Heterogeneous beliefs and the Phillips curve*

Roland Meeks† and Francesca Monti‡

November 28, 2021

Abstract

Heterogeneous beliefs modify the New Keynesian Phillips curve by introducing additional terms that depend on cross-section distributions of inflation expectations. We establish a set of novel empirical facts showing that time variation in distributions of heterogeneous survey expectations may be summarized by three factors we call disagreement, skew, and shape. We then estimate models of inflation that allow for heterogeneity via functional principal component regression, and establish the importance of heterogeneous beliefs for inflation dynamics. Our findings are robust to lagged inflation, trend inflation, and supply factors and hold in similar form across two major economies. The models remain stable after the onset of the coronavirus pandemic.

Keywords:
Inflation dynamics; New Keynesian Phillips curve; Survey expectations; Functional principal components; Functional regression

*We are indebted to John Barrdear, Andy Filardo, Ferre De Graeve, Marco Del Negro, Paul Hubert, Mike Kiley, Emmanuel Moench, Giovanni Ricco, Paolo Surico, Mark Watson, and especially to Ricardo Reis for helpful discussions. We also thank conference and workshop participants at the Bank of England, the European Central Bank, Glasgow Business School, Lancaster Business School, London Business School, the Oxford-Indiana macroeconomic policy conference, the Sciences-Po/OFCE workshop on empirical monetary economics, the FRB St. Louis Applied Time Series Econometrics workshop, the Royal Economic Society annual conference, and Warwick Business School. The views expressed in this paper are those of the authors, and should not be taken to represent the views of the IMF, its Executive Board, or IMF management.

†Corresponding author. International Monetary Fund and CAMA. Email: `rmeeks@imf.org`

‡King's College London and UC Louvain. Email: `francesca.monti@kcl.ac.uk`

1 Introduction

The idea that expectations are among the principal influences on inflation has a long and distinguished pedigree (Friedman, 1968). The aggregate supply relationship, or Phillips curve, that underpins the behavior of inflation in widely-used New Keynesian macroeconomic models, micro-founds this forward-looking property.[1] A cornerstone of the New Keynesian description of inflation is that price setters form their beliefs using complete information and perfect knowledge of the structure of the economic environment (full information rational expectations, or FIRE).[2] But a substantial body of recent work that examines direct observations on beliefs elicited from surveys has found extensive evidence that contradicts these assumptions. Whereas empirical estimates of the Phillips curve once rested heavily upon the assumption of FIRE, work that instead uses survey forecasts has enjoyed notable successes, helping to resolve otherwise puzzling shortcomings of the New Keynesian Phillips curve, to circumvent the econometric traps that can plague FIRE estimation, and to improve inflation forecasts; see amongst many others Roberts (1995), Ang, Bekaert, and Wei (2007), Brissimis and Magginas (2008), Faust and Wright (2013), Mavroeidis, Plagborg-Møller, and Stock (2014), Binder (2015), Coibion and Gorodnichenko (2015), and Coibion, Gorodnichenko, and Kamdar (2018).[3]

Despite the advantages that survey forecasts bring to the study of inflation, there are complications. Individuals' forecasts often differ substantially from one another. It is widely appreciated that when beliefs are heterogeneous in this way, a standard Phillips curve that involves aggregates alone may lack proper microfoundations. Nevertheless, empirical studies generally proceed with a consensus (or average) forecast—an aggregate—standing in for the forward-looking terms, even though the required conditions are highly restrictive, and challenging to test. Perhaps more prosaically, even determining the consensus amongst survey respondents is often less than straightforward. To illustrate, during 2008 and 2009, the mean year-ahead forecast for US inflation in the Michigan survey never fell below 2%. But the distribution of forecasts, displayed in Fig. 1, indicates that beliefs were clustered in as many as three distinct modes, the highest of which was at 0% for over a year. Many other respondents saw inflation running at 10% or more.[4] Yet researchers seldom report how their estimates are affected by alternative measures of 'the' consensus expectation.

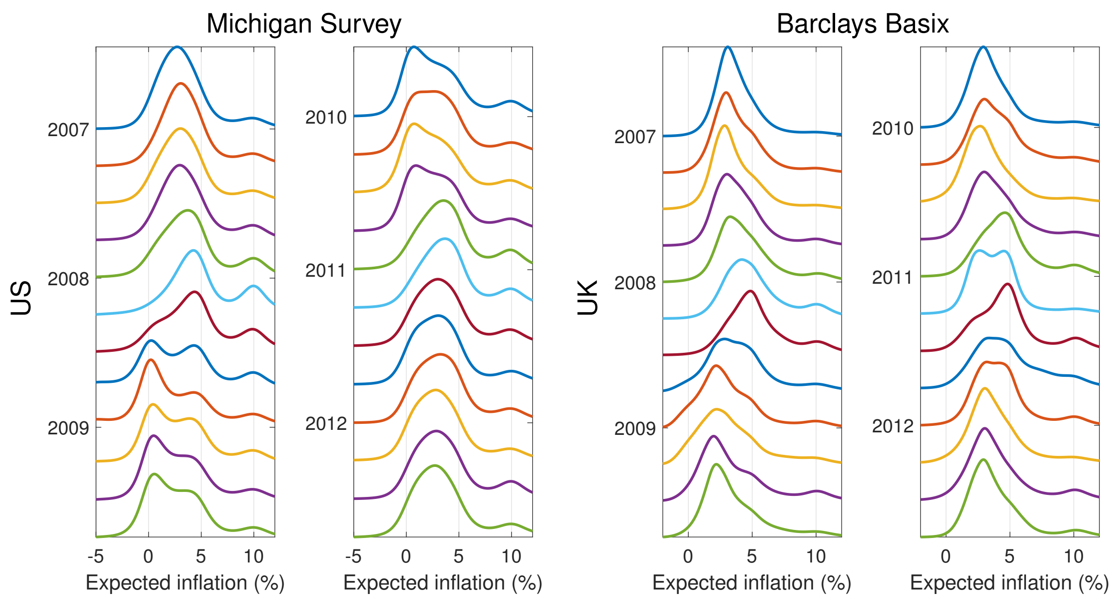
[1] A specification for the Phillips curve in which inflation depends on expected inflation can be underpinned by a variety of underlying microfoundations, including time- and state-dependent pricing frictions (see Coibion, Gorodnichenko, and Kamdar, 2018, Tab. 5).

[2] We use the terms 'beliefs' (about the future), 'forecasts', and 'expectations' interchangeably in this text. For an application of GMM to the FIRE New Keynesian Phillips curve, see Galí and Gertler (1999).

[3] The referenced econometric problems relate mainly to the weakness of the available instruments for future inflation in GMM estimation approaches. The advantages that the use of survey data bring compared to alternative estimation procedures led Mavroeidis, Plagborg-Møller, and Stock (2014, p. 151) to describe the survey approach as having gained a 'commanding' presence in the literature.

[4] The prevalence of forecasts of 0%, 5% and 10% inflation are not artifacts, but a result of known biases towards reporting round numbers when respondents become uncertain (Binder, 2017). Underlying the distributions are thousands of individual observations per quarter, making it unlikely that such features would be 'averaged away' in ever-larger samples.

Figure 1. Cross-section distributions of inflation forecasts during the financial crisis and its aftermath



Note: Panels show time series of distributions of individual survey respondents' year-ahead point forecasts from each of the named surveys. Dates reflect when forecasts were made. Details of the density estimation method may be found in Part II of the Supplementary Material.

In this paper we aim to make progress in understanding the heterogeneity in expectations, and its relevance for inflation dynamics, along several related fronts. First, we introduce a novel approach to structuring the information present in distributions of survey point forecasts, such as those in Fig. 1. Following Kneip and Utikal (2001), we apply functional principal component analysis to time series of distributions, and establish a novel set of stylized facts that add to an extensive literature documenting time variation in the cross-section dispersion of expectations—known as 'disagreement' following the early work of Mankiw, Reis, and Wolfers (2003).[5] We show that a handful of factors, which correlate with disagreement, skew, and a third factor we call 'shape', can jointly characterize much of the variation present in beliefs.[6] Although some studies have looked beyond summary measures of disagreement to complete distributions of forecasts (Filardo and Genberg, 2010; Pfajfar and Santoro, 2010), none that we know of have characterized the structure of those distributions as we do here.

Second, our paper proposes a new econometric approach to using survey expectations in heterogeneous beliefs models of inflation. The specification we propose is motivated by a version of the New Keynesian model that is grounded in the same sticky-price microfoundations encountered in the literature, except in respect of its assumptions on expectations. We avoid

[5]Disagreement can persist even in the long-run (Patton and Timmermann, 2010; Andrade, Crump, Eusepi, and Moench, 2016), and is present even in surveys of sophisticated agents such as professional forecasters (Andrade and Le Bihan, 2013). Information frictions have played a prominent role in explaining these disagreement, see Coibion and Gorodnichenko (2012).

[6]As explained later, the functional factors that we identify in our descriptive statistical analysis are used to estimate the generalized Phillips curve as a functional linear model via principal component regression (Reiss and Ogden, 2007).

imposing a priori conditions on beliefs that would permit us to replace the usual rational expectations term with the cross-section average forecast (for example, those made by Adam and Padula, 2010). The result is a modified Phillips curve which contains an additional term that depends on the distribution of beliefs held by individual firms relative to the average. We propose a flexible empirical approach that can capture the effects of heterogeneous beliefs on inflation, making prior aggregation judgements unnecessary, but that allows the standard model to emerge as a special case. To preview, consider a Phillips curve in which the scalar index $\bar{\pi}^e_{t,h}$ summarizes the state of h-step ahead inflation expectations:

$$\pi_t = \alpha(u_t - u_t^*) + \bar{\pi}^e_{t,h} + \varepsilon_t \quad \text{where} \quad \bar{\pi}^e_{t,h} = \int \gamma(\pi^e) \mathrm{d}\mathsf{P}_{t,h}(\pi^e) \tag{1}$$

Here we let π_t denote inflation, u_t be a measure of business cycle activity with a star denoting its natural rate, and $\mathsf{P}_{t,h}$ be the probability distribution of h-step ahead point forecasts π^e. We think of γ as an *aggregator function*, as it enters Eq. (1) under the integral and so determines how the distribution of beliefs influences price-setting behaviour. Conventional expectations averaging is recovered when $\gamma(x) = \beta x$, so we refer to that standard case as 'linear aggregation'. However, rather than impose linear aggregation prior to estimation, we will allow the γ function to appear along with the other parameters to be estimated in the model, up-weighting parts of the distribution that correlate strongly with inflation, and down-weighting parts that don't.

Third, we demonstrate that the payoff to our new approach is the discovery of an enhanced role for expectations in the inflation process. To our knowledge, the effects that we report have not been previously documented. We estimate the expectations-augmented Phillips curve using complete sets of household inflation forecasts reported in the US Michigan survey, and in a newly-collated UK household survey. We show that in both regions, signals about future inflation contained in the distribution of beliefs—not just its mean—affect current inflation. The effects can be quantitatively substantial, particularly in periods of heightened uncertainty concerning the conduct of monetary policy. Our tests of the hybrid (forward- and backward-looking) Phillips curve show that fully accounting for expectations entirely eliminates lag terms in inflation. Moreover, after accounting for long-run inflation trends, estimates from our FLM retain a very strong role for near-term expectations, even in periods when monetary policy is 'well run', in contrast to Cecchetti, Feroli, Hooper, Kashyap, and Schoenholtz (2017). Our findings are robust to lagged inflation, lagged consensus forecasts, trend inflation, and supply factors.

Lastly, our paper provides a novel application of the techniques of functional data analysis to a problem in macroeconomics (Ramsay and Silverman, 2005; Horváth and Kokoszka, 2012). The association between a scalar quantity (inflation) and a functional quantity (the distribution of expectations) makes Eq. (1) an example of a functional linear model (FLM). Functional data analysis (FDA) deals with infinite-dimensional random variables, and is particularly suited to the analysis of big data such as the large sets of survey responses studied here (Tsay, 2016).

Previous applications of FDA in econometrics include the work on yield curve forecasting in Bowsher and Meeks (2008), the model of relative price dispersion and inflation in Chaudhuri, Kim, and Shin (2016), and the investigation of cross-market dependence in stock returns in Park and Qian (2012).

Roadmap

The rest of this paper is organized as follows. Section 2 summarizes the main sources of variation in the time series of distributions of survey data using functional principal component analysis. Section 3 sets out our heterogeneous beliefs Phillips curve model, and the econometric approach we adopt to estimate the effects of heterogeneity on inflation. Section 4 contains our headline results, with separate treatment of the US and UK Phillips curves. The economic implications of the heterogenous beliefs model including those on the hybrid Phillips curve, and those on modeling the gap between inflation and its trend also feature there. We further discuss regression on distributional moments, and show that this alternative approach is inferior to the FLM. In Section 5 we explore the stability of our estimated models to the shock caused by the coronavirus (COVID-19) pandemic. Finally, Section 6 offers concluding comments.

2 Structure in the distribution of expectations

In this paper, we study inflation expectations in the United States and the United Kingdom, two countries for which long-running household inflation surveys exist.[7] This section investigates the distributional properties of hundreds of thousands of survey responses, reported over several decades. Since much has been written about forecast disagreement—the dispersion of individuals' subjective beliefs—in the context of inflation surveys, one of our tasks will be to assess the extent to which that attention is warranted, and to establish what else the data have to say. In what follows we confirm that time variation in disagreement is, on average, an important source of belief dynamics, but also that: (a) it is not always the principal factor; (b) several additional belief factors also matter, on average; and (c) the relative importance of disagreement, compared to other factors, is itself time dependent.

2.1 Data sources

Our analysis uses individual point forecasts recorded in two household surveys of inflation expectations. For the US we have the Michigan Survey of Consumer Attitudes (MSC), and for the UK the Barclays survey of inflation expectations (Basix). To the best of our knowledge, we are the first to make research use of the full Basix data set. The surveys ask similar questions about 'prices in general' or 'inflation', without specifying a particular measure. Each asks respondents to report their expectation for inflation over the following year, and their expectations for at

[7] In the supplementary material, we detail much of the same analysis for professional forecasters of US inflation, and record noteworthy differences as they arise below. Unfortunately, firm surveys of comparable length are not generally available.

least one other horizon.[8] Quarterly data is available spanning a period from the late 1970s (US) or mid-1980s (UK). A summary of the main features survey data used in this study is given in Tab. I.1 of the Supplementary Material.

2.2 Estimating distributions of survey forecasts

The first step in our analysis to transform the discrete cross-section of point expectations reported by survey respondents into continuous distribution functions.[9] Dealing with functions is one way to overcome the problem of dimensionality, and allows a degree of smoothing—or regularization—that proves to be helpful in the subsequent analysis. In each survey quarter, we adopt a nonparametric technique to obtain consistent estimates of that distribution.[10] The notation $\mathsf{p}_{t,h}(\cdot)$ will denote the distribution of h-step ahead point forecasts made at date t. The sequence $\{\mathsf{p}_{t,h}(\cdot)\}_0^T$ is then a functional time series (Bowsher and Meeks, 2008; Tsay, 2016), and a sub-sample of that time series was displayed in Fig. 1. Our approach to analysing the complex patterns of temporal and cross-sectional functional variation is set out in the next section.

2.3 Average distributions

What shape does the distribution of expectations take, on average? An interpretable answer requires us to align the distributions shown in Fig. 1 around some common feature (a process known as 'registration'; Ramsay and Silverman, 2005, Ch. 7). The most obvious such feature is the mean forecast, and so we center (i.e. horizontally translate) each distribution by subtracting from the h-step ahead inflation forecasts π_h^e made in each period the quantity $\int \pi_h^e \, \mathrm{d}\mathsf{P}_h$. The sample average distribution, or functional mean, of h-step ahead point forecasts is then given by:[11]

$$\overline{\mathsf{p}}_h(x) = \frac{1}{T}\sum_{t=1}^{T} \mathsf{p}^{\mathsf{c}}_{t,h}(x) \tag{2}$$

where the reader will have seen that $\mathsf{p}^{\mathsf{c}}_{t,h}$ represents the distribution of centered forecasts. We take the functional median—a robust measure of central tendency—to be the function with maximal band depth, as in López-Pintado and Romo (2009).[12] Given an empirical distribution

[8] In the Michigan survey, respondents are asked: "During the next 12 months, do you think that prices in general will go up, or go down, or stay where they are now?" and "By about what percent do you expect prices to go (up/down), on average, during the next 12 months?" In the Basix survey, respondents are asked: "From this list [below zero, about zero, about 1%, about 2%, . . . , about 10%, greater than 10%], can you tell me what you expect the rate of inflation to be over the next 12 months – i.e. to [date]?" The same question is asked for "the following 12 months", and (since the third quarter of 2008) "in five years time".

[9] Some form of initial data processing is typical in the analysis of functional data (Ramsay and Silverman, 2005, Ch. 1.5), as observations are seldom continuous even if the underlying processes are best thought of that way.

[10] Details of the penalized maximum likelihood (pML) approach we adopt are described in Part II of the supplementary material. In the case of the Michigan survey, we discard extreme observations prior to density estimation, using the same truncation rule as those who construct the commonly-used set of summary statistics associated with the data set. For further details on working with Michigan survey data, see Curtin (1996).

[11] The expectation of a random function $\mathsf{p}(x)$ is defined as the ordinary expectation taken pointwise for $x \in [a, b]$. For discussion on the concept of functional expectation, see Cuevas (2014, Section 3.1).

[12] Our depth calculation sets the number of curves used to form each band to three, as in López-Pintado and Romo. In practice, we truncate the range of the density functions before computing band depth to avoid regions of

of functional objects $\mathbb{P}_T$ and a particular function p, depth is a function $D(\mathbb{P}_T, \mathsf{p}) \geq 0$ indicating how far 'inside' that distribution p lies. A measure of depth therefore provides an ordering of the data, with the usual notion of the median being the function that lies the 'deepest' within the set.[13]

The average shapes of the distribution functions display remarkable similarities across the two regions. Fig. 2 displays the time averages of the centered density functions for both surveys (bold lines), overlaid with the cross-sectional densities for every time period (thin lines). For the latter, lighter colours correspond to observations further (in the sense of band depth) from the median. The standard deviation of the belief distribution is 4.2 percent in the US sample (Fig. 2, Col. 1), somewhat higher than the 2.3 percent seen in the UK (Fig. 2, Col. 2), since the former includes observations from the high-inflation period of the late 1970s while the latter does not, owing to the shorter sample at our disposal. The standardized third moment of the Michigan distribution is .96, and for the Basix distribution is 1.3, indicating that inflation beliefs are skewed quite strongly to the right.[14] The average distributions have excess kurtosis of 3.7 and 3.2, for the US and UK respectively, indicating that they possess fatter-than-normal tails.

2.4 Principal component analysis

Cross-sectional distributions of survey forecasts display considerable variation around their means (Fig. 2). A natural question to ask is whether that variation can be effectively summarized using a smaller number of functions. Functional principal component analysis (FPCA) is a standard technique for dimension reduction in functional data sets, and was applied to probability density functions by Kneip and Utikal (2001). The representation of a function in terms of its principal component functions (synonymously 'eigenfunctions') is known as the Karhunen-Loève expansion. The principal component functions form an optimal basis for the observations to hand.[15] Optimality in this context means that, for a given K, the linear approximation $\hat{\mathsf{p}}_{h,t}^{(K)}$ minimizes the integrated squared error criterion:

$$\mathsf{ISE}_{t,h}^{(K)} = \int \left\{ \left(\hat{\mathsf{p}}_{t,h}^{(K)} - \bar{\mathsf{p}}_h \right) - \left(\mathsf{p}_{t,h}^{\mathsf{c}} - \bar{\mathsf{p}}_h \right) \right\}^2 \mathrm{d}x, \quad \text{where} \quad \hat{\mathsf{p}}_{t,h}^{(K)} = \bar{\mathsf{p}}_h + \sum_{k=1}^{K} s_{kt} \mathsf{e}_k \tag{3}$$

averaged over all t, subject to the constraint that the functions $\mathsf{e}(\cdot)$ satisfy $\langle \mathsf{e}_k, \mathsf{e}_k \rangle = 1$ and $\langle \mathsf{e}_k, \mathsf{e}_j \rangle = 0$, $k \neq j$ where $\langle \cdot, \cdot \rangle$ denotes the usual inner product for square-integrable functions.

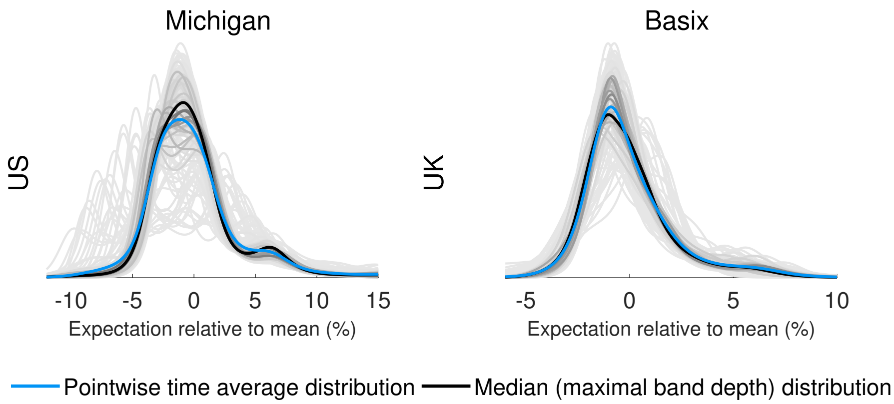
the tails which are close to zero. This prevents multiple small curve crossings in regions of near-zero density which would tend to reduce the depth of all functions.

[13]The concept of band depth is based on the graph of a function on the plane. A band can be thought of as the envelope delimited by n such graphs. The band depth of a given curve p_0 is given by the proportion of times that it falls inside the bands formed by taking all possible combinations of n curves. For example, if $n = 2$ and $T = 10$, there would be 45 pairs of curves (bands), and if the graph of p_0 lay entirely inside 9 of those bands its depth would be 0.2. See Cuevas (2014, Section 4.3) for further discussion.

[14]By contrast, the average distribution of professional forecaster beliefs are almost perfectly symmetric about the mean; see the supplementary material, Part III.

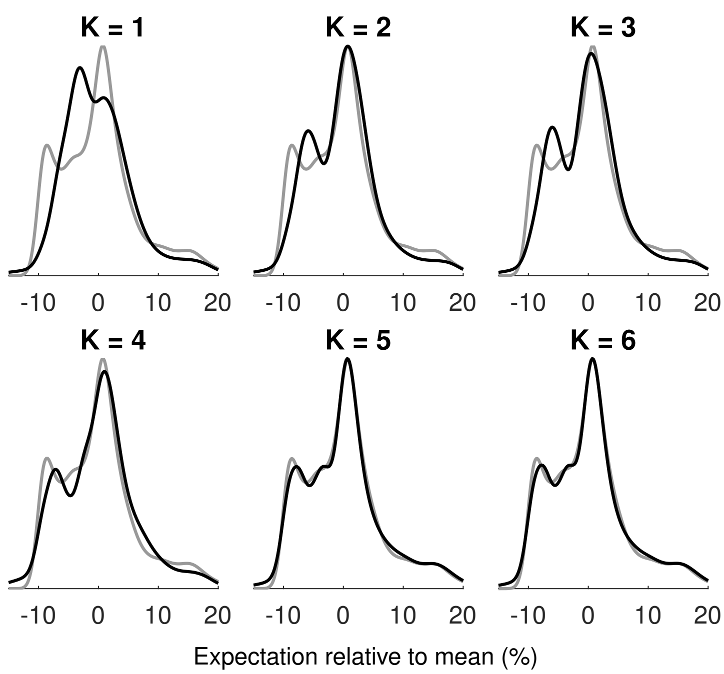
[15]FPCA will also be central to the approach we adopt for the estimation of the functional linear model, in Section 3.2. For an even-paced introduction to FPCA that sets out the correspondences with PCA on multivariate data, see Ramsay and Silverman (2005, Ch. 8).

Figure 2. Mean and median cross-section distributions for year-ahead inflation forecasts



Note: For each survey, panels overlay distributions of responses for all dates. The average expectation at each date has been subtracted to ensure every distribution is mean zero. Darker shaded curves are closer to the median distribution, where the median is the distribution that lies inside the most three-curve bands. For further details, see Table I.1.

Figure 3. Density function expansion in the empirical basis for a selected date



Note: Michigan survey, 1979-Q3. Grey line–observed distribution of expectations. Black line–approximation given by $\hat{\mathsf{p}}^{(K)}_{1979\text{-Q}3,4}$, $K = 1, \ldots, 6$, defined in Eq. (3). The magnitudes of the associated integrated squared errors are $\log_{10}[\mathsf{ISE}^{(K)}_{1979\text{-Q}3,4}] = \{-2.22, -2.71, -2.71, -3.06, -3.83, -3.84\}$.

The principal component scores are given by $s_{kt} = \langle \mathsf{p}_t, \mathsf{e}_k \rangle$. Although exact solutions to the principal component problem are not generally available, computational approximations are, the details of which are summarized in Appendix B (see also Tsay, 2016, Section 3.3).

It is helpful to gain a qualitative sense for how an approximation to the observed cross section varies with K by examining one particular case. Fig. 3 plots the distribution of forecasts reported by respondents to the Michigan survey in 1979-Q3, in gray, along with its approximation in terms of the sum of $K = 1, \ldots, 6$ principal components. Recall that the distribution has been centered on the average respondent's year-ahead expected inflation rate, which in that quarter was 9.4 percent. As additional components are added, the degree of approximation error declines, eventually by one-and-a-half orders of magnitude. Five components appear to provide a reasonable approximation to what is a highly complex functional shape, with the third and sixth having negligible loadings (and so providing negligible reductions in ISE).

Adding more principal components naturally leads to lower approximation errors in every time period. Fig. 4 displays the complete time series of approximation errors for both surveys. For the Michigan survey (left panel), there is something of a downward trend in the errors between 1978 and 1985, as the observed distributional shapes go from complex and multi-modal, as in Fig. 3, towards being close to average, as in Fig. 2. Capturing shapes that are closer to the functional mean naturally requires fewer components. It can be seen that there are some periods—for example, in 1995—where one component alone produces approximately the same magnitude of error as three components. But there are also periods where the two additional components reduce the approximation error by more than an order of magnitude—for example, in 2012. Similar observations apply for the Basix survey (right panel). Finally, the average share of variation explained by K components across all time periods is shown in Fig. 5. The scree plot displays the ten largest normalized eigenvalues associated with each e_k (left panel) and their cumulative sums (right panel). It can be seen that to explain 90, 95 or 99 percent of variation in either survey requires 2, 3, or 6 components respectively.[16]

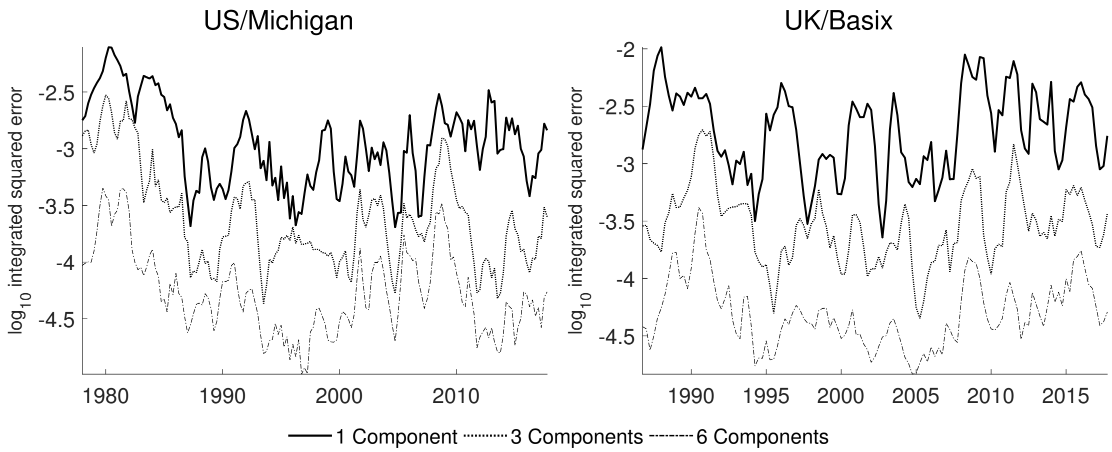
2.5 Disagreement, skew, and shape factors

The three leading principal components $\{s_{1t}, s_{2t}, s_{3t}\}$ identified above may be readily interpreted in terms of features of the belief distributions. The scores are correlated with empirical measures of disagreement (d_t), skew (κ_t), and a third factor we call 'shape' (which we will denote τ_t) that is explained below. The correlations between the scores and empirical disagreement and skew are $\rho^{MSC}(s_{1t}, d_t) = .97$ and $\rho^{MSC}(s_{2t}, \kappa_t) = .88$ for the Michigan survey. For the Basix survey, the first score correlates with skew and the second with disagreement, with $\rho^{BBS}(s_{1t}, \kappa_t) = .93$ and $\rho^{BBS}(s_{2t}, d_t) = .97$.[17]

[16] An alternative to the simple threshold criterion uses a Hellinger distance based cross-validation approach, see Tsay (2016).

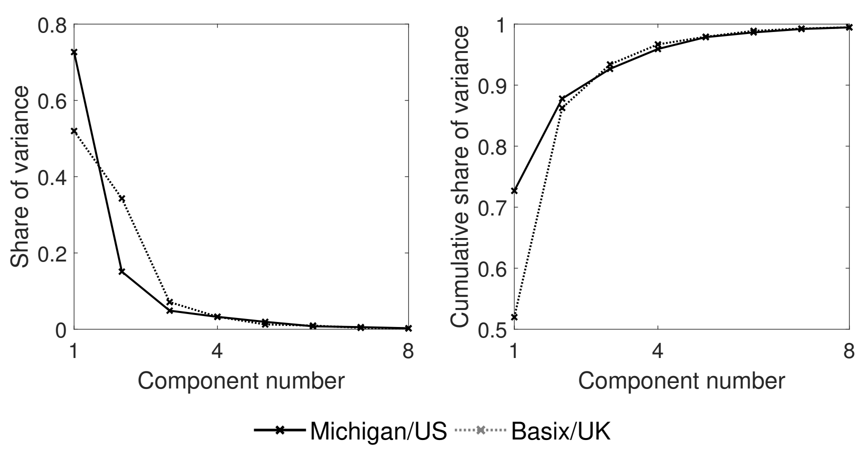
[17] We derive numerical values for standardized central moments and (combinations of) quantiles directly from the time series of distribution functions. Alternative measures of the same quantity are typically very similar: for example, 'disagreement' as the square root of the second moment or as the inter-quartile or -decile range; or skew

Figure 4. Integrated square errors of *K*-component approximations to the distribution of inflation forecasts



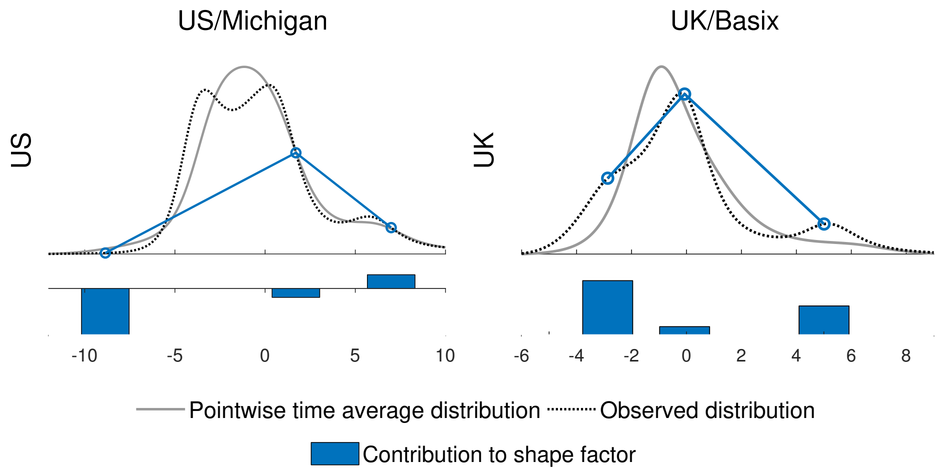
Note: The $\log_{10}$ integrated square error, Eq. (3), associated with a $K = \{1, 3, 6\}$ component expansion of the observed distributions of forecasts. US data is from the Michigan survey; UK data from the Basix survey. Centered three-quarter moving average.

Figure 5. Shares and cumulative shares of functional variation explained by the leading *K* principal components



Note: Scree plot gives the normalized sums of eigenvalues of the covariance operator.

Figure 6. The shape factor



Note: The top panel of the Figure shows the average distribution given in Fig. 2 and observed distributions for 1985-Q2 (Michigan) and 2011-Q3 (Basix). The three points marked by ∘ summarize distributional shape. Their contributions to the shape factor given in Eq. (4) with $\mathbf{x} = (-8.9, 1.7, 7)$ (Michigan) and $\mathbf{x} = (-2.9, 0, 5)$ (Basix) are shown in the bottom panel. Scale is omitted as units have no interpretation.

The third major factor—one that accounts for around 5 percent of functional variation—is related to the behavior of the *shape* of the distribution. The name arises from the combination of three points forming a tent shape summary of the distribution, as shown in Fig. 6. The shape factor is given by:

$$\tau_t(x_1, x_2, x_3) = \left[\mathsf{p}^{\mathsf{c}}_{t,h}(x_1) - \bar{\mathsf{p}}_h(x_1)\right] + \left[\mathsf{p}^{\mathsf{c}}_{t,h}(x_2) - \bar{\mathsf{p}}_h(x_2)\right] + \left[\mathsf{p}^{\mathsf{c}}_{t,h}(x_3) - \bar{\mathsf{p}}_h(x_3)\right] \quad (4)$$

where $x_1 < x_2 < x_3$. The correlations between the third score and the shape factors are $\rho^{MSC}(s_{3t}, \tau_t) = .88$ and $\rho^{BBS}(s_{3t}, \tau_t) = .87$ for the Michigan and Basix respectively.

Summary

Beliefs about future inflation are highly heterogeneous, but variation in them can be summarised by a few interpretable factors. In the US data, disagreement emerges as a central factor—other than the mean forecast—driving variation in belief distributions. It accounts for close to 80 percent of the variance in the mean-centered data, with only around 15 percent due to the skew factor, and 5 percent due to shape. But in the UK data, the primary factor turns out to be skew. The relative importance of the two principal factors is closer than in the US data, but the UK case serves to highlight the potential for important cross-country differences in the drivers of belief dynamics. In the time dimension, there are periods where (in addition to the average expectation) a single component—disagreement or skew—fares about as well in approximating observed beliefs as does a three-component model. But it is more often the case that capturing the shifting distribution of beliefs requires us to go beyond a single factor.

as the standardized third moment or as Pearson's median-based non-parametric statistic. We report maximum correlations between scores and similarly-defined measures of disagreement and skew in the text.

3 Heterogeneous beliefs and inflation dynamics

3.1 A heterogeneous beliefs Phillips curve

A Phillips curve derived from firms' optimizing behavior lies at the core of widely-used New Keynesian macroeconomic models. But the standard micro-foundations for that Phillips curve are not generally valid in the presence of heterogeneous beliefs (Mavroeidis, Plagborg-Møller, and Stock, 2014, p. 135). In this section, we derive a Phillips curve that is consistent with general forms of heterogeneous beliefs, before examining the issue of estimation in the following section. We focus on the case where there are differences of opinion between agents in the spirit of Harris and Raviv (1993), rather than differences in information. We take such differences of opinion as given.

Suppose there are N firms operating under monopolistic competition seeking to maximize their profits under time-dependent pricing. For an individual firm j, familiar calculations lead to a log-linear expression for the optimal reset price, $p^{\star(j)}$, in terms of current and expected future nominal marginal costs (log real marginal cost φ plus the log of the aggregate price level p):

$$p_t^{\star(j)} = (1-\beta\theta)\mathbb{E}_t^{(j)}\left[\sum_{\tau=t}^{\infty}(\beta\theta)^{\tau-t}(\varphi_\tau + p_\tau)\right] \tag{5}$$

where β is the discount factor applied to future profits and θ is the per-period probability that a firm's nominal price remains fixed. Firms hold diverse expectations, as indicated by the j superscript on the expectations operator.

To express the infinite sum on the right of Eq. (5) in the form of a difference equation, it is sufficient to assume the law of iterated expectations applies to the expectation operators $\mathbb{E}_t^{(j)}[\cdot]$ used by each firm j.[18] Applying this assumption, then subtracting the current aggregate price level from both sides gives an expression for the individual firm's optimal real, or relative, price:

$$q_t^{\star(j)} = (1-\beta\theta)\varphi_t + \beta\theta\mathbb{E}_t^{(j)}\left[q_{t+1}^{\star(j)} + \pi_{t+1}\right] \tag{6}$$

where $q_t^{\star(j)} := p_t^{\star(j)} - p_t$, and $\pi_{t+1} := p_{t+1} - p_t$ is the rate of price inflation.

To obtain an aggregate relationship, we sum Eq. (6) over the cross-section of firms:

$$q_t = (1-\beta\theta)\varphi_t + \beta\theta\overline{\mathbb{E}}_t[\pi_{t+1} + q_{t+1}] + \beta\theta\mathsf{E}_N\left\{\mathbb{E}_t^{(j)}\left[q_{t+1}^{\star(j)} - q_{t+1}\right]\right\} \tag{7}$$

where, using an obvious notation for the sum over N units, the future average relative price $q_{t+1} := \mathsf{E}_N\left\{q_{t+1}^{\star(j)}\right\}$ has been added and subtracted on the right hand side. We also use the notation:

$$\overline{\mathbb{E}}_t(x) := \mathsf{E}_N\left\{\mathbb{E}_t^{(j)}(x)\right\}$$

for the average expectation of variables x that do not depend on j ('aggregates').

[18] Although not innocuous, this assumption retains tractability without requiring stronger assumptions on the dependence (or lack thereof) of beliefs in the cross-section (see Coibion, Gorodnichenko, and Kamdar, 2018, p. 1466-7; Branch and McGough, 2009, Section 2.1).

We can now obtain an expression for the *heterogeneous beliefs* Phillips curve. Under the sticky price assumption, the aggregate price level is a combination of past and current prices. It follows that inflation is proportional to the average relative reset price:

$$\pi_t = \frac{(1-\theta)}{\theta} q_t$$

with the coefficient of proportionality depending on the frequency of price resets. Substituting into Eq. (7), we find:

$$\pi_t = \frac{(1-\beta\theta)(1-\theta)}{\theta}\varphi_t + \beta\overline{\mathbb{E}}_t(\pi_{t+1}) + \beta(1-\theta)\Delta_t \tag{8}$$

The term Δ_t, that does not appear in the rational expectations version of the Phillips curve, is defined by:

$$\Delta_t := \mathsf{E}_N\left\{\mathbb{E}_t^{(j)} q_{t+1}^{\star(j)} - \mathbb{E}_t^{(j)} q_{t+1}\right\}$$

It captures the extent to which, on average, differences of opinion between firms lead them to expect their own optimal price to diverge from the average of everyone else's.[19]

3.2 Estimation and testing with heterogeneous beliefs

We now consider approaches to estimating and testing the heterogeneous beliefs Phillips curve. If surveys elicited information on individual firms' price plans along with their expectations about general inflation, it would be possible to estimate Eq. (8) directly. However, such data has not generally been collected. To make progress, we must therefore make an approximating assumption based on the information most commonly to hand.

In the absence of any particular guidance from theory, we propose a flexible and data-driven approach to capture potential effects from heterogeneous beliefs. Suppose that the distance an individual firm expects its reset price to be from the average of new prices is related to the gap between its expectation and consensus. That is:

$$\mathbb{E}_t^{(j)} q_{t+1}^{\star(j)} - \mathbb{E}_t^{(j)} q_{t+1} \approx \gamma\left(\mathbb{E}_t^{(j)} q_{t+1} - \overline{\mathbb{E}}_t q_{t+1}\right)$$

where γ is a function to be estimated. Adopting the shorthand notation π^e and $\overline{\pi}^e$ for time-t individual and consensus expectations about the rate of aggregate inflation respectively, and with a slight abuse of notation as inflation π substitutes for real reset prices q, we may therefore write:

$$\Delta_t \approx \mathsf{E}_N\left\{\gamma\left(\pi_t^e - \overline{\pi}_t^e\right)\right\}$$

or in the limit as N becomes large:

$$\lim_N \Delta_t \approx \int \gamma(\pi_t^e - \overline{\pi}_t^e) \mathrm{d}\mathsf{P}_t^{\mathsf{c}}(\pi^e) \tag{9}$$

[19] Conditions under which $\Delta_t = 0$ for all t include full information rational expectations and those given by Andrade, Gaballo, Mengus, and Mojon (2019, pp. 13–14). Conditions that produce variation in Δ_t include those given by Kurz, Piccillo, and Wu (2013, Section 4).

with P^{c} denoting the distribution of mean-centered beliefs as above.

The estimable version of the heterogeneous beliefs model Eq. (8) is given by:

$$\pi_t = \alpha\varphi_t + \beta\overline{\pi}^e_{t,h} + \int \gamma(\pi^e_t - \overline{\pi}^e_t)\mathrm{d}\mathsf{P}^{\mathsf{c}}_t(\pi^e) + \varepsilon_t \tag{10}$$

Where $|\gamma|$ is large for some value of π^e (the expected rate of inflation, relative to consensus), expectations in that region of the distribution have greater influence on inflation. One way to think about the resulting estimates (β, γ) is that they indicate how to construct a single index of expectations that is most closely related to actual inflation from a large set of survey responses.

As noted in the Introduction, Eq. (10) is known as a functional linear model (Ramsay and Silverman, 2005, Ch. 15). A variety of estimation approaches have been proposed for the functional linear model (see Reiss, Goldsmith, Shang, and Ogden, 2017). We adopt the popular functional principal component regression approach, under which the functional regression Eq. (10) is recast as a multiple regression problem. To understand the procedure, recall that the functional data $\{\mathsf{p}^{\mathsf{c}}_{t,h}\}_0^T$ can be expressed in terms of its Karhunen-Loève expansion in the orthonormal basis $\{\mathsf{e}_k\}$ as $\mathsf{p}^{\mathsf{c}}_{t,h} = \mu_{\mathsf{p}} + \sum_{k=1}^{\infty}\langle \mathsf{p}^{\mathsf{c}}_{t,h}, \mathsf{e}_k\rangle \mathsf{e}_k$. Expanding the functional coefficient in the same basis allows us to write $\gamma = \sum_{k=1}^{\infty}\langle \gamma, \mathsf{e}_k\rangle \mathsf{e}_k$. Then using the properties of the e_k, see Eq. (A.1), the functional linear model of Eq. (10) can be rewritten as:

$$\pi_t = \alpha\varphi_t + \beta\overline{\pi}^e_{t,h} + \sum_{k=1}^{K}\gamma_k s_{k,t} + \varepsilon_t \tag{11}$$

where the γ_k are scalar coefficients to be estimated, and the functional principal component scores $s_{k,t}$ obtained in Section 2 appear as covariates.[20]

Having recast the functional linear model Eq. (10) as the multiple regression model Eq. (11), estimation proceeds as follows. Denote the $(T \times 1)$ vector formed by stacking the dependent variable by $\boldsymbol{\pi}$, and the $(T \times K)$ matrix of orthogonal principal component scores $s_{k,t}$ by $\mathbf{M}$. The N additional (scalar) regressors, including a vector of mean expectations, are collected in the $(T \times N)$ matrix $\mathbf{Z}$. Then conditional on the truncation level K and the true principal component scores, the heterogeneous beliefs Phillips curve model Eq. (10) is written compactly as:

$$\boldsymbol{\pi} = \mathbf{M}\boldsymbol{\gamma} + \mathbf{Z}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon} \sim \mathsf{N}(0, \sigma^2\mathbf{I})$$

where with a slight abuse of notation $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_K)^\top$. Let $\mathbf{X} = [\mathbf{Z}, \mathbf{M}]$ be the $T \times (N + K)$ matrix of regressors, and define the idempotent matrices:

$$\mathbf{P}_X = \mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top \qquad\qquad \mathbf{P}_Z = \mathbf{Z}(\mathbf{Z}^\top\mathbf{Z})^{-1}\mathbf{Z}^\top$$

Then the maximum likelihood estimator of the coefficients on the functional principal component scores is:

$$\hat{\boldsymbol{\gamma}} = \mathbf{Q}^{-1}\mathbf{M}^\top(\mathbf{I} - \mathbf{P}_Z)\boldsymbol{\pi} \tag{12}$$

[20] Additional details, along with references to the literature, are given in Appendix A.

where $\mathbf{Q} := (\mathbf{\Lambda} - \mathbf{M}^\top \mathbf{P}_Z \mathbf{M})$ is the Schur complement of $(\mathbf{Z}^\top \mathbf{Z})$ in $(\mathbf{X}^\top \mathbf{X})$, and $\mathbf{\Lambda} = \text{diag}(\lambda_1, \ldots, \lambda_K)$ contains the first K size-ordered eigenvalues corresponding to the scores arrayed in the columns of $\mathbf{M}$.

To establish whether an association exists between current inflation and the distribution of inflation forecasts, we employ the classical testing procedure of Kong, Staicu, and Maity (2016). A natural null hypothesis is that $\boldsymbol{\gamma}(\pi^e) = 0$, which recalling that the distributions p^{c} are mean zero by construction, corresponds to the special case where only the average forecast matters for inflation. As the distribution functions that appear in the model are mean zero, testing that null amounts to testing for the absence of a functional effect on inflation. A test of the hypothesis $H_0 : \boldsymbol{\gamma}(\pi^e) = 0$ for all π^e is equivalent to:

$$H_0 : \gamma_1 = \gamma_2 = \cdots = \gamma_K = 0 \quad \text{vs.} \quad H_a : \gamma_j \neq 0 \quad \text{for at least one } j,\ 1 \leq j \leq K$$

Then H_0 can be tested using the F-statistic:

$$T_F = \frac{\boldsymbol{\pi}^\top (\mathbf{P}_X - \mathbf{P}_Z)\boldsymbol{\pi}/K}{\boldsymbol{\pi}^\top (\mathbf{I} - \mathbf{P}_X)\boldsymbol{\pi}/(T-K-N)} \overset{approx.}{\sim} F_{K,T-K-N} \tag{13}$$

where $F_{K,T-K-N}$ denotes the F distribution with degrees of freedom depending on the number of functional principal components K and the number of scalar regressors N (Kong, Staicu, and Maity, Theorem 3.1).

An outstanding question is how to select the truncation level K. One simple approach is to select only those components for which the cumulative share of variance (in the functional explanatory variable) is below some threshold value, often set at 95% or 99%. But a low variance share for a particular component does not necessarily imply that it is unimportant in the regression model (see the discussion in Jolliffe, 2002, Section 8.2).[21] In the subsequent analysis, we select two values of K, one based on the simple cumulative eigenvalue criterion, and one based on the Bayes Information Criterion (BIC), which takes account of both fit and parameterization.

4 Quantitative implications of heterogeneous beliefs

The preceding section set out the micro-foundations of a New Keynesian Phillips curve that allows agents to hold heterogeneous beliefs, along with a novel strategy for estimating such a model. We now turn to the empirical assessment of this more-general version of the standard Phillips curve. We ask whether heterogeneity 'matters' for inflation dynamics or whether, as is commonly assumed, the average survey forecast captures all relevant effects of expectations on inflation. Our answer to this question takes the form of tests of the linear aggregation

[21] Kneip and Utikal (2001) develop asymptotic inference for selecting principal components of density functions, and Tsay (2016) proposes a cross-validation procedure based on the Hellinger distance. Faraway states in his comment on Kneip and Utikal (2001) that: "In other situations, selection of dimension [the number of components] is a secondary consideration to some [primary] purpose—typically prediction. The dimension should be chosen to obtain good predictions ... It is important to optimize the secondary selection with respect to the primary objective and not some criterion associated with the secondary objective". His arguments motivate our use of the BIC.

assumption discussed above. In subsequent sections, we dissect the relationship between beliefs and inflation in greater detail, and document its robustness to variations in the baseline model. We consider identical models and estimation methods for the United States and United Kingdom.

4.1 Do heterogeneous beliefs enter the Phillips curve?

Inflation in the United States

The baseline specification for the US Phillips curve relates the annualized quarter-on-quarter percent change in the seasonally-adjusted consumer price index (CPI) to the CBO measure of the unemployment gap, the survey average one-year-ahead expected inflation rate from the Michigan survey, and an outlier dummy variable.[22] The importance of the average survey expectation that has been documented in other studies is confirmed by the results (Tab. 1, Col. 1). The slope of the Phillips curve ($\hat{\alpha}$) is around -0.3, and is significant at 1%. The substance of these results is very similar to that reported in Coibion, Gorodnichenko, and Kamdar (2018), as they are based on an equivalent specification and a sample that is only modestly extended.

While average expectations hold considerable explanatory power for inflation, heterogeneity in expectations matters too. Tab. 1 (Cols. 2–3) reports that the aggregation function γ is strongly significant in our Phillips curve regressions. The BIC selects for three principal components of the belief distributions, but remarkably the penalty for the model with six components is no larger than that for the model with none.[23] The p-values of the functional T_F-statistic are below 0.1%, both when three components are used and when six are used. At the same time, the estimated coefficient on the average expectation remains highly significant, although its point estimate is somewhat affected by the specification of the functional effect. Our results are robust to including supply factors (Col. 4).[24]

Inflation in the United Kingdom

We estimated identically-specified models on UK data, again using year-ahead expectations. Because no official measures of the natural rate of unemployment exist for the UK for the sample period in question, we compute one by fitting a cubic spline to the raw unemployment data using OLS (Poirier, 1973). Our measure of the unemployment gap is the residual from

[22]The average value of the CPI in a given quarter is used to compute the quarter-on-quarter inflation rate. We use expectations reported in the first month of the quarter, which may incorporate information about last quarter's inflation rate, but cannot incorporate any data for the current quarter. This practice helps to ameliorate concerns over endogeneity bias in the expectations data, but results based on full-quarter responses are very similar.

[23]In models with SPF data, the first component is selected. It has a high correlation with disagreement, and is significant at the 1% level. Overall we find that the household model encompasses the professional forecaster model, in line with the findings reported in Coibion and Gorodnichenko (2015). For further details, see the supplementary material, Part III.

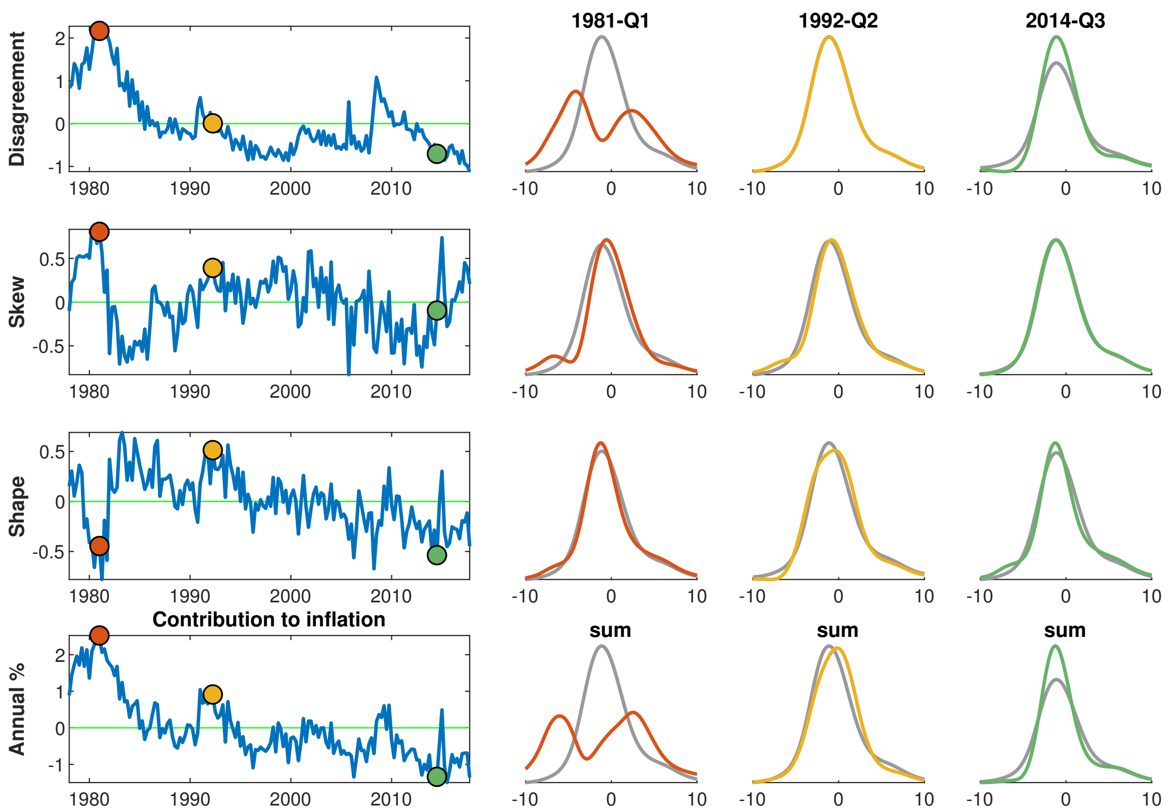
[24]Because supply shocks have at times driven inflation and demand—summarized by the unemployment gap—in opposite directions, if omitted they may impart a downward bias to the coefficient on slack. We include distributed lags in the supply factors in our regressions, and eliminate those variables/lags that are statistically insignificant. For the US, this leads us to retain only the contemporaneous change in the oil price; for the UK, the change in the sterling price of oil and its first lag are retained, along with the change in the relative price of imported goods.

Table 1. Baseline heterogeneous beliefs Phillips curve

	US/Michigan				UK/Basix			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable	CPI	CPI	CPI	CPI	CPI	CPI	CPI	CPI
Unemployment gap	$-.267^{**}$ (.103)	$-.283^{***}$ (.095)	$-.266^{**}$ (.104)	$-.354^{***}$ (.066)	.048 (.131)	$-.233^{**}$ (.109)	$-.237^{**}$ (.108)	$-.176^{*}$ (.100)
Average expectation	1.71^{***} (.104)	1.54^{***} (.131)	1.79^{***} (.168)	1.23^{***} (.095)	1.06^{***} (.129)	$.732^{***}$ (.199)	$.756^{***}$ (.239)	$.890^{***}$ (.187)
Number of FPCs	–	3	6	3	–	3	6	3
T_F-statistic	–	8.32 [.000]	5.26 [.000]	19.22 [.000]	–	8.48 [.000]	5.51 [.000]	5.56 [.001]
Supply factors	n	n	n	y	n	n	n	y
Outlier dummy	y	y	y	y	y	y	y	y
Sample	1978Q1–2017Q4				1986Q4–2017Q4			
R^2	.773	.805	.813	.870	.674	.733	.743	.767
BIC	.914	.858	.914	.486	.652	.571	.645	.509
Number of obs.	160	160	160	160	125	125	125	125

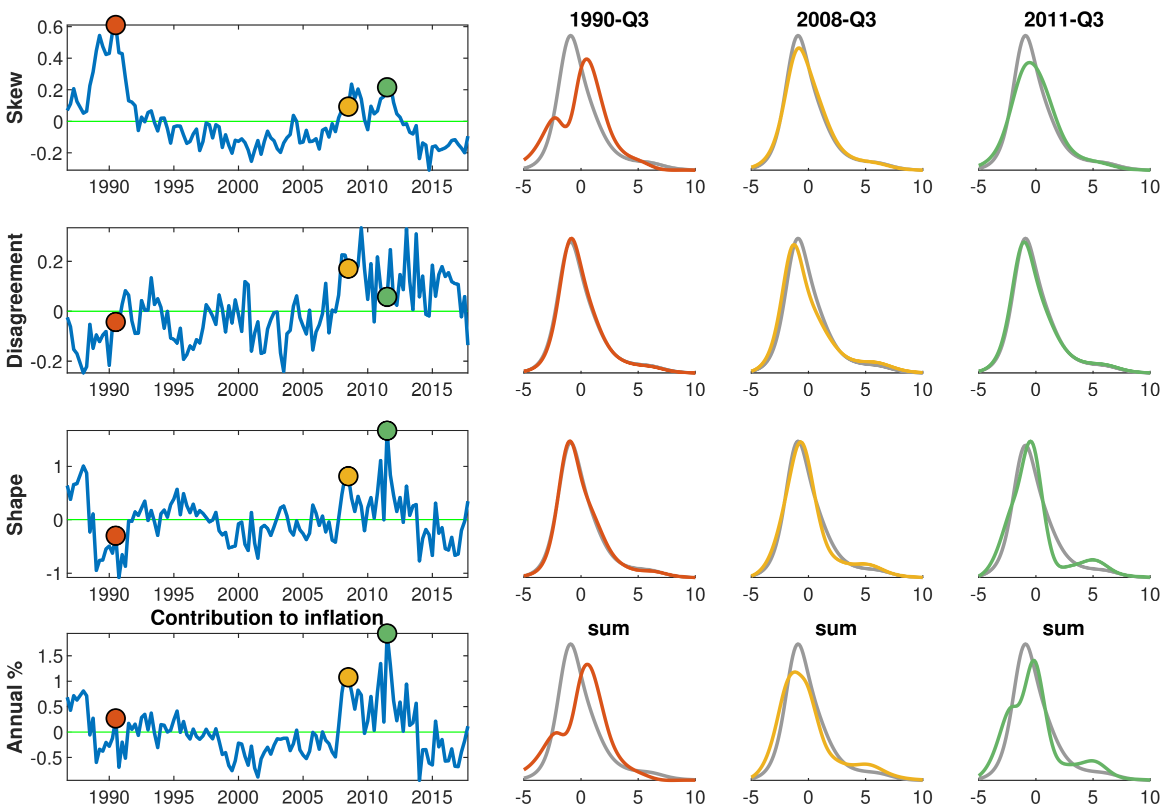
Note: Estimates of Eq. (11). Dependent variable is the seasonally adjusted annualized quarter-on-quarter percentage change in the consumer price index. Newey-West adjusted (5 lags) standard errors for t test (scalar covariates) appear in parentheses. p-values for F test (functional covariate) appear in brackets. All regressions include a constant. Outlier dummies equal 1 in 2008-Q4 (US and UK) and in 1991-Q2 (UK). Supply factors are the quarterly percentage change in the oil price (US) or the sterling oil price lagged one quarter (UK), and the quarterly percentage change in import prices (UK). Asterisks denote significance at the 10% (*), 5% (**), and 1% (***) levels.

Figure 7. Effect of belief factors on US inflation



Note: The kth row of each panel shows information corresponding to the kth largest eigenvalue of the covariance operator, $k = 1, 2, 3$. Component labels (left, rotated) are as described in Section 2.5. Col. 1 shows the kth score multiplied by $\hat{\gamma}_k$ (see Eq. 12) taken from the regression reported in Tab. 1 Col. 4. The ordinate value indicates the partial contribution of component k to the fitted value of that model. In Cols. 2–4, the gray curve is the sample average mean-centered density of household year-ahead inflation forecasts. The colored curves are the kth term in the Karhunen-Loève expansion of the date t mean-centered probability density function, $\mathrm{f}_t^{(k)} = \mu_{\mathrm{p}} + \langle \mathrm{p}_{t,h}, \mathrm{e}_k \rangle \mathrm{e}_k$. The bottom row shows the sum of the preceding rows in each column: Col. 1 indicates the combined contribution of the first three FPCs to the fitted value of the indicated regression; Cols. 2-4 show the approximation to the observed density based on the first three FPCs.

Figure 8. Effect of belief factors on UK inflation



Note: Refer to Fig. 7 for explanation. Panels are based on the regression reported in Tab. 1 Col. 8.

that regression.[25] Estimates of the Phillips curve that exploit our newly-constructed household survey data series (Basix) are also reported in Tab. 1 (Cols. 5–8). The standard variant featuring average expectations only (Col. 5) has a positive ('incorrect') but insignificant slope. The coefficient on average expectations is almost identical to unity.

To the standard specification, we once more add functional principal components from the full distribution of survey responses. The BIC selects three components, but even the model with six components is preferred to that with none. Estimates given in Cols. (6–7) show values for the T_F-statistic that indicate a high level of statistical significance for both three and six components. The additional information also yields a more interpretable model: When the distribution is included, the coefficient on the unemployment gap becomes sizeable, correctly-signed, and significant. Including supply factors does not change the nature of the results (Col. 8).[26]

4.2 How do expectations affect inflation?

The results reported above lend support to a heterogeneous beliefs specification for the Phillips curve. They demonstrate that the association between current inflation and inflation expectations is more complex than previously recognized. To help understand how expectations taken in the round shape inflation dynamics, it is helpful to relate changes in the shapes taken by distribution functions to changes in actual inflation. To that end, Fig. 7 (Col. 1, Rows 1–3) displays the partial contributions of the three named factors identified in Section 2.5 to the fitted value of inflation. The total contribution of the factors (the sum of the preceding rows) to model fit appears in the final row. Subsequent columns in the figure show terms in the Karhunen-Loève expansion of the mean-centered distributions of point forecasts at dates selected to display notable functional variation (Cols. 2–4). The final row gives the three-term approximation to the distribution (which again is the sum of the preceding rows). We adopt estimates $\hat{\gamma}_k$ from the models featuring three components and controls for supply factors throughout. Fig. 8 shows identical information for the UK.

Several broad points stand out from the figures. First, changes in the distribution of expectations can have a material impact on inflation. For example, in the UK in 2011-Q3 (Fig. 8, Col. 4), actual annual inflation was running at around 5%, following a substantial devaluation of the British pound. The dominant feature of the distribution was an unusual second mode, captured by the shape factor (Row 3). The presence of pushed up on inflation by around 1.5ppt. Second, it can happen by chance that just one mode of variation (or none) is relevant for inflation in a given period. But generally, changes in distributions have complex and occasionally

[25] Unemployment gap measures based on natural rates estimates constructed using more sophisticated methods, including filter-based methods, were closely comparable to those produced via our spline approach. Moreover, constructing the unemployment gap using a spline-interpolated version of the OECD's annual natural rate series, and using that in our regressions, produced estimates of the Phillips curve slope that were very similar to those reported in Tab. 1.

[26] The aggregation function is shown in the Supplemental Material, Fig. III.5 (right panel).

off-setting effects on inflation. For example, in the US in 1981-Q1 (Fig. 7, Col. 2), the skew and shape factors have a roughly equal and opposite effects. Likewise a given contribution to above- or below-average inflation may be associated at different times with very different belief distributions. A final and overriding observation is that episodes when beliefs are particularly diverse—that is, where the values of s_{kt} are furthest from zero—coincide with sizeable macroeconomic disturbances, such as those attended by changes in the monetary regime (the early 1980s in the US, the early 1990s in the UK), or the financial crisis and Great Recession of the late 2000s.

For the US, the overall effect of expectations on inflation, seen through the lens of this model, was considerably less supportive of inflation following the Great Recession of 2007-9 than is commonly thought. Fig. 7 (bottom left panel) plots the estimated effect of expectations. The contributions of average beliefs and of their distribution around the average are shown separately. Although the average household expectation held up well after 2011, expectations as a whole imparted a substantial disinflationary impulse. This observation changes the conclusion reached in Coibion and Gorodnichenko (2015), who used the same underlying data, but summarized expectations using the cross-section average alone in their model.[27]

For the UK, the impact of heterogeneous expectations on fitted inflation appears particularly sizeable around the time of the financial crisis, Fig. 8 (bottom panel). This leads us to revise our narrative of inflation drivers over the period of devaluation-driven of inflation in 2011-12; in particular, the reassuring stability of the average inflation expectation masked the positive contribution to inflation being made by the skew and shape factors. Indeed, expectations contributed far more to CPI inflation that did the direct effects from oil and other import prices at that time (see decompositions in Section X of the Supplementary Material).

4.3 Is inflation backward-looking?

An important question in monetary economics is the extent to which inflation depends on its own past values. In a purely backward-looking model, disinflating the economy is costly, because unemployment must be driven high enough for long enough to 'wring out' inflation from the system. But in a purely forward-looking model, anticipated disinflations need not be costly at all. Backward-looking inflation behaviour is commonly identified with one of two potential mechanisms. The first is simply that expectations themselves are formed in a backward-looking manner. The second mechanism relates to the intrinsic persistence of the inflation process, rather than the persistence of expectations (or indeed, any of the other determinants of inflation), for example due to price indexation (Fuhrer, 2011).

We investigate the extent and sources of backward-looking behaviour using the Phillips curve framework set out above. To our baseline specification, we add an additional term in

[27] A replication of Coibion and Gorodnichenko's results for the Michigan survey and the Survey of Professional Forecasters is reported in the supplementary material, Section VI. Formal tests provide no evidence against the stability of parameters on the expectations terms, see supplementary material, Section V.

Table 2. Backward- and forward-looking components in inflation

	US/Michigan			UK/Basix		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	CPI	CPI	CPI	CPI	CPI	CPI
Unemployment gap	−.059 (.093)	−.177** (.079)	−.309*** (.062)	−.267* (.119)	−.027 (.101)	−.180* (.095)
Lagged inflation	.701*** (.044)	.231*** (.067)	.083 (.063)	.449*** (.072)	.111 (.074)	.073 (.059)
Average expectation	–	1.20*** (.143)	1.61*** (.192)	–	.918*** (.133)	.814*** (.188)
Number of FPCs	–	–	3	–	–	3
T_F-statistic	–	–	14.28 [.000]	–	–	4.99 [.003]
Supply factors	y	y	y	y	y	y
Outlier dummy	y	y	y	y	y	y
Sample	1978Q1–2017Q4			1986Q4–2017Q4		
R^2	.743	.836	.872	.612	.740	.770
BIC	1.07	.652	.497	.903	.544	.536
Number of obs.	160	160	160	125	125	125

Note: Estimates of Eq. (14). For further explanatory notes, see Tab. 1.

lagged inflation to produce a hybrid Phillips curve:

$$\pi_t = \alpha\varphi_t + \beta\overline{\pi}^e_{t,h} + \delta\pi_{t-1} + \int \gamma(\pi^e)\mathsf{p}^{\mathsf{c}}_{t,h}(\pi^e) + \varepsilon_t \tag{14}$$

In Tab. 2 we show the results of adding the expectation terms $\overline{\pi}^e_{t,4}$ and $\mathsf{p}^{\mathsf{c}}_{t,4}$ one at a time to a purely backwards-looking model.

When lagged inflation appears without any forward looking terms in the Phillips curve, its coefficient is large and significant for both the Michigan and Basix models (Tab. 2, Cols. 1 and 4). However, this result is not robust. In both cases, the coefficient on π_{t-1} is upward biased because of its positive correlation with the omitted variable $\overline{\pi}^e_{t,4}$. Adding the average survey expectation substantially reduces the magnitude of the coefficient, consistent with the findings reported by Fuhrer (2017). For the US (Col. 2), the weight on the backward-looking term falls by a factor of three, although it remains significant. The result is a hybrid Phillips curve, with forward- and backward-looking components both having a statistically significant role. For the UK (Col. 5), adding average expectations results in backward-looking terms becoming economically and statistically indistinguishable from zero. As a result, the other parameter estimates are close to those in Tab. 1 (Col. 5).

Allowing for heterogeneity in inflation expectations eliminates the backward-looking component from the Michigan regression (Tab. 2, Col. 3). The T_F-statistic is large and significant with three FPCs. Omitting the information contained in the distribution of beliefs about future inflation leads to an upward bias in the backward-looking coefficient δ in Eq. (14) even after adding average expectations. For the Basix regression (Col. 6), lagged inflation remains irrelevant, and the distribution function is strongly significant. We also observe that the version with forward-looking terms is preferred by the BIC over the purely backwards-looking version in both regions. Taken in the round, these results imply that intrinsic persistence is not an important feature of the inflation process, over the periods covered here, consonant with the micro evidence.[28]

4.4 Do trends drive out expectations?

The recent literature recognizes the importance of accounting for time-varying trend inflation when thinking about cyclical inflation dynamics. Cogley and Sbordone (2008) present a micro-founded Phillips curve that features inflation trends, and fit it to US data; and leading statistical approaches to modeling and forecasting inflation formulate the inflation process in 'gap' form, that is, in terms of deviations from trend (Stock and Watson, 2007; Faust and Wright, 2013). A potential concern is that the apparent importance of expected inflation found using the type of models studied in Section 4.1 may be down to an association between expectations and trend. For example, Cecchetti, Feroli, Hooper, Kashyap, and Schoenholtz (2017) argue for the unimportance of short run expectations on this basis, at least in periods where monetary policy was well run.

We modify the baseline heterogeneous beliefs Phillips curve Eq. (10) to remove the trend component of inflation τ_t, measured using exponential smoothing, as follows:

$$\pi_t - \tau_t = \alpha\varphi_t + \beta(\overline{\pi}^e_t - \tau_{t-1}) + \int \gamma(\pi^e)\mathsf{p}^{\mathsf{c}}_{t,h}(\pi^e) + \varepsilon_t \tag{15}$$

The inflation gap depends on an average expectation gap, which is the difference between average expectations and trend, along with the unemployment gap, and the distribution of expectations summarized by functional principal components.[29] To align as well as possible with the information available to form near-term expectations, and to avoid biasing our estimates, we form the expectations gap using the trend at time $t-1$.[30]

The inflation gap in the US

Estimates for the standard inflation gap model with linear aggregation indicate that average near-horizon forecasts have a moderate effect on inflation, after accounting for trend. Tab. 3

[28] Theories of inflation that are based on information rigidities predict that average lagged expectation terms should appear in the Phillips curve. Although the match with these theories is not perfect, we nevertheless checked whether a term in $\pi^e_{t-1,4}$ was significant in our regressions. It was not in either region, and in the US, its coefficient was below 0.1.

[29] Model (15) is similar in spirit to Models (8) and (9) of Faust and Wright (2013). Those authors use lagged inflation to proxy forward-looking behaviour rather than directly including survey expectations as a covariate.

[30] The ES trend computed recursively using $\tau_t = \rho\tau_{t-1} + (1-\rho)\pi_t$, where π_t is the relevant inflation measure and $\rho = 0.9$ is a parameter. Because the trend is formed using current-quarter inflation, the t-dated expectations gap is correlated with the regression errors, whereas the $t-1$-dated gap is not.

(Col. 1) reports that the average expectation gap has a coefficient of 0.45. When the distribution of beliefs about near-term inflation are added to the model (Col. 2) the coefficient on the expectation gap rises to 0.87 with a t-statistic above 12. Notably, this parsimonious model now also appears well-specified having no apparent residual autocorrelation. The distribution of beliefs has a T_F-statistic above 25. We observe a marked improvement in overall fit, as measured by R^2, and large reductions in the BIC, which selects for three components. Allowing the slope on average expectations to change in the Great Moderation (Col. 3), we observe some attenuation in the response of inflation, in line with the arguments in Cecchetti, Feroli, Hooper, Kashyap, and Schoenholtz (2017). However, the overall effect of the average remains sizeable and significant, and the first two components of the distribution of beliefs retain their explanatory power for inflation gap dynamics.

The inflation gap in the UK

Estimates for the UK inflation gap are very similar to the baseline results for inflation in levels. The estimates are shown in Tab. 3 (Cols. 4–6). The similarity likely follows from the fact that the coefficient on the expectations gap is close to unity. In that case, as the trend is relatively smooth, with τ_t being close to τ_{t-1}, terms in the trend then roughly cancel from the two sides of Eq. (15). That said, the T_F-statistic continues to reject linear aggregation, and the heterogeneous beliefs model is free from autocorrelation problems. Our final result (Col. 6) indicates that the responsiveness of inflation to the average expectations gap during the 15-year 'NICE' period between the adoption of inflation targeting and the onset of the global financial crisis (1992-Q4 through 2007-Q4) may have been slightly smaller than at other times (around .9 rather than 1.16).[31] However, the break is imprecisely estimated, with a t-statistic of 1.3, and indeed we found no strong evidence to suggest breaks in the coefficient on any decadal sub-sample.

4.5 Can moments substitute for functional principal components?

In Section 2 we associated the three leading principal components of the belief distributions to empirical measures of disagreement, skew, and shape. We pointed out that the principal component score most closely related to disagreement was the primary factor driving the dynamics of the US belief distributions, while the score associated with the skew played a more prominent role in the UK. Noting that in some cases a probability distribution can be determined from knowledge of its' moments—as in such familiar cases as the exponential family—we investigate whether straightforward regression on moments provides an alternative to principal component regression for capturing the effect of shifting beliefs on inflation.[32]

[31] The term NICE was coined by former Bank of England Governor Mervyn King, and stands for 'Non-Inflationary Consistently Expansionary'. It is the UK equivalent of the Great Moderation, and is taken to commence with the adoption of inflation targeting as the monetary regime after the UK's exit from the European Exchange Rate Mechanism.

[32] The quoted inversion is what is known as the 'problem of moments'. A correspondence between moments and distributions need not exist, or be unique. We consider standardized moments, which are not nested in the FLM. However, regression on the raw moments is equivalent to the restriction that γ lie in the space of polynomials.

Table 3. Inflation gaps and heterogeneous beliefs

	US/Michigan			UK/Basix		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	CPI gap	CPI gap	CPI gap	CPI gap	CPI gap	CPI gap
Unemployment gap	$-.321^{***}$ (.104)	$-.386^{***}$ (.055)	$-.415^{***}$ (.058)	.059 (.123)	$-.234^{**}$ (.095)	$-.223^{**}$ (.096)
Average expectation gap	$.449^{***}$ (.100)	$.870^{***}$ (.067)	$.849^{***}$ (.062)	1.14^{***} (.179)	1.02^{***} (.124)	1.16^{***} (.165)
Average expectation gap × Great Moderation	–	–	$-.263^{*}$ (.115)	–	–	−.271 (.202)
Number of FPCs	–	3	2	–	5	5
T_F-statistic	–	25.62 [.000]	38.43 [.000]	–	5.60 [.000]	5.54 [.000]
Supply factors	y	y	y	y	y	y
Outlier dummy	y	y	y	y	y	y
Sample	1978Q1–2017Q4			1986Q4–2017Q4		
R^2	.643	.763	.764	.638	.710	.713
BIC	.613	.298	.294	.366	.338	.364
DW test (p-value)	.000	.271	.273	.001	.181	.171
Number of obs.	160	160	160	125	125	125

Note: Estimates of Eq. (15). Dependent variable is the seasonally adjusted annualized quarter-on-quarter percentage change in the consumer price index less the exponentially-smoothed inflation trend. Newey-West adjusted (5 lags) standard errors for t test (scalar covariates) appear in parentheses. p-values for F test (functional covariate) appear in brackets. All regressions include a constant. Outlier dummies equal 1 in 2008-Q4 (US and UK) and in 1991-Q2 (UK). Supply factors are the quarterly percentage change in the oil price (US) or the sterling oil price lagged one quarter (UK), and the quarterly percentage change in import prices (UK). In functional models, the number of principal components is selected using BIC. DW test: the Durbin-Watson test, null of no residual autocorrelation. Great Moderation dummy is 1 for 1984-Q1 through 2007-Q4 (US) and 1992-Q4 through 2007-Q4 (UK). Asterisks denote significance at the 10% (*), 5% (**), and 1% (**) levels.

Table 4. Proxy regressions using moments of the belief distributions

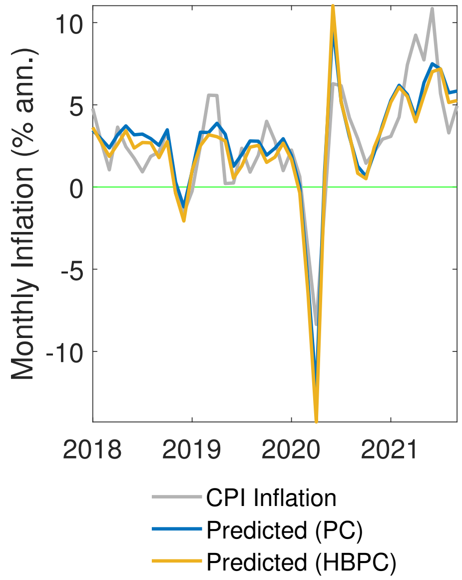
	US/Michigan			UK/Basix		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	CPI	CPI	CPI	CPI	CPI	CPI
Unemployment gap	−.267** (.103)	−.493*** (.104)	−.370*** (.101)	.048 (.481)	−.017 (.122)	−.216** (.109)
Average expectation	1.71*** (.104)	1.14*** (.178)	1.32*** (.161)	1.06*** (.129)	1.03*** (.302)	.859*** (.252)
Second moment	–	1.02*** (.327)	1.23*** (.095)	–	1.21** (.519)	.659 (.935)
Third moment	–	−.574 (.482)	−.441 (.518)	–	.044 (.684)	.656 (.722)
T_F-statistic	–	–	func [.006]	–	–	func [.000]
Outlier dummy	y	y	y	y	y	y
Sample	1978Q1–2017Q4			1986Q4–2017Q4		
Number of FPCs	–	–	3	–	–	3
R^2	.773	.795	.811	.674	.695	.738
BIC	.914	.878	.891	.652	.663	.628
Number of obs.	160	160	160	125	125	125

Note: Estimates of Eq. (10) augmented with moments of the distributions $\mathsf{p}^c_{t,4}$. Dependent variable is the seasonally adjusted annualized quarter-on-quarter percentage change in the consumer price index. Newey-West adjusted (5 lags) standard errors for t test (scalar covariates) appear in parentheses. p-values for F test (functional covariate) appear in brackets. All regressions include a constant. Outlier dummies equal 1 in 2008-Q4 (US and UK) and in 1991-Q2 (UK). Supply factors are the quarterly percentage change in the oil price (US) or the sterling oil price lagged one quarter (UK), and the quarterly percentage change in import prices (UK). Asterisks denote significance at the 10% (*), 5% (**), and 1% (**) levels.

We re-estimated our baseline Phillips curve model using moments of the distribution of beliefs as regressors instead of functional principal components. Focussing first on US data, Tab. 4 (Cols. 1–3) reports the results for the regression including second and third moments.[33] The second moment is significant in the regression, and where present reduces the coefficient on the average expectation, much as observed in Tab. 1. The third moment does not appear to be significant in the regression. When FPCs also appear in the model (Col. 3), the p-value of the functional T_F-statistic is well below 1%. Similar results are found for the UK (Cols. 4–7). The regressions confirm that only the second moment is significant, but, unlike for the US, it becomes insignificant when the FPCs are also included. In summary, functional components above the second are highly correlated with inflation, but weakly correlated with (linear combinations of) moments. Functional principal components therefore capture information in cross-sections of expectations that is relevant for inflation, even after including the moments of the distribution.

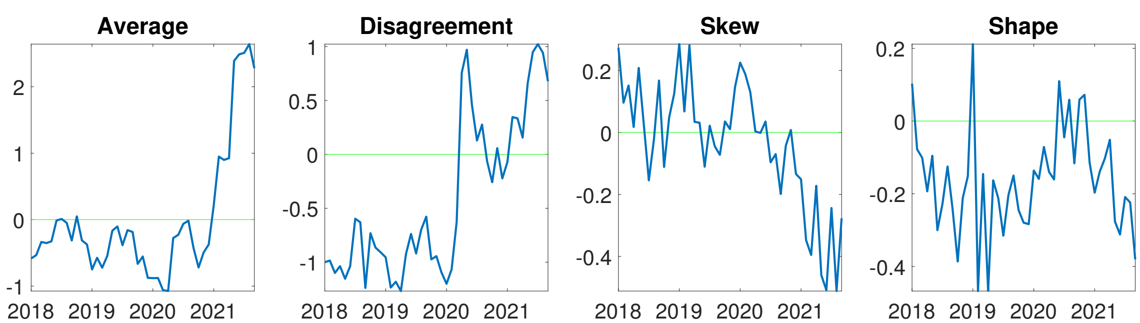
[33]We experimented with including moments up to the sixth, but all moments above the third were statistically insignificant.

Figure 9. Monthly CPI and out of sample fitted values



Note: CPI inflation is the month-on-month percentage change in the seasonally-adjusted CPI, expressed at an annual rate. Predicted values result from parameters estimated using Eq. (11) with $K = 0$ (PC) and $K = 3$ (HBPC) over the period January 1978-December 2017. Regression includes current and lagged month-on-month percent change in oil prices, and outlier dummies for September 2005 and November 2008. All covariates take their observed values for January 2018-September 2021.

Figure 10. Contributions of belief factors to inflation before and during the pandemic



Note: Each panel shows the contribution of the named component of beliefs about year-ahead inflation to predicted inflation, January 2018-September 2021. Component titles are as described in Section 2.5.

5 Inflation and the pandemic

In the 18 months or so following the onset of the COVID-19 pandemic in early 2020, US inflation was more volatile that at any time since the global financial crisis. A number of factors underlay its movements: the headline rate of US unemployment increased by more than 10 percentage points in the space of two months; prices of some primary commodities, notably oil, collapsed; consumers' expenditure patterns underwent rapid shifts in response to changes in both the demand and supply of goods and services, which complicated measurement (Cavallo, 2021); and average survey measures of household inflation expectations, which were initially stable, then increased notably as the economy recovered in 2021. The unusual nature and size of the shock, coupled with its far-reaching impact on the economy, naturally raise concerns that the relationships reported so far could be subject to breaks.

To investigate the stability of our estimated Phillips curves and the impact of expectations on inflation during the COVID-19 pandemic, we examine how well they track the data out of sample. Because policymakers needed to monitor macroeconomic out-turns at high frequency during the pandemic, we re-estimated our models at the monthly frequency, for the same 1978-2019 sample period as in the main analysis. (Results for models estimated on monthly data appear in the supplementary material; results for the stability of the quarterly model are similar to those of the monthly model.) Fig. 9 plots actual inflation alongside its predicted value when the covariates take their observed values for January 2018 through September 2021. Both the baseline model (PC) and the heterogeneous beliefs model (HBPC) track the data without conspicuous error. Both tended to over-predict the decline and subsequent rebound in inflation as the pandemic started to bite in 2020, as a consequence of over-weighting the impact of oil price changes in this period. But the direction and magnitude of the predicted changes in inflation were qualitatively correct. A formal statistic test of the stability of the models across pre- and post-COVID periods confirms this visual impression.[34]

The distribution of beliefs around their mean had off-setting overall effects on inflation in the COVID period, which explains why the PC and HBPC model variants perform similarly despite substantial shifts in disagreement, skew, and shape. Figure 10 shows the contribution of the mean and the scores of the FPC decomposition of the distribution of beliefs to the HBPC model's fitted values. It is notable that whereas the contribution of the average expectation moved within a 1ppt range in 2020, heightened disagreement raised predicted inflation by close to 2ppt early in the pandemic.[35] Moving into 2021, the combined negative effects from skew and shape were sufficient to outweigh the positive effect from disagreement. The expectations components of the two models were therefore rather similar. This episode is another illustration

[34] We compute the Chow breakpoint test, setting the break date to be February, 2020. The F-statistic is 1.67, compared to a 1% critical value of 2.3. We are therefore unable to reject the null of no structural change.

[35] There is an offsetting effect coming from the steeper slope estimated by the HBPC (-0.37 versus -0.21 for the PC). Because the unemployment gap is so large in early 2020, the overall predicted value for the HBPC model is below that of the PC model in April 2020, in spite of the positive expectation effects mentioned in the text.

of the complexity in the relationship between expectations and inflation that becomes apparent when heterogeneity is accounted for.

6 Conclusion

Expectations are widely believed to be a core driver of inflation, and survey expectations are frequently used in empirical work. Our analysis of household survey data for US and UK showed that although beliefs about future inflation held by different agents can be highly heterogeneous, they can nonetheless be described by an interpretable factor structure. Disagreement, skew, and distributional 'shape' emerged as the principal forces driving the evolution of beliefs over time. We set out how under general forms of heterogeneity in beliefs the New Keynesian Phillips curve is modified by an additional term that depends on the cross-section distribution of inflation expectations, and proposed a straightforward approach to estimating and testing an extended version of it. Our principal findings lend support to a heterogeneous beliefs Phillips curve. As a corollary, they demonstrate that usual approach of summarizing survey beliefs by a single 'consensus' or average forecast is insufficient. We report that a robust association exists between the distribution of household beliefs about future inflation and actual inflation, even after accounting for average expected inflation, trend inflation, and the usual controls for supply factors. Allowing for heterogeneity eliminates any economically or statistically relevant role for lagged inflation in the Phillips curve. The contribution of heterogeneous beliefs to inflation is estimated to be largest at times of macroeconomic disruption, such as changes in monetary regimes, and the global financial crisis. Work to develop models that can account both for time-varying distributions of beliefs and for the influence of those distributions on inflation is needed to refine our understanding of the findings we report.

References

Adam, K., and M. Padula (2010): "Inflation dynamics and subjective expectations in the United States," *Economic Inquiry*, 49(1), 13–25.

Andrade, P., R. K. Crump, S. Eusepi, and E. Moench (2016): "Fundamental disagreement," *Journal of Monetary Economics*, 83, 106–128.

Andrade, P., G. Gaballo, E. Mengus, and B. Mojon (2019): "Forward guidance and heterogeneous beliefs," *American Economic Journal: Macroeconomics*, 11(3), 1–29.

Andrade, P., and H. Le Bihan (2013): "Inattentive professional forecasters," *Journal of Monetary Economics*, 60(8), 967–982.

Ang, A., G. Bekaert, and M. Wei (2007): "Do macro variables, asset markets, or surveys forecast inflation better?," *Journal of Monetary Economics*, 54, 1163–1212.

Binder, C. C. (2015): "Whose expectations augment the Phillips curve?," *Economics Letters*, 136, 35–38.

——— (2017): "Measuring uncertainty based on rounding: New method and application to inflation expectations," *Journal of Monetary Economics*, 90, 1–12.

Bowsher, C. G., and R. Meeks (2008): "The dynamics of economic functions: Modeling and forecasting the yield curve," *Journal of the American Statistical Association*, 103(484), 1419–1437.

Branch, W. A., and B. McGough (2009): "A New Keynesian model with heterogeneous expectations," *Journal of Economic Dynamics and Control*, 33, 1036–1051.

Brissimis, S. N., and N. S. Magginas (2008): "Inflation forecasts and the New Keynesian Phillips curve," *International Journal of Central Banking*, 4(2), 1–22.

Cavallo, A. (2021): "Inflation with Covid consumption baskets," *Covid Economics*, 72, 121–135.

Cecchetti, S. G., M. E. Feroli, P. Hooper, A. K. Kashyap, and K. L. Schoenholtz (2017): "Deflating inflation expectations: The implications of inflation's simple dynamics," Discussion paper, University of Chicago Booth School of Business.

Chaudhuri, K., M. Kim, and Y. Shin (2016): "Forecasting distributions of inflation rates: the functional auto-regressive approach," *Journal of the Royal Statistical Society, Series A*, 179(1), 65–102.

Cogley, T., and A. M. Sbordone (2008): "Trend inflation, indexation, and inflation persistence in the New Keynesian Phillips curve," *American Economic Review*, 98(5), 2101–2126.

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

——— (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., Y. Gorodnichenko, and R. Kamdar (2018): "The formation of expectations, inflation and the Phillips curve," *Journal of Economic Literature*, 56(4), 1447–1491.

Cuevas, A. (2014): "A partial overview of the theory of statistics with functional data," *Journal of Statistical Planning and Inference*, 147, 1–23.

Curtin, R. T. (1996): "Procedure to estimate price expectations," Surveys of Consumers, University of Michigan.

Faraway, J. (2001): "Comment on Kneip and Utikal," *Journal of the American Statistical Association*, 96(454), 537.

Faust, J., and J. H. Wright (2013): "Forecasting inflation," in *Handbook of Economic Forecasting*, ed. by G. Elliott, and A. Timmermann, vol. 2A, pp. 2–56. Elsevier.

Filardo, A., and H. Genberg (2010): "Targeting inflation in Asia and the Pacific: Lessons from the recent past," in *Twenty Years of Inflation Targeting: Lessons Learnt and Future Prospects*, ed. by D. Cobham, O. Eitrheim, S. Gerlach, and J. Qvigstad, pp. 144–171. Cambridge.

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

Fuhrer, J. C. (2011): "Inflation persistence," in *Handbook of Monetary Economics*, vol. 3A, chap. 11, pp. 423–486. Elsevier.

——— (2017): "Expectations as a source of persistence: Evidence from survey expectations in a dynamic macro model," *Journal of Monetary Economics*, 86, 22–35.

Galí, J., and M. Gertler (1999): "Inflation dynamics: A structural econometric analysis," *Journal of Monetary Economics*, 44(2), 195–222.

Harris, M., and A. Raviv (1993): "Differences of opinion make a horse race," *Review of Financial Studies*, 6(3), 473–506.

Horváth, L., and P. Kokoszka (2012): *Inference for Functional Data with Applications*, Springer Series in Statistics. Springer-Verlag, New York.

Jolliffe, I. T. (2002): *Principal Component Analysis*, Springer Series in Statistics. Springer-Verlag, New York, second edn.

Kneip, A., and K. J. Utikal (2001): "Inference for density families using functional principal component analysis," *Journal of the American Statistical Association*, 96(454), 519–532.

Kong, D., A.-M. Staicu, and A. Maity (2016): "Classical testing in functional linear models," *Journal of Nonparametric Statistics*, 28(4), 813–838.

Kurz, M., G. Piccillo, and H. Wu (2013): "Modeling diverse expectations in an aggregated New Keynesian model," *Journal of Economic Dynamics and Control*, 37(8), 1403–1433.

López-Pintado, S., and J. Romo (2009): "On the concept of depth for functional data," *Journal of the American Statistical Association*, 104(486), 718–734.

Mankiw, N. G., R. Reis, and J. Wolfers (2003): "Disagreement about inflation expectations," in *NBER Macroeconomics Annual*, ed. by M. Gertler, and K. Rogoff, vol. 18, pp. 209–248. University of Chicago Press.

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

Park, J. K., and J. Qian (2012): "Functional regression of continuous state distributions," *Journal of Econometrics*, 167, 397–412.

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

Pfajfar, D., and E. Santoro (2010): "Heterogeneity, learning and information stickiness in inflation expectations," *Journal of Economic Behavior & Organization*, 75(3), 426–444.

Poirier, D. J. (1973): "Piecewise regression using cubic spline," *Journal of the American Statistical Association*, 68(343), 515–524.

Ramsay, J. O., and B. W. Silverman (2005): *Functional Data Analysis*, Springer Series in Statistics. Springer-Verlag, New York, second edn.

Reiss, P. T., J. Goldsmith, H. L. Shang, and R. T. Ogden (2017): "Methods for scalar-on-function regression," *International Statistical Review*, 85(2), 228–249.

Reiss, P. T., and R. T. Ogden (2007): "Functional principal component regression and functional partial least squares," *Journal of the American Statistical Association*, 102(479), 984–996.

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

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

Tsay, R. S. (2016): "Some methods for analysing big dependent data," *Journal of Business and Economic Statistics*, 34(4), 673–688.

A An introduction to functional regression

This section provides a condensed primer on functional regression. The literature on estimation of the functional linear model is extensive. An excellent treatment of functional principal component regression may be found in Reiss and Ogden (2007), with Reiss, Goldsmith, Shang, and Ogden (2017) providing an up-to-date survey. A textbook treatment of estimation and inference in the functional linear model is given by Horváth and Kokoszka (2012), while the particular approach to inference we adopt is due to Kong, Staicu, and Maity (2016).

Although various formalizations of functional data are found in the literature (Cuevas, 2014, Section 2.3), we follow common practice and take X to be a measurable function in a sample space $L^2(\mathcal{I}), \mathcal{I} \subset \mathbb{R}$ defined on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$. The real-valued scalar random variable Y is defined on the same probability space as X. We have a sample $(y_t, \mathsf{x}_t), t = 1, \ldots T$ drawn from (Y, X). The scalar-on-function (SOF) regression model is defined as:

$$y_t = m_y + \int \gamma(\iota)\mathsf{x}_t(\iota)\mathrm{d}\iota + \varepsilon_t, \qquad \varepsilon_t \sim \text{i.i.d.}(0, \sigma^2)$$

where γ is a square integrable function, $\|\gamma^2\| < \infty$, and ε is independent of x. Here and elsewhere integration is over $\mathcal{I}$. We express the functional regressor in terms of its Karhunen-Loève expansion, truncated at the Kth term:

$$\mathsf{x}_t(i) = \sum_{k=1}^{K} s_{kt}\mathsf{e}_k(i)$$

where the principal component scores $s_{kt} = \langle \mathsf{x}_t, \mathsf{e}_k \rangle$ satisfy $\mathbb{E}[s_{kt}] = 0$, $\mathbb{E}[s_{kt}^2] = \lambda_k$, and $\mathbb{E}[s_{kt}s_{k't}] = 0, k \neq k'$. As we observe only T curves, there are at most $T-1$ non-zero eigenvalues, so we must choose $K \leq T-1$. Expand the coefficient function in the same basis to obtain:

$$\gamma(i) = \sum_{k'=1}^{K} \gamma_{k'}\mathsf{e}_{k'}(i)$$

We may then express the integral in the SOF model as:

$$\begin{aligned} \int \left(\sum_{k'=1}^{K} \gamma_{k'}\mathsf{e}_{k'}(i) \right) \left(\sum_{k=1}^{K} s_{kt}\mathsf{e}_k(i) \right) \mathrm{d}i &= \sum_{k=1}^{K} \gamma_k s_{kt} \int \mathsf{e}_k(i)^2 \mathrm{d}i \\ &= \sum_{k=1}^{K} \gamma_k s_{kt} \end{aligned}$$

where the first line follows from $\langle \mathsf{e}_k, \mathsf{e}_{k'} \rangle = 0, k \neq k'$, and the second line follows from $\|\mathsf{e}_k\| = 1$. Making the above substitution, the SOF model may be written as a multiple regression:

$$y_t = m_y + \sum_{k=1}^{K} \gamma_k s_{kt} + \varepsilon_t \tag{A.1}$$

The normal equations for the γs are then immediately seen to be:

$$0 = \sum_{t=1}^{T} s_{jt} \left\{ (y_t - m_y) - \sum_{k=1}^{K} \gamma_k s_{kt} \right\}, \quad j = 1, \ldots, K$$

Recalling that the scores are orthogonal, and that the variance of the jth score is equal to the jth eigenvalue, it is easy to see that:

$$\hat{\gamma}_j = \frac{c_{y,s_k}}{\lambda_j} \tag{A.2}$$

where $c_{y,s_k} = \sum_t (y_t - m_y) s_{jt}$ is the sample covariance between the dependent variable and the jth score. It follows that our estimate of the functional coefficient will be given by:

$$\hat{\gamma}(i) = \sum_{k=1}^{K} \frac{c_{y,s_k}}{\lambda_j} \mathsf{e}_k(i) \tag{A.3}$$

As we have seen, SOF regression using FPCs reduces to multiple regression, so extending the model to include scalar covariates, as in our application, is rather routine.

B Computing functional principal components

This section gives the computational results necessary to compute the functional principal components used throughout this paper. The basic approach is to replace functions with linear combinations of basis functions. The material, which is standard, draws on Ramsay and Silverman (2005, Section 8.4).

The FPCA problem is to find to minimize the integrated squared error criterion:

$$\mathsf{ISE}_{t,h}^{(K)} = \int \left\{ \left(\hat{\mathsf{p}}_{t,h}^{(K)} - \overline{\mathsf{p}}_h \right) - \left(\mathsf{p}_{t,h}^{\mathsf{c}} - \overline{\mathsf{p}}_h \right) \right\}^2 \mathrm{d}\iota, \quad \text{where} \quad \hat{\mathsf{p}}_{t,h}^{(K)} = \overline{\mathsf{p}}_h + \sum_{k=1}^{K} s_{kt} \mathsf{e}_k \tag{B.1}$$

averaged over all t, subject to the constraint that the functions $\mathsf{e}(\cdot)$ satisfy $\|\mathsf{e}_k\| = 1$ for all k, and $\langle \mathsf{e}_k, \mathsf{e}_j \rangle = 0, k \neq j$.

Let the functions $\{\mathsf{x}_t(i)\}_1^T$ be defined as in Appendix A. The eigenequation of the covariance operator $V(\mathsf{x})(\cdot)$ is:

$$\int v(i,j) \mathsf{e}_k(i) \mathrm{d}j = \lambda_k \mathsf{e}_k(i) \tag{B.2}$$

Now let the basis expansion of the x_t be:

$$\mathsf{x}_t(i) = \sum_{k=1}^{K} c_{tk} \phi_k(i)$$

or, stacking by t:

$$\mathbf{x}(i) = \mathbf{C}\boldsymbol{\phi}(i), \qquad \underset{(T\times K)}{\mathbf{C}} = [c_{tk}] \quad \text{and} \quad \underset{(K\times 1)}{\boldsymbol{\phi}} = [\phi_k]$$

We may then express the sample covariance function as:

$$v(i,j) = (T-1)^{-1} \boldsymbol{\phi}(i)^\top \mathbf{C}^\top \mathbf{C} \boldsymbol{\phi}(j) \tag{B.3}$$

Assume that the eigenfunctions have the basis expansion:

$$\mathsf{e}(i) = \sum_{k=1}^{K} b_k \phi_k(i) = \boldsymbol{\phi}(i)^\top \mathbf{b}, \qquad \underset{(K\times 1)}{\mathbf{b}} = [b_k]$$

Then substituting (B.3) into (B.2), the eigenequation may be written:

$$(T-1)^{-1}\boldsymbol{\phi}(i)^\top \mathbf{C}^\top \mathbf{C}\mathbf{W}\mathbf{b} = \lambda \boldsymbol{\phi}(i)^\top \mathbf{b} \tag{B.4}$$

where the symmetric $(K \times K)$ matrix $\mathbf{W} = \int \boldsymbol{\phi}(i)\boldsymbol{\phi}(i)^\top$ is a matrix of inner products of the basis functions $\boldsymbol{\phi}_k(\cdot)$, and λ is the eigenvalue corresponding to e. Observing that (B.4) must hold for all i implies that a solution to (B.2) may be obtained from the solution to the symmetric matrix eigenvalue problem:

$$(T-1)^{-1}\mathbf{W}^{1/2}\mathbf{C}^\top \mathbf{C}\mathbf{W}^{1/2}\mathbf{u} = \lambda \mathbf{u}, \qquad \mathbf{u} = \mathbf{W}^{-1/2}\mathbf{b}$$

using standard methods. For an alternative approach that applies standard PCA to the grid of G values $\{\mathsf{p}_{t,h}(x_i) | i = 1, \ldots, G; t = 1, \ldots, T\}$, see Tsay (2016, Section 3.3).

C A New Keynesian Phillips Curve with heterogeneous beliefs

This section briefly sets out the steps that lead to a micro-founded NKPC of the form given in Eq. (8). Standard assumptions of monopolistic competition and time-dependent pricing lead to a log-linear expression for firm j's optimal reset price, $p^{\star(j)}$, in terms of current and expected future nominal marginal costs (log real marginal cost φ plus the log of the aggregate price level p):

$$p_t^{\star(j)} = (1-\beta\theta)\mathbb{E}_t^{(j)}\left[\textstyle\sum_{\tau=t}^{\infty}(\beta\theta)^{\tau-t}(\varphi_\tau + p_\tau)\right] \tag{C.1}$$

where θ is the per-period probability that a firm's price cannot be reset, and β is the discount factor applied to future profits. Firms hold diverse expectations, as indicated by the superscript on the expectations operator; see for example Kurz, Piccillo, and Wu (2013, Section 3), Mavroeidis, Plagborg-Møller, and Stock (2014, Section 2.1), and Andrade, Gaballo, Mengus, and Mojon (2019, Appendix 2.3). Note that differences between firms' optimal price $p^\star$ arise only because of differences in expectations.

To express the infinite sum on the right of Eq. (C.1) in the form of a difference equation, it is sufficient to assume the law of iterated expectations applies to the expectation operators $\mathbb{E}_t^{(j)}[\cdot]$ used by each firm j. Although not innocuous, this assumption retains tractability without requiring stronger assumptions on the dependence (or lack thereof) of beliefs in the cross-section (see Coibion, Gorodnichenko, and Kamdar, 2018, p. 1466-7; Branch and McGough, 2009, Section 2.1). Applying this assumption, then subtracting the current aggregate price level from both sides gives an expression for the optimal real, or relative, price:

$$q_t^{\star(j)} = (1-\beta\theta)\varphi_t + \beta\theta\mathbb{E}_t^{(j)}\left[q_{t+1}^{\star(j)} + p_{t+1} - p_t\right] \quad \text{for} \quad q_t^{\star(j)} := p_t^{\star(j)} - p_t$$

Thinking of the integral notation as representing the limit of the sum over the dimension N cross section of beliefs (alternatively denoted $\lim_N \mathbb{E}_N$) we aggregate using the notation that $q := \int q^{\star(j)}$ and that for a random variable x not indexed by j, $\overline{\mathbb{E}}[x] := \int \mathbb{E}^{(j)}[x]$ to obtain:

$$
\begin{aligned}
q_t &= (1-\beta\theta)\varphi_t + \beta\theta \int \left(\mathbb{E}_t^{(j)}[q_{t+1}^{\star(j)}] + \mathbb{E}_t^{(j)}[p_{t+1} - p_t] \right) \\
&= (1-\beta\theta)\varphi_t + \beta\theta \overline{\mathbb{E}}_t[\pi_{t+1} + q_{t+1}] + \beta\theta \int \left(\mathbb{E}_t^{(j)}[q_{t+1}^{\star(j)} - q_{t+1}] \right)
\end{aligned}
$$

where π denotes inflation and a multiple of q (the aggregate reset price) has been added and subtracted in the second line. Exploiting the relationship between reset prices and inflation $q = \theta/(1-\theta)\pi$, which arises is the usual way from the linearized constant substitution elasticity price aggregator, allows us to write an expression in terms of marginal cost, average expected inflation, and a new term in beliefs:

$$
\pi_t = \frac{(1-\beta\theta)(1-\theta)}{\theta}\varphi_t + \beta\overline{\mathbb{E}}_t(\pi_{t+1}) + \beta\theta\Delta_t, \quad \Delta_t = \int \left(\mathbb{E}_t^{(j)}[\pi_{t+1}^{\star(j)} - \pi_{t+1}] \right) \tag{C.2}
$$

In the expression for Δ, the term $\pi^{\star(j)}$ bears the same relation to $q^\star$ as π does to q.[36] This term vanishes only when firms' beliefs about their optimal reset price coincide on average with their beliefs about the economy-wide reset price.[37] In the text we refer to expressions of the form Eq. (C.2) as the *heterogeneous beliefs* Phillips curve because compared to the standard NKPC there is an additional term Δ that represents time-varying beliefs.

[36]The substitution in terms of inflation is made only for convenience, and Δ could equally be expressed in terms of reset prices as on a preceding line.

[37]Conditions under which $\Delta_t = 0$ for all t include full information rational expectations and those given by Andrade, Gaballo, Mengus, and Mojon (2019, pp. 13–14). Conditions that produce variation in Δ_t include those given by Kurz, Piccillo, and Wu (2013, Section 4).