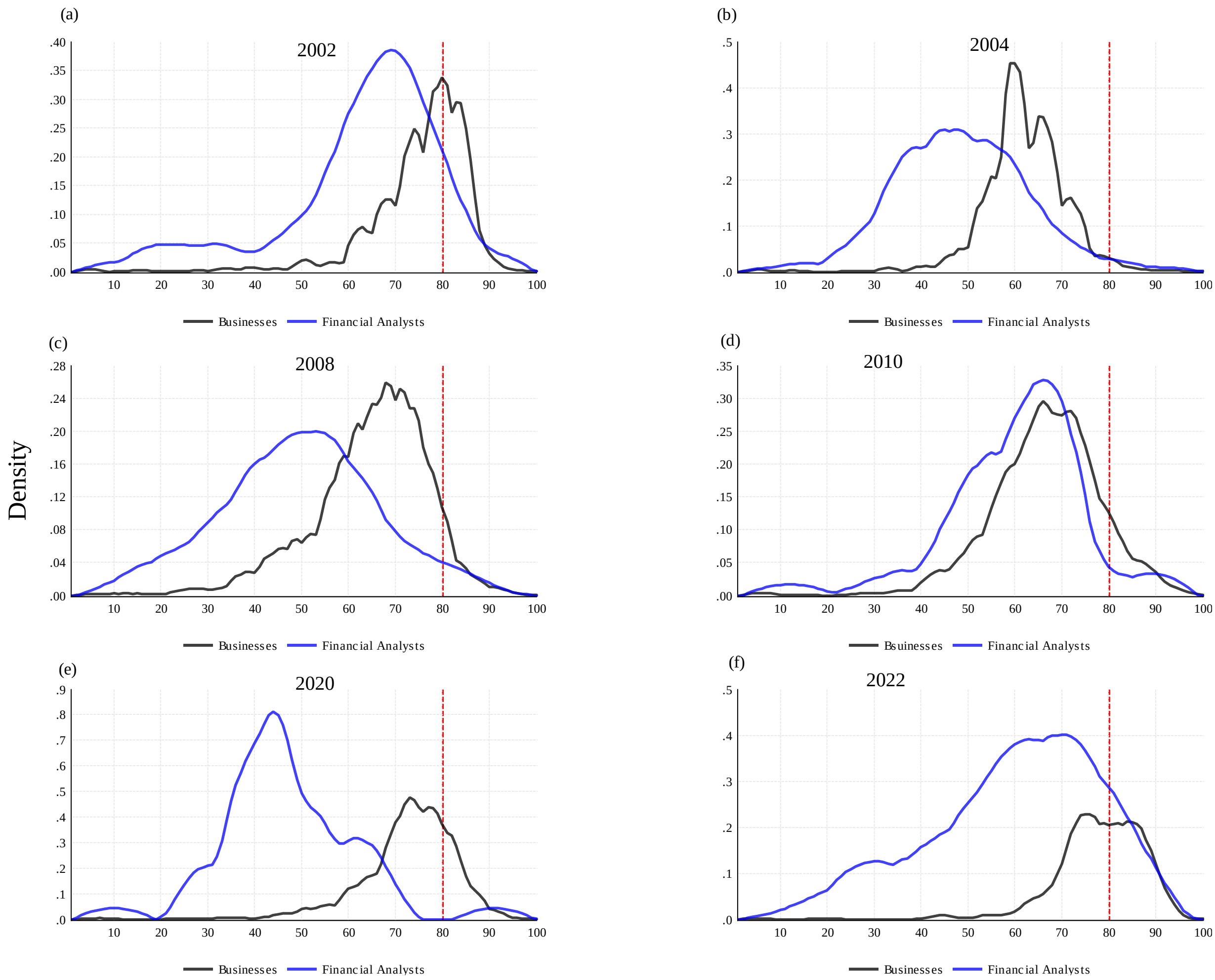
Figure 1 shows how much the distribution can differ across respondent type and across time. The vertical axis plots the frequency of responses for different levels of inflation expectations while the horizontal axis plots the range of responses, from lowest to highest, normalised to range from 0 to 100.[15] To provide a benchmark, the vertical dashed line in all the figures uses the peak value in the 2002 business sector survey of inflation forecasts one year ahead. As can be seen from the figure, the frequency of

[14] The graph is the kernel distribution, which essentially replaces the bars or boxes in a histogram by a smooth line. The horizontal axis is the proportion of the sample.

[15] The range remains consistent across the panels of Figure 1 (i.e. across time).

responses peaks at around 35%, which is at the mean of the distribution of inflation expectations responses. In 2002, average one-year-ahead inflation forecasts stood at 7.61% (S.D. of 0.53%), which is at the high end of the South African inflation experience over the sample covered by this study. If the frequency of responses for the same level of inflation expectations stays the same, then the peak of the other distributions shown should not change. In contrast, if the share of responses shifts, suggesting lower inflation expectations, then the benchmark vertical line drawn should be to the right of the distribution shown. If the bulk of inflation expectations views rises relative to the 2002 experience, then the benchmark vertical line shown ought to be to the left of the peak of the distribution shown.

Figure 1: The distribution of inflation across time and sources: some examples



Note: The data are based on annual averages of observations for the calendar years shown. The plots are the kernel densities (parabolic kernel). The vertical dashed line uses the peak in the distribution of the business sector survey in 2002 as a benchmark (Figure (a)). It is where the one year (T1) is at its peak (see Figure (a)). For example, in the case of the financial analysts' survey (Figure (b)), the peak in the distribution (i.e. where inflation expectations response values are most frequent) is lower in that group than the one for the business sector. Stated differently, if we were to superimpose the distribution graph for the business sector, the peak in the distribution from this group surveyed would be to the right of the financial analysts' distribution shown above. Comparing 2020 to 2022, the peak of the distribution has shifted to the right, indicating rising average inflation expectations.

Figure 1 shows that the distribution of one-year-ahead inflation expectations among firms surveyed shifted downward considerably after 2002, with the sharpest fall taking place in 2004. However, this is reversed in 2020, and by 2022 the distribution looks similar to the one in 2002, although there is more disagreement in 2020 as the area around the peak is considerably 'fatter' than in 2002. There is, however, considerable contrast between the firms surveyed and financial analysts, as can be seen from Figure 1 where the distribution of the inflation expectations of financial analysts are superimposed. Although the peak in the distributions shown are always to the left of the vertical dashed line, suggesting as before that financial analysts' inflation expectations are relatively lower than the mean of the firms in 2002, the gaps appear to be larger over the subsequent years shown. However, what is more striking is that, since 2010, the distributions reveal growing disagreement between financial analysts even if the differences seem quantitatively smaller than for the other groups in the BER survey (see Table 1). The foregoing results provide a further indication that useful information can be lost depending on the form of aggregation of inflation expectations.

2.1 Bloomberg asset price-based inflation expectations data

The recent proliferation of surveys of different groups in different countries attests to the benefits of survey data. That said, it is occasionally argued by analysts (e.g. Sandbu (2024)) that inflation expectations derived from bond market participants are superior because traders have 'skin in the game'. However, markets have been prone to miss turning points in inflation and, since these asset prices also contain largely unobserved risk premia, the resulting inflation expectations may be biased (e.g. Bianco and Haubrich (2010); Faust and Wright (2013); Gagnon and Sarsenbayev (2021)). Nonetheless, it is valuable to consider the performance of this data in this study too.

3. Aggregation choices

While there exists no formal guidance about how to aggregate micro-level inflation expectations data, it is natural to start with simple averages of the different groups surveyed, or some combination of these. Other forms of aggregation are also conceivable. As seen from Table 1, the BER data can be readily broken down according to various sub-categories of respondents too. For example, firm responses can be analysed by size of firm surveyed or by the position of the respondents in each group (i.e. from CEO to economist, to give two examples), and the sector or industry represented by respondents. However, to motivate the main results of the paper and in the interest of succinctness[16] we will only consider the groups presented below.

We evaluate four types of aggregation:

(1) The BER aggregate (aggregation across groups)

Each quarter, the BER releases the aggregate inflation expectations for each of the three social groups it surveys – financial analysts (FIN), businesses[17] (BUS) and trade unions (LAB). They also release an aggregate measure, which is the arithmetic average of these three group averages. In the rest of this paper, we will refer to this as the 'BER aggregate' measure.

This is the measure published by the BER and the measure most typically used (perhaps even exclusively used) by forecasters or modellers in South Africa when they need a measure of inflation expectations. It is also the measure used by the SARB modellers, as noted above. We must, however, recognise that even this form of aggregation involves choices. We are giving equal weights to each of the three groups surveyed without considering the relative size of the contributions of each of the groups. Without any rigorous reason to weight the groups differently, this simple approach is convenient and is followed elsewhere.[18] However, there is considerable heterogeneity in this data, which can bias the aggregated measure, so there is reason to rethink this starting point.

---

16 Results for various sub-groups were tested and are available upon request.

17 Henceforth, in line with the language used in the relevant literature, we shall refer to the respondents in the business survey as 'firms'.

18 For example, the much publicised mean forecasts published by Consensus Economics are also found by taking an arithmetic average, although all members of that survey are considered financial analysts. There is a long tradition, since at least Granger and Ramanathan (1984), of averaging forecasts, often relying on some weighted average. Also, see Timmermann (2006).

(2) Aggregation at the group level

The first alternative we propose is to look at the three groups surveyed individually and compare these (each of which represents an aggregate of the microdata of that particular group).[19] Research has repeatedly shown that financial analysts are relatively more rational (Ehlers and Steinbach 2007; Crowther-Ehlers 2019; Reid and Siklos 2023), and therefore some people argue that using this group is most likely to offer a good forecast. But in practice the forecasters are typically not using this group on its own.

More recently, a lot of the international literature has focused on household- and firm-level expectations as they are theoretically important price setters. Surveys of these groups have generally been far less common and, when conducted, they were often snapshots rather than regular time series. While the BER does also offer a survey of households, this part of the survey is conducted by a market research firm rather than by the BER. Moreover, the other three groups are not easily compared to the survey of households, so we do not include it in this paper.[20] We are, however, able to analyse the inflation expectations of firms as we are privileged in South Africa to have time series data since 2000.

Unlike the other groups surveyed by the BER, firms are the only institutional form that must simultaneously consider both input and output prices in reaching economic decisions that ensure their survival and profitability. Indeed, a variety of practices, including deciding on margins in setting prices, will influence prices they can expect to set in the marketplace. Accordingly, firms are well suited to provide insights about future inflationary developments. That said, unlike financial analysts or financial markets that understand inflation better and are likely to consider the official CPI numbers with relative accuracy when asked to report their inflation expectations, many firms of different sizes or in different sectors of the economy are more likely to think about their

---

[19] Although group-level forecasts can be further disaggregated (e.g. by firm size, industry and occupation of the respondent), space limitations prevent a more exhaustive, micro-level analysis.

[20] For further details about the household data, see Reid, Siklos and Du Plessis (2021).

own experience of inflation (subjective expectations)[21] rather than the aggregate CPI numbers.

Figure 2 plots the time series of CPI expectations for the four horizons considered by the BER surveys, namely the current year (T0), the year ahead (T1), two years ahead (T2) and the average over five years ahead (5Y).[22] The first three are calendar or fixed-event forecasts while the long-term inflation expectations generated by the surveys (only available since 2011) are akin to fixed-horizon forecasts, so we would expect 5Y to be relatively smoother. The figures also highlight the SARB's inflation target (IT) range of 3% to 6%, as well as plotting observed CPI inflation.[23] To provide some additional context, we shift observed inflation by one or two years depending on the forecast horizon to better gauge the potential differences between actual and expected inflation.

It is worth noting that in 2009 the SARB changed the measure of inflation that it officially targeted from CPIX (CPI excluding mortgages) to CPI. Where CPI strongly breaches the lower end of the target range in 2003/2004, CPIX (the measure officially targeted at the time) remained within the target band. However, in this analysis we consistently use CPI and survey respondents' forecasts of CPI, which is possible because the BER surveyed respondents about their expectations of both CPI and CPIX until CPIX was abandoned in 2009. Thereafter, they were only asked to forecast CPI.[24]

---

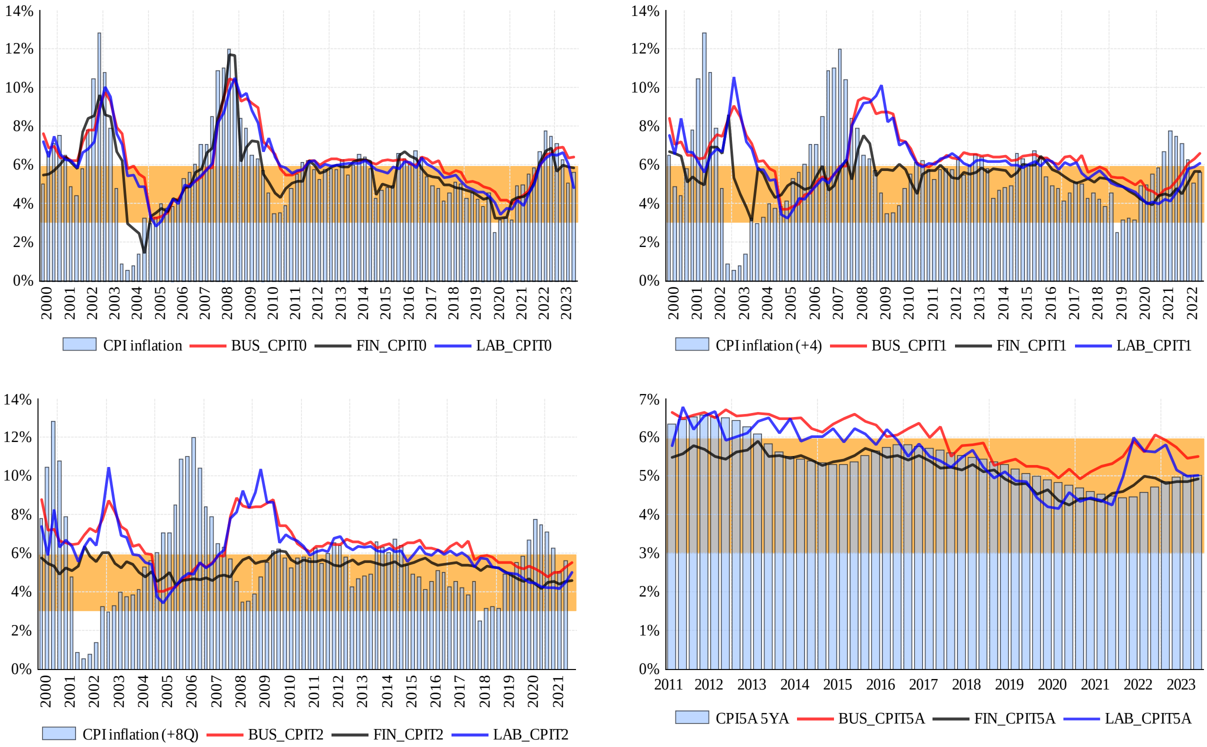
21 Weber et al. (2022).

22 Not shown here, but see below, are the SARB's forecasts. As noted above, forecasts that enter the SARB's model are the arithmetic mean of the three BER groups surveyed.

23 In Reid and Siklos (2024), we transformed the fixed-event forecasts into fixed-horizon forecasts by relying on a widely used formula (e.g. see Siklos (2013)). The impact on the data is very modest. The annexures provide some graphs that support this contention.

24 An interesting alternative (left for future research) would be to create a 'targeted inflation' variable which consists of CPIX before 2009 and CPI thereafter. This would then be compared with the matching expectations values. There is a chance that the groups that are not professional forecasters (labour and firms) might be less informed about the CPI before 2009 when it was not the official target and therefore received less attention, but the impact of this is likely to be reduced by the fact that respondents were asked to forecast both CPI and CPIX before 2009, implicitly bringing this distinction to their attention.

Figure 2: Inflation forecasts compared



Note: T0, T1, T2 and T5a are the forecast horizons (h). Also, see Table 1. CPI inflation is the observed inflation rate. The numbers in parenthesis indicate the number of quarters ahead actual inflation is shifted to match the one-year (+4) and two-year (+8) horizons of the inflation forecasts. CPI forecasts are for the mean of business, financial analysts and trade union surveys, respectively. CPI5A 5YA is the five years moving average of observed inflation while CPIT5a is the forecast for inflation five years ahead. The shaded area is the SARB's inflation target (3–6%). Since the inflation target range is defined on a calendar year basis, we did not add shading to the five-year-horizon case. Bloomberg are five-year-ahead inflation forecasts implied by market interest rate futures. BUS, FIN and LAB are business, financial analysts and trade union forecast surveys. Bloomberg_5YR are the five-year-ahead inflation expectations derived from the asset-backed securities five years return (SAGGBE05 Index). To ensure comparability, the left-hand side scales are the same for all four figures.

The data reveal a few notable features. All groups surveyed tend to over-estimate actual CPI inflation, especially after the first three years of the IT regime. Both observed and expected inflation rates up to the two-year horizon are noticeably less volatile since the 2008–09 GFC, although there is a resurgence once the COVID-19 pandemic erupts with early signs of a renewed decline in inflation volatility beginning in 2022. Generally, financial analysts expect lower inflation than the other two groups. The inflation expectations of firms and labour are more alike than those of financial analysts and trade places from time to time in expecting highest rates of future inflation. It is also notable that one- and two-year-ahead expectations react with a lag but in the same direction as a previous surge in inflation. This phenomenon is especially noticeable in the early years of the IT regime and in response to the GFC. One can also clearly see the emergence of another rise in inflation expectations as the fallout from the COVID-

19 crisis begins to raise inflation from 2021. Long-run inflation expectations show a consistent decline since these were surveyed in 2011, which offers a first indication of the credibility of the IT regime. However, the sharp reversals around the time of COVID and its aftermath are striking.

(3) Asset price data (Bloomberg)

Asset price data, which capture the implicit expectations of financial markets, are also often used (e.g. see Stock and Watson (2003) and Baumeister (2021)). The main reasons for this are that these data are easily available, have a high frequency and are relatively cheap to collect. Another benefit is that the data constitute a measure of market-wide behaviour rather than opinions collected in surveys (Armantier et al. 2015).[25] Theoretically, financial asset participants should be forward-looking in response to the profit motive they face (they have 'skin in the game') and therefore the evolution of asset prices over time can hold clues about the future course of the economy more generally and inflation in particular.

We use the break-even inflation rates available from Bloomberg as our measure of the inflation expectations of the financial markets. Break-even rates are the difference between nominal and real (inflation-indexed) bond yields (inflation compensation). However, in line with the Fisher equation, this difference between nominal and real rates includes both expected inflation as well as the risk premium associated with holding the underlying bond. As long as these risk premia are small and stationary (i.e. zero mean and constant variance) they should not distort our interpretation of what financial markets think will happen to inflation in future, but more precise measures of market-based inflation expectations would be preferable if they were available.

It is also important to note that this financial market group is distinct from the financial analysts surveyed by the BER. In Figure 2, we can see that the implied inflation

---

[25] Additionally, expectations derived from financial markets can, in principle, be derived for a much larger set of future time horizons than the commonly examined ones in the literature (e.g. current year to a few years ahead). Indeed, a term structure of inflation expectations can be derived from such data. An example is the Federal Reserve Bank of Cleveland's estimate of US expected inflation over a 30-year horizon. See https://www.clevelandfed.org/indicators-and-data/inflation-expectations.

expectations from Bloomberg data are far more volatile than the expectations of the financial analysts surveyed by the BER.

(4) Factor models

The three social groups surveyed by the BER are quite distinct and there are reasonable arguments that can be made that each of them may influence the actual inflation outcome in some way.[26] Central banks, including the SARB, typically rely on a single proxy for inflation expectations as an input into their models used for policy analysis and forecasting. In the South African case, the BER has always combined the inflation expectations of the three groups into a simple arithmetic mean, but placing equal weight on the views of the three groups may not be the most efficient approach. As shown in Figure 2, there are various views (and disagreement) on the future course of inflation. There is no reason, a priori, to believe that a simple arithmetic average of existing indicators is representative of the public's and markets' views about inflation expectations.

We therefore offer one alternative way of aggregating the information from these three survey groups and investigate whether it delivers better forecasts of inflation. This final form of aggregation involves a linear combination of forecasts that relies on factor models to generate a single forecast from all available forecasts (e.g. see Stock and Watson (2003) and Baumeister (2021)). Factor and dynamic factor models have been used for some time to provide improved forecasts of inflation (e.g. see Gosselin and Tkacz (2001); Hall, Tavlas and Wang (2023); and Aysun and Wright (2024)). As shown in Figure 3, we apply a data-driven method to obtain an aggregate measure of inflation expectations relying on the factor model methodology to obtain a single indicator. The factor model is written:

$$\tilde{\pi}^e_{t,h} = \pi^e_{t,h} - \bar{\pi}^e_{t,h} = \sum_{i=1}^{m} \lambda_{i,h} f_{i,h} + \varepsilon_{i,h} \qquad (1)$$

[26] The financial analyst and trade union components of the BER surveys are both small sample sizes and intuitively these groups are relatively homogenous. In contrast, firm-level data consist of a far larger, more diverse sample of respondents.

where $\tilde{\pi}^e_{t,h}$ are deviations of inflation expectations from their mean value, that is, $\bar{\pi}^e_{t,h}$ over the forecast horizon (*h*) in question (i.e. T0, T1, T2).[27] There are *m* series used, each of which contributes to the estimate of aggregate inflation expectation with an effective weight given by $\lambda_{i,h}$. These are referred to as the factor loadings. Hence, a factor model represents a best fit consisting of the linear combination of the available inflation forecasts at different horizons. The three separate survey forecasts (BUS, FIN, LAB) make up one factor model. In the case of long-term inflation expectations (i.e. *h*=5Y), we also consider a second model by adding financial market forecasts (Bloomberg) to the survey forecasts. Hence, *m*=3 in one model and 4 in the model for long-term (i.e. 5Y) forecasts. The factors are estimated using Bai and Ng's (2002) ICp2 criterion recommended by Stock and Watson (2016: 436), except for the 5Y case where the Ahn and Horenstein (2013) method was used. In all cases, one factor was found to be adequate. Demeaning was done according to equation 1.[28]

We restrict attention to the case where the factor loadings remain constant throughout the sample. We also experimented with time-varying loadings, but the marginal benefit was judged to be low (results not shown) and so we retain the full sample version of the factor loadings.

There are three takeaways from Figure 3, which compares the factor model-derived forecasts with the BER aggregate forecasts. First, the factor model-derived forecasts appear to track observed inflation expectations more closely than the forecasts from the individual surveys shown in Figure 2. Second, the most visible gaps between observed and expected inflation are for the long-term horizon. It is worth noting that these are

---

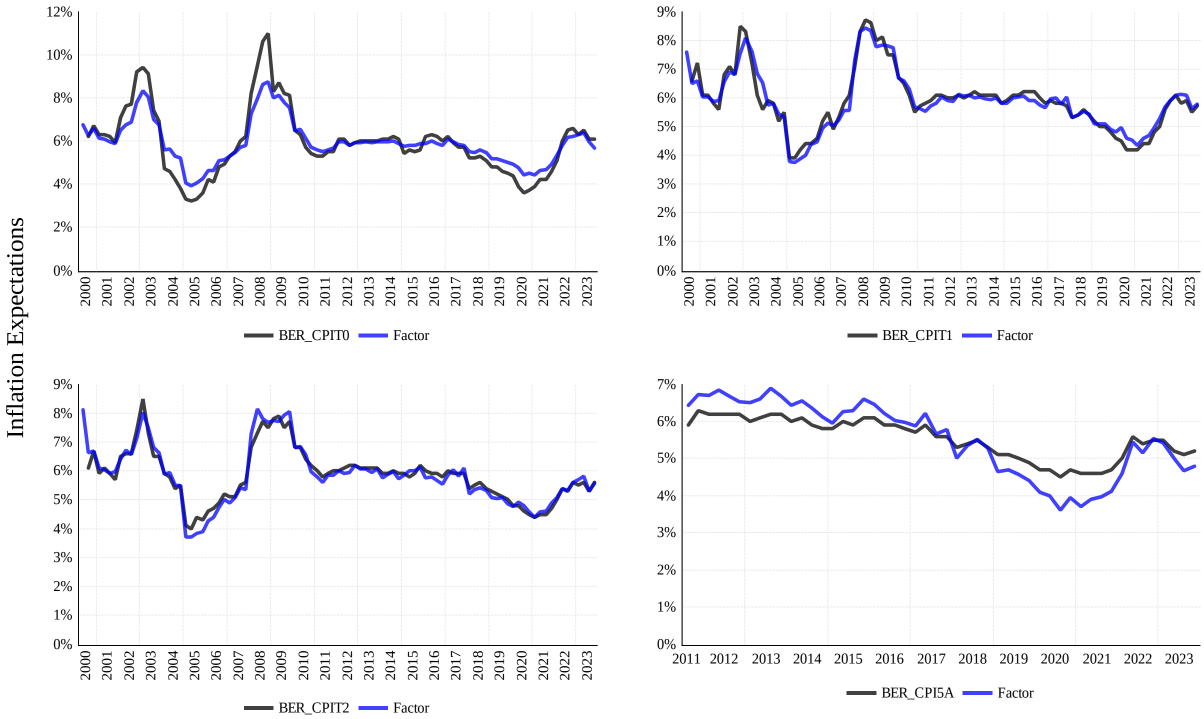
27 Factor models are typically estimated as deviations from their mean. When the variables differ from each other, unlike the present situation, series are often standardised. For a recent application, see Hamilton and Xi (2024).

28 The US Federal Reserve uses a similar methodology to estimate its index of common inflation expectations. However, since their index includes as many as 21 separate inflation forecasts that differ considerably from each other, they resort instead to a dynamic factor model. In our case we are able to generate four distinct indicators of common inflation expectations. For the US Federal Reserve's approach, see https://www.federalreserve.gov/econres/notes/feds-notes/index-of-common-inflation-expectations-20200902.html.

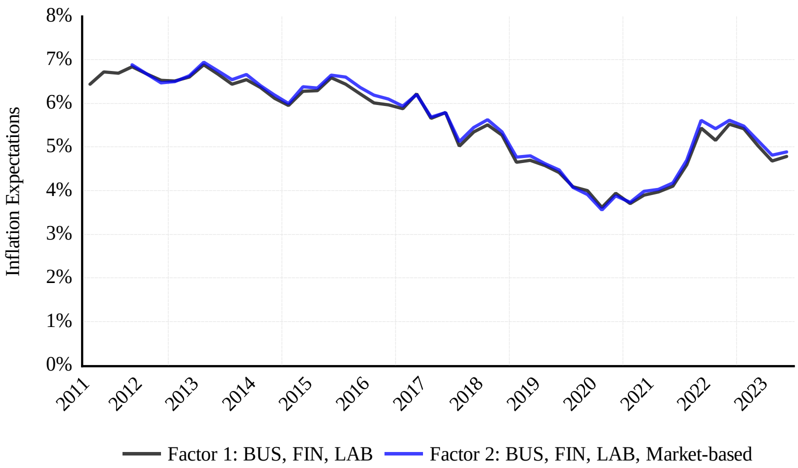
fixed-horizon forecasts covering a five-year period, as opposed to the forecasts provided for the other forecast horizons.

Figure 3: Factor model-derived forecasts versus the BER aggregate forecasts

(a) Comparisons across forecast horizons



(b) Adding market-based forecasts to factor model-derived forecasts

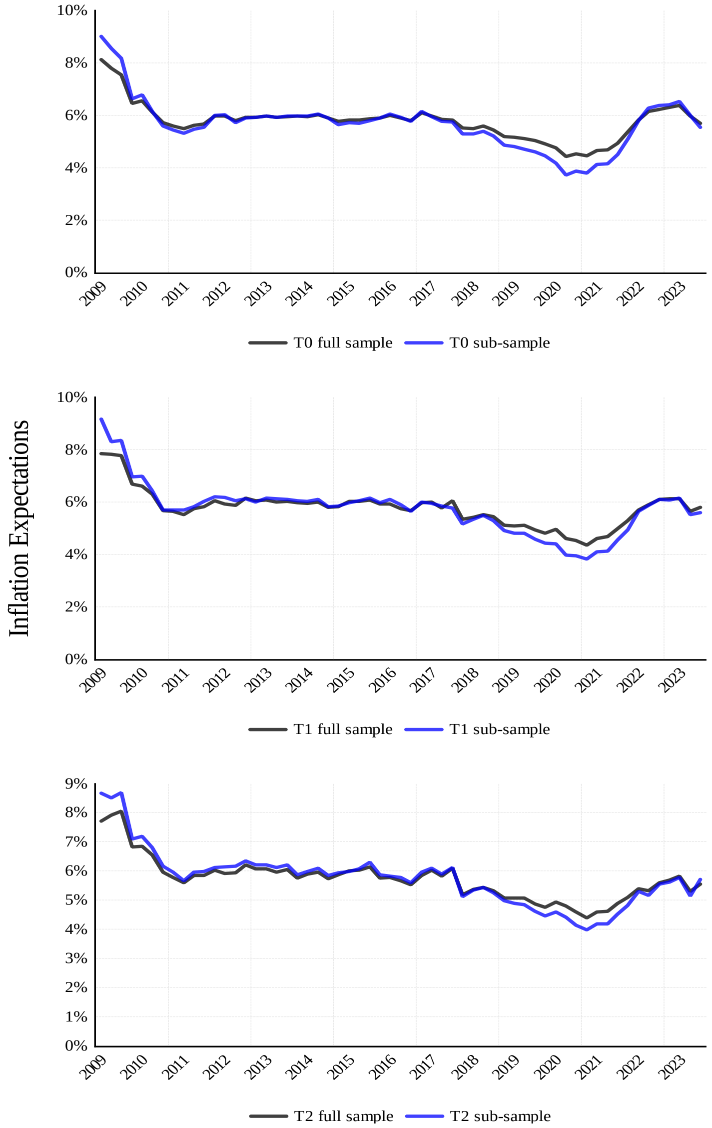


Note: Factor model 1 consists of business, financial analysts and trade union forecasts while factor model 2 adds the market-derived five-year-ahead inflation forecasts. The sample is 2011Q3–2023Q4.

In Figure 4, we display the revised factor model estimates as shown in Figure 3 against the estimates for the post-GFC period alone. We observe that, until 2018, sub-sample estimates result in an upward revision of the aggregate inflation forecasts, but thereafter there is a reversal. For the remainder of the sample, sub-sample estimates are revised downward relative to the full sample estimates. Since there do not appear to have been major international events that might explain this result, it is plausible that the change in the behaviour of inflation expectations was driven by domestic events – most likely the communication that the SARB would target the midpoint of the target range from mid-2017.[29] The difference in the two estimations also suggests that the factor model is sensitive to the use of different sample periods. The general narrative from the two is similar and the differences are not very large, although these are more notable during more volatile periods, such as the GFC and COVID-19 pandemic.

[29] 'The MPC would prefer expectations to be anchored closer to the mid-point of the target range' (SARB 2017).

Figure 4: Sample sensitivity of factor model-derived inflation forecasts: post-GFC period



Note: The sub-sample is defined as 2009Q2–2023Q4. Five-year-ahead forecasts begin in 2011Q3. No sub-sample factor model estimates were generated for the long-horizon forecasts.

In the next section, we discuss how to compare the forecasting performance of these alternative aggregate measures of expected inflation.

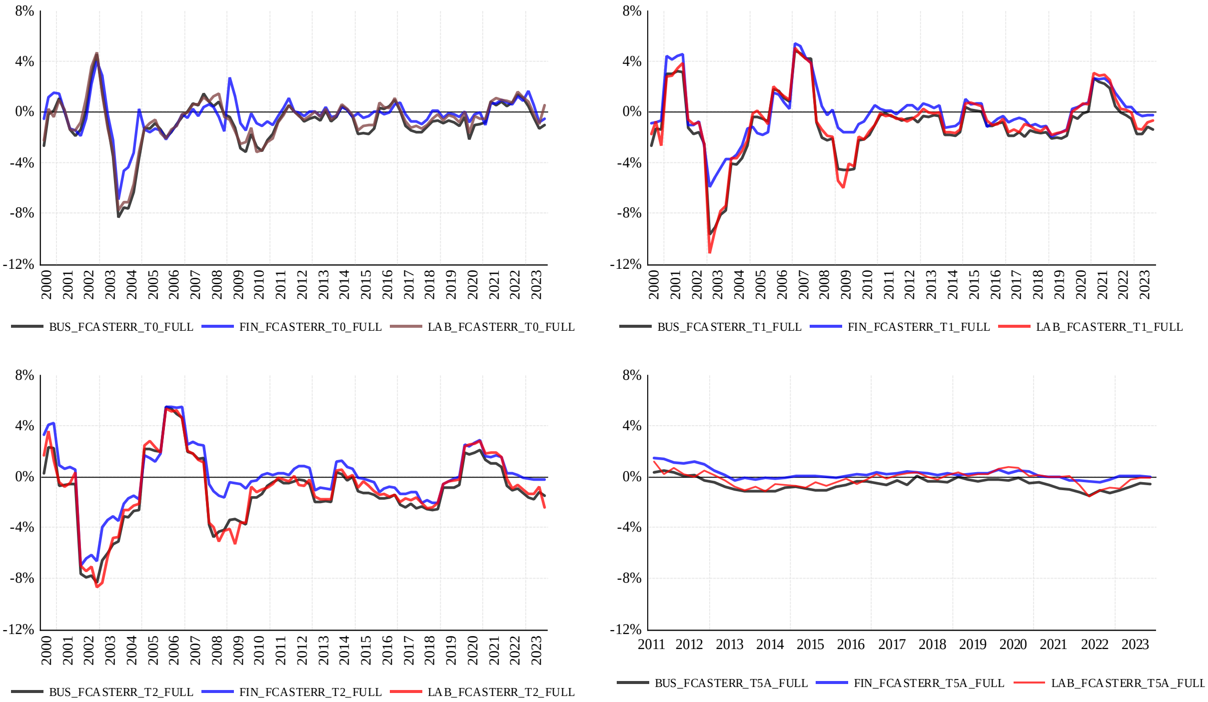
4. Forecast errors and measures of forecast performance

Forecast errors (FE) are defined as follows:

$$FE_{t,h} = \pi_{t,h} - \pi^{e}_{t,h} \qquad (2)$$

where $\pi_{t,h}$ is actual or observed inflation and $\pi^{e}_{t,h}$ is expected inflation over horizon h (also see equation 1). A positive FE signifies that respondents under-estimated inflation while an over-estimate follows when the FE is negative.

Figure 5: Forecast errors from the BER survey compared



Note: FULL refers to the full sample (2000Q2–2023Q4) for T0, T1 and T2 horizons, 2011Q3–2023Q4 for the five-year-ahead horizon. BUS, FIN and LAB are explained in the previous figure.

Figure 5 plots the FE at all horizons. Generally, FE is approximately stationary for the short-horizon forecasts (i.e. T0, T1 and T2).[30] The earlier period of the IT regime

[30] Unit root test results can, as usual, be sensitive to the test and specification. Nevertheless, if we take the example of one-year-ahead forecasts, all forecast errors were found to be stationary albeit with the possibility of a structural break. Interestingly, the timing of the break differs at this horizon depending on the forecast in question. A structural break occurs early in the IT era for the FIN and LAB survey forecast errors, and in 2007Q1 for the BUS survey. If we rely on the arithmetic mean of the BER surveys, then a break is found in 2005Q2. At the longer horizon, one generally still rejects the unit root finding for all three groups surveyed and the Bloomberg series even if a

included a large exchange rate crisis, the GFC and a period of learning by the public after the adoption of inflation targeting. After 2012, the FEs become more stationary, particularly for the short-horizon forecasts (i.e. T0, T1 and T2). FEs seem far more persistent for the long-term forecasts even if the errors are smaller in magnitude. Whether these differences reflect the SARB's credibility, communication or other sources is beyond the scope of this paper.

Forecast evaluation is often based on the root mean squared error criterion (RMSE) defined as:[31]

$$RMSE_{t,h} = \sqrt{\sum_{j=1}^{k}(\pi_{t,h} - \pi_{t,h}^{e})^2} \qquad (3)$$

where all the variables were previously defined and *k* represents the number of observations in the sample and the term in brackets is the FE (as presented in equation 2) squared.

However, Hyndman and Koehler (2006) persuasively argue that RMSE is scale dependent, which means that if the variable being evaluated (in this case, inflation) varies in size considerably over time, this can undermine the validity of the RMSE as a way of evaluating forecast performance. They also show that it is sensitive to outliers, although this problem is mitigated by taking the square root of the MSE. Finally, the RMSE lacks a benchmark, with the choice of best forecast reliant on the value of the statistic alone. For example, many economic phenomena are thought to contain a random walk component. Hence, one natural benchmark is the naïve forecast, defined in the present context as the last period's inflation forecast. As the survey-based and

structural break is considered, although in a few cases the results are overturned when a trend is added to the test specification.

[31] To conserve space, we focus on two straightforward forecast evaluation test statistics. A potential drawback with tests of the kind shown here is that if the loss from making an incorrect forecast is not a quadratic function (a very common assumption made by economists when specifying a loss function) then the metrics summarised below may not be ideal. Accordingly, in addition to the results shown below, we performed some forecast encompassing tests (results available in the annexures) (Chong and Hendry 1986); in an earlier paper, for the same data set, we applied a version of the well-known Diebold-Mariano test that allows for structural breaks (see Reid and Siklos (2023)). The conclusions presented below are unchanged.

market-based expectations are arguably substantively different from each other, it is plausible that RMSE may not be the best criterion on which to judge forecast performance in the case where we compare the survey-based inflation expectations with the asset market-based measure. Therefore, Hyndman and Koehler (2006) propose a scale-free superior criterion called mean absolute scaled errors (MASE), calculated as follows:

$$q_{t,h} = \frac{FE_{t,h}}{1/_{k-1} \sum_{t=2}^{k} |\Delta\pi_{t,h}|} \qquad (4)$$

and

$$MASE_{t,h} = mean(|q_{t,h}|) \qquad (5)$$

where $[\Delta\pi_{t,h}]$ is the absolute value of the change in observed inflation and all the other terms are as previously defined. The benchmark, as seen from the denominator, are the accumulated changes in inflation. If an economic agent were to rely solely on the naïve forecast, then the next period's forecasted inflation rate would be this period's inflation. When MASE<1, the forecast in question outperforms the naïve forecast and vice versa when MASE>1. When MASE=1, both forecasts perform equally well.

5. Forecast evaluation: results

Table 2 provides some estimates for both MASE and RMSE, as the former is our preferred measure but the latter is more commonly encountered in the literature. The top three performing results for each horizon are identified in the table using the exponents on the top right-hand side of the statistics in the table to facilitate interpretation. To help identify the patterns, the cells in which these top three performers appear are also shaded grey and the top performers appear in bold outline.

We focus on the MASE criterion represented in Table 2 first. For the full sample (Table 2(a)), at the first three horizons the factor model consistently performs best, with the performance of the financial analysts (FIN) usually in second place. The superior performance of the factor model is in line with principle 4 of good inflation forecasting identified by Faust and Wright (heavy shrinkage in the use of information improves inflation forecasts). The good performance of the financial analysts is also the typical finding that they are more rational than firms or trade unions (using the information available to them more efficiently). The BER aggregate comes in third place for the one- and two-year horizons, suggesting again that aggregation of diverse views is worthwhile.

The results for the five-year expectations look a little different from the other horizons, a distinction which persists when we consider different economic conditions and use the RMSE. The performance of the factor model, Bloomberg forecasts and trade unions are all very close, and notably superior to the alternatives. The good performance of labour at this horizon is interesting. Collective bargaining councils in South Africa generally negotiate multi-year agreements of between one and three years, so they may focus on longer horizons than the other survey groups, but they are not typically as long as five years.

Table 2: Mean absolute scaled errors

(a) Full sample

Type	CPIT0	CPIT1	CPIT2	CPI5a
BER	0.515	0.402 [3]	0.333 [3]	0.451
BUS	0.618	0.633	0.583	0.990
FIN	0.324 [2]	0.026 [2]	0.076 [2]	0.743
LAB	0.494 [3]	0.512	0.491	0.339 [3]
Factor model	0.003/0.088* [1]	0.003/0.070* [1]	0.016[1]/0.057*	0.423/0.290* [2]
Bloomberg	NA	NA	NA	0.315[1]

(b) 'Calm' sub-sample: 2009Q4–2019Q4

Type	CPIT0	CPIT1	CPIT2	CPI5a
BER	0.847	0.796 [3]	0.763 [3]	0.287 [2]
BUS	0.847	0.918	0.916	0.984
FIN	0.533 [2]	0.432 [2]	0.314 [2]	0.872
LAB	0.760 [3]	0.832	0.826	0.364 [3]
Factor model	0.072 [1]	0.164 [1]	0.246 [1]	0.154 [1]
Bloomberg	NA	NA	NA	0.491

(c) 'Volatile' sub-sample: 2000Q2–2009Q3 and 2020Q1–2023Q4

Type	CPIT0	CPIT1	CPIT2	CPI5a
BER	0.411	0.248 [3]	0.161 [2]	0.732
BUS	0.525	0.505	0.437	1
FIN	0.258 [2]	0.059 [2]	0.176 [3]	0.330 [3]

LAB	0.394 [3]	0.384	0.350	0.299 [2]
Factor model	0.026 [1]	0.055 [1]	0.060 [1]	0.749
Bloomberg	NA	NA	NA	0.100 [1]

Note: CPIT0 is the current year inflation forecast, CPIT1 is the forecast of the year ahead and CPIT2 is the forecast of inflation two years ahead. CPI5a is the forecast of average inflation expectations five years ahead.

*Post-GFC sample only (i.e. post-2009Q3 period). The numbered estimates (superscripts 1 to 3) represent a ranking from best performance on down. Only the top three performers are highlighted.

If financial analysts are consistently more rational and therefore forecast inflation more effectively, is there any reason to include the inflation expectations of the other groups? The performance of the factor model-based measure suggests that there is. We conjecture that financial analysts perform well in good times, when models generally perform well too, but that it is possible that those at the ‘coal face’ might have important insights that become particularly valuable when the experts face times of notable uncertainty.[32] To investigate this, we distinguish between ‘calm’ and ‘volatile’ inflation periods.[33] The former is defined as the period from 2009Q4 to 2019Q4 (see Figure 2) while the remaining data define the ‘volatile’ inflation period.

As Table 2(b) and (c) reveal, the results for both the calm and volatile periods reveal a pattern similar to those of the full sample results, although those from the calm period are stronger. At the first three horizons, the factor model and financial analysts (FIN) perform best. At the one- and two-year-ahead horizons, the BER aggregate also performs relatively well, and at the current-year and five-year horizons labour performs comparatively well. It is worth noting that the one- and two-year horizons are most consistent with the policy horizon.

32 This would be in line with the ideas in von Hayek’s Nobel lecture, ‘The pretence of knowledge’ (1974).

33 We also experimented with separate samples for rising and falling inflation periods, as well as recession and no recession samples (results not shown). The asymmetry highlighted here also holds for the other cases considered.

Table 3: Root mean squared errors

(a) Full sample

Type	CPIT0	CPIT1	CPIT2	CPI5a
BER	1.766 [3]	2.033 [2]	2.214 [3]	0.411 [1]
BUS	2.085	2.210	2.362	0.688
FIN	1.449 [2]	2.162	2.426	0.414 [2]
LAB	1.955	2.063 [3]	2.184 [2]	0.541 [3]
Factor Model	0.864 [1]	1.085 [1]	1.304 [1]	1.060
Bloomberg	NA	NA	NA	0.917

(b) Selected root mean squared errors: calm sample

Type	CPIT0	CPIT1	CPIT2	CPI5a
BER	0.959 [3]	1.117 [3]	1.233 [3]	0.827 [2]
BUS	1.291	1.447	1.562	1.112
FIN	0.555 [1]	0.854 [1]	0.967 [1]	0.720 [1]
LAB	1.162	1.269	1.355	0.868
Factor model	0.833 [2]	0.966 [2]	1.134 [2]	0.851 [3]
Bloomberg	NA	NA	NA	0.893

(c) RMSE: volatile sample

Type	CPIT0	CPIT1	CPIT2	CPI5a
BER	2.196 [3]	2.253 [2]	2.741 [3]	1.389
BUS	2.526	2.646	2.821	1.397
FIN	1.860 [2]	2.769	3.106	1.614
LAB	2.387	2.503 [3]	2.645 [2]	1.246 [2]
Factor model	0.934 [1]	1.317 [1]	1.626 [1]	1.378 [3]
Bloomberg	NA	NA	NA	0.890 [1]

Turning to the RMSE criterion (Table 3), the factor model-based forecast again performs consistently best in the full sample for the first three horizons, followed by the BER aggregate survey measure.

When the calm versus volatile sub-samples are considered, the factor model's advantage emerges in the volatile sub-sample while it is often in second or third place in the calm sample. When inflation fluctuates relatively little, then financial analysts outperform all other forecasters, including financial markets, at all horizons except, again, at the five-year horizon. In contrast, volatile inflation means less predictable inflation, in which case the factor model picks the relatively superior forecasts and combines them to outperform all others save the financial markets. While the results for the RMSE are not exactly the same as those of the MASE, the primary conclusions across both metrics are quite robust. The differences between the RMSEs may be due to their sensitivity to the changing scale of inflation itself.

6. Conclusions

The recent global surge in CPI has led to a renewed focus on central bank forecast accuracy in inflation. Indeed, many analysts (e.g. Giles (2024)) have suggested that the events of the past three years have eroded the public's trust in the forecasting performance of the monetary authorities, with a consequent loss of credibility.

One of the choices modellers have to make is which measure of inflation expectations to use in their models. South Africa is in the privileged position that the BER has surveyed businesses, financial analysts and trade unions for well over two decades. In so doing, a long time series of highly disaggregated expectations data are available. To date, the arithmetic mean of the BER's aggregate inflation forecasts have entered the SARB's Quarterly Projection Model. Therefore, given the availability of a rich set of inflation expectations data, we revisit the question of which forecast is best in a statistical test by considering a variety of ways of aggregating the data. In addition to the BER forecasts, we examine the individual survey forecasts, add financial market forecasts and combine all these forecasts in factor models.

We find that averaging the information from the different survey groups is valuable, but that the simple arithmetic averaging of firms, financial analysts and labour inflation expectations (i.e. the BER aggregate) rarely yields the best forecast. Instead, a simple factor model seems to consistently perform well (confirming principles 1 and 4 from Faust and Wright), regardless of whether we use the MASE or RMSE measure of forecast performance.[34]

Whether inflation is relatively calm or volatile does potentially affect forecast performance. Across these different conditions, the relatively consistent performance of the factor model is the clearest result. Beyond that, we conclude that the BER aggregate

[34] That is, whether we look at the record of accumulated forecast errors or compare individual or aggregated forecasts relative to a benchmark defined as a naïve forecast.

also performs quite well (confirming that aggregating data from the different sources is worthwhile). Depending on which metric we use and which horizon we evaluate, the financial analysts and trade unions each seem to contribute. At times, the firm-level results are not notably weaker than those for trade unions. Our conclusion is that these different groups each contain relevant information that can be productively used through some form of aggregation.

The main policy implications are then two-fold. First, a simple arithmetic mean of survey forecasts can be improved upon and this can have implications for the data the SARB forecasters should include in their forecasting model. Second, forecast performance is conditional on the overall state of the economy.[35] In practice, this would mean that policymakers should pay more attention to the information from the other sub-groups and apply judgement when they believe they are in a volatile period. Testing sensitivity would also be useful.

There are also at least four extensions to the research presented here. First, we may obtain additional insights into forecast performance by providing a more detailed statistical analysis of forecast errors over time. For example, if we find structural breaks in these errors, we may be able to better pinpoint the kinds of large shocks that lead to a deterioration of forecast performance. Second, as BER survey respondents provide forecasts of many other critical macrofinancial variables, we can use this additional data not only to find out what drives inflation forecasts but also to further improve them. Third, there may be other forecasts that were omitted (e.g. Consensus, a broader array of financial market-determined inflation forecasts) that could be added to the factor model to generate still better forecasts. Finally, there are other techniques that are available to aggregate individual forecasts (e.g. machine learning). We leave these extensions for future research.

[35] Here we focused on the distinction between calm and volatile inflation samples. However, other distinctions may also matter (e.g. recessions versus expansions, rising or falling inflation).

Annexures

A.1 Descriptive tables

Number of observations by quarter and source of survey

	Survey		
Date	**Business sector**	**Financial analysts**	**Trade union**
2000Q2	823	26	17
2000Q3	674	23	10
2000Q4	579	21	14
2001Q1	555	15	9
2001Q2	487	21	13
2001Q3	483	19	16
2001Q4	459	20	13
2002Q1	436	15	15
2002Q2	441	19	11
2002Q3	393	16	9
2002Q4	428	16	12
2003Q1	368	15	12
2003Q2	694	25	28
2003Q3	538	20	27
2003Q4	532	17	26
2004Q1	483	17	28
2004Q2	456	18	24
2004Q3	411	18	19
2004Q4	426	17	19
2005Q1	421	16	18
2005Q2	484	18	18
2005Q3	425	15	16
2005Q4	358	15	11
2006Q1	369	13	15
2006Q2	324	11	12
2006Q4	322	14	10
2006Q4	328	22	13
2007Q1	309	16	34
2007Q2	318	17	24
2007Q3	415	23	22
2007Q4	365	19	21
2008Q1	364	19	25
2008Q2	385	23	25
2008Q3	333	20	20
2008Q4	336	19	19
2009Q1	349	19	15
2009Q2	345	19	19
2009Q3	362	20	14
2009Q4	322	20	13
2010Q1	397	19	12
2010Q2	371	15	12
2010Q3	376	30	9
2010Q4	355	15	11
2011Q4	438	19	16
2011Q4	394	18	13
2011Q4	378	15	11
2011Q4	388	17	12
2012Q1	391	17	11
2012Q2	350	17	13
2012Q3	380	17	13

2012Q4	334	17	14
2013Q1	242	17	16
2013Q2	339	13	10
2013Q3	249	13	11
2013Q4	309	10	11
2014Q1	265	16	15
2014Q2	286	15	10
2014Q3	254	15	10
2014Q4	206	12	8
2015Q1	202	17	16
2015Q2	213	13	17
2015Q3	363	19	13
2015Q4	258	15	12
2016Q1	223	19	15
2016Q2	200	14	9
2016Q3	211	15	19
2016Q4	168	14	11
2017Q1	177	14	13
2017Q2	172	15	10
2017Q3	151	11	12
2017Q4	172	13	12
2018Q1	159	17	12
2018Q2	132	12	13
2018Q3	131	15	10
2018Q4	148	17	9
2019Q1	130	19	13
2019Q2	139	17	9
2019Q3	123	15	13
2019Q4	141	18	12
2020Q1	131	14	9
2020Q2	89	15	9
2020Q3	82	18	12
2020Q4	117	14	12
2021Q1	97	16	11
2021Q2	161	18	10
2021Q3	120	12	13
2021Q4	112	14	13
2022Q1	110	18	12
2022Q2	142	19	11
2022Q3	98	13	15
2022Q4	288	14	10
2023Q1	141	18	10
2023Q2	117	17	10
2023Q3	270	28	10
2023Q4	134	16	10
Total	***30 261***	***1 651***	***1 423***

Note: The number of observations refer to the CPIT0 series (current calendar year forecast for CPI). Normally, respondents provide forecasts for all horizons (T0, T1, T2 and five years ahead) but occasionally some respondents fail to do so. Accordingly, the cumulative sum of each column may not add up to the total, which provides the potential number of available observations. Data for the five-year-ahead average inflation forecasts begin in 2011Q3.

Source: Authors’ calculations from data provided by the BER.

Observations by firm size

Code	Name	# Employees	Business	Financial analysts	Trade union
1	Micro	E< 21	8 440	691	687
2	Small	21<E <50	5 991	109	109
3	Medium	51< E<200	8 696	295	292
4	Large	E > 200	6 844	292	280

Note: Raw data is available at a higher level of disaggregation but we adopted the aggregation used by the BER. Not all respondents respond by providing firm size data so the sum of each column may not add up to total shown in the preceding table. The code is the one used by the BER.

Source: Same as previous table.

Position of survey respondents: business and financial analyst surveys

Code	Title	Business	Financial analysts
0	CEO	18 435	29
1	Financial manager/accountant	8 422	18
2	Senior sales/production manager	898	0
3 or 7	Other	384	2
4	Economist	24	1 420
5	Investment analyst	0	60
6	Fund manager	0	80
8	Trade union	5	0
9	Employer organisation	1	0

Note: See notes to preceding tables. Trade union survey does not contain the foregoing information. The code is the one used by the BER.

Source: Same as first table.

SIC classification: business survey only

Sector name	SIC code (2-digit)	Mnemonic	Observations
Agriculture	11	AGR	2 489
Mining	13	MIN	559
Manufacturing	30–39	MFG	11 196
Electricity & water	42	ELE	13
Construction	5	CON	1 423
Wholesale & retail	61–64	RET	9 792
Transportation & communication	71–75	TRA	518
Finance & real estate	82–88	FIN	2 933
Community & social services	91–99	COM	1 256

Note: See notes to preceding tables.

Source: Same as first table.

Sector classification: financial analysts only

Name	Sector code	Observations
Banks	200	471
Advisers/brokers	210	859
Insurers	220	138
Other	230	182

Note: See notes to preceding tables. Labour is classified as belonging to sector 300.

A.2 Econometric and other statistical results: unless otherwise noted, the full sample is used (2000Q2–2023Q4; 2011Q3–2023Q4 for five-year-ahead forecasts)

Persistence: Business sector, no structural breaks permitted

Variable	AR(1)-Full	AR(1)-Q75	AR(1)-Q25
CPI inflation	0.89 (.05)		0.88 (.05)
CPIT0	0.92 (.04)	0.90 (.05)	0.92 (.04)
CPIT1	0.90 (.04)	0.87 (.05)	0.88 (.05)
CPIT2	0.88 (.04)	0.86 (.05)	0.85 (.05)
CPI5a	0.90 (.06)	0.93 (.06)	0.83 (.08)
SARB forecasts			
T0	0.87 (.05)		
T1	0.68 (.08)		
T2	0.69 (.06)		
Market forecasts	0.89 (.07)		

Note: Estimates via OLS. Heteroskedastic consistent standard errors in parenthesis. Unless otherwise noted, all coefficients are statistically significant at the 1% level. Market forecasts are from Bloomberg for the five-year-ahead horizon. CPI inflation is the annualised quarterly CPI headline inflation rate (i.e. $100*\Delta^4 \log P_t$). Q75 is the right tail of the inflation forecast distribution (75th percentile and above) while Q25 is the left-tail of the same distribution (25th percentile and below).

Persistence: Business sector, allowing for structural breaks

Variable	Break dates (MAX = 2)	AR(1) estimates
CPI inflation	2003Q1	1.26 (.18) T=10,
	2005Q2	0.72 (.08) T=9, 0.89 (.07) T=75
CPIT0	2005Q3	0.99 (.11) T=20,
	2008Q4	1.18 (.03) T=13, 0.88 (.04) T=61
CPIT1	2008Q1	0.87 (.08) T=30,
	2010Q2	0.26 (.06) T=9, 0.88 (.07) T=55
CPIT2	2008Q1	0.83 (.08) T=30,
	2010Q2	*0.11 (.09) T=9,* 0.86 (.06) T=55
CPI5a	2018Q1	0.64 (.10) T=25,
	2019Q2	-0.42 (.06) T=5, 0.76 (.08) T=19
SARB forecasts	2004Q3	0.13 (.13) T=11, 0.93 (.04) T=78
T0	2016Q3	0.44 (.11) T=63, 0.66(.06) T=30
T1	2006Q1	0.89 (.11) T=21, 1.37 (.04)
T2	2008Q4	T=11, 0.73 (.07) T=61

Variable	Break dates (MAX = 1)	AR(1) estimates
CPI inflation	None	Same as no break
CPIT0	None	Same as no break
CPIT1	2002Q4	*0.34 (.16) T=9*, 0.92 (.04) T=75,
CPIT2	2002Q4	0.26 (.08) T=9, 0.91 (.04) T=75,
CPI5a	2018Q1	0.65 (.09) T=25, 0.57 (.14) T=24

Notes to persistence table: Estimated via OLS. T is the number of observations. Break test relies on the Bai-Perron structural break test. The global test of M=2 or 1 against zero breaks is employed. Sample is trimmed by 10%, and a 5% significance level is used to detect a statistically significant structural break. Standard errors (Newey-West) are in parenthesis. The sample is 2000Q3–2023Q4 except for CPI5a, which begins 2011Q4. SARB forecasts begin 2000Q4 and end 2023Q4. Market forecasts (Bloomberg five-year-ahead implied inflation forecast) begin 2012Q3. CPIT0, CPIT1, CPIT2 and CPI5a are for the business sector survey. See below for the comparable results for the financial analysts and trade union portions of the survey. Estimates in *italics* are not statistically significant. The remaining are statistically significant at least at the 1% level.

Persistence: Financial analysts and trade unions

Variable	AR(1) -Full	AR(1)-Q75	AR(1)-Q25
Financial analysts			
CPIT0	0.92 (.04)	0.88 (.05)	0.90 (.05)
CPIT1	0.88 (.05)	0.84 (.08)	0.87 (.05)
CPIT2	0.83 (.06)	0.78 (.06)	0.87 (.05)
CPI5a	0.87 (.07)	0.80 (.06)	0.80 (.09)
Trade unions			
CPIT0	0.92 (.04)	0.90 (.05)	0.88 (.05)
CPIT1	0.88 (.05)	0.87 (.05)	0.84 (.06)
CPIT2	0.83 (.06)	0.78(.06)	0.87 (.05)
CPI5a	0.87 (.07)	0.80 (.09)	0.80 (.09)

Variable	Break dates (MAX = 1)	AR(1) estimates
Financial analysts		
CPIT0	None	Same as no break
CPIT1	None	Same as no break
CPIT2	2003Q1	-0.39 (.14) T=10, 0.86 (.07) T=84,
CPI5a	2016Q4	-0.08 (.15) T=20, 0.76 (.08) T=29
Trade unions		
CPIT0	No break	Same as no break
CPIT1	No break	Same as no break
CPIT2	2003Q1	*-0.39 (.27)* T=10, 0.86 (.07) T=84
CPI5a	2016Q4	*-0.08 (.15)* T=24, 0.76 (.08) T=29

Note: See notes to preceding tables. The case of 1 break only is shown. The case for 2 breaks has been estimated and results are available on request.

Unconditional correlations between forecast errors

T0 Horizon

Source	BUS	FIN	LAB
BUS – Full BUS – 'Calm' BUS – 'Volatile'			
FIN – Full FIN – 'Calm' FIN – 'Volatile'	0.864 0.830 0.873		
LAB – Full LAB – 'Calm' LAB – 'Volatile'	0.985 0.966 0.988	0.846 0.816 0.856	
SARB – Full SARB – 'Calm' SARB – 'Volatile'	0.815 0.748 0.844	0.792 0.635 0.823	0.812 0.839 0.829

T1 Horizon

Source	BUS	FIN	LAB
BUS – Full BUS – 'Calm' BUS – 'Volatile'			
FIN – Full FIN – 'Calm' FIN – 'Volatile'	0.878 0.838 0.886		
LAB – Full LAB – 'Calm' LAB – 'Volatile'	0.978 0.972 0.978	0.833 0.809 0.978	
SARB – Full SARB – 'Calm' SARB – 'Volatile'	0.893 0.773 0.906	0.969 0.921 0.973	0.859 0.767 0.871

Note: Unless otherwise noted, all unconditional correlations are statistically significant at the 1% level. 'Calm' refers to the 2009Q4–2010Q4 period; 'Volatile' refers to the 2000Q2–2009Q3 and 2020Q1–2023Q4 periods.

T2 Horizon

Source	BUS	FIN	LAB
BUS – Full BUS – 'Calm' BUS – 'Volatile'			
FIN – Full FIN – 'Calm' FIN – 'Volatile'	0.948 0.922 0.955		
LAB – Full LAB – 'Calm' LAB – 'Volatile'	0.972 0.965 0.973	0.900 0.886 0.907	
SARB – Full SARB – 'Calm' SARB – 'Volatile'	0.924 0.861 0.932	0.983 0.959 0.985	0.877 0.827 0.885

Five-year-ahead horizon

Source	BUS	FIN	LAB
BUS – Full BUS – 'Calm' BUS – 'Volatile'			
FIN – Full FIN – 'Calm' FIN – 'Volatile'	0.787 0.734 0.870		
LAB – Full LAB – 'Calm' LAB – 'Volatile'	0.822 0.813 0.923	0.679 0.675 0.839	
SARB – Full SARB – 'Calm' SARB – 'Volatile'	0.743 0.798 0.835	0.740 0.806 0.917	0.782 0.728 0.833

Note: See note to the previous table.

Factor model estimates: factor loadings – full sample

(a) Benchmark

Forecast horizon	Business		Financial analysts		Trade unions	
	Full	Sub	Full	Sub	Full	Sub
T0	0.985	0.975	0.829	0.798	0.983	0.971
T1	0.974	0.970	0.564	0.691	0.964	0.970
T2	0.964	0.971	0.730	0.773	0.932	0.928
5a	0.966	NA	0.928	NA	0.941	NA

(b) Extended model: full sample

Forecast horizon	Business	Financial analysts	Trade unions	SARB
	Full	Full	Full	Full
T0	0.965	0.893	0.966	0.898
T1	0.934	0.683	0.927	0.595
T2	0.952	0.785	0.924	0.485
5a	0.959	0.928	0.953	0.856

Note: The Bai-Ng method is used to estimate factor models with the exception of the five-year-ahead horizon (sample: 2011Q3–2023Q4), which relies on the Ahn-Horenstein method. Estimates are based on principal factors (maximum likelihood yielded similar results). In all cases, one cannot reject the null that there is a single factor. When a second factor is found, its explanatory power was found to be very small (less than 15%). Results are available on request. Full is the full sample (2000Q2–2023Q4); Sub are the sub-sample estimates (2009Q2–2023Q4). NA means not applicable because differences between Full and Sub-sample are too small to affect estimates. Results for the extended model estimated for the sub-sample are almost the same as for the full sample; these results are not shown but are available on request.

Correlation matrix of RMSE: full sample

Probability	BUST0	BUST1	BUDT2	BUS5A	FINT0	FINT1	FINT2	FIN5A	LABT0	LABT1	LABT2	LAB5A
BUST1	**0.94**	1.00										
	18.67	-----										
	0.00	-----										
BUST2	**0.90**	**0.97**	1.00									
	13.96	28.08	-----									
	0.00	**0.00**	-----									
BUS5A	**0.03**	**0.02**	**0.05**	1.00								
	0.20	0.14	0.35	-----								
	0.84	0.89	0.73	-----								
FINT0	**0.33**	**0.26**	**0.22**	**0.10**	1.00							
	2.44	1.83	1.58	0.68	-----							
	0.02	**0.07**	0.12	0.50	-----							
FINT1	**0.48**	**0.46**	**0.49**	**0.21**	**0.59**	1.00						
	3.81	3.57	3.90	1.48	5.07	-----						
	0.00	**0.00**	**0.00**	0.15	**0.00**	-----						
FINT2	**0.46**	**0.42**	**0.46**	**0.26**	**0.60**	0.92	1.00					
	3.57	3.17	3.63	1.88	5.16	15.90	-----					
	0.00	**0.00**	**0.00**	**0.07**	**0.00**	**0.00**	-----					
FIN5A	**-0.06**	**0.05**	**0.08**	**0.19**	**-0.12**	**-0.20**	**-0.22**	1.00				
	-0.44	0.33	0.54	1.36	-0.83	-1.39	-1.54	-----				
	0.66	0.74	0.59	0.18	0.41	0.17	0.13	-----				
LABT0	**0.81**	**0.67**	**0.62**	**0.04**	**0.52**	**0.60**	**0.57**	**-0.05**	1.00			
	9.74	6.28	5.54	0.29	4.24	5.16	4.79	-0.35	-----			
	0.00	**0.00**	**0.00**	0.78	**0.00**	**0.00**	**0.00**	0.73	-----			
LABT1	**0.78**	**0.71**	**0.69**	**0.05**	**0.44**	**0.71**	**0.67**	**-0.06**	**0.90**	1.00		
	8.68	6.99	6.65	0.33	3.41	6.92	6.23	-0.44	14.72	-----		
	0.00	**0.00**	**0.00**	0.74	**0.00**	**0.00**	**0.00**	0.66	**0.00**	-----		
LABT2	**0.77**	**0.74**	**0.78**	**0.16**	**0.44**	**0.69**	**0.71**	**-0.02**	**0.77**	**0.87**	1.00	
	8.36	7.60	8.67	1.15	3.36	6.54	6.98	-0.14	8.41	12.21	-----	
	0.00	**0.00**	**0.00**	0.25	**0.00**	**0.00**	**0.00**	0.89	**0.00**	**0.00**	-----	
LAB5A	**0.11**	**0.08**	**0.13**	**0.52**	**0.16**	**0.47**	**0.43**	**0.16**	**0.26**	**0.33**	**0.30**	1.00
	0.79	0.59	0.89	4.25	1.15	3.66	3.32	1.11	1.85	2.42	2.20	-----
	0.44	0.56	0.38	**0.00**	0.26	**0.00**	**0.00**	0.27	**0.07**	**0.02**	**0.03**	-----

Encompassing tests: full and sub-samples

Forecast horizon	BER vs BUS		BER vs FIN		BER vs TU		BER vs SARB	
	Full	Sub	Full	Sub	Full	Sub	Full	Sub
T0	.00	.00	.00	.00	.00	.00	.00	.00
T1	.00	.00	.00	.01	.00	.00	.00	.00
T2	.00	.72	.00	.00	.03	.00	.00	.00
5a	.69	.69	.00	.00	.00	.00	.00	.00

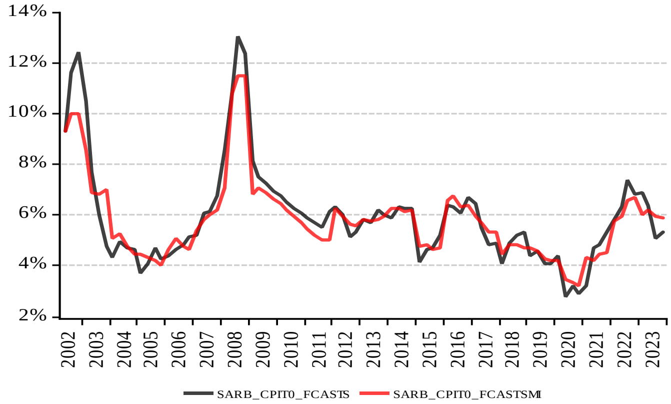
Multiple forecasts encompassing tests

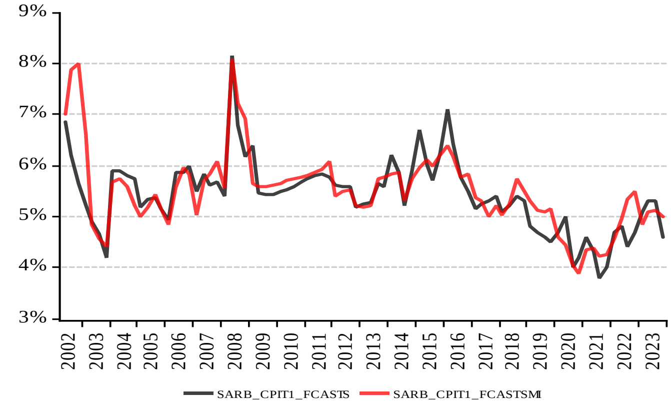
Forecast horizon	BUS		FIN		TU		SARB	
	Full	Sub	Full	Sub	Full	Sub	Full	Sub
T0	.00	.00	.00	.00	.00	.00	.00	.00
T1	.00	.00	.00	.06	.00	.00	.02	.06
T2	.00	.00	.01	.44	.00	.00	.03	.56
5a	.00	.00	.00	NA	.00	NA	.00	NA

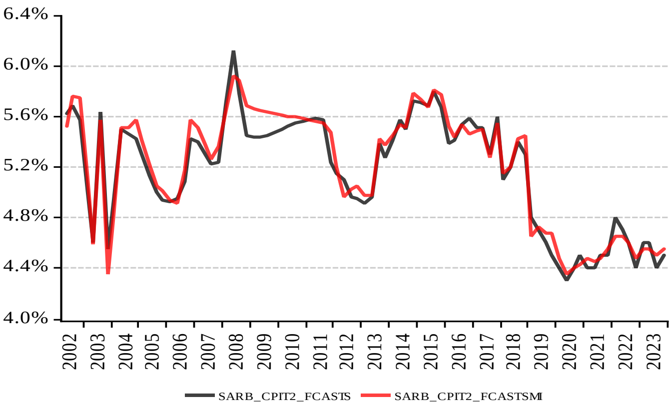
Note: Values shown are p-values for the test of the statistical significance of the γ coefficient in the regression $e^{BER}_{t+h,t} = \beta + \gamma\left(F^{j}_{t+h,t} - F^{BER}_{t+h,t}\right) + \varepsilon'_{t+h}$ (i.e., H_0.: $\gamma = 0$). BER is the BER forecast (average of BUS, FIN and TU forecasts. *j*= BUS, FIN, TU and SARB forecasts. *h* is the forecast horizon (T0, T1, T2 and 5a). NA means insufficient data for the sample in question. The full sample is 2000Q2–2023Q4 and the sub-sample is 2009Q2–2023Q4. The multiple forecasts encompassing test regresses forecast errors from the sources shown (i.e. BUS, FIN, TU and SARB) on forecast differences of the remaining forecasts. Accordingly, there are *k*=4 models estimated, and the null hypothesis is whether the F-statistic with dependent variable e^{k}_{t}, where *k*=BUS, FIN, TU, SARB is insignificant. P-values for the F-statistic are shown in the table.

A.3. Selected figures: unless otherwise noted, the full sample is used (2000Q2–2023Q4; 2011Q3–2023Q4 for five-year-ahead forecasts)

Fixed-event forecasts versus fixed-horizon SARB forecasts

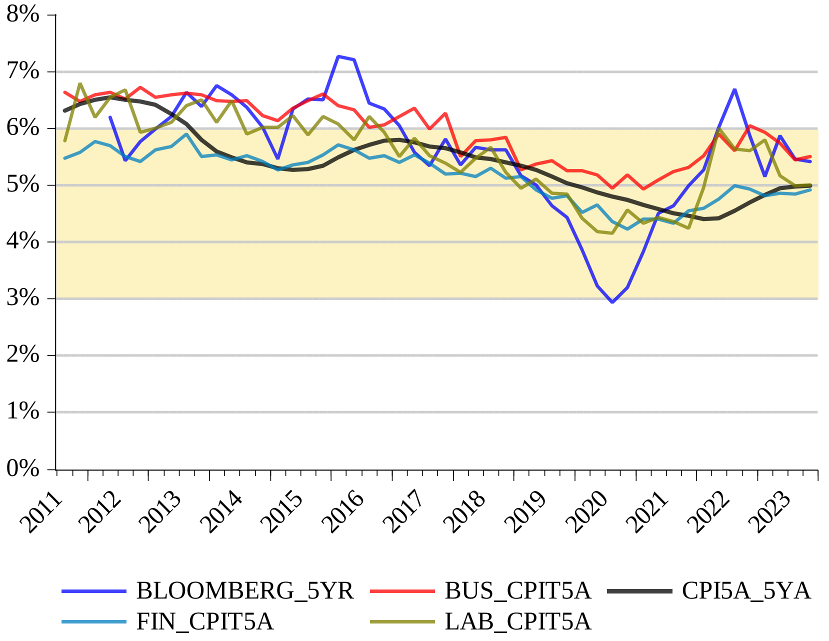






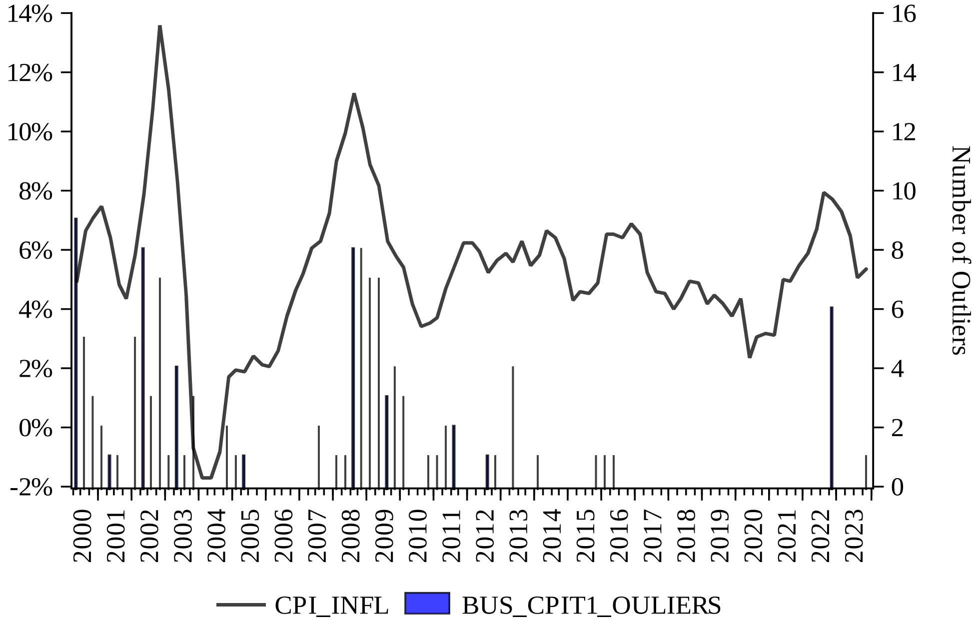
Note: SARB are SARB inflation forecasts. CPIT0, CPIT1 and CPIT2 are current, one- and two-year-ahead horizons, respectively. *_FCASTS and *_FCASTSMI are, respectively, calendar year forecasts that come close to the forecast provided in the BER survey (i.e. fixed-event forecasts) versus fixed-horizon forecasts equivalents. See Siklos (2013) for the methodology and calculation details.

Five-year-ahead forecasts compared



Note: Bloomberg are five-year-ahead inflation forecasts implied by market interest rate futures. CPI5A_5YA is the five years moving average of future inflation. BUS, FIN and LAB are business, financial analyst and trade union forecast surveys.

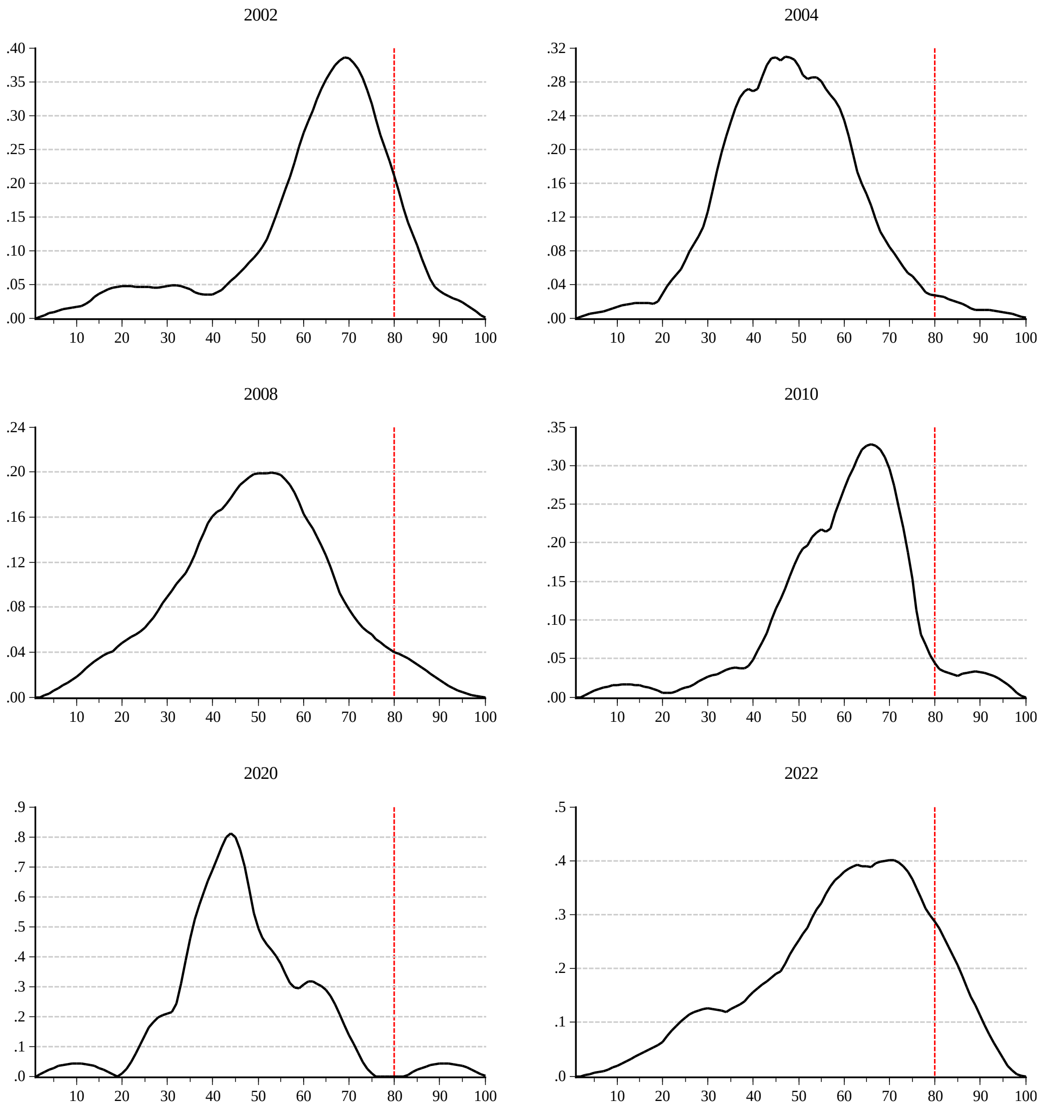
Observed inflation and outliers in business sector survey: the case of one-year-ahead forecasts



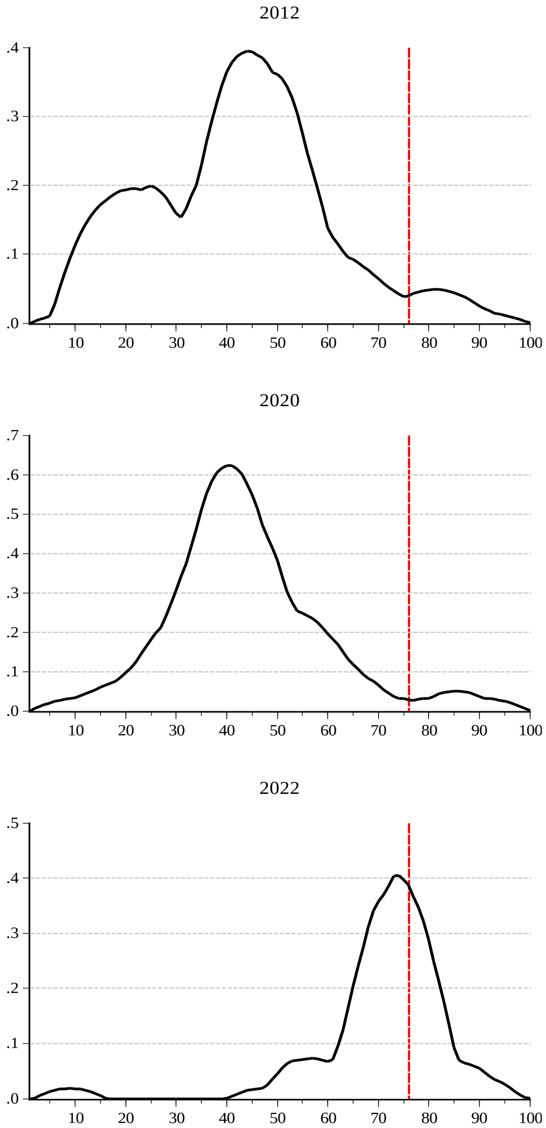
Note: An outlier is identified if two tests agree. They are: (1) thresholds defined by the distance from the mean such that the range of values a series can take that is not an outlier is: $\mu \pm m\sigma$. m is set to 2.7; (2) a wavelet outlier detection, following Bilen and Huzurbazar (2002).

Selected distribution graphs

Trade unions: one-year-ahead forecasts

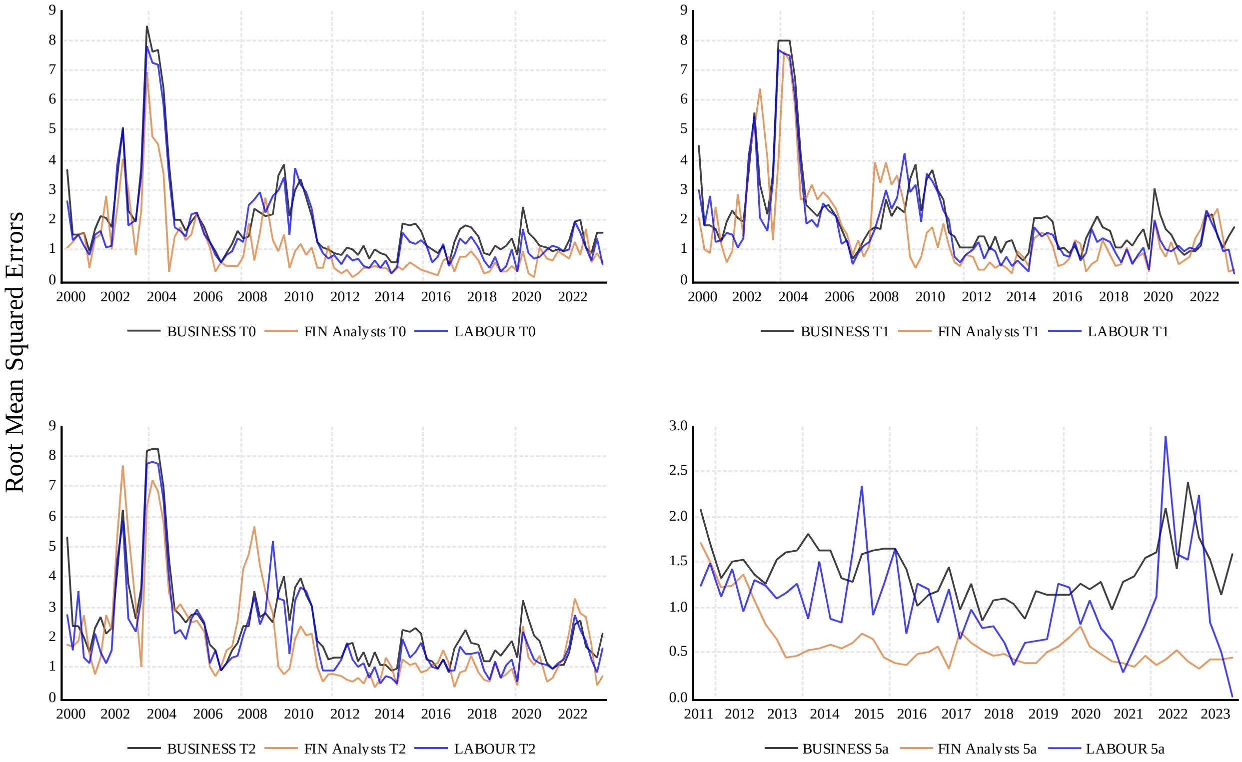


Trade unions: five-year-ahead moving average forecast



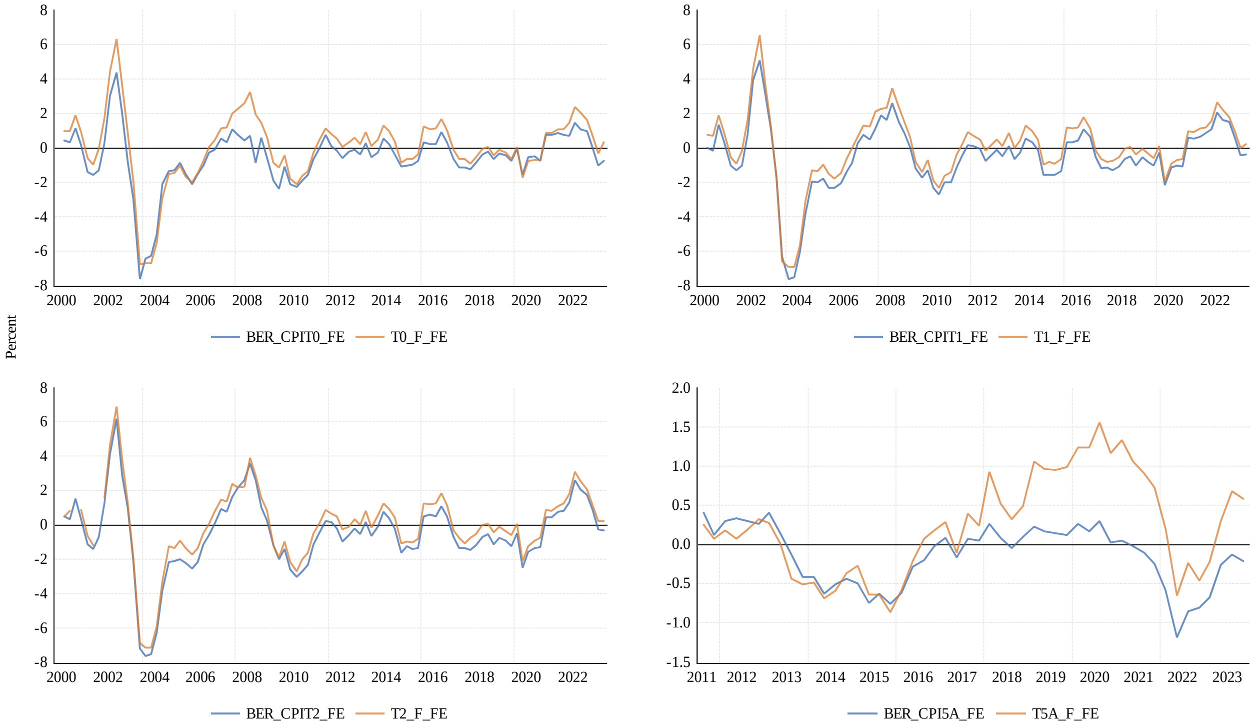
Note: The data are based on annual averages of quarterly forecast data for the calendar years shown above each figure. Each plot represents the kernel density (Epanechnikov or parabolic kernel). The vertical dashed line identifies where the one year (T1) or average over five years ahead (5a) distributions from the business sector survey are at their peak. The vertical line is used as a benchmark. The peak in the distribution (i.e. where inflation expectations responses are most frequent) is lower for each year shown in the trade union sector than in the business sector (peak is the vertical dashed line). Stated differently, if we were to superimpose the distribution graph for the business sector, the peak in that distribution would be to the right of the one shown above for the trade union group. Comparing 2020 to 2022, the peak of the distribution has shifted to the right, indicating rising average inflation expectations.

Root mean squared errors over time: selected results



Note: BUSINESS is business sector survey, FIN Analysts are the financial analysts and LABOUR are the trade unions surveyed. T1, T1, T2 and 5a are the forecast horizons for which respondents are asked to provide forecasts (T0, T1 and T2 are calendar year forecasts, 5a are five-year average forecasts).

Forecast errors compared: BER aggregate versus full sample factor model



Note: BER are the aggregate BER forecast errors (i.e. FE). T0, T1, T2 and T5A are the full sample forecast errors from the factor model as described in the main body of the paper.

References

Armantier, O, Bruine de Bruin, W, Topa, G, van der Klaauw, W and Zafar, B. 2015. 'Inflation expectations and behavior: do survey respondents act on their beliefs?' *International Economic Review* 56(2): 505–536.

Aysun, U and Wright, C. 2024. 'A two-step dynamic factor model approach for forecasting inflation in small open economies'. *Emerging Market Review* 62 (September): 101188.

Baumeister, C. 2021. 'Measuring market expectations'. National Bureau of Economic Research (NBER) working paper 29232.

Bernanke, B. 2024. 'Forecasting for monetary policy making and communication at the Bank of England: a review'. *Bank of England*, 12 April 2024.

Bianco, T and Haubrich, G. 2010. 'Inflation: noise, risk, and expectations'. Federal Reserve Bank of Cleveland, *Economic Commentary* 2010-05.

Bilen, C and Huzurbazar, S. 2002. 'Wavelet-based detection of outliers in time series.' *Journal of Computational and Graphical Statistics* 11(2): 311–327.

Botha, B, de Jager, S, Ruch, F and Steinbach, R. 2017. 'The Quarterly Projection Model of the SARB'. *SARB Working Paper Series*, WP/17/01.

Chong, Y Y and Hendry, D F. 1986. 'Econometric evaluation of linear macroeconometric models'. *Review of Economics and Statistics* 53(4): 671–690.

Coibion, O and Gorodnichenko, Y. 2015. 'Is the Phillips Curve alive and well after all? Inflation expectations and the missing disinflation'. *American Economic Journal: Macroeconomics* 7(1): 197–232.

Coibion, O, Gorodnichenko, Y, Kumar, S and Pedemonte M. 2020. 'Inflation expectations as a policy tool?' *Journal of International Economics* 124 (May): 103297.

Coletti, D. 2023. ‘A blueprint for the fourth generation of Bank of Canada Projection and Policy Analysis Models’. Bank of Canada, Staff Discussion Paper 2023-23.

Crowther-Ehlers, N. 2019. ‘Inflation expectations in South Africa: non-rational, intertemporal and idiosyncratic heterogeneity represented by a term structure approach’. PhD thesis, Stellenbosch University.

Ehlers, N and Steinbach, R. 2007. ‘The formation of inflation expectations’. *SARB Working Paper Series*, WP/07/06.

Faust, J and Wright, J H. 2013. ‘Forecasting inflation’. In *Handbook of Economic Forecasting*, Volume 2A, edited by G Elliott and A Timmermann (Amsterdam: Elsevier), 2–56.

Forbes, K, Gagnon, J and Collins, C. 2021. ‘Low inflation bends the Phillips curve around the world’. NBER Working Paper 29323.

Gagnon, J and Collins, C. 2019. ‘Low inflation bends the Phillips curve’. Peterson Institute for International Economics, Working Paper No. 19-6.

Gagnon, J and Sarsenbayev, M. 2021. ‘Who are the better forecasters of inflation, bond traders or economists?’. Peterson Institute of International Economics, Realtime Economics blog, 24 June.

Giles, C. 2024. ‘Learning to trust the forecasts’. *Financial Times*, 30 July.

Gosselin, M-A and Tkacz, G. 2001. ‘Evaluating factor models: an application to forecasting inflation in Canada’. Bank of Canada, Staff Working Paper 2001-18.

Granger, C W J and Ramanathan, R. 1984. ‘Improved methods for combining forecasts’. *Journal of Forecasting* 3 (April/June): 197–204.

Hall, S, Tavlas, G and Wang, Y. 2023. ‘Drivers and spillover effects of inflation: the United States, the euro area, and the United Kingdom’. *Journal of International Money and Finance* 131(March): 102776.

Hamilton, J and Xi, J. 2024. ‘Principal component analysis for nonstationary series’. NBER working paper 32068.

Hyndman, R and Koehler, A B. 2006. ‘Another look at measures of forecast accuracy’. *International Journal of Forecasting* 22(4): 679–688.

Niu, X and Harvey, N. 2023. ‘Are lay expectations of inflation based on recall of specific prices? If so, how and under what conditions?’ *Journal of Economic Psychology* 98: 102662.

Pirozhkova, E, Rakgalakane, J, Soobyah, L and Steinbach, R. 2023. ‘Enhancing the Quarterly Projection Model’. *SARB Working Paper Series*, WP/23/05.

Reid, M and Siklos, P. 2022. ‘How firms and experts view the Phillips curve: evidence from individual and aggregate data from South Africa’. *Emerging Markets Finance and Trade* 58(12): 3355–3376.

Reid, M and Siklos, P. 2023. ‘Rationality and biases: insights from disaggregated firm-level inflation expectations data’. *SARB Working Paper Series*, WP/23/07.

Reid, M and Siklos, P. 2024. ‘Firm-level expectations and macroeconomic conditions: underpinnings and disagreement’. *SARB Working Paper Series*, WP/24/05.

Reid, M B, Siklos, P and Du Plessis, S A. 2021. ‘What drives household inflation expectations in South Africa? Demographics and anchoring under inflation targeting’. *Economic Systems* 45(3): 100878.

Sandbu, M. 2024. ‘A self-conglatulatory inflation narrative at Jackson Hole’. *Financial Times*, 29 August.

Siklos, P. 2013. ‘Sources of disagreement in inflation forecasts: an international empirical investigation’. *Journal of International Economics* 90(1): 218–231.

South African Reserve Bank. 2017. *Monetary Policy Statement.* July.

Stock, J and Watson, M. 2003. ‘Forecasting output and inflation: the role of asset prices’. *Journal of Economic Literature* 41(3): 788–829.

Svensson, L E O. 1997. ‘Inflation forecast targeting: implementing and monitoring inflation targets’. *European Economic Review* 41(6): 1111–1146

Timmermann, A. 2006. ‘Forecast combinations’. In *Handbook of Economic Forecasting*, edited by G Elliott, C W J Granger and A Timmermann (Amsterdam: Elsevier), 135–196.

Weber, M, D’Acunto, F, Gorodnichenko, Y and Coibion, O. 2022. ‘The subjective inflation expectations of households and firms: measurement, determinants, and implications’. *Journal of Economic Perspectives* 36(3): 157–184.