Do professional forecasters believe in the Phillips curve?

International Journal of Forecasting 40 (2024) 1238–1254

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Do professional forecasters believe in the Phillips curve?[☆]

Michael P. Clements

ICMA Centre, Henley Business School, University of Reading, Reading RG6 6BA, United Kingdom

ARTICLE INFO

Keywords:
Inflation forecasting
Phillips curve
Model heterogeneity
Forecast disagreement
Theory consistency

ABSTRACT

The expectations-augmented Phillips curve (PC) is a cornerstone of many macroeconomic models. We consider the extent to which professional forecasters' inflation and unemployment rate forecasts are 'theory consistent', and find much heterogeneity. Perceptions about the responsiveness of inflation to the unemployment rate are shown to depend on whether the respondent was active earlier or later during the period 1981–2019, and on whether the respondent happened to forecast at times of tight labour markets.

Theory consistency is related to more accurate forecasts at the shortest horizon but not significantly so at longer horizons. At longer horizons PC-model heterogeneity accounts for the lion's share of the observed disagreement in reported inflation forecasts.

1. Introduction

At the end of the 'General Theory', Keynes states: '...the ideas of economists and political philosophers, both when they are right and when they are wrong, are more powerful than is commonly understood. Indeed, the world is ruled by little else.' (Keynes, 1936).[1] This paper looks at the extent to which professionals' forecasts of inflation and the unemployment rate conform to the Phillips curve and addresses the normative question of whether they ought to, in terms of enhanced accuracy. As suggested by Keynes, forecasters' perceptions will affect government policy and the macroeconomy, and expectations play a key role in modern macroeconomics.

We consider the Phillips curve because it has been a mainstay of academic research since Phillips (1958) first drew attention to the inverse relationship between UK wage inflation and unemployment. The Phillips curve is usually cast as a relationship between price inflation (rather than wage inflation) and an 'activity' variable, although the unemployment rate remains a popular choice. Since the late 1960s, the importance of expectations has been recognized (see, e.g., Friedman (1968)), and it has been suggested that the 'curve' may be vertical in the long run, suggesting the trade-off cannot be exploited to permanently reduce output below it's 'natural rate'.

Research in the 1970s culminating in Calvo (1983) provided 'micro-foundations' for the forward-looking expectations Phillips curve. Rather than just being an association or correlation between the real side of the economy (the unemployment rate or an activity variable more generally) and price or wage inflation, the equation was shown to arise from the optimizing behaviour of monopolistically competitive firms subject to 'sticky prices'. This cemented the role of the New Keynesian Phillips Curve (NK-PC) as the key determinant of inflation.[2]

Given the central role the Phillips curve (hereafter PC) has played in macroeconomics over the last half a

☆ This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

E-mail address: m.p.clements@reading.ac.uk.

1 This is followed immediately by the well-known quotation, 'Practical men, who believe themselves to be quite exempt from any intellectual influences, are usually the slaves of some defunct economist.'

2 See, e.g., Gali and Gertler (1999) and Coibion et al. (2018) for a review of the historical development.

https://doi.org/10.1016/j.ijforecast.2023.11.004

century, the question we address in this paper is whether professional forecasters' expectations embody a belief in the PC. Suppose the economics profession believes inflation expectations and the level of slack in the economy are key determinants of inflation. In that case, it may seem evident that survey expectations would embody this relationship. We argue that it is an interesting question for many reasons. Firstly, the extent to which professional forecasters base their expectations on theoretical models of the economy, with fully articulated PC linkages between prices and activity, is unclear.[3] An alternative is that they might use simple (possibly univariate) time-series models which do not embody such linkages.[4]

Secondly, evidence has accumulated against the PC, with the 'missing disinflation' in the US following the 2008 Financial Crisis (see, e.g., Stock (2011) and Coibion and Gorodnichenko (2015b)), and the recent low rates of inflation despite low rates of unemployment (see, e.g., Ball and Mazumder (2020)). These may be viewed as prominent failures of the model from the Crisis period onwards, although as early as Atkeson and Ohanian (2001), the out-of-sample forecast performance of the PC had been questioned.[5] We consider whether the professional forecasters' expectations reflect the apparent fall from grace of the PC.

Thirdly, the precise definitions of the inflation and unemployment rate variables may matter. For example, Ball and Mazumder (2020) stress the importance of distinguishing between core and headline inflation, excluding food and energy prices. Food and energy prices will affect the headline figure but may not be closely related to the level of activity. They go further and argue that large relative price changes also occur in industries other than food and energy. They argue for measuring inflation using the weighted median of price changes across industries, proposed as a measure of core inflation by Bryan and Cecchetti (1993). Ball and Mazumder (2020) show that the evidence for a PC is stronger if this measure of core inflation is used. The PC relationship between inflation and unemployment may be hidden if the econometrician only has access to headline CPI. In addition, others (e.g., Gali and Gertler (1999)) suggest the forcing variable should be a measure of marginal cost rather than of the output gap or unemployment rate (or an adjusted unemployment rate).

Fourthly, as well as possibly 'flattening' over time, the slope may depend on the tightness of the labour market, that is, the relationship may be non-linear (see, e.g., Hooper et al. (2019)).

For all these reasons, it is an open question whether belief in the PC underpins survey expectations of inflation and unemployment: whether the variables elicited by the survey are the variables linked by a PC, as well as whether forecasters' beliefs in the PC have changed over time (given episodes such as the missing disinflation), or depend on the tightness of the labour market.

At the individual level, short data samples with missing observations make it challenging to estimate time-varying and/or non-linear models. Our methodological contribution is an approach to determining whether non-linearities and time-variation play a role. We estimate simple linear PC relationships for each individual, but then test whether respondents who were active later in the sample or during tighter labour market conditions, differ systematically from earlier participants or those who made more of their responses during normal times. This approach might be useful whenever the researcher is interested in individual-level survey data since the sample sizes are often small relative to the complexity of the models one might wish to entertain.

Our empirical contributions relate to a number of strands of the literature on the perceptions or beliefs of agents about how the economy operates and the theory-consistency of agents' expectations. The literature considers a variety of agents, and for the most part, we consider professional forecasters (see, e.g., Clements et al. 2022).[6] (Krane, 2011) and Bluedorn and Leigh (2018) investigate the beliefs of professional forecasters regarding the permanency of shocks to output and whether there are long-term costs to recessions, following evidence provided by Aguiar and Gopinath (2007) for emerging market economies that the 'cycle is the trend'. Clements

[3] The focus of much of the recent literature on expectations formation has been on departures from full-information rational expectations, which stress informational rigidities and frictions: see, *inter alia*, Mankiw and Reis (2002), Woodford (2002), Sims (2003), and Mankiw et al. (2003), and more recently Coibion and Gorodnichenko (2012, 2015a)). There is some literature on forecasters' underlying beliefs about how the economy operates and how this informs their expectations, some of which are discussed below. A recent contribution by Kontny (2019) considers the impact of the presence of information frictions in survey forecasts of inflation and unemployment on estimates of the slope of the *average* forecaster's perceived Phillips Curve model. We consider the heterogeneity in the individuals' PC slope estimates and do not consider the approach of Kontny (2019) here, but this issue warrants further research.

[4] There is little direct (self-reported) information on how survey respondents form their expectations. A special survey of the US SPF respondents (see Stark 2013) reported: 'We found that almost all respondents use a combination approach to forecasting: Twenty of 25 respondents said they use a mathematical/computer model plus subjective adjustments to that model in reporting their projections. (One respondent reported using pure model-generated forecasts, and four said they use only their experience and intuition.) One interpretation of these results is that SPF panellists, like many macroeconomists in general, think models are useful but should not be fully trusted to deliver reasonable results in every circumstance.' Unfortunately, what is meant by a model and the extent to which such models embody relationships, such as the PC, is unclear.

An interesting possibility suggested by Malmendier and Nagel (2016) is that expectations are in part formed by "lived experiences" as opposed to learning by doing, but we cannot address these issues without knowledge of the forecasters' ages. The survey data we consider allows us to track individuals over time, but they remain anonymous, and we are ignorant of their personal characteristics.

[5] For forecasting inflation a year ahead over the period 1958–1997, Stock and Watson (1999) recorded a more positive assessment of the value of the PC but suggested that there might be better measures of real economic activity than the unemployment rate in terms of forecasting.

There are also theoretical arguments against the PC. McLeay and Tenreyro (2019) argue that the monetary authorities will act to increase inflation when there is slack in the economy. As a result, the PC will not be apparent in the data.

[6] See Bachmann et al. (2022) for detailed analyses of various agents' expectations. A prominent example of household expectations is the study of the Michigan Survey data of Carvalho and Nechio (2014).

(2020) also considers perceptions about output-growth persistence but looks at differences in perceptions of persistence across forecasters rather than aggregate perceptions. Jain (2019) considers inflation persistence at an individual level.

Fisher et al. (2023) consider individual long-run inflation expectations using a trend-cycle model that builds on Stock and Watson (2007) and Chan et al. (2018), and others. Relative to this strand of literature, and to Jain (2019), our focus is the consistency of inflation expectations with the PC.

Devereux et al. (2012) is close to our paper in considering whether a particular relationship holds between forecasts of the relevant variables. They use forecasts of these quantities to explore the relationship between relative consumption growth and real exchange rate depreciations across countries. However, they use aggregate expectations, and their motivation is that the use of forecasts may provide a better answer to the Backus-Smith puzzle (Backus & Smith, 1993) than using actual data. Hence, although (Devereux et al., 2012, p. 40, eqn. (14)) uses survey data alone to investigate the 'puzzle', the expectations are aggregate, and whether individual survey respondents have different perceptions of the putative relationship is not addressed. Dräger and Nghiem (2021) also consider consumers' expenditure, specifically whether spending decisions are consistent with an Euler equation model.

Fendel et al. (2011, p. 286) is similar to our paper in that it asks 'whether professional economic forecasters believe in and, thus, apply the wage and price Phillips curves for their forecasts'. They use Consensus Economics data and consider a number of countries. Their data is monthly and has a fixed-event structure because each month's forecasts are made for the current year and the following year. Our forecasts are fixed horizon, which seems preferable, and we can consider whether the PC relationships hold at a range of horizons. More importantly, we investigate the individual heterogeneity in the PC model slopes (similar in spirit to Jain (2019)), whereas Fendel et al. (2011) assume slope homogeneity and use a panel estimator. A question we ask is whether the heterogeneity we find depends on participation times.

Dräger et al. (2016) use individual survey data on consumers and professionals to determine whether expectations are consistent with a number of macro-theory relationships, including the PC. Mostly, they consider the number of times inflation and unemployment forecasts move in opposite directions (at each point in time, across the cross-section of forecasters). They focus on a simple bivariate association in neglecting the expected response of current inflation expectations to lagged and future expected inflation. We estimate PC models for each individual and, in so doing, consider not just the directions of changes but also the relationships between the expected magnitudes of the changes.[7] Moreover, we consider the individual survey data *by individual*, in that we use all the forecasts by a given respondent to determine the theory-consistency of that respondent's expectations.

Our exercise complements that of Casey (2020), who also considers whether forecasts are theory-consistent using individual-level forecasts. Casey (2020) uses different formulations of the PC compared to those described in Sections 2 and 3,[8] and for the U.S. SPF uses forecasts of annual average realizations rather than the quarterly forecasts that we exploit.[9] We use systems of equations per individual to provide more precise estimates of the PC model parameters, given the relatively small number of forecasts made by some respondents. We also quantify the relationship between model heterogeneity and forecaster disagreement. Despite some of these differences in approach, some of the key findings are broadly in line, as we note in Section 4.

Our paper is also related to a literature that uses survey expectations to test theories that involve expectations of future values of variables. The PC is a prime example, as discussed below. That literature uses survey expectations as an external source of expectations that is potentially superior to instrumenting future values of variables. The expectations are usually aggregate, and neither the cross-sectional dispersion of expectations nor how or whether the expectations are related to expectations of other series made by the survey respondents are addressed.

The plan for the rest of the paper is as follows. Section 2 begins with a brief review of the PC and the use of survey-based inflation expectations. Section 3 describes our approach to modelling the inflation and unemployment rate forecasts of survey respondents, and Section 3.1 how we test for parameter variation and non-linearity, given the relatively short samples of forecasts available for some respondents. Section 4 presents our empirical findings. Section 4.1 considers whether the observed heterogeneity in PC beliefs is attributable to respondents being active early or late in the period or in tighter than normal labour-market conditions. Section 4.2 answers

[7] However, as pointed out by a referee, in some circumstances, the simpler approach may provide more robust tests of the theory underpinning forecaster behaviour than the possibly misspecified PC models estimated here.

[8] Casey (2020) considers three formulations of the PC relationship. He regresses the h -step ahead forecasts of inflation on the unemployment rate forecasts, for individual j (his eqn(6), p.1443):

$$E_{j,t}\pi_{t+h} = \alpha_j + \beta_j E_{j,t}u_{t+h} + \varepsilon_{j,t}.$$

This does not allow for either future or lagged inflation expectations as explanatory variables.

The same relationship is then estimated but in terms of the expected *changes* in the inflation rate and the unemployment rate (eqn(7), p.1443). (That is, π_{t+h} and u_{t+h} in the above equation are replaced by $\Delta\pi_{t+h}$ and Δu_{t+h}, respectively). This is described by Casey as 'an expectations-augmented Phillips curve'. The third formulation is as the above but with time-fixed effects and an allowance for serial correlation in the error term.

Relative to Casey (2020), our hybrid PC (based on an expectations-augmented PC with a lagged inflation term) includes the time t -expectations of inflation at $t+h+1$ as the 'expectations-augmentation' term and inflation at $t+h-1$, as explained in the subsequent sections.

[9] Casey (2020) also considers whether the survey expectations are consistent with Okun's law, as we are reporting findings for the ECB's Survey of Professional Forecasters and the UK Survey of External Forecasters. Our findings are solely for the PC for the U.S. SPF. The U.S. SPF's longer duration facilitates investigating changes over time, which is one of our key focuses.

whether a belief in the PC is associated with more accurate forecasts. Section 4.3 provides an answer to the question of how much of the observed disagreement in inflation expectations can be explained by model heterogeneity. Section 5 reports a number of robustness checks, suggesting the findings are broadly unchanged for a range of alternative specifications. Section 6 offers some concluding remarks. The online Appendix provides full details of the actual and forecast data, as well as a discussion of PC models of the forecasts versus the revisions.

2. The Phillips curve and inflation expectations

The forward-looking PC can be written as:

$$\pi_t = \beta E_t \pi_{t+1} + \gamma u_t + \delta' w_t + \varepsilon_t \quad (1)$$

where under the full information rational expectations (FIRE) assumption, the inflation expectations term is the expectation of π_{t+1} given the information available at t. The unemployment rate u_t is sometimes replaced with other measures of activity or a measure of marginal cost. w_t is a vector that may contain oil prices and other supply-side shocks, and ε_t is the random error term. The fit of Eq. (1) can often be improved by including a lag of inflation relative to time t, giving the so-called hybrid PC:

$$\pi_t = \beta_b \pi_{t-1} + \beta_f E_t \pi_{t+1} + \gamma u_t + \delta' w_t + \varepsilon_t \quad (2)$$

The inclusion of the lagged inflation rate may exacerbate the 'fragility' of the model when it is estimated by replacing $E_t\pi_{t+1}$ by the actual value π_{t+1}, and instrumenting this term. Mavroeidis et al. (2014) discuss the problem of weak identification and the resulting high level of estimation uncertainty.

Instead of making the assumption of FIRE inflation expectations, a number of authors have instead used survey expectations (denoted by $\hat{E}_t\pi_{t+1}$, say). Coibion et al. (2018) provide a recent review of studies using 'real-time' or survey expectations to estimate PCs. They conclude that using survey expectations yields more stable and robust forward-looking PCs than estimating the PC under the assumption of FIRE. As well as avoiding the problem of weak identification, survey expectations would be preferable because they provide a better approximation to the expectations of the relevant economic agents (firms) when macroeconomic conditions change: see also (Coibion & Gorodnichenko, 2015b). Studies such as (Adam & Padula, 2011) find that a forward-looking NKPC for the US using SPF expectations for expected inflation is largely insensitive to the measure of the slack variable, contrary to PCs estimated under the FIRE assumption. Roberts (1995) compares the use of survey expectations (Livingston and Michigan) with FIRE as in McCallum (1976), when future inflation is assumed to have a unit coefficient ($\beta = 1$ in (1)). He finds a 'sensible' PC using survey expectations, whether detrended output or the unemployment rate is used as the activity variable, although both measures are statistically insignificant using McCallum's approach.[10] Rather than using survey expectations, Barnichon and Mesters (2020) propose using identified monetary shocks as instruments for inflation expectations.[11]

Coibion and Gorodnichenko (2015b) argue that using firms' expectations instead of those of professional forecasters allows the PC to explain the 'missing disinflation' following the 2008-9 Financial Crisis. Fig. 1 shows (headline) CPI inflation and the UR over the period 1981 to 2019, as well as the current-quarter[12] median forecasts. Despite a sharp fall in 2008:Q4 of −9.2%, inflation returns to fluctuate around a level barely lower than before the Crisis, notwithstanding an approximate doubling in the unemployment rate between 2008:Q1 and 2009:Q4 (from 4.9 to 10.0). Fig. 2 shows year-ahead ($h = 4$) median inflation forecasts are flat over this period, signifying no change in the longer-horizon inflation outlook.[13] At the same time, according to Coibion and Gorodnichenko (2015b), household expectations (taken to proxy firms' expectations) rose from a low of 2 $\frac{1}{2}$ % in 2009 to around 4% by 2013.

In summary, the recent literature suggests the use of real-time professional Forecasters' inflation expectations provide reasonable estimates of a PC relationship at least up to the Financial Crisis of 2008–9, but after that there is evidence that firms/consumers' inflation expectations diverge from those of professional forecasters, and that the use of the latter in a PC fails to explain the recent course of inflation. Our focus is somewhat different from the recent literature: we look at the extent to which a belief in a PC underpins professional survey respondents' *forecasts* of inflation and the unemployment rate over the last 40 years, including over the Financial Crisis period, and whether perceptions have been constant over time.

3. Phillips curves for individual respondents

We do not observe individuals' PC models, but we can estimate these from respondents' reported inflation forecasts and the unemployment rate. Model heterogeneity can arise for a number of reasons. The forecasts of a particular respondent j may be derived from j's model of the economy, and that model may be misspecified, for

[10] Although the size of the estimated coefficient is comparable, the insignificance reflects the lower precision. The R^2 is markedly higher when survey expectations are used. In all cases, he includes the current and lagged change in the real crude price of oil. See McCallum (1976, Table 1, p. 982).

[11] Relative to conventional methods, Barnichon and Mesters (2020) find a larger slope parameter on the activity variable and a reduced role for forward-looking inflation expectations.

[12] We also refer to these as zero-horizon ($h = 0$) forecasts. Some monthly and higher frequency data will be available for the target quarter when the forecasts are reported to the survey (around the middle of the quarter). In contrast, the latest quarterly data will be the advance estimate for the previous quarter.

[13] Note that the horizontal axis denotes the survey — when the forecast was made and the corresponding realization. Hence, the 2008:4 fall in inflation is associated with 2007:4, the period when the 4-step forecast was made.

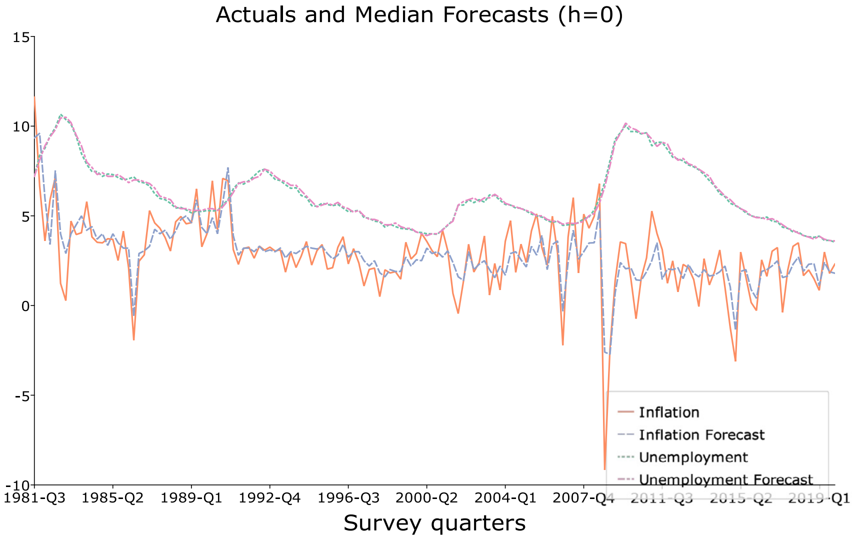

Fig. 1. Inflation and the unemployment rate, and consensus (median) current-quarter forecasts. The horizontal axis denotes the survey quarter (here also the forecast target period).

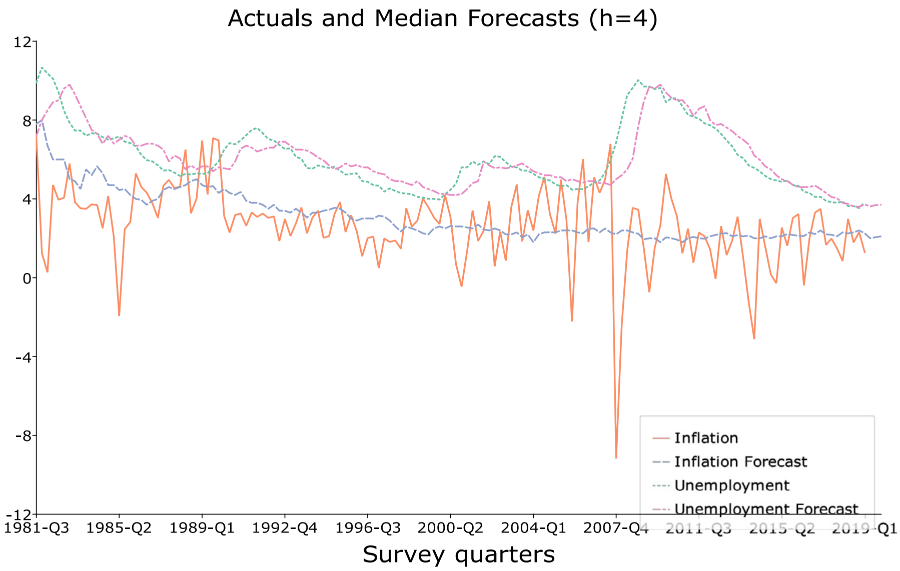

Fig. 2. Inflation and the unemployment rate, and consensus (median) four-quarter ahead forecasts. The horizontal axis denotes the survey quarter.

example, by omitting relevant variables.[14] Alternatively, the respondent may not have a formal model (based on observed data) but may use an informal model of how the economy works. Either way, we codify the implied PC relationship for that respondent by estimating a PC-model based on their survey responses. If a respondent's informal 'model' is not PC-theory consistent, this would show up as an insignificant coefficient on the unemployment rate, or the unemployment rate having a positive sign, say.

For each respondent j, we estimate a hybrid-PC model (2) using as data the *forecasts* of that respondent:

$$E_{j,t}\pi_{t+h} = \beta_{b,j}E_{j,t}\pi_{t+h-1} + \beta_{f,j}E_{j,t}\pi_{t+h+1} + \gamma_j E_{j,t}u_{t+h} + e_{j,t,h} \quad (3)$$

where $E_{j,t}\pi_{t+h}$ are j's h-step ahead forecasts of π_{t+h} made at time t, and so on. Relative to (2), the coefficients $\{\beta_b, \beta_f, \gamma\}$ are subscripted by j, to indicate that they may

14 The reader may wonder why a survey respondent would use a misspecified model of the PC to forecast. One might expect the respondent to become aware that the model generated poor forecasts over time. But as shown by Clements and Hendry (2002, p. 550–2), for example, model misspecification of itself will not be readily apparent in that it will not cause forecast failure — forecasts will not be worse than expected judged by the past fit of the model. Allowing for the model specification is a reasonable alternative to assuming all respondents know the true underlying relationship.

We have suggested that differences in models across respondents may result from differences in the set of included variables. Equally, such differences could result from some respondents imposing theory restrictions (such as that $\beta_b + \beta_f = 1$, implying a long-run PC), for example.

vary across respondents.[15] As noted in the Introduction, respondents' perceptions of the slope (γ) may depend on time, or may exhibit non-linearity, and in Section 3.1 we explain how we test for those features.

As well as estimating a PC model for the forecasts of each respondent, we also fit a PC model to the *revisions* to the forecasts of each respondent, as explained below.

The PC relationship in terms of forecasts suggests that the h-step forecast at time t depends on the $h+1$-step forecast at time t (from the 'expectations-augmentation term' of the usual PC) as well as the 'backward-looking' inflation term — the forecast of $t+h-1$ (made at time t).

The U.S. Survey of Professional Forecasters (SPF) forecast data (described in the Appendix) allows an equation such as (3) to be estimated for left-hand-side forecasts of $h = 0, 1, 2$ and 3. Notice $h = 0$ denotes a current-quarter forecast — a forecast of inflation in period t (π_t) made in survey quarter t. $E_{j,t}\pi_{t-1}$ denotes a 'forecast' of the previous quarter, or a $h = -1$ forecast. At time t, the advance estimate of the previous period has been released by the statistics office, and respondents usually but not always report this number as the $h = -1$ forecast. We use the respondent's $h = -1$ forecasts. The longest-horizon quarterly forecast is $h = 4$, and this restricts us to a system of four equations. We estimate the parameters of j rate forecasts:

$$\begin{aligned}
E_{j,t}\pi_t &= \zeta_j + \beta_{b,j}E_{j,t}\pi_{t-1} + \beta_{f,j}E_{j,t}\pi_{t+1} + \gamma_j E_{j,t}u_t + e_{j,t,0} \\
E_{j,t}\pi_{t+1} &= \zeta_j + \beta_{b,j}E_{j,t}\pi_{t} + \beta_{f,j}E_{j,t}\pi_{t+2} + \gamma_j E_{j,t}u_{t+1} + e_{j,t,1} \\
E_{j,t}\pi_{t+2} &= \zeta_j + \beta_{b,j}E_{j,t}\pi_{t+1} + \beta_{f,j}E_{j,t}\pi_{t+3} + \gamma_j E_{j,t}u_{t+2} + e_{j,t,2} \\
E_{j,t}\pi_{t+3} &= \zeta_j + \beta_{b,j}E_{j,t}\pi_{t+2} + \beta_{f,j}E_{j,t}\pi_{t+4} + \gamma_j E_{j,t}u_{t+3} + e_{j,t,3}
\end{aligned} \tag{4}$$

We have included constant and error terms because the relationship does not hold exactly. Notice that the parameters are the same across the equations. Using a system of equations ought to improve the precision of the parameter estimates (as in Jain (2019)). The j-subscripts on the parameters indicate that the system is estimated separately for each forecaster.

We also consider forecast *revisions* rather than the forecasts themselves, as in the studies by, e.g., Krane (2011), Jain (2019) and Clements (2020).[16] In the online Appendix we set out some of the relative advantages of the two approaches and highlight the potential importance of measurement issues. As we explain there, we think it is likely that the formulation using forecasts may be preferable. Nevertheless, comparing the findings of the two approaches is interesting.

The revision between the forecasts made at time $t-1$ and t is defined as $r_{j,t,h} = E_{j,t}\pi_{t+h} - E_{j,t-1}\pi_{t+h}$. Using (3), the revision in the current-quarter forecast is given by:

$$\begin{aligned}
r_{j,t,0} &= E_{j,t}\pi_t - E_{j,t-1}\pi_t \\
&= \left(\beta_{b,j}E_{j,t}\pi_{t-1} + \beta_{f,j}E_{j,t}\pi_{t+1} + \gamma_j E_{j,t}u_t\right) \\
&\quad - \left(\beta_{b,j}E_{j,t-1}\pi_{t-1} + \beta_{f,j}E_{j,t-1}\pi_{t+1} + \gamma_j E_{j,t-1}u_t\right) \\
&= \beta_{b,j}\left(E_{j,t}\pi_{t-1} - E_{j,t-1}\pi_{t-1}\right) \\
&\quad + \beta_{f,j}\left(E_{j,t}\pi_{t+1} - E_{j,t-1}\pi_{t+1}\right) \\
&\quad + \gamma_j\left(E_{j,t}u_t - E_{j,t-1}u_t\right) \\
&= \beta_{b,j}r_{j,t,-1} + \beta_{f,j}r_{j,t,1} + \gamma_j r_{U,j,t,0}
\end{aligned}$$

where $r_{U,j,t,0}$ is the revision in the forecasts of the current quarter unemployment rate (signified by the 'U' subscript). Similarly, the revision to the h-step ahead forecast is:

$$\begin{aligned}
r_{j,t,h} &= E_{j,t}\pi_{t+h} - E_{j,t-1}\pi_{t+h} \\
&= \left(\beta_{b,j}E_{j,t}\pi_{t+h-1} + \beta_{f,j}E_{j,t}\pi_{t+h+1} + \gamma_j E_{j,t}u_{t+h}\right) \\
&\quad - \left(\beta_{b,j}E_{j,t-1}\pi_{t+h-1} + \beta_{f,j}E_{j,t-1}\pi_{t+h+1}\right. \\
&\quad \left. + \gamma_j E_{j,t-1}u_{t+h}\right) \\
&= \beta_{b,j}r_{j,t,h-1} + \beta_{f,j}r_{j,t,h+1} + \gamma_j r_{U,j,t,h}.
\end{aligned} \tag{5}$$

Given the longest horizon quarterly CPI inflation and the unemployment rate forecasts in the US SPF are for $h = 4$ (that is, of the same quarter in the next year), and because the revision to the horizon h forecast depends on the revision to the $h+1$ horizon inflation forecast, the longest-horizon revisions we are able to model are given by $r_{j,t,2}$. From (5), $r_{j,t,2}$ depends on $r_{j,t,3}$, and $r_{j,t,3}$ uses the longest-available forecast, $E_{j,t-1}\pi_{t+3}$.

This means that we have the following system of three equations for forecast revisions:

$$\begin{aligned}
r_{j,t,0} &= \beta_{b,j}r_{j,t,-1} + \beta_{f,j}r_{j,t,1} + \gamma_j r_{U,j,t,0} \\
r_{j,t,1} &= \beta_{b,j}r_{j,t,0} + \beta_{f,j}r_{j,t,2} + \gamma_j r_{U,j,t,1} \\
r_{j,t,2} &= \beta_{b,j}r_{j,t,1} + \beta_{f,j}r_{j,t,3} + \gamma_j r_{U,j,t,2}.
\end{aligned} \tag{6}$$

This system allows for the identification of β_j and γ_j from estimating the following set of equations:

$$\begin{aligned}
r_{j,t,0} &= \kappa_j + \beta_{b,j}r_{j,t,-1} + \beta_{f,j}r_{j,t,1} + \gamma_j r_{U,j,t,0} + \upsilon_{j,t,1} \\
r_{j,t,1} &= \kappa_j + \beta_{b,j}r_{j,t,0} + \beta_{f,j}r_{j,t,2} + \gamma_j r_{U,j,t,1} + \upsilon_{j,t,2} \\
r_{j,t,2} &= \kappa_j + \beta_{b,j}r_{j,t,1} + \beta_{f,j}r_{j,t,3} + \gamma_j r_{U,j,t,2} + \upsilon_{j,t,3}
\end{aligned} \tag{7}$$

As for (4), the derivation of (6) suggests that the population parameters $\left\{\beta_{b,j}, \beta_{f,j}, \gamma_j\right\}$ are the same for the different horizon revision. So whether we use forecasts or forecast revisions, the equations are estimated by pooling the data over h and t (for a given j) to obtain more efficient parameter estimates. We estimate the equations by GLS assuming the following error structure: $E\left(\upsilon_{j,t,i}\upsilon_{j,s,k}\right) = \sigma_{ik}$ when $t = s$, but is zero otherwise. That is, the errors in the equations for the revisions to the forecasts

15 Either reflecting different beliefs/informal models or that the forecasts have been generated from models that are misspecified in different ways.

We do not attempt to capture all the variables that might have been included in the models used to generate the forecasts (given by the w_t variables in (2), for example), and differences in the parameter estimates across respondents might, in part, reflect this.

16 Krane (2011) argues that forecast revisions are superior to using forecast errors for identifying shocks to information sets because forecast errors will be affected by data revisions. We consider forecasts and forecast revisions and avoid actual values and data-vintage considerations.

made by individual j in response to the same surveys are allowed to be correlated. This reflects the impact of unmodelled factors at time t on the revisions at the three horizons. (The same error structure is assumed for the forecast equations errors, the e's.) To make matters concrete, consider the equations for revisions in (7). If the data are pooled such that the vector of errors for forecaster j is $v_j' = \left[v_{j,1,1}\ v_{j,2,1}\ \ldots v_{j,T,1};\ v_{j,1,2}\ v_{j,2,2}\ \ldots v_{j,T,2};\ v_{j,1,3}\ v_{j,2,3}\ \ldots v_{j,T,3}\right]$, then defining $\Omega_{v_j} = E\left(v_j v_j'\right)$, we have $\Omega_{v_j} = \Sigma_{v_j} \otimes I_T$, where Σ_{v_j} is a 3 by 3 matrix with elements estimated from the OLS residuals of the pooled regression: the i, k element is estimated from $T^{-1}\hat{v}_{j,i}'\hat{v}_{j,k}$, where $v_{j,i}' = \left[v_{j,1,i}\ v_{j,2,i}\ \ldots v_{j,T,i}\right]$, etc. Data will be missing in practice, so the number of (non-missing) forecasts will differ over j. Nevertheless, when an individual responds to a survey, the respondent almost always supplies forecasts of inflation and unemployment for the current quarter and each of the next four quarters. So, for a respondent to the $t-1$ and t surveys, we are generally able to calculate r_{jth} and $r_{U,jth}$ for $h = 0, \ldots, 4$. In the few instances when this is not true, we delete the revisions from the sample.

3.1. Testing for parameter variation and non-linearity

The literature suggests a 'flattening' in the PC relationship in the actual data in recent times.[17] Respondents' estimated PC slopes might simply reflect when they were active — either earlier or later in the period 1981–2019.

In principle, one might estimate the PC on rolling windows of forecast data for each respondent and track whether the PC's slope becomes less pronounced over time. In practice, many forecasters do not provide enough forecasts over a long enough historical period to allow us to determine reliably whether their PCs flatten over time. An individual may make too few forecasts to allow for the estimation of the model on sub-samples. Instead, we investigate whether perceptions have changed by considering whether the variation across respondents' PC estimates is related to the variation in terms of participation times. We might consider any aspects of the models. Still, we focus on (i) the explanatory power of the individuals' PC models, as measured by the R^2, and (ii) the effect size of the unemployment rate variable. Are these correlated with when, in the period 1981—2019, an individual was an active survey respondent? Forecasters do not necessarily join the survey and record contiguous responses to n surveys before exiting. Given the prevalence of non-responses, this means we cannot necessarily rely on the start (or end) date of their participation to determine when they were active.

Based on the literature, we suppose that the perception of the PC slope may be flatter in the last two decades (2000:1 to 2019:4) relative to the period 1981:3 to 1999:4, although alternative assumptions about the change point can easily be accommodated. For each individual respondent, we calculate the proportion of their forecasts made in the earlier two decades (1981:3 to 1999:4),[18] and call this $Period_j$ for respondent j. Larger values of *Period* indicate the respondent made a greater number of their forecasts in the earlier period. A zero value suggests an individual was only active in the 21st century.

The rationale for our approach is the following. For simplicity of exposition, suppose the expectations of all respondents are underpinned by the same PC-model:

$$\pi_t = 1_{(t\leq\tau)}\beta_1' x_t + \left(1 - 1_{(t\leq\tau)}\right)\beta_2' x_t + v_t \tag{8}$$

where $x_t = \left[\pi_{t-1}, E_t\pi_{t+1}, u_t, w_t'\right]'$, and β_1 and β_2 contain the coefficients, which change abruptly at time τ in (8). Allowing for a smooth change over time to allow γ to flatten gradually (i.e., approach 0 from a negative value) would not affect the argument.[19] If we estimate a *linear* PC on observations $t = 1, 2, \ldots, \tau, \tau+1, \ldots, T$, when the data have been generated by (8), the estimate of the parameter vector $\hat{\beta}$ will be a weighted average of β_1 and β_2 (in expectation), with the weights depending on the number of pre- and post-break observations, τ and $T-\tau$. Hence, individuals' $\hat{\beta}$'s (and especially $\hat{\gamma}$'s) will differ because their forecasts consist of different proportions from the pre- and post-τ periods, where τ corresponds to 1999:4 in our empirical work. Hence, heterogeneous slope parameters arise in this example because respondents are active to different degrees in the pre and post τ. In practice, they may not have a shared perception of the PC, of course, but below, we consider the importance of participation times in explaining the observed heterogeneity.

An alternative explanation of the PC-slope heterogeneity is that non-linearity in the relationship is important — that price inflation will only respond to reductions in the unemployment rate when the unemployment rate is already low, that is, in 'hot' labour markets,[20] and that this is embodied in respondents' PCs. That is, respondents who were primarily active when the economy was close to full employment are more likely to report forecasts consistent with a *linear* PC with a large negative slope, than respondents who were active at times of high unemployment. In principle, one might estimate a non-linear model for each respondent, which includes a term allowing the slope to depend on the unemployment rate, but in practice, there might be too few observations at the individual level for this effect to be accurately captured. But we can proceed as above, but now we define a variable Hot_j, which records the proportion of individual j linear PC-slope parameter for individual j will depend on the proportion of forecasts made in the two regimes (low unemployment versus not-low unemployment, as opposed to pre- and post-τ).

We calculate Spearman's rank correlation coefficient of whether there is a statistically significant relationship between $Period_j$ and $\hat{\gamma}_j$, and $Period_j$ and R_j^2, as well as

17 See, for example, Hooper et al. (2019).

18 More precisely, the set of t for which we have an observation for the system of equations. So $Period_j$ will differ (typically to a small extent) for the systems of forecasts and revisions.

19 See Lundbergh et al. (2003) for a discussion of time-varying smooth transition models.

20 See, e.g., Albuquerque and Baumann (2017) for recent evidence for the U.S.

between Hot_j and $\hat{\gamma}_j$, and Hot_j and R_j^2.[21] The rank correlation does not require that the relationship between the variables is linear because it works off the ranks of the variables.

It is possible to combine parameter variation and non-linearity. For example, to test whether respondents who were active in tight labour-market conditions pre-2000 have steeper slopes, that is, whether there are interaction effects.

The tests of rank correlation between Hot_j (or $Period_j$) and $\hat{\gamma}_j$ do not indicate the magnitude of the differences in the $\hat{\gamma}$-estimates across *Hot* or *Period*. For this reason, we report the (average) $\hat{\gamma}$ estimates of individuals in different quartiles of the distribution of *Period* values (say). This indicates how effect sizes vary with the time when the forecasts were made and the state of the labour market.

4. Empirical findings

The forecast data are from the US SPF from 1981:Q3 to 2019:Q4. The Appendix provides full details, see also Croushore (1993). Forecasts of our measure of inflation, the headline CPI, were first collected in 1981:Q3, although the survey began in 1968:Q4. Section 4.1 considers the individual-level heterogeneity in PC beliefs, and the extent to which this reflects when the individual was an active respondent, including labour-market conditions at the time. Section 4.2 asks whether a belief in the PC is associated with more accurate inflation forecasts and Section 4.3 whether differences in respondents' models account for the observed disagreement in inflation forecasts.

4.1. Individual heterogeneity

We consider the subset of 67 forecasters for whom we have at least twenty observations to estimate the four-equation system of forecasts (4) and the three-equation system of revisions (7).[22] Because revisions require forecasts from adjacent surveys, say $t-1$ and t, this means that we have more observations to estimate the system of forecasts (4). For some forecasters, we have more than 100 observations to estimate (4).

Table 1 summarizes the results on forecaster heterogeneity by reporting summary statistics for estimating (4) and (7) separately for each of the 67 respondents. We report the cross-sectional mean, standard deviation, and the lower and upper quartiles for each estimated parameter. We also report the proportion of respondents for whom we reject the null that the coefficient is zero, at the 1%, 5% and 10% levels (in two-sided tests). We also report summary statistics for the R^2's, and the sum of β_b and β_f.

Generally, respondents attach a greater weight to the forward-looking inflation term than the lagged rate of inflation. The cross-sectional means (for Forecasts, top panel) are 0.635 and 0.235 for β_f and β_b, respectively, and half the respondents' estimates of β_f are between 0.595 and 0.716 and between 0.226 and 0.339 for β_b. The findings for the inflation coefficients are similar if the model is estimated on revisions instead. The coefficients of the inflation terms are statistically significant for the vast majority of respondents. However, although the unemployment coefficients are predominantly negative, they are only statistically significant for just under a quarter of the respondents using forecasts (at the 5% level). The finding of statistical significance for a quarter of respondents is broadly comparable to Casey (2020, Table 3, p. 1445). Casey reports that 32% of respondents have a negative and statistically significant unemployment rate coefficient in his version of the expectations-augmented PC.

For revisions, the unemployment coefficients are only statistically significant for 1 in 10 respondents. Moreover, the unemployment rate coefficient is more sensitive to the use of forecasts or revisions than the inflation terms: the cross-sectional mean is -0.023 for forecasts and -0.063 for revisions. The cross-sectional standard deviation for revisions is 0.221, suggesting a good deal of variability around this larger (more negative) average response. The main change from using revisions (rather than the forecasts themselves) is then the greater dispersion in the perceptions of responses to the activity variable.[23]

Table 1 also records statistics relating to the R^2 of the individual regressions. The R^2 can be regarded as a measure of the extent to which an individual's forecasts of inflation (lagged, current and future) and the unemployment rate conform to a PC.[24] The mean R^2 for forecasts is nearly 50%, and greater than 70% for those in the upper quartile of the distribution. For revisions, the corresponding figures are lower, at around 34% (for the mean) and 40% (upper quartile).

Whereas Table 1 reports results for forecasts (4) and revisions (7), all the subsequent empirical results are based on the forecasts. As noted earlier, the Appendix

[21] The Spearman rank correlation r lies between -1 and 1, where 0 indicates no relationship. The rank correlation is given by:

$$r = 1 - \frac{6R}{N\left(N^2-1\right)}$$

where R is the sum of squared differences between the ranks (of the forecasters by sample size, and by the value of $\hat{\gamma}_j$). It is common to calculate the Fisher transformation,

$$F(r) = \frac{1}{2}\ln\frac{1+r}{1-r}$$

such that $z = F(r) . \sqrt{\frac{N-3}{1.06}} \sim N(0,1)$ under the null of statistical independence. As well as reporting r, we report the probability of the test statistic z being at least as large as we obtained if the null hypothesis (of a zero correlation) is true. Probabilities less than 0.025 or greater than 0.975 indicate rejections of the null in a two-sided test at the 5% level. (High probabilities suggest a negative relationship, and low probabilities a positive relationship.)

[22] In her study of perceived inflation persistence, Jain (2019) looked at forecasts from 1984:Q1 to 2010:Q1 and found 80 forecasters who had submitted enough forecasts to calculate at least ten revisions.

[23] The constant term is smaller for revisions because these measures change the expected inflation rate between two periods.

[24] With the proviso that the signs of the estimated parameters accord a PC interpretation. For example, a negative coefficient on the unemployment rate.

Table 1
Summary of hybrid PC estimates for 67 respondents.

		Constant		Lagged, β_b		Forward, β_f		UR, γ		R^2	$\beta_b + \beta_f$
Forecasts											
Mean	1%	0.307	0.188	0.266	0.986	0.637	0.986	−0.023	0.072	0.486	0.902
s.d.	5%	0.591	0.333	0.103	0.986	0.201	0.986	0.055	0.232	0.256	0.255
l.q.	10%	−0.012	0.391	0.226	1.000	0.595	0.986	−0.026	0.333	0.235	0.916
u.q.		0.326	.	0.339	.	0.723	.	0.000	.	0.703	1.009
Revisions											
		Constant		Lagged, β_b		Forward, β_f		UR, γ		R^2	$\beta_b + \beta_f$
Mean	1%	−0.010	0.014	0.272	0.899	0.621	0.928	−0.059	0.058	0.339	0.893
s.d.	5%	0.060	0.072	0.109	0.899	0.213	0.928	0.221	0.101	0.234	0.267
l.q.	10%	−0.035	0.130	0.231	0.913	0.559	0.942	−0.120	0.145	0.192	0.837
u.q.		0.023	.	0.342	.	0.736	.	0.022	.	0.400	1.036

The estimates for Forecasts are based on individual systems of 4 equations, and for Revisions on individual systems of 3 equations.
For each parameter, we present in the first column summary statistics of the cross-section distribution over j (mean, s.d., lower (l.q.) and upper (u.q.) quartiles), and in the second column, the proportion of the 67 regressions for which we reject the null hypothesis of the parameter equalling zero at three different significance levels.

discusses the relative merits of estimating individuals' PCs from forecasts versus revisions to those forecasts (between adjacent forecast origins), and concludes that the former is likely to provide more accurate estimates.

Next, we investigate participation time. As discussed in Section 3.1, participation time is a possible explanation of the heterogeneity in individuals' estimated PCs, both in terms of whether they were active earlier or later in the period, and the extant labour-market conditions.

We suppose that the perception of the PC slope may differ in the last two decades (2000:1 to 2019:4) relative to the period 1981:3 to 1999:4 and during 'tight' as opposed to normal labour markets.[25] Table 2 provides the rank correlation tests between certain features of the individuals' PC models ($\hat{\gamma}_j$ or R_j^2) and either the proportion of a respondent's forecasts made prior to 2000 (denoted *Period*) or the proportion made during tight labour markets (denoted *Hot*).

The results for *Period* suggest statistically positive correlations with both $\hat{\gamma}$ and R^2. Respondents who made a greater proportion of their forecasts during 1981–1999 are likely to have *less* steep negative slopes and higher R^2's. Hence, in recent times, the proportion of the variation in inflation believed to be explained by the PC has fallen. Still, at the same time, the responsiveness of inflation to perceived slack in the economy has increased. This finding suggests respondents' beliefs do not reflect the view that the PC has recently flattened. To shed some light on the extent of the variation explained by *Period*, we report the values of $\hat{\gamma}$ and R^2 at various points in the cross-sectional distribution of *Period*. Specifically, we report the average values of $\hat{\gamma}$ and R^2 of the individuals in the four quartiles of the distribution sorted by their *Period* scores. The average $\hat{\gamma}$ values of the 1st and 4th quartiles are −0.042 and −0.015, whereas the difference between the 2nd and 3rd quartile averages of −0.024 and −0.007 is relatively modest. The average R^2 value of the 4th-quartile respondents is roughly twice that of those in the 1st quartile: pre-2000 respondents' forecasts of inflation and the unemployment rate more closely conform to a PC relationship, albeit the unemployment rate effect is more muted.

The rank correlation between *Hot* and $\hat{\gamma}$ and R^2 is negative and statistically significant for both $\hat{\gamma}$ and R^2, suggesting the responsiveness of inflation to the unemployment rate is perceived to be greater when the unemployment rate is low (less than 5%), and that the PC explains less of the variability of inflation at these times. The table gives the averages of the estimates of $\hat{\gamma}$ of respondents in the 3rd and 4th quartiles by *Hot* as −0.058 and −0.020, compared to values of essentially zero in the first two quartiles.

We have found that the time of participation and the level of unemployment rate are both significantly related to the cross-sectional distributional of beliefs about the PC-slope, with the unemployment rate non-linearity suggests respondents expect inflation to respond more in tight labour markets. However, the finding that respondents active in more recent periods have a steeper slope is not consistent with the view that the (actual) PC has flattened over time.

Finally, we consider interaction effects and calculate the proportion of observations made by each respondent, pre-2000 and during hot labour markets (Panel C) in the table. We find the variation in $\hat{\gamma}$ and R^2 is more muted over the quartiles. Panel D considers the proportion of forecasts made in the last two decades (2000–2019) at times of hot labour markets and finds an enhanced $\hat{\gamma}$ effect — a 1st quartile average of −0.003, and a 4th-quartile value of −0.051. These two interaction effects are consistent with the results in Panels A and B: the steepness of the slope is greater if the respondent was more active in the later period and during tighter labour markets.

25 Specifically, a response to survey-t was made during a 'hot' labour market if the unemployment rate at time $t-1$ was less than or equal to 5. Just over a quarter of the periods between 1981:3 and 2019:4 satisfied this condition (28%). Setting the threshold to 5.5 did not materially alter the findings.

Table 2
Correlation between the PC slope parameter and R^2 and *Period* and *Hot*.

Spearman	p-value	Q1	Q2	Q3	Q4
[A] *Period*-Slope					
0.362	0.001	−0.042	−0.023	−0.008	−0.015
Period-R^2					
0.564	0.001	0.372	0.321	0.539	0.713
[B] *Hot*-Slope					
−0.243	0.975	−0.007	−0.008	−0.053	−0.020
Hot-R^2					
−0.606	1.000	0.739	0.513	0.410	0.312
[C] *Period* AND *HOT*-Slope					
0.288	0.010	−0.022	−0.037	−0.010	−0.019
Period AND *HOT*-R^2					
0.091	0.236	0.480	0.426	0.498	0.533
[D] 1-*Period* AND *HOT* - Slope					
−0.353	0.998	−0.003	−0.024	−0.013	−0.051
1-*Period* AND *HOT*-R^2					
−0.661	1.000	0.723	0.550	0.349	0.304

The first two columns report the Spearman rank correlation test and p-value; the entries in the Qi columns are the averages of the slope or R^2 estimates of respondents in the ith quartile by *Period* or *Hot*, or by the interaction of the two. Q1 to Q4 denotes most recent to earliest in terms of *Period*, and least to most *Hot* in terms of labour markets.

4.2. Does belief in the PC improve forecast accuracy?

We present two types of evidence to determine whether it pays to use theory-consistent expectations. The first uses individual-level data and the second repeated cross-sections.

4.2.1. Individual-level forecast accuracy analysis

A key difficulty in comparing individuals who were active survey respondents at different times is that the different times might have been characterized by very different economic conditions. Failure to control for this might camouflage effects of interest, such as whether respondents whose expectations are consistent with the PC make more (less) accurate forecasts. We control for economic conditions by dividing forecast errors at time t by the cross-sectional average accuracy (measured by RMSFE) of all forecasts made at time t (see D'Agostino et al. (2012), and Clements (2014, 2022)).

Letting $e_{i,t+h|t}$ denote the forecast error made by individual i in response to forecast survey t, for period $t+h$, we calculate the normalized forecast errors as:

$$\widetilde{e}_{i,t+h|t} = \frac{e_{i,t+h|t}}{\sqrt{\frac{1}{N_{t,h}} \sum_{j=1}^{N_{t,h}} e_{j,t+h|t}^2}} \tag{9}$$

where $N_{t.h}$ is the number of respondents to survey t, so that the denominator is the cross-section RMSFE. Then, the scaled-error MSFE for respondent i (at horizon h) is:

$$\frac{1}{n_i} \sum_{t \in N_i} \widetilde{e}_{i,t+h|t}^2 \tag{10}$$

where the summation is over all the surveys to which i responded, given by the set N_i, and n_i is the number of elements in N_i.

The actual values used to calculate forecast errors are, again, the estimates published one quarter after the reference quarter, although, as noted earlier, revisions to the CPI inflation rate were small.

Table 3
Spearman rank correlations for individual-level forecast accuracy and size of unemployment rate coefficient, or R^2.

	UR coefficient		R^2	
h	Statistic	p-value	Statistic	p-value
Correlation with scaled MSFE				
0	0.234	0.030	0.159	0.103
1	0.102	0.209	−0.111	0.811
2	0.010	0.469	−0.153	0.889
3	0.029	0.411	−0.118	0.826
4	−0.023	0.571	−0.190	0.936

The table is based on the hybrid PC.

We test whether there is a systematic relationship between belief in the PC (given by the value of $\hat{\gamma}_i$, or the R^2) and forecast accuracy across individuals. Table 3 reports the rank correlation coefficients and their p-values. There is a statistically significant correlation between forecast accuracy and $\hat{\gamma}$ for the current quarter horizon, in that a more negative value of $\hat{\gamma}$ is associated with a smaller MSFE for the current quarter forecasts. The probability of obtaining a larger rank correlation under the null is 3%, so we would reject at the 6% level in a two-sided test. At all the other horizons, there is no relationship between accuracy and $\hat{\gamma}$, nor is accuracy related to R^2 at any horizon.

The takeaway is that, across individuals, there is weak evidence that the reporting of expectations consistent with a PC is conducive to greater accuracy for the shortest horizon current-quarter forecasts.

4.2.2. Percentiles of cross sections

We supplement the respondent-level comparisons above by constructing an artificial series of forecasts, comprising at each point in time the inflation forecast of the respondent at a given percentile in the cross-sectional distribution of $\hat{\gamma}$-estimates. Here, the size of $\hat{\gamma}$ is a measure of belief in the PC. Alternatively, we could use

Table 4
Forecast Accuracy of Selected Percentiles of Cross-sections based on Respondents' Unemployment Rate Responses (measured by γ estimates). The unemployment rate estimates are obtained from the hybrid Phillips curve estimated using forecasts (as opposed to revisions).

h	Percentile				
	0.15	0.25	0.50	0.75	0.85
0	0.756	0.812	0.848	0.904	1.076
	0.892	0.957	1.000	1.065	1.268
1	0.979	0.869	0.903	0.985	1.045
	1.084	0.963	1.000	1.091	1.158
2	0.950	0.939	0.892	0.972	0.953
	1.066	1.053	1.000	1.090	1.069
3	0.941	0.904	0.929	1.035	1.069
	1.013	0.973	1.000	1.115	1.151
4	0.993	0.919	0.783	0.978	1.064
	1.268	1.173	1.000	1.248	1.358

For each forecast horizon, the first row reports the MSFEs and the second row the same divided by the MSFE of the median forecaster.

the R^2 estimates. To be clear, at time t, we consider the set of available inflation forecasts. Of the active forecasters at time t, we then look at their $\hat{\gamma}$ estimates and select the respondent at a given percentile p of the distribution of $\hat{\gamma}$-estimates. That respondent's inflation forecast becomes the percentile-p forecast. We repeat each period t. As the set of active respondents changes (with exit and entry, and temporary non-participation), so does the identity of the respondent whose inflation forecast is chosen, but we always choose the inflation forecast of the individual with the p-percentile $\hat{\gamma}$ -estimate (of the pool of active respondents). We consider percentiles from $p = 0.15$ to $p = 0.85$, with the $\hat{\gamma}$ estimates varying from large negative to large positive with p. The resulting series do not correspond to any given respondent in the SPF. Still, they could be regarded as typical of a forecaster who 'strongly believes in the PC' when $p = 0.15$ or typical of a forecaster who does not believe in a negative relationship between unemployment and inflation when $p = 0.85$.[26] The use of percentiles of cross sections allows forecast accuracy comparisons based on all 154 surveys (1981:3 to 2019:4), lessening concerns about forecasts being made during very different conditions. And we can still use scaled forecast errors based on (9) in this setting.

Results are shown in Table 4. For the current-quarter inflation forecasts, forecast accuracy at the .15 percentiles (corresponding to the larger negative unemployment rate effects) is around 10% more accurate than for the median. At the longer horizons, the median is at least as good as the lower (or upper) percentile. This supports the finding of the previous sub-section that belief in the PC only improves forecast accuracy at the shortest horizon.

4.3. Does model heterogeneity account for inflation forecast disagreement?

In this section, we consider the extent to which differences in respondents' models account for the observed disagreement in inflation forecasts. Disagreement may arise for a variety of reasons. Our interest is in whether heterogeneous beliefs about the PC relationship between activity and inflation play an important role.

We consider the hybrid PC. For respondent j, (3):

$$E_{j,t}\pi_{t+h} = \beta_{b,j}E_{j,t}\pi_{t+h-1} + \beta_{f,j}E_{j,t}\pi_{t+h+1} + \gamma_j E_{j,t}u_{t+h} \quad (11)$$

When $h = 0$, corresponding to a current-quarter forecast, the first term on the right-hand side is the respondent's estimate of $t-1$, for which the advance estimate is data. Whether we use the respondents' forecasts of $t-1$ or the advance estimates makes little difference.

Eq. (11) can be estimated for $h = 0, 1, 2$ and 3. For $h = 3$ the forward term of the right-hand side is a 4-quarter ahead forecast; the longest quarterly forecast supplied by the SPF. Estimation of the 4-variable system for respondent j gives $\left\{\hat{\beta}_{b,j}, \beta_{f,j}, \hat{\gamma}_j\right\}$. These are the econometrician's estimates of respondent j's 'model' or beliefs about how the economy operates, as revealed by the respondent's forecasts. Respondent $\hat{j}$'s model forecasts of inflation h-steps ahead are given by:

$$\widehat{E_{j,t}\pi_{t+h}} = \hat{\beta}_{b,j}E_{j,t}\pi_{t+h-1} + \hat{\beta}_{f,j}E_{j,t}\pi_{t+h+1} + \hat{\gamma}_j E_{j,t}u_{t+h} \quad (12)$$

where $E_{j,t}\pi_{t+h-1}$ and $E_{j,t}\pi_{t+h+1}$ are the $h-1$ and $h+1$-step ahead inflation forecasts, and $E_{j,t}u_{t+h}$ is the reported h-step ahead forecast of the unemployment rate (all made to the time t-survey). In (12), we suppress the constant term, which is estimated in all the models.

To determine the importance of model heterogeneity in accounting for forecast disagreement, calculate the model cross-sectional disagreement $\sigma_{M,t,h}$ as the standard deviation of $\widehat{E_{j,t}\pi_{t+h}}$ over j, and compare this with the series of cross-sectional standard deviations of the reported forecasts (for the same j), and denote this series $\sigma_{t,h}$. The averages of the cross-sectional standard deviations $\sigma_{M,t,h}$ and $\sigma_{t,h}$ across t are reported in Table 5 panel A for each horizon. Disagreement (of the reported forecasts) is around 15% lower at $h = 3$ than at $h = 0$ (0.717 compared to 0.852), but does not decline between $h = 2$ and $h = 3$, and the majority of the decline has already occurred by $h = 1$. The suggestion in Patton and Timmermann (2010) that respondents may have different views about the long-run values of variables like inflation and output growth accords with the analysis developed here of agents having different models and is a rationale for observed disagreement among forecasters remaining high as the horizon increases (and the effect of possibly heterogeneous signals lessens). The second column of Table 5 panel A records the proportion of the disagreement in reported forecasts that is consistent with agents having different models. For the current-quarter forecasts, this is nearly two thirds, rising to nearly 80% for $h = 1$ and thence to close to 90%. Beyond the current and next-quarter forecasts, heterogeneous inflation-forecasting models account for the lion's share of inflation forecast disagreement.

A possible caveat to the finding that different PC models across respondents play an important role in accounting for disagreement is the following. Suppose that the

[26] We do not consider values more in the tails than the 15th and 85th percentiles to avoid extreme values.

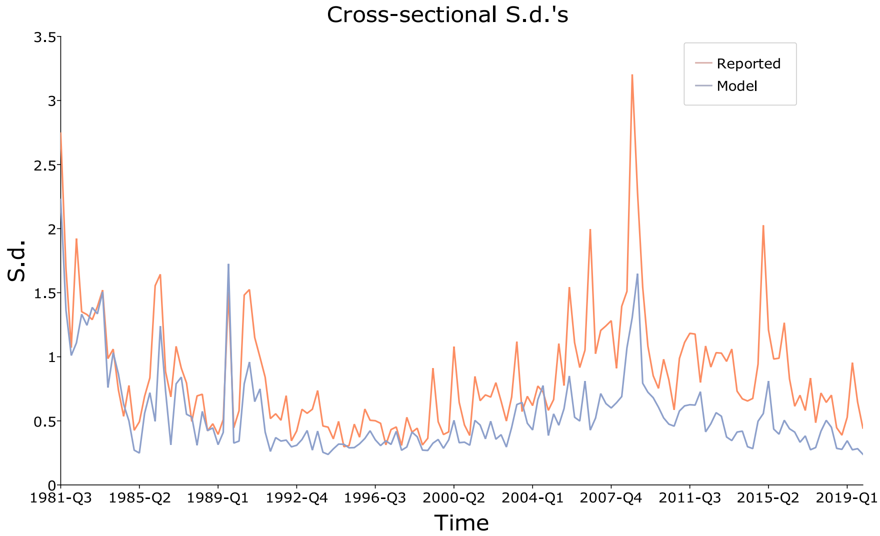

Fig. 3. Reported and hybrid-PC cross-sectional standard deviations, current-quarter forecasts.

Table 5
Forecast disagreement accounted for by model heterogeneity.

Panel A		
h	Reported	Model/Reported
0	0.852	0.646
1	0.775	0.778
2	0.710	0.880
3	0.717	0.884

Panel B						
	Reported			Model		
	LQ	Median	UQ	LQ	Median	UQ
0	1.154	1.456	1.717	0.673	0.864	1.133
1	0.646	0.841	1.178	0.598	0.735	0.949
2	0.562	0.750	0.995	0.441	0.620	0.804
3	0.482	0.673	0.939	0.402	0.576	0.809

Panel A: The column for 'Reported' is the averages (across time) of the cross-sectional standard deviations of the respondents' forecasts that is, the time-average of $\sigma_{t,h}$. The column 'Model/Reported' is the ratio of the average of the model forecast standard deviations, $\sigma_{M,t,h}$, to the time-average of $\sigma_{t,h}$.
Panel B: We report statistics of the cross-sectional distribution of the individuals' standard deviations of their Reported and Model forecasts. The statistics are the 25th percentile (LQ), the median, and the 75th percentile (UQ). The forecast standard deviation for respondent j is the square root of the sample variance of all forecasts made by that respondent.
The model underlying these figures is the hybrid PC, estimated separately for each respondent.

right-hand side variables in (12) are all relatively unimportant and that the (average) cross-sectional differences described in Table 5 panel A simply reflect different estimated constant terms (reflecting in turn different biases across forecasters). If that explanation were true, the variance of an individual's model forecasts over time would be close to zero, or at any rate, small compared to the variance of the respondent's reported forecasts. Table 5 panel B provides evidence against this. The table reports summary statistics of the cross-sectional distribution of the individuals' standard deviations of their Reported and Model forecasts. For each respondent, we calculate the forecast standard deviation of her reported forecasts and of her model forecasts and denote these $\sigma_{j,h}$ and $\sigma_{M,j,h}$. The table reports the cross-sectional medians of $\sigma_{j,h}$ and $\sigma_{M,j,h}$. These are quite different for the current-quarter forecasts — the latter is around 60% of the value of the former (0.862 compared to 1.435). However, beyond $h = 0$, the ratio is over 80%, suggesting that the model forecasts are not simply reflecting individual-level biases but different beliefs about the PC. The table also reports the lower and upper quartiles of $\sigma_{j,h}$ and $\sigma_{M,j,h}$, showing that the same is true for the quartiles as for the median.

A more important caveat is that the estimated PC unemployment rate coefficients need not be statistically significant or even negative (as shown in 1), so do not correspond to what is normally understood as a PC. In many instances, the proportion of the variation explained by the unemployment rate is small and the lagged and forward inflation terms play a more prominent role.

Finally, Figs. 3 and 4 show the evolution of $\sigma_{M,t,h}$ and $\sigma_{t,h}$ over time for a given h: for $h = 0$ in the first figure, and for $h = 4$ in the second figure. (The time axis refers to when the forecasts were made, not the quarter being forecast, although these coincide for $h = 0$). Fig. 3 shows a larger gap between the model and reported forecast disagreement series over the last 15 years or so. This is consistent with the finding of a negative correlation between $Period_j$ and R^2_j reported in Section 4.1, whereby the proportion of the variation in the inflation forecasts explained by the PC is lower in recent times.

Nevertheless, the main message for both the current-quarter forecasts and 4-quarters ahead is how closely related the model and reported-forecast disagreement are over time. Equally as apparent is the oft-observed increase in disagreement at the time of the Financial Crisis and ensuing recession (see, e.g., the classic paper on uncertainty and disagreement by Zarnowitz and Lambros (1987), and

Table 6
Summary of hybrid PC estimates for 67 respondents by equation for each horizon.

		Constant		Lagged, β_b		Forward, β_f		UR, γ		R^2	$\beta_b+\beta_f$
$h=0$ Forecasts											
Mean	1%	0.321	0.074	0.216	0.494	0.731	0.679	−0.053	0.025	0.475	0.946
s.d.	5%	1.990	0.173	0.116	0.667	0.439	0.778	0.248	0.136	0.243	0.439
l.q.	10%	−0.719	0.259	0.148	0.728	0.550	0.840	−0.102	0.173	0.270	0.809
u.q.		0.750	.	0.291	.	1.015	.	0.050	.	0.651	1.185
$h=1$ Forecasts											
Mean	1%	0.692	0.173	0.175	0.531	0.618	0.778	−0.054	0.074	0.629	0.793
s.d.	5%	1.140	0.284	0.144	0.593	0.268	0.864	0.151	0.185	0.240	0.265
l.q.	10%	0.116	0.346	0.076	0.654	0.475	0.889	−0.090	0.247	0.441	0.679
u.q.		1.110	.	0.246	.	0.796	.	0.013	.	0.821	0.983
$h=2$ Forecasts											
Mean	1%	0.543	0.148	0.279	0.679	0.544	0.753	−0.029	0.086	0.708	0.822
s.d.	5%	1.266	0.259	0.213	0.815	0.310	0.877	0.115	0.111	0.260	0.395
l.q.	10%	−0.097	0.333	0.212	0.840	0.408	0.914	−0.049	0.198	0.547	0.786
u.q.		0.598	.	0.407	.	0.714	.	0.012	.	0.910	0.997
$h=3$ Forecasts											
Mean	1%	0.470	0.111	0.336	0.630	0.458	0.741	−0.004	0.000	0.713	0.794
s.d.	5%	0.856	0.235	0.234	0.778	0.298	0.852	0.063	0.049	0.296	0.387
l.q.	10%	−0.004	0.296	0.211	0.827	0.356	0.877	−0.027	0.111	0.446	0.705
u.q.		0.618	.	0.477	.	0.621	.	0.019	.	0.946	0.985

The estimates are for forecasts estimated separately for each horizon.
For each parameter, we present in the first column summary statistics of the cross-section distribution over j (mean, s.d., lower (l.q.) and upper (u.q.) quartiles), and in the second column, the proportion of the 67 regressions for which we reject the null hypothesis of the parameter equalling zero at three different significance levels.

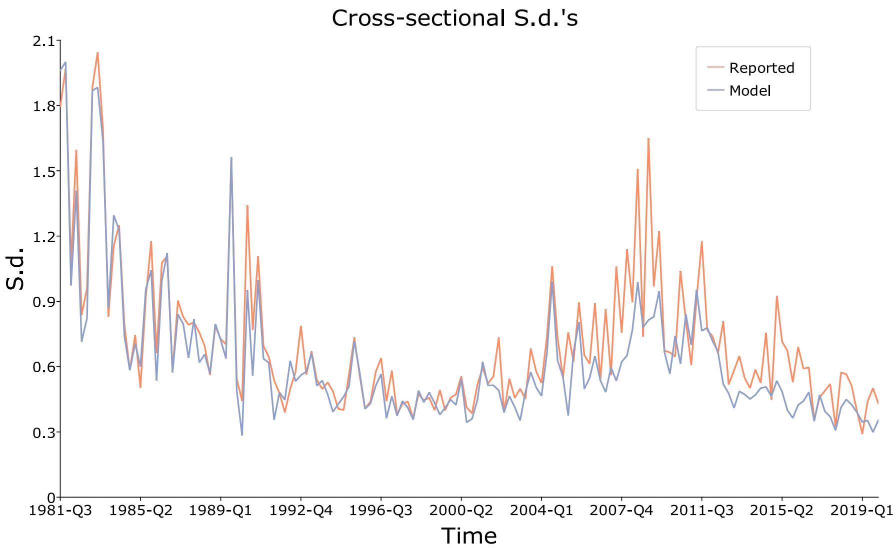

Fig. 4. Reported and hybrid-PC Cross-sectional standard deviations, 4-quarter forecasts.

more recently Rich and Tracy (2010) and Bachmann et al. (2013)).

5. Robustness checks

In this section, we assess whether our findings are robust to a number of modelling choices we have made. Specifically, (i) the assumption that the parameters of the PC are the same across horizons, (ii) potential mismeasurement of the PC activity variable, (iii) the omission of relevant variables such as oil prices, and (iv) the use of the GDP deflator in place of the CPI deflator.

5.1. Estimating the PC for each horizon

We have restricted the estimates of the PC to be the same for each horizon. Estimating the PC separately for each horizon does not change the results in certain key respects. Table 6 shows that the cross-sectional average estimate of the effect of the unemployment rate is smaller (closer to zero) for the $h = 3$ forecasts but that even at the shortest horizons, the coefficient is only statistically significantly different from zero for less than one in five respondents. The effect of forward inflation declines, and

Table 7
Summary of hybrid PC estimates for 67 respondents by equation for each horizon, using the unemployment rate Gap.

		Constant		Lagged, β_b		Forward, β_f		UR, γ		R^2	$\beta_b+\beta_f$
$h=0$ Forecasts											
Mean	1%	0.058	0.123	0.216	0.494	0.726	0.679	−0.062	0.037	0.475	0.942
s.d.	5%	1.087	0.247	0.116	0.654	0.442	0.778	0.271	0.123	0.242	0.443
l.q.	10%	−0.537	0.346	0.152	0.741	0.534	0.827	−0.151	0.160	0.272	0.797
u.q.		0.481	.	0.290	.	1.012	.	0.065	.	0.650	1.181
$h=1$ Forecasts											
Mean	1%	0.442	0.185	0.172	0.531	0.613	0.778	−0.066	0.111	0.631	0.785
s.d.	5%	0.621	0.272	0.144	0.580	0.271	0.852	0.177	0.173	0.238	0.273
l.q.	10%	0.004	0.333	0.073	0.642	0.459	0.877	−0.108	0.296	0.461	0.676
u.q.		0.657	.	0.241	.	0.797	.	0.010	.	0.822	0.987
$h=2$ Forecasts											
Mean	1%	0.398	0.198	0.278	0.654	0.544	0.753	−0.039	0.074	0.708	0.822
s.d.	5%	0.913	0.296	0.211	0.815	0.307	0.877	0.134	0.123	0.259	0.391
l.q.	10%	−0.041	0.358	0.206	0.852	0.401	0.914	−0.059	0.173	0.547	0.791
u.q.		0.504	.	0.410	.	0.706	.	0.014	.	0.910	1.002
$h=3$ Forecasts											
Mean	1%	0.456	0.198	0.336	0.617	0.459	0.728	−0.010	0.000	0.713	0.795
s.d.	1%	0.827	0.284	0.234	0.778	0.290	0.852	0.076	0.074	0.296	0.376
l.q.	5%	0.015	0.296	0.214	0.827	0.359	0.889	−0.039	0.111	0.445	0.707
u.q.	10%	0.592	.	0.478	.	0.630	.	0.021	.	0.946	0.988

The unemployment rate gap is the forecast of the quarterly unemployment rate minus the NAIRU for time $t-1$ when the survey is dated t. The estimates are for Forecasts estimated separately for each horizon.
For each parameter, we present in the first column summary statistics of the cross-section distribution over j (mean, s.d., lower (l.q.) and upper (u.q.) quartiles), and in the second column, the proportion of the 67 regressions for which we reject the null hypothesis of the parameter equalling zero at three different significance levels.

that of the backward term increases as we increase the horizon h, as does the average R^2. The finding that the unemployment rate only plays a role for a small fraction of respondents holds up when the PC coefficients vary across horizons.

5.2. Measurement of the activity variable

In this section, we check whether our results are robust to using the 'gap' rather than the unemployment rate as the activity variable. Hitherto, we have used the unemployment rate forecasts, although the forecasters might expect inflation to respond to the difference between the unemployment rate and the natural rate. Forecasts of the natural rate were only collected by the SPF from 1996:Q3 onwards and only for the third quarters of the year. Hence, the use of forecasts of the natural rate would entail a loss of 15 of the 39 years of data, and require a method for dealing with the missing values for all but the third quarters. Instead, we use the estimate of the natural rate at the time of the survey (specifically, of the quarter immediately before the survey quarter) to calculate the gap.[27]

As Table 7 shows, the evidence of a negative slope parameter is strengthened to a small extent, but for the most part, the results are qualitatively unchanged (compare Table 7 with Table 6). We interpret this as suggesting our focus on the unemployment rate as the activity variable (rather than the gap) is largely inconsequential.

Fig. 5 shows the unemployment rate gap. Over our sample period, the natural rate has not changed much, increasing a little around 2010/11 after a gradual trend decline, and the profiles of the unemployment rate and gap match closely.

5.3. The omission of oil prices

Coibion and Gorodnichenko (2015b) stress the role of oil prices in driving up household/firm inflation expectations relative to those of professionals after the 2008–9 recession. More generally, there has been a growing recognition that inflation 'globalisation' suggests external sources of inflation may play a role, and it is reasonable to suppose professionals might consider such factors.[28] The upshot is that commodity prices, including oil, may affect expectations. Ideally, we would like to include forecasts of oil prices and revisions to these forecasts in PC specifications. However, the SPF does not elicit forecasts of oil prices, and instead, we use lagged changes of oil prices in our PC specifications. Adding the percentage change in the Spot Crude Oil Price at time $t-1$ (relative to survey quarter t) had little discernible effect on the unemployment rate coefficients, and these results are not reported. Relative to Table 6, the coefficient on the backward inflation term approximately halved for the current quarter forecasts, but besides this, the only noticeable effect was a small increase in the average R^2 values.[29]

27 Formally, one could think of this as a 'no-change' forecast of the gap in future quarters.

28 The globalisation of inflation refers to the finding that a common factor accounts for nearly 70% of the variance of inflation in 22 OECD countries — Ciccarelli and Mojon (2010).

29 The lagged change in the oil price and the $t-1$ actual (or forecast) $t-1$ inflation rate is naturally correlated.

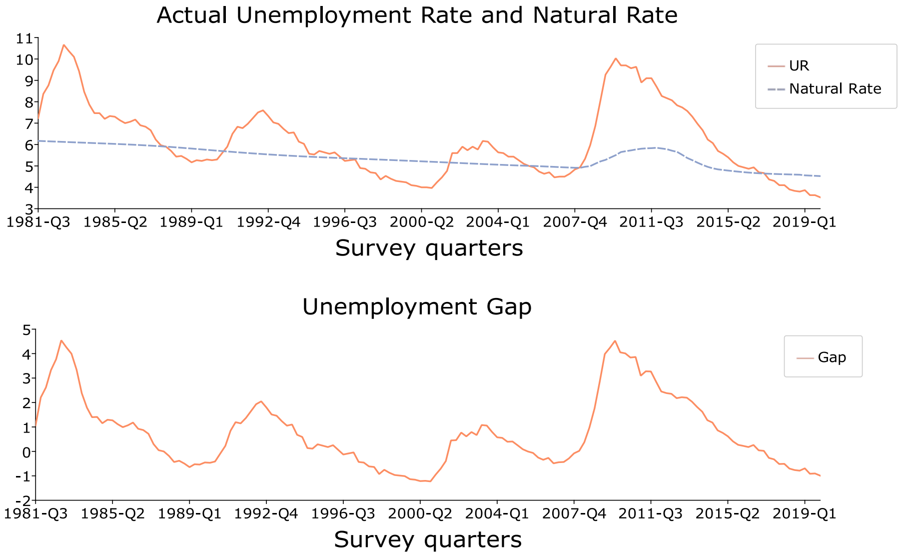

Fig. 5. Actual and natural rates of unemployment.

Table 8
Summary of hybrid PC estimates, using the forecasts of the GDP deflator in place of the CPI forecasts.

Forecasts											
		Constant		Lagged, β_b		Forward, β_f		UR, γ		R^2	$\beta_b+\beta_f$
Mean	1%	0.619	0.449	0.190	0.826	0.528	0.855	−0.001	0.333	0.455	0.719
s.d.	5%	0.906	0.507	0.118	0.884	0.291	0.942	0.086	0.449	0.243	0.316
l.q.	10%	0.017	0.565	0.109	0.913	0.348	0.957	−0.045	0.551	0.236	0.500
u.q.		0.944	.	0.273	.	0.733	.	0.011	.	0.654	0.950
Revisions											
		Constant		Lagged, β_b		Forward, β_f		UR, γ		R^2	$\beta_b+\beta_f$
Mean	1%	−0.050	0.087	0.092	0.478	0.421	0.812	−0.029	0.072	0.225	0.513
s.d.	5%	0.116	0.203	0.123	0.594	0.332	0.826	0.255	0.159	0.201	0.392
l.q.	10%	−0.072	0.304	0.041	0.623	0.213	0.841	−0.131	0.217	0.070	0.304
u.q.		−0.005	.	0.151	.	0.632	.	0.052	.	0.295	0.800

The estimates for forecasts are based on individual systems of four equations and for revisions on individual systems of three equations.
For each parameter, we present in the first column summary statistics of the cross-section distribution over j (mean, s.d., lower (l.q.) and upper (u.q.) quartiles), and in the second column, the proportion of the 67 regressions for which we reject the null hypothesis of the parameter equalling zero at three different significance levels.
The table is comparable to Table 1: Table 1 uses CPI inflation, and this table uses the GDP deflator. Because the forecasters do not report the previous period's GDP deflator inflation rate, we use the CPI inflation forecasts for the $h-1$ horizon instead.

5.4. An alternative inflation measure

Our estimates up to this point have been based on the CPI inflation forecasts. The CPI has recently been used by Casey (2020), amongst others, to study whether expectations conform to a Phillips curve. However, other inflation expectations series are available. Table 8 reproduces Table 1 but uses the forecasts of the GDP deflator as the dependent variable, and as the lag and lead of the inflation rate in the Phillips curve. We use the same sample period as for CPI inflation, that is, the surveys from 1981 to 2019.

The results for the GDP deflator match those for CPI inflation in that respondents continue to attach greater weight to the forward-looking inflation term than the lagged rate of inflation. However, there is less persistence. The sum of the β_f and β_b coefficients is lower. The average unemployment coefficients are smaller (less negative). Still, they are predominantly negative judging by the interquartile range (−0.045 to 0.011 for Forecasts) and statistically significant for roughly twice as many respondents using forecasts (at the 5% level).

We conclude that the results using the GDP deflator are broadly in line with those obtained using the deflator in that there is little evidence of a negative relationship between inflation and activity for a majority of respondents (in a linear model).

6. Conclusions

There is considerable heterogeneity in survey participants' perceptions of the Phillips curve relationship between inflation and the unemployment rate. We estimated hybrid Phillips curves for each respondent (who made more than a minimum number of forecasts) and found that the coefficient on the unemployment rate was statistically significantly negative for only around a quarter of professional forecasters. There was, however, less heterogeneity regarding the backward and forward-looking inflation coefficients. A possible explanation is that belief in the relationship has not been constant — for example, the view that the PC has flattened over time, perhaps because of monetary policy (see McLeay and Tenreyro (2019)). Or because the relationship between inflation and the unemployment rate is non-linear. There are not always enough forecasts for non-constancy over time and non-linearities at the individual level. Our approach is to adopt a simple linear PC for each forecaster and then consider whether the cross-sectional distribution of slope estimates is correlated with when the respondent was an active participant or the state of the labour market when the respondent was active. Such correlations would suggest that beliefs are sensitive to these factors.

We find that there is a systematic relationship between when respondents were active over the 1981–2019 period, and both the R^2's and estimated unemployment rate coefficients of their PC regressions. In more recent times the proportion of the variation in inflation that can be attributed to the PC is lower, but at the same time, the responsiveness of inflation to perceived slack in the economy has increased. We also find that respondents who were more active at times of tight labour markets, on average, have steeper (negative) slopes, suggesting perceptions are consistent with a non-linear PC.

We find that a belief in the PC is associated with a greater forecast accuracy, but only at the shortest horizon.

We find that differences in respondents' PC models can account for a large share of the observed disagreement in reported inflation forecasts at all but the shortest forecast horizon. Thus, our explanation of the persistence of forecast disagreement beyond the shortest horizons stresses heterogeneous beliefs about the PC relationship between activity and inflation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ijforecast.2023.11.004.

References

Adam, K., & Padula, M. (2011). Inflation dynamics and subjective expectations in the United States. *Economic Inquiry*, *49*(1), 13–25.

Aguiar, M., & Gopinath, G. (2007). Emerging market business cycles: The cycle is the trend. *Journal of Political Economy*, *115*, 69–102.

Albuquerque, B., & Baumann, U. (2017). Will US inflation awake from the dead? The role of slack and non-linearities in the Phillips curve. *Journal of Policy Modeling*, *39*(2), 247–271.

Atkeson, A., & Ohanian, L. (2001). Are phillips curves useful for forecasting inflation? *Federal Reserve Bank of Minneapolis Quarterly Review*, *25*(1), 2–11.

Bachmann, R., Elstner, S., & Sims, E. R. (2013). Uncertainty and economic activity: Evidence from business survey data. *American Economic Journal: Macroeconomics*, *5*(2), 217–249.

Bachmann, R., Topa, G., & Klaauw, W. (2022). *Handbook of economic expectations*. Elsevier: Academic Press.

Backus, D. K., & Smith, G. W. (1993). Consumption and real exchange rates in dynamic economies with non-traded goods. *Journal of International Economics*, *35*(3–4), 297–316.

Ball, L., & Mazumder, S. (2020). The nonpuzzling behavior of median inflation. In G. Castex, J. Galí, & D. Saravia (Eds.), *Volume 27 of central banking, analysis, and economic policies book series*, *Changing inflation dynamics, evolving monetary policy* (pp. 49–70). Central Bank of Chile, Chapter 3.

Barnichon, R., & Mesters, G. (2020). Identifying modern macro equations with old shocks. *Quarterly Journal of Economics*, *135*(4), 2255–2298.

Bluedorn, J. C., & Leigh, D. (2018). *Is the cycle the trend? evidence from the views of international forecasters*: *Working paper 18/163*, International Monetary Fund.

Bryan, M., & Cecchetti, S. (1993). *Measuring core inflation*: *Technical report*, Federal Reserve Bank of Cleveland, Working Paper no. 93-04.

Calvo, G. A. (1983). Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics*, *12*, 383–398.

Carvalho, C., & Nechio, F. (2014). Do people understand monetary policy? *Journal of Monetary Economics*, *66*(C), 108–123.

Casey, E. (2020). Do macroeconomic forecasters use macroeconomics to forecast? *International Journal of Forecasting*, *36*(4), 1439–1453.

Chan, J. C., Clark, T. E., & Koop, G. (2018). New model of inflation, trend inflation, and long run inflation expectations. *Journal of Money, Credit and Banking*, *50*(1), 5–53.

Ciccarelli, M., & Mojon, B. (2010). Global inflation. *The Review of Economics and Statistics*, *92*(3), 524–535.

Clements, M. P. (2014). Forecast uncertainty - ex ante and ex post: US inflation and output growth. *Journal of Business & Economic Statistics*, *32*(2), 206–216. http://dx.doi.org/10.1080/07350015.2013.859618.

Clements, M. P. (2020). Individual forecaster perceptions of the persistence of shocks to GDP. *Journal of Applied Econometrics*, *37*, 640–656.

Clements, M. P. (2022). Forecaster efficiency, accuracy and disagreement: Evidence using individual-level survey data. *Journal of Money, Credit and Banking*, *54*(2–3), 537–567.

Clements, M. P., & Hendry, D. F. (2002). Explaining forecast failure in macroeconomics. In M. P. Clements, & D. F. Hendry (Eds.), *A companion to economic forecasting* (pp. 539–571). Blackwells: Oxford.

Clements, M. P., Rich, R., & Tracy, J. (2022). Surveys of professionals, chapter 3. In Ruediger Bachmann, Giorgio Topa, & Wilbert Klaauw (Eds.), *Handbook of economic expectations* (pp. 71–106). Elsevier: Academic Press.

Coibion, O., & Gorodnichenko, Y. (2012). What can survey forecasts tell us about information rigidities? *Journal of Political Economy*, *120*(1), 116–159.

Coibion, O., & Gorodnichenko, Y. (2015a). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review*, *105*(8), 2644–2678.

Coibion, O., & Gorodnichenko, Y. (2015b). 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., & Kamdar, R. (2018). The formation of expectations, inflation, and the Phillips curve. *Journal of Economic Literature*, *56*(4), 1447–1491.

Croushore, D. (1993). *Introducing: the survey of professional forecasters* (pp. 3–15). Federal Reserve Bank of Philadelphia Business Review, November.

D'Agostino, A., McQuinn, K., & Whelan, K. (2012). Are some forecasters really better than others? *Journal of Money, Credit and Banking*, *44*(4), 715–732.

Devereux, M. B., Smith, G. W., & Yetman, J. (2012). Consumption and real exchange rates in professional forecasts. *Journal of International Economics*, *86*(1), 33–42.

Dräger, L., Lamla, M. J., & Pfajfar, D. (2016). Are survey expectations theory-consistent? The role of central bank communication and news. *European Economic Review*, *85*(C), 84–111.

Dräger, L., & Nghiem, G. (2021). Are consumers' spending decisions in line with a Euler equation? *The Review of Economics and Statistics*, *103*(3), 580–596.

Fendel, R., Lis, E. M., & Rülke, J.-C. (2011). Do professional forecasters believe in the phillips curve? evidence from the G7 countries. *Journal of Forecasting*, *30*(2), 268–287.

Fisher, J. D. M., Melosi, L., & Rast, S. (2023). *Long-run inflation expectations*: *Technical report*, FRB Chicago.

Friedman, M. (1968). The role of monetary policy. *American Economic Review*, *58*(1), 1–17.

Gali, J., & Gertler, M. (1999). Inflation dynamics: A structural econometric analysis. *Journal of Monetary Economics*, *44*, 195–222.

Hooper, P., Mishkin, F. S., & Sufi, A. (2019). *Prospects for inflation in a high pressure economy: is the phillips curve dead or is it just hibernating?*: *Working paper 25792*, National Bureau of Economic Research.

Jain, M. (2019). Perceived inflation persistence. *Journal of Business & Economic Statistics*, *37*(1), 110–120.

Keynes, J. M. (1936). *The general theory of employment*. London, Mcmillan: Interest and Money.

Kontny, M. (2019). *Phillips curves in noisy information forecasts*: *Technical report*, Germany: Department of Economics, University Hohenheim.

Krane, S. D. (2011). Professional forecasters' view of permanent and transitory shocks to GDP. *American Economic Journal: Macroeconomics*, *3*(1), 184–211.

Lundbergh, S., Teräsvirta, T., & van Dijk, D. (2003). Time-varying smooth transition autoregressive models. *Journal of Business & Economic Statistics*, *21*, 104–121.

Malmendier, U., & Nagel, S. (2016). Learning from inflation experiences. *Quarterly Journal of Economics*, *131*(1), 53–87.

Mankiw, N. G., & Reis, R. (2002). Sticky information versus sticky prices: A proposal to replace the new Keynesian Phillips curve. *Quarterly Journal of Economics*, *117*, 1295–1328.

Mankiw, N. G., Reis, R., & Wolfers, J. (2003). *Disagreement about inflation expectations*. Cambridge MA: mimeo, National Bureau of Economic Research.

Mavroeidis, S., Plagborg-Møller, M., & Stock, J. H. (2014). Empirical evidence on inflation expectations in the new Keynesian Phillips curve. *Journal of Economic Literature*, *52*(1), 124–188.

McCallum, B. T. (1976). Rational expectations and the natural rate hypothesis: Some consistent estimates. *Econometrica*, *44*(1), 43–52.

McLeay, M., & Tenreyro, S. (2019). Optimal inflation and the identification of the Phillips curve. In *NBER chapters*, *NBER macroeconomics annual 2019, volume 34* (pp. 199–255). National Bureau of Economic Research, Inc.

Patton, A. J., & Timmermann, A. (2010). Why do forecasters disagree? Lessons from the term structure of cross-sectional dispersion. *Journal of Monetary Economics*, *57*(7), 803–820.

Phillips, A. W. H. (1958). The relation between unemployment and the rate of change of money wage rates in the United Kingdom 1861–1957. *Economica*, *25*, 283–299.

Rich, R., & Tracy, J. (2010). The relationships among expected inflation, disagreement, and uncertainty: Evidence from matched point and density forecasts. *Review of Economics and Statistics*, *92*(1), 200–207.

Roberts, J. M. (1995). New Keynesian economics and the Phillips curve. *Journal of Money, Credit and Banking*, *27*(4), 975–984.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, *50*, 665–690.

Stark, T. (2013). *SPF panelists' forecasting methods: A note of the aggregate results of a november 2009 special survey*. Research Department, Federal Reserve Bank of Philadelphia, Available at: http://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters/.

Stock, J. H. (2011). Discussion of Ball and Mazumder, inflation dynamics and the great recession. In *Brookings panel on economic activity, spring 2011*, *Brookings papers on economic activity* (pp. 387–402).

Stock, J. H., & Watson, M. W. (1999). Forecasting inflation. *Journal of Monetary Economics*, *44*, 293–335.

Stock, J. H., & Watson, M. W. (2007). Why has U.S. inflation become harder to forecast? *Journal of Money, Credit and Banking. Supplement to*, *39*, 3–33.

Woodford, M. (2002). Imperfect common knowledge and the effects of monetary policy. In P. Aghion, R. Frydman, J. Stiglitz, & M. Woodford (Eds.), *Knowledge, information, and expectations in modern macroeconomics: in honor of edmund phelps* (pp. 25–58). Princeton University Press.

Zarnowitz, V., & Lambros, L. A. (1987). Consensus and uncertainty in economic prediction. *Journal of Political Economy*, *95*(3), 591–621.