Heterogeneous beliefs in the Phillips curve*

Roland Meeks† and Francesca Monti‡

February 13, 2019

Abstract

We propose a new approach to estimating forward-looking models of inflation using survey expectations that overcomes the problem of response aggregation. We capture the information in individual-level responses via smooth distributions over entire sets of forecasts. In our generalized Phillips curve, distributions of expectations are aggregated using a flexible function in a functional linear model. We show that variation in the distributions may be summarized by disagreement, skew, and shape factors, which leads to estimation via functional principal component regression. By using the information in micro data more effectively, our approach reveals an enhanced role for expectations in inflation dynamics that is robust to lagged inflation, trend inflation, and supply factors. Our findings are shown to hold in similar form across two major economies.

Keywords:
Survey expectations; Inflation dynamics; Density function; Functional regression; Functional principal components

1 Introduction

According to standard models of price-setting behaviour, what agents believe about where inflation is headed helps to determine inflation today (Friedman, 1968).[1] The predicted dependence of out-turns on beliefs finds support in tests of forward-looking models of inflation that include direct observations on expectations obtained from surveys (Roberts, 1995; Nunes, 2010; Fuhrer, 2017; Pfajfar and Roberts, 2018). Using survey forecasts has proved empirically successful, as Coibion et al. (2018) document, and circumvents—at least partially—the econometric traps that plague full information rational expectations implementations of the Phillips curve. These advantages have given the survey approach a 'commanding' presence in the literature (Mavroeidis et al., 2014, p. 151). However, some difficulties remain. Amongst these is the 'aggregation problem'—whose beliefs to use, and how to summarize them (Yellen, 2016).

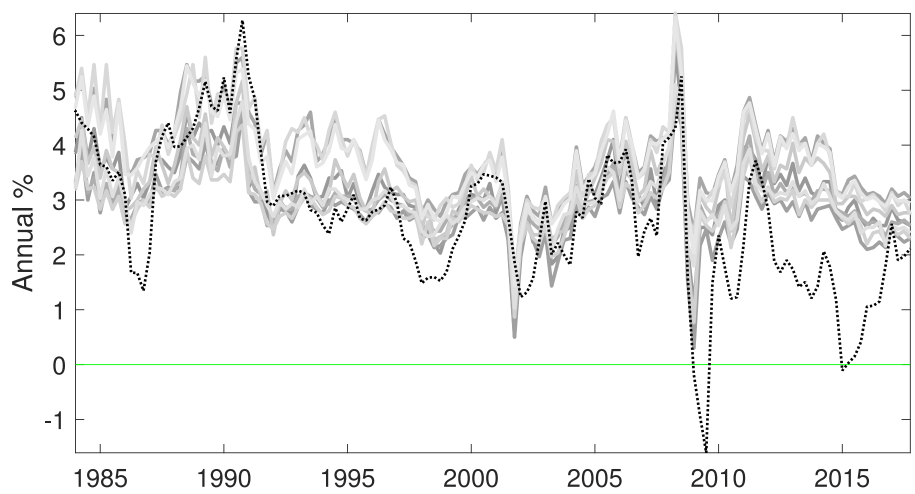
*We are indebted to John Barrdear, Ferre De Graeve, Paul Hubert, Mike Kiley, Emmanuel Moench, Paolo Surico, and especially to Ricardo Reis for helpful discussions on this paper. We also thank conference and workshop participants at Lancaster Business School, London Business School, the University of Sussex, the University of Oxford, Sciences-Po/OFCE, and the Bank of England. The views expressed in this paper are those of the authors, and should not be taken to represent the official views of either the Bank of England or the International Monetary Fund.

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

‡Bank of England and CfM. Email: francesca.monti@bankofengland.co.uk

[1] We use the terms 'beliefs' (about the future), 'forecasts', and 'expectations' interchangeably in this text.

Figure 1. Inflation and Michigan survey inflation expectations aggregated using alternative assumptions



Note: The annual percentage change in CPI inflation ($\cdots$) is shown against the year-ahead expected inflation rate for responses aggregated by: (i) All respondents; (ii) Working age only; (iii) Male only; (iv) Highest income tercile only; (v) Truncating responses at $(-5, 30)$ percent rather than $(-10, 50)$; means and medians plotted for all measures. Sub-sample shown is 1984-Q1 to 2017-Q4. Dates shown reflect when forecasts were made.

Fig. 1 illustrates that the apparently simple question: 'What is expected inflation?' can have a broad range of answers, even within a *single* survey of expectations. The answer may depend, for example, on what type of average is chosen, the socioeconomic characteristics of individual respondents, or their housing tenure (Lombardelli and Saleheen, 2003). Such choices matter for Phillips curve estimates (Binder, 2015), but the literature has yet to settle on a single approach, most likely because theory provides little in the way of guidance.

In this paper we propose a straightforward approach to using survey forecast data in models of inflation that, we argue, mostly overcomes these problems. It does so by avoiding prior aggregation of the underlying micro data altogether. In our functional linear model (FLM; see Ramsay and Silverman, 2005) smooth cross-section distributions of survey forecasts enter the Phillips curve, and are aggregated into a scalar index of beliefs using a flexible aggregator function. The function appears along with the other parameters to be estimated in the model. Functional principal components capture variation in the distributions of beliefs, and are used to estimate the FLM via principal component regression (Reiss and Ogden, 2007). Our method has several appealing properties. It brings all the available micro data to bear in estimation. We retain the same New Keynesian model of inflation studied in the literature, and the same expectations data, yet improve on the standard estimation approach by using that data more effectively. A simple method follows for testing whether a particular summary of expectations captures all the information relevant to current inflation. For example in our applications, we nest as a special case the standard approach of using the mean forecast. And last, because it does not require any ancillary information on the characteristics of individual respondents, it is applicable to any set of survey expectations.

The payoff to our method is the discovery of an enhanced role for expectations in the inflation process. We estimate the generalised expectations-augmented Phillips curve on household inflation forecasts reported in the US Michigan survey, and on a newly-collated UK household survey. We show that signals about future inflation contained in belief distributions affect current inflation even after accounting for average expected inflation, lagged inflation, trend inflation, and supply factors. Our tests of the hybrid (forward- and backward-looking) Phillips curve show that fully accounting for expectations eliminates intrinsic persistence in inflation (Fuhrer, 2011). Moreover, omitting distributions of beliefs leads to poorly-specified models of trend inflation that are inconsistent with theory. But our estimates of the functional model indicate a very strong role for near-term expectations, even after accounting for beliefs about long-run inflation.

The paper also documents a set of novel facts about the distribution of survey forecasts. Beliefs about future inflation are highly heterogeneous, particularly during periods of macroeconomic turmoil. However, we show that a handful of factors, which correlate with the cross-section dispersion of forecasts—known in the literature as 'disagreement'—along with their skew, and a third factor we call 'shape', can jointly characterize much of the variation present in the distribution of beliefs. To date, most studies have focused on disagreement, following the early work of Mankiw et al. (2003). The observed dispersion in expectations is understood to result from imperfect or sticky information (Coibion and Gorodnichenko, 2012), or from combinations of the two (Andrade and Le Bihan, 2013), and can exist even at long horizons thanks to differences in prior beliefs (Patton and Timmermann, 2010), or problems in disentangling short- and long-run shifts in the inflation process (Andrade et al., 2016). A handful of studies have looked beyond summary measures of disagreement to complete distributions of forecasts (Filardo and Genberg, 2010; Pfajfar and Santoro, 2010), but none that we know of have characterised their structure as we do here.

Finally, 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). 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 application to yield curve forecasting in Bowsher and Meeks (2008), the model of relative price dispersion and inflation in Chaudhuri et al. (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 briefly introduces the survey data that we use in our main analysis, and details how we construct estimates of belief distributions. Section 3 summarizes the main sources of variation in the expectations data using functional principal component analysis. Section 4 sets out our heterogeneous beliefs Phillips curve model, and the econometric approach we adopt to estimate the effects of heterogeneity on inflation. It also provides our headline results, with separate treatment of the US and UK Phillips curves. The economic implications of the heterogenous beliefs model appear in Section 5, including those on the hybrid Phillips curve, and those on modeling the gap between inflation and its trend. Finally, Section 6 offers concluding comments.

2 Summarizing survey forecasts with time series of distributions

In this paper, we study inflation expectations in the United States and the United Kingdom, two countries for which relatively long-running household inflation surveys exist. In the supplementary material, we detail much of the same analysis for professional forecasters, and record noteworthy differences as they arise below. Unfortunately, producer surveys of comparable length and quality are not available.

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, 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 (point forecast) for inflation over the following year, and their expectations for at least one other horizon.[2] Quarterly data is available, with the longer time series—spanning a period from the late 1970s or mid 1980s—available for the US and the UK respectively. A summary of the main features survey data used in this study is given in Tab. A.1 of the Appendix.

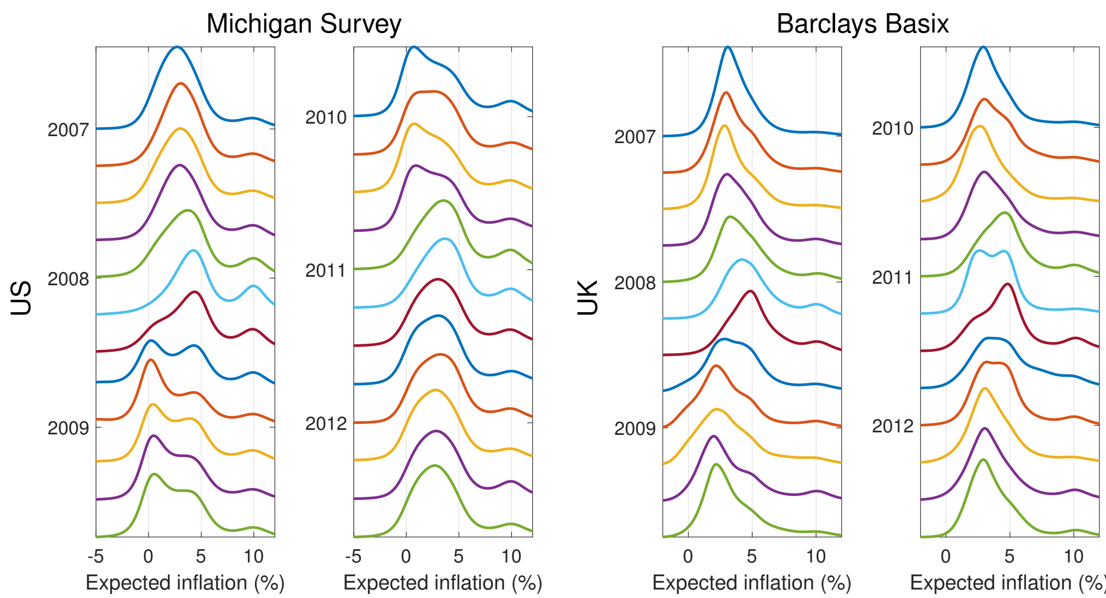
2.2 Estimating distributions of survey forecasts

The first step in our analysis to transform the discrete cross-section of point expectations reported by survey respondents into continuous distribution functions.[3] Dealing with functions is one way to overcome the problem of dimensionality, and allows a degree of smoothing—or regularization—that will later prove to be helpful. In each survey quarter, we adopt a

[2] 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".

[3] 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.

Figure 2. Cross-sections of beliefs about future inflation during the financial crisis and its aftermath



Note: Panels show time series of distributions of individual survey respondents' point forecasts from each of the named surveys. Dates reflect when forecasts were made. Details of the estimation method may be found in the main text and in Appendix B.

nonparametric technique to obtain consistent estimates of that distribution.[4] 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).

Our first hint that there is rich information contained in the cross-section of beliefs—information that we will go on to show is exploitable for modelers—is given in Fig. 2. We plot the functional time series data obtained using our method over the sub-sample 2007-Q1 to 2012-Q4, which includes the financial crisis. Notable changes in the location and scale of belief distributions are observed across both surveys during 2008. But other variation is apparent also, for example the emergence of distinct modes around 5% and 10% inflation. These are not artifacts of the density estimation method, but a result of known biases towards reporting round numbers when respondents are uncertain (Binder, 2017). The following section sets out an approach to analysing these complex patterns of temporal and cross sectional functional variation.

3 The structure of survey expectations

The distributions described in the preceding section capture the information in hundreds of thousands of survey responses, reported over several decades. This section investigates

[4]Details of the penalized maximum likelihood (pML) approach we adopt are described in Appendix B. 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).

the statistical properties of those distributions. 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.

3.1 Average distributions

What shape does the distribution of expectations take, on average? An interpretable answer requires us to align the distributions shown in Fig. 2 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 of h-step ahead point forecasts is then given by:[5]

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

where the reader will have seen that $\mathsf{p}_{t,h}^{\mathsf{c}}$ represents the centered distribution of 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).[6] Given an empirical distribution 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.[7]

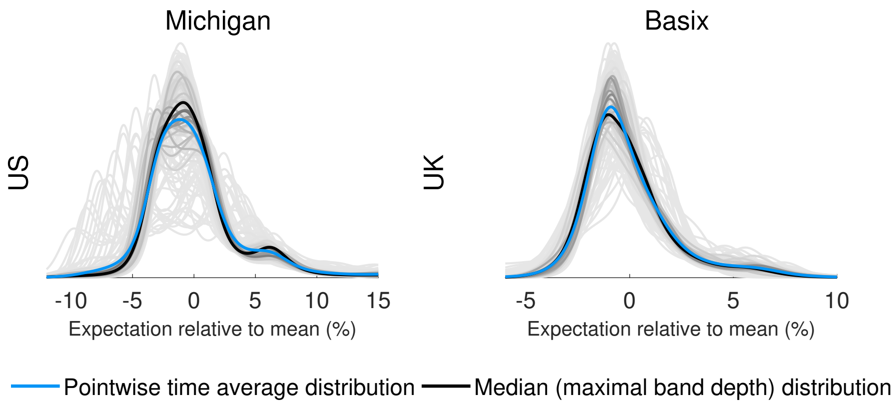
The average shapes of the distribution functions display remarkable similarities across the two regions. Fig. 3 displays the time averages of the centered density functions for both surveys (bold lines), overlaid with the cross-sectional densities for every time period (thin lines). For the latter, lighter 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. 3, Col. 1), somewhat higher than the 2.3 percent seen in the UK (Fig. 3, Col. 2), since

[5]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).

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

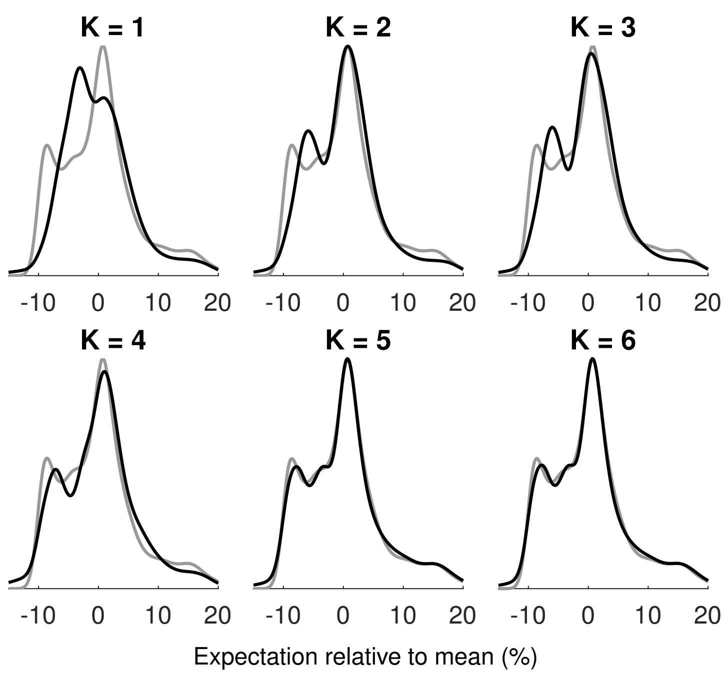
[7]The concept of band depth is based on the graph of a function on the plane. A band can be thought of as the envelope delimited by n such graphs. The band depth of a given curve p_0 is given by the proportion of times that it falls inside the bands formed by taking all possible combinations of n curves. 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.

Figure 3. 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 A.1.

Figure 4. 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 $\hat{p}^{(K)}_{1979\text{-Q3},4}$, $K = 1, \ldots, 6$, defined in Eq. (2). The magnitudes of the associated integrated squared errors are $\log_{10}[\mathsf{ISE}^{(K)}_{1979\text{-Q3},4}] = \{-2.22, -2.71, -2.71, -3.06, -3.83, -3.84\}$.

the former includes observations from the high-inflation period of the late 1970s while the latter does not. 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.[8] 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. We have often encountered the view that the presence of a substantial group of households with inflation beliefs that, say, lie far above the average should raise doubts over the usefulness of the data. However, we are persuaded by several sophisticated studies that household expectations are neither the product of gross irrationality (Pfajfar and Santoro, 2010; Malmendier and Nagel, 2016), nor irrelevant for actual inflation (Coibion and Gorodnichenko, 2015; Pfajfar and Roberts, 2018).

3.2 Principal component analysis

Cross-sectional distributions of survey forecasts display considerable variation around their means (Fig. 3). A natural question to ask is whether that variation can be effectively summarized using a smaller number of functions. Functional principal component analysis is a standard technique for dimension reduction in functional data sets, and may be applied to our distributions (Kneip and Utikal, 2001). The representation of distribution functions in terms of their 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.[9] 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:

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

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 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 D (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. 4 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

[8] By contrast, the average distribution of professional forecaster beliefs are almost perfectly symmetric about the mean.

[9] FPCA will also be central to the approach we adopt for the estimation of the functional linear model, in Section 4.2. For an even-paced introduction to FPCA that sets out the correspondences with PCA on multivariate data, see Ramsay and Silverman (2005, Ch. 8).

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, or better approximations, in every time period. Fig. 5 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. 4, towards being close to average, as in Fig. 3. Capturing shapes that are closer to the 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. 6. 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.[10]

3.3 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. Fig. 7 shows these belief factors along with empirical measures of disagreement (d_t), skew (κ_t), and a third factor we call 'shape' (which we will denote τ_t), 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 $\rho^{BBS}(s_{2t}, d_t) = .97$.[11]

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. 8. The shape factor is given by:

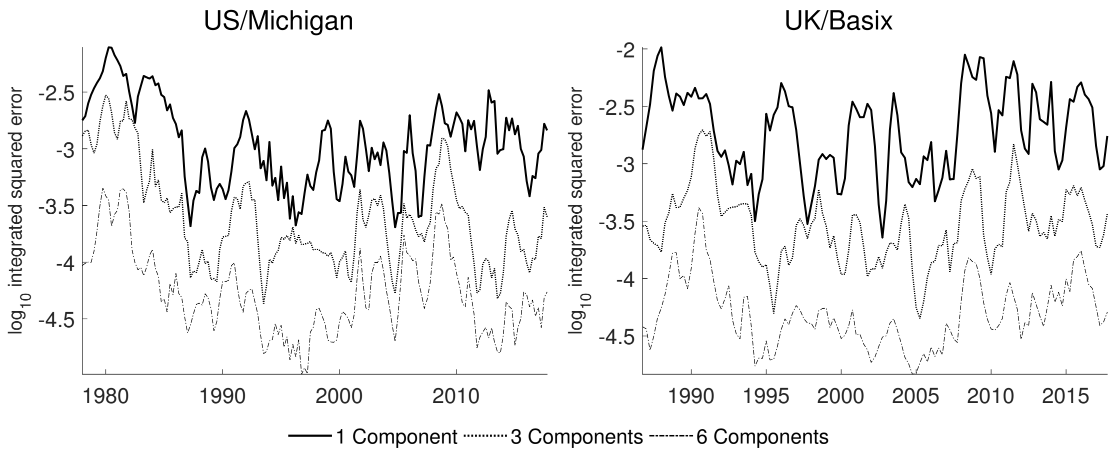
$$\tau_t(x_1, x_2, x_3) = \left[\mathsf{p}^{\mathsf{c}}_{t,h}(x_1) - \overline{\mathsf{p}}_h(x_1)\right] + \left[\mathsf{p}^{\mathsf{c}}_{t,h}(x_2) - \overline{\mathsf{p}}_h(x_2)\right] + \left[\mathsf{p}^{\mathsf{c}}_{t,h}(x_3) - \overline{\mathsf{p}}_h(x_3)\right] \tag{3}$$

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.

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

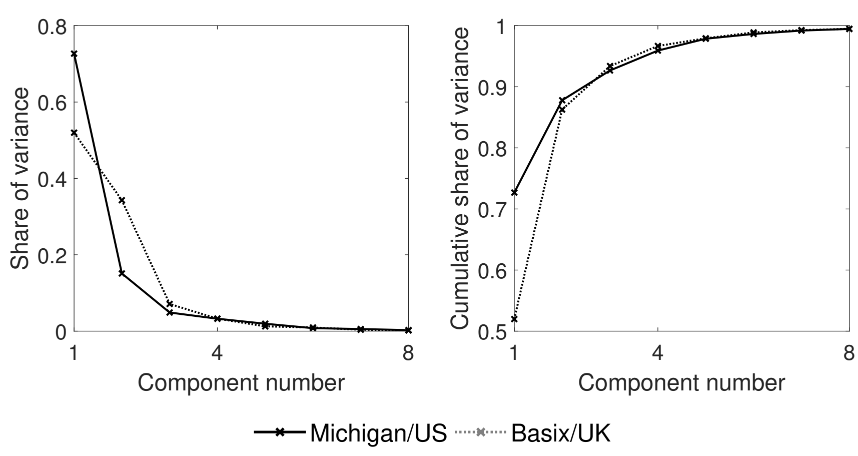
[11] We derive numerical values for standardized central moments and (combinations of) quantiles directly from the time series of distribution functions. Alternative measures of the same quantity are typically very similar: for example, 'disagreement' as the square root of the second moment or as the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

Figure 5. Integrated square errors of *K*-component approximations over time



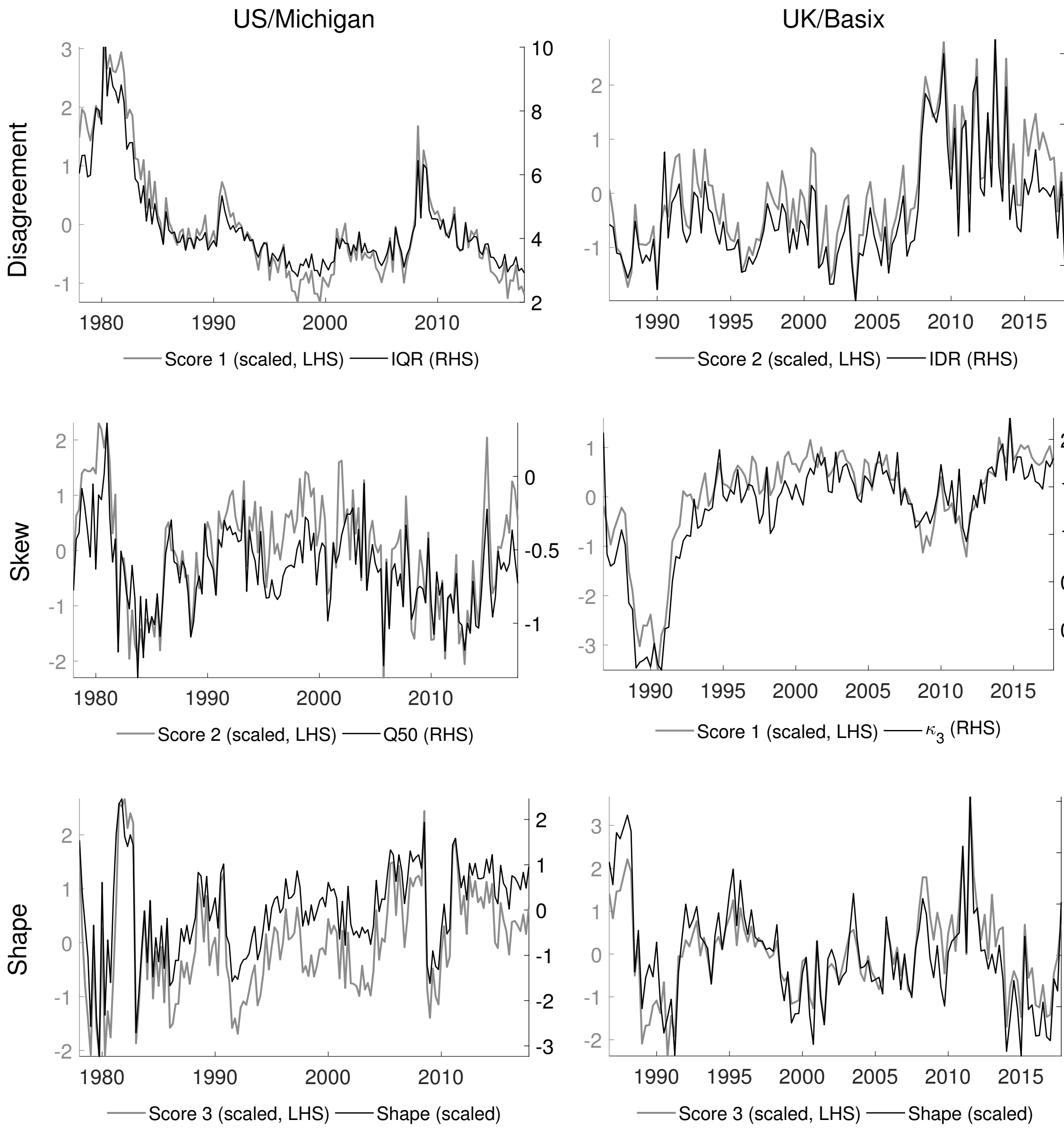
Note: The $\log_{10}$ integrated square error, Eq. (2), associated with a $K = \{1, 3, 6\}$ component expansion of the observed distributions of forecasts. US data is from the Michigan survey; UK data from the Basix survey. Centered three-quarter moving average.

Figure 6. Shares and cumulative shares of functional variation explained by the leading *K* principal components



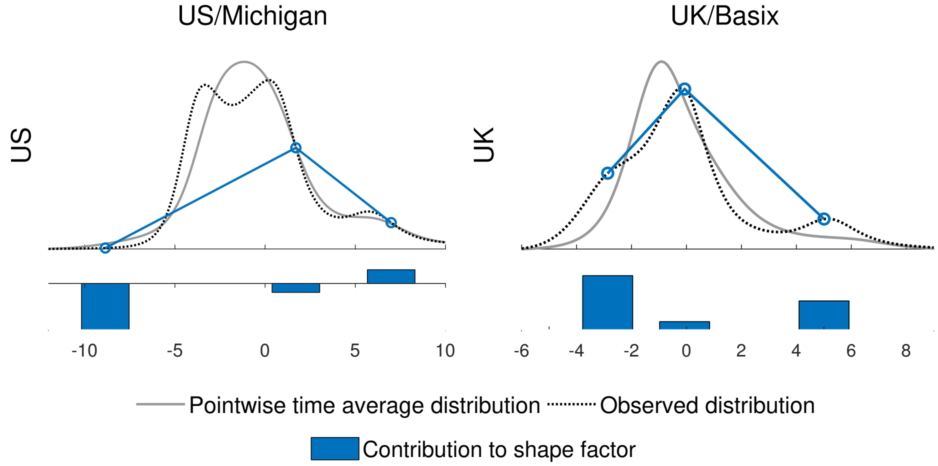
Note: Scree plot gives the normalized sums of eigenvalues of the covariance operator.

Figure 7. Belief factors and data-based disagreement, skew, and shape



Note: Belief factors are the leading functional principal component scores $\{s_{1t}, s_{2t}\}$. IQR: the inter-quartile range. IDR: the inter-decile range. Q50: the .5 quantile. κ_3: the standardized 3rd central moment of the distribution. Shape: $\tau_t(\mathbf{x})$ as given by Eq. (3), where $\mathbf{x} = (-8.9, 1.7, 7)$ (Michigan) and $\mathbf{x} = (-2.9, 0, 5)$ (Basix). Moments and quantiles are obtained from the estimated density functions via quadrature methods.

Figure 8. The shape factor



Note: The top panel of the Figure shows the average distribution given in Fig. 3 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. (3) 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.

Summary

In our analysis, disagreement emerges as a central factor driving the dynamics of belief distributions. For the US data, disagreement is the primary factor, accounting for close to 80 percent of the variance in the data, with only around 15 percent due to the skew factor, and 5 percent due to shape. But for 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.

4 Heterogeneous beliefs and inflation dynamics

In much of the macroeconomics literature, nominal rigidities in price-setting underpin the aggregate supply relation. When prices are sticky, expectations about future inflation matter for inflation today, as price setters anticipate how variation in the future relative price of their goods will affect their profitability. A popular empirical implementation of this variety of Phillips curve, first formulated in Roberts (1995), uses survey forecasts as a direct measure of expectations. Survey forecast-based models, which do not assume full information rational expectations, have been shown to address a number of empirical shortcomings present in standard implementations, that do (Coibion et al., 2018). The model takes the form:

$$\pi_t = \beta \overline{\pi}^e_{t,h} + \alpha(u_t - u^*_t) + \varepsilon_t, \qquad \text{where} \qquad \overline{\pi}^e_{t,h} := \int \pi^e_{t,h} \mathrm{dP}_{t,h} \tag{4}$$

with π denoting the annualized quarter-on-quarter percentage change in the price level, and π^e again denoting expected inflation between t and $t+h$.[12]

4.1 The heterogeneous beliefs model

A leading question when estimating the survey-based Phillips curve is whose whose expectations are most relevant for inflation. As Yellen (2016) notes, theory does not provide clear guidance on this point. The problem is often framed in terms of whether the beliefs of households or professional forecasters better represent those of producers. Coibion and Gorodnichenko (2015) argue that the matter can be settled by estimating a version of Eq. (4) that contains both types of forecast.[13] But there is another problem. Even within a population of respondents, some agents' beliefs seem to matter more than others for inflation (Binder, 2015).

To overcome the problem of aggregating responses into a single index of expectations, we propose a straightforward generalization of the Roberts formulation that accounts flexibly for heterogeneity in beliefs about future inflation. Our approach uses a functional linear model (Ramsay and Silverman, 2005, Ch. 15), in which interest centers on estimates of the function γ appearing in the generalized expectational Phillips relation:

$$\pi_t = \beta\overline{\pi}^e_{t,h} + \int \gamma(\pi^e)\mathsf{p}^{\mathsf{c}}_{t,h}(\pi^e)\,\mathrm{d}\pi^e + \alpha(u_t - u^*_t) + \varepsilon_t \tag{5}$$

We think of γ as an aggregator function, as it enters Eq. (5) under the integral and determines how the distribution of beliefs $\mathsf{p}^{\mathsf{c}}_{t,h}$—beliefs in this case centered on each period's average forecast—influences current price setting behaviour. Where $|\gamma|$ is large for some value of π^e, expectations in that region of the distribution have greater influence on inflation. Note that as the survey mean has been the focus of previous enquiries, we prefer to account for it as a separate scalar regressor in our estimation. We could equally have written the model in terms of the distributions in levels of expected inflation ($\mathsf{p}_{t,h}$), in which case the standard model would be given by $\gamma(x) = \beta x$.[14] For that reason, we will sometimes refer to the conventional expectations averaging approach as 'linear aggregation'.

4.2 Estimation

A variety of estimation approaches have been proposed for the functional linear model (see Reiss et al., 2017). We adopt the popular functional principal component regression approach, under which the functional regression Eq. (5) is recast as a multiple regression problem. To

[12]In commonly-adoped formulations with time-dependent pricing, h is set to 1. In empirical Phillips curves, data availability typically means h is set to 4 (quarters). Some authors impose a constant natural rate, which appears as an estimated intercept in Eq. (4); others replace the unemployment gap with the output gap, following the logic of Okun's law. Our results will not depend on that choice.

[13]They report that for the US, the average SPF forecast is statistically insignificant when the average Michigan survey expectation is present in the regression, consistent with household expectations being more representative of price setters.

[14]The first principal component of the uncentered distributions is close, but not equal, to the distribution mean. The conclusions we present continue to hold in the alternative formulation of the model in terms of levels.

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. (C.1), the functional linear model of Eq. (5) can be rewritten as:

$$\pi_t = \beta \overline{\pi}^e_{t,h} + \sum_{k=1}^{K} \gamma_k s_{k,t} + \alpha(u_t - u_t^*) + \varepsilon_t \tag{6}$$

where the γ_k are scalar coefficients to be estimated, and the functional principal component scores $s_{k,t}$ obtained in Section 3 appear as covariates.[15]

Having recast the functional linear model Eq. (5) as the multiple regression model Eq. (6), 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. (5) 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:

$$\hat{\boldsymbol{\gamma}} = \mathbf{Q}^{-1}\mathbf{M}^\top(\mathbf{I} - \mathbf{P}_Z)\boldsymbol{\pi} \tag{7}$$

where $\mathbf{Q} = (\boldsymbol{\Lambda} - \mathbf{M}^\top\mathbf{P}_Z\mathbf{M})$ is the Schur complement of $(\mathbf{Z}^\top\mathbf{Z})$ in $(\mathbf{X}^\top\mathbf{X})$, and $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_K)$ 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 et al. (2016). A natural null hypothesis is that $\gamma(\pi^e) = 0$, which recalling that the distributions p are mean zero by construction, corresponds to the special case where only the average forecast matters for inflation, just as in Eq. (4). 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 $H_0 : \gamma(\pi^e) = 0$ for all π^e is equivalent to:

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

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

Then H_0 can be tested using the F-statistic:

$$T_F = \frac{\boldsymbol{\pi}^\top(\mathbf{P}_X - \mathbf{P}_Z)\boldsymbol{\pi}/K}{\boldsymbol{\pi}^\top(\mathbf{I} - \mathbf{P}_X)\boldsymbol{\pi}/(T-K-N)} \overset{approx.}{\sim} F_{K,T-K-N} \tag{8}$$

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 et al., 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).[16] In the subsequent analysis, we select two values of K, one based on the simple cumulative eigenvalue test, and one based on the Bayes Information Criterion (BIC), which takes account of both fit and parameterization.

5 Economic implications of heterogeneous beliefs

5.1 How to aggregate expectations?

How important is the aggregation problem for survey expectations-augmented Phillips curves? We provide an answer to this question in the form of tests of the linear aggregation assumption, based on comparing results from the standard estimation approach with those from the variant with flexibly aggregated expectations. We consider identical models and estimation methods for the United States and United Kingdom.

Inflation in the United States

We estimate the conventional expectations-augmented Phillips curve given by Eq. (4) for the US using the CBO measure of the unemployment gap, and the survey average one-year-ahead expected inflation rate from the Michigan survey.[17] 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 is around 0.27, and is significant at 1%. The substance of these results is very similar to that reported in recently published work by Coibion et al. (2018), as they are based on an equivalent specification and a modestly extended sample.

Estimates for our heterogeneous beliefs model Eq. (5) indicate that information about

[16] 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.

[17] 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.

current inflation is contained in the distribution of beliefs, but that it is missed by using the simple average alone. Tab. 1 (Cols. 2–3) reports that the aggregation function is strongly significant in our Phillips curve regressions, a rejection of linear aggregation. The BIC selects three components, but remarkably the penalty for the model with six components is no larger than that for the model with none.[18] 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 sensitive to the specification of the functional effect. This result suggests that the information contained in these variables is not orthogonal, consistent with findings elsewhere that higher average expected inflation has a positive association with disagreement (Rich and Tracy, 2010). Our results are robust to including supply factors (Col. 4).[19]

The shape of the estimated aggregator function—the functional coefficient on $\mathsf{p}^{\mathsf{c}}_{t,h}$ in Eq. 5—indicates that shifts in the mass of respondents around the consensus expectation tend to be amplified. For example, Fig. 9 (left panel) indicates that when more forecasts than typical lie in the interval $[-4, 0]$, imparting a rightward skew to the belief distribution, expectations impart a greater-than-usual downward force on inflation. The opposite is true when the mass of forecasts lies in $[0, 4]$. It is important to bear in mind that these effects are in addition to the effect of the average expectation on inflation.

The overall effect of expectations on inflation, seen through the lens of this model, has been considerably less supportive of US inflation over the past decade than is commonly thought. Fig. 10 plots the contribution of the average expectation to the fit of the heterogeneous beliefs model (Tab. 1, Col. 2) along against the total expectation effect. The gap between the lines represents the contribution of the estimated functional component $\langle \hat{\gamma}, \mathsf{p}^{\mathsf{c}}_{t,h} \rangle$. As expected, during the Great Moderation that gap is generally small (since the distributions of beliefs were closer to their average shape). But both during the late 1970s and after 2007, a notable divergence can be observed. In the more recent period, household expectations imparted a substantial disinflationary impulse, contrary to the arguments in Coibion and Gorodnichenko (2015).[20]

Inflation in the United Kingdom

We estimated identically-specified models on UK data, again using year-ahead expectations data. Because no official measures of the natural rate of unemployment exist for the UK for the

[18]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 A.

[19]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.

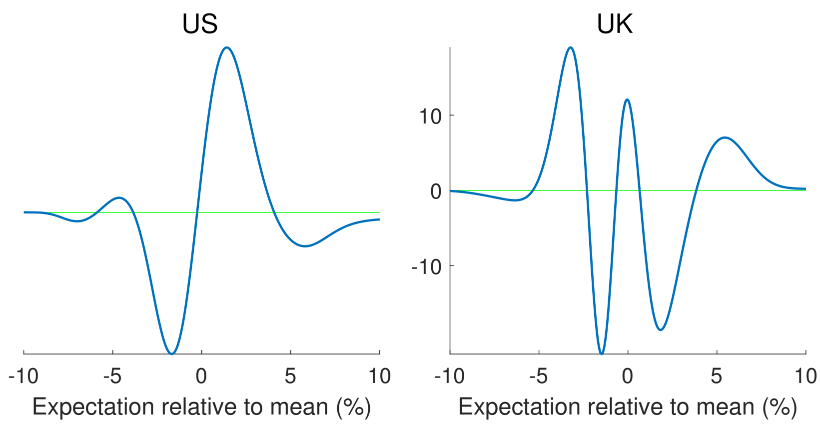
[20]A replication of Coibion and Gorodnichenko's results for the Michigan survey and the Survey of Professional Forecasters confirms this impression. Details are reported in the supplementary material, Section D. The issue of structural breaks affecting expectations around the time of the financial crisis is examined in Section 5.3.

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)
Distribution	–	func [.000]	func [.000]	func [.000]	–	func [.000]	func [.000]	func [.000]
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			
Number of FPCs	–	3	6	3	–	3	6	3
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

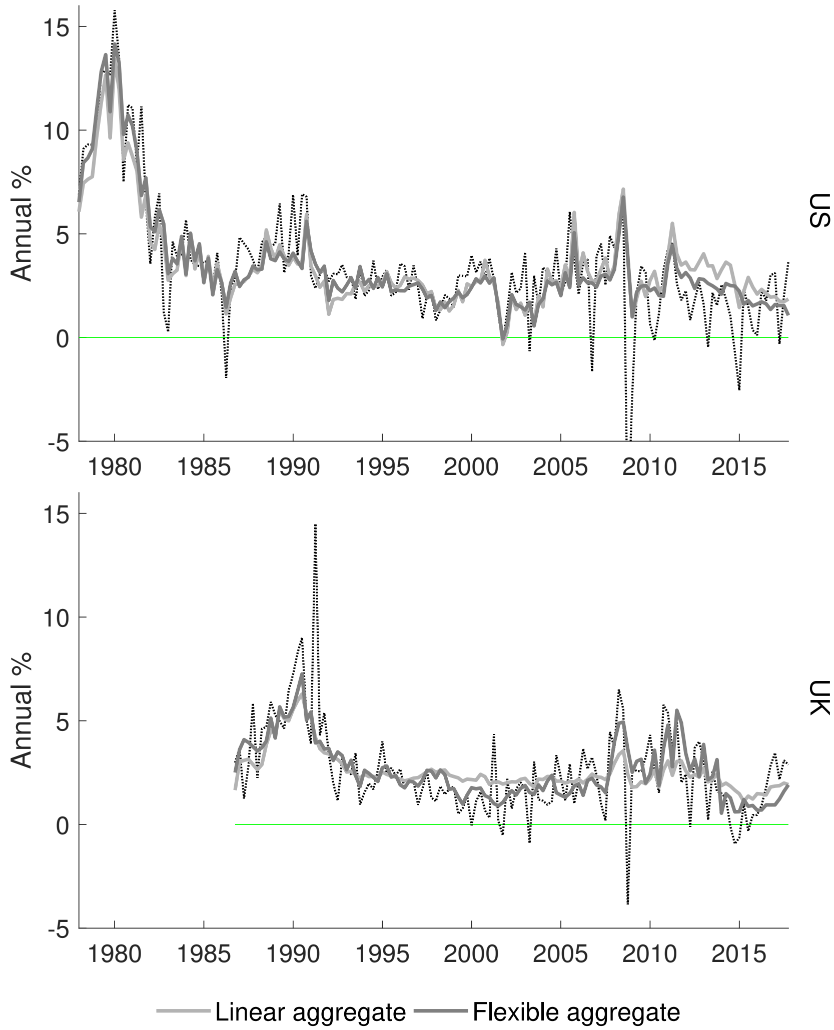
Note: Estimates of Eq. (6). Dependent variable is the seasonally adjusted annualized quarter-on-quarter percentage change in the consumer price index. Newey-West adjusted (5 lags) standard errors for t test (scalar covariates) appear in parentheses. p-values for F test (functional covariate) appear in 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 9. Estimated aggregation functions in the heterogeneous beliefs model



Note: Panels show the estimated coefficient function $\hat{\gamma}(x) = \sum_k \hat{\gamma}_k \mathsf{e}_k(x)$ for the models given in Tab. 1, Cols. (2) and (4). See Appendix C for further details.

Figure 10. Estimated inflation expectations indexes under linear and flexible aggregation



Note: Panels show actual inflation against partial fitted values, or expectations indexes, from the estimation of model (6) with $K = 3$ components as reported in Tab. 1. Shown are the: (a) contribution of average expectations (a linear aggregate), $\hat{\beta}\overline{\pi}_t^e$; (b) the overall contribution of expectations under flexible aggregation, $\hat{\beta}\overline{\pi}_t^e + \langle \hat{\gamma}, \mathsf{p}_t^c \rangle$.

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 that regression.[21] Estimates of the Phillips curve that exploit our newly-constructed household survey data series (Basix) are reported in Tab. 1, Cols. (5–8). The standard variant (Col. 5) has a positive ('incorrect') but insignificant slope. The coefficient on average expectations is almost identical to unity.

To the standard specification, we again add functional principal components from the full distribution of survey responses. The BIC again selects three components, but also the model with six components is preferred to that with none. Estimates given in Cols. (6–7) show *p*-values on the aggregation functions that indicate a high level of statistical significance for both three and six components. Again, the distribution of beliefs provides information relevant to price-setting behaviour, over and above the information contained in the average expectation. Moreover, this additional information 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). The aggregation function shown in Fig. 9 (right panel) is harder to interpret than the equivalent for the US. But the impact on the contribution made by expectations to inflation dynamics, Fig. 10, appear to be larger in the UK case. Expectations track inflation rather more closely in the heterogeneous beliefs model than in the standard model, particularly in the period after the late 1990s and including the financial crisis.

5.2 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, announced 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.

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

[21] 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.

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	$-.092$ (.104)	$-.220^{**}$ (.089)	$-.246^{**}$ (.094)	$-.270^{*}$ (.139)	$-.030$ (.125)	$-.237^{**}$ (.105)
Lagged inflation	$.733^{***}$ (.049)	$.179^{**}$ (.075)	.083 (.063)	$.439^{***}$ (.081)	.038 (.184)	.053 (.065)
Average expectation	–	1.39^{***} (.158)	1.61^{***} (.192)	–	$.935^{***}$ (.165)	$.668^{***}$ (.207)
Distribution	–	–	func [.001]	–	–	func [.000]
Outlier dummy	y	y	y	y	y	y
Sample	1978Q1–2017Q4			1986Q4–2017Q4		
Number of FPCs	–	–	3	–	–	3
R^2	.651	.783	.807	.550	.679	.734
BIC	1.34	.903	.881	.982	.675	.604
Number of obs.	160	160	160	125	125	125

Note: Estimates of Eq. (9). 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.

lagged inflation to produce a hybrid Phillips curve:

$$\pi_t = \beta\overline{\pi}^e_{t,h} + \sum_{k=1}^{K} \gamma_k s_{k,t} + \alpha(u_t - u^*_t) + \delta\pi_{t-1} + \varepsilon_t \tag{9}$$

In Tab. 2 we show the results of adding the expectation terms $\overline{\pi}^e_{t,h}$ and $\mathsf{p}^{\mathsf{c}}_{t,h}$ 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 $\overline{\pi}^e_{t,h}$. 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 four, although it remains significant. For the UK (Col. 5), it becomes economically and statistically indistinguishable from zero. As a result, the other parameter estimates are close to those in Tab. 1 (Col. 5).

Adding the distribution of inflation expectations, along with the average belief, eliminates the backward-looking component from the Michigan regression (Col. 3). For the Basix regression (Col. 6), lagged inflation is also irrelevant, and the distribution function is strongly significant. 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. (9) even after adding average expectations. 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. The finding that survey expectations—and especially cross-sectional heterogeneity in expectations—wholly drive out lagged inflation suggest that the latter serves only as a second-rate proxy for agents' underlying forward-looking beliefs.

5.3 Inflation gaps and heterogeneous beliefs

The recent literature recognizes the importance of accounting for trend inflation when thinking about cyclical inflation dynamics. Cogley and Sbordone (2008) present a micro-founded Phillips curve that features time-varying trend inflation, 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). Accounting for trend in an expectations-augmented Phillips curve may also be important because average near-term expectations—those pertaining to changes in prices at horizons of a year or two—often seem to track trend inflation closely. There is a risk that the apparent importance of expected inflation may actually be down to its association with trend. This type of concern was used by Cecchetti et al. (2017) to argue for the unimportance of short run expectations, at least in periods where monetary policy was well run.

We modify the baseline heterogeneous beliefs Phillips curve Eq. (5) to remove the trend component of inflation τ_t, measured using long-horizon inflation expectations as described in Faust and Wright (2013), as follows:[22]

$$\pi_t - \tau_t = \beta(\overline{\pi}_t^e - \tau_{t-1}) + \sum_{k=1}^{K} \gamma_k s_{k,t} + \alpha(u_t - u_t^*) + \varepsilon_t \tag{10}$$

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.[23] 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 $t-1$ to accommodate the periods in which the exponentially-smoothed (ES) trend stands in for long-run expectations.[24] Because the trend is formed using current-quarter inflation, the t-dated expectations gap would be correlated with the regression errors.

The inflation gap in the US

Estimates for the standard inflation gap model, with linear aggregation, do not support an economically or statistically important role for average near-horizon forecasts in explaining inflation, after accounting for long-horizon forecasts. Tab. 3 (Col. 1) reports that the average expectation gap has a coefficient only slightly above 0.3, and is not significant at the 10% level, thanks to large Newey-West adjusted standard errors. This result is something of a surprise from the perspective of New Keynesian versions of the Phillips curve. For parameterizations of price rigidity that accord best with the evidence from micro data, the average time between price changes is less than a year. That observation implies that the expected near-term rate of inflation should be an important influence on current price setting.

Matters change when the distribution of beliefs about near-term inflation are added to the model. Tab. 3 (Col. 2) indicates that near-horizon expectations—expectations plural—matter a great deal for US inflation, even after accounting for trend. The functions have T_F-statistics above 55, with corresponding p-values of zero. We observe a marked improvement in overall fit, as measured by R^2, and large reductions in the BIC, which selects for two components. But

[22]We adopt the 5-to-10 year ahead inflation expectation reported in the Michigan survey, which has the earliest start date of the available long-run inflation surveys. Respondents are asked: 'By about what percent per year do you expect prices to go (up/down) on the average, during the next 5 to 10 years?'. The question has been asked monthly since the early 1990s, and intermittently before that. For the quarters where the question was not asked, we use cubic spline interpolation to fill in the gaps. Before 1979, we use the exponentially-smoothed CPI inflation rate. The SPF has had a 'next ten years' question since the early 1990s; the Blue Chip survey (used by Faust and Wright) has asked about 5-10 year ahead inflation since the mid-1980s. Annex B contains results using this alternative series, with the exponentially-smoothed inflation series being spliced to the Blue Chip data for the early part of the sample.

[23]Model (10) 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.

[24]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. Lagging is not strictly necessary for the Michigan data, for which trend is mostly based on reported expectations. For the Basix data it is essential. We chose to treat the two surveys symmetrically.

more importantly, we see that some role for average expectations is restored. The expectation gap now has a t-statistic above 7. Notably, this parsimonious model now also appears well-specified. The Durbin-Watson test for residual autocorrelation is passed. This was not the case for the model with linear aggregation. In Col. (3), we show that there was apparently no statistically significant change in the importance of the average expectations gap during the Great Moderation, somewhat contrary to the arguments in Cecchetti et al. (2017). It is worth stressing once again that we reach this conclusion using precisely the same underlying expectations data that others have used to argue for the irrelevance of expectations for the inflation gap.

The inflation gap in the UK

We turn now to the experience of the UK. In the absence of a adequate series on far-horizon expectations, we opt to form the UK inflation gap using the exponentially-smoothed (ES) inflation trend.[25] The estimates in Tab. 3 (Cols. 4–5) are very similar to the baseline results for inflation in levels given in Tab. 1 (Col. 5–7). This is likely the result of using ES to remove the trend component of inflation. In that case, τ_t is close to τ_{t-1}, and as the coefficient on the expectations gap is close to unity, terms in the trend then roughly cancel from the two sides of Eq. (10). 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).[26] 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.

5.4 Regression on moments

In Section 3 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, we investigate whether straightforward regression on moments provides a alternative to principal component regression for capturing the effect

[25] For the UK, the longest-running source of 5-to-10 year ahead expectations comes from a survey by Yougov/Citigroup, but this starts only in the mid-2000s. For periods where the Yougov/Citigroup average 5-10 year ahead expectations is available, the ES trend tracks the data reasonably well.

[26] 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.

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
Unemployment gap	$-.228^{*}$ (.135)	$-.392^{***}$ (.060)	$-.378^{***}$ (.061)	.059 (.123)	$-.234^{**}$ (.095)	$-.223^{**}$ (.096)
Average expectation gap	.308 (.228)	$.659^{***}$ (.091)	$.782^{***}$ (.150)	1.14^{***} (.179)	1.02^{***} (.124)	1.16^{***} (.165)
Average expectation gap × Great Moderation	–	–	$-.188$ (.183)	–	–	$-.271$ (.202)
Distribution	–	func [.000]	func [.000]	–	func [.000]	func [.000]
Supply factors	y	y	y	y	y	y
Outlier dummy	y	y	y	y	y	y
Sample	1978Q1–2017Q4			1986Q4–2017Q4		
Number of FPCs	–	2	2	–	5	5
R^2	.488	.703	.704	.638	.710	.713
BIC	.979	.495	.522	.366	.337	.364
DW test (p-value)	.000	.236	.297	.001	.181	.171
Number of obs.	160	160	160	125	125	125

Note: Estimates of Eq. (10). Dependent variable is the seasonally adjusted annualized quarter-on-quarter percentage change in the consumer price index less the mean household 5–10 year ahead average inflation rate from the Michigan survey (US) or the exponentially-smoothed inflation trend (UK). 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.

of shifting beliefs on inflation.[27]

We re-ran our baseline Phillips curve regressions using moments as proxies for changes in the distribution of beliefs, instead of functional principal components. Focussing first on US data, Tab. 4 (Cols. 2-3) reports the results for the regression including second and third moments.[28] 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 distributions also appear in the model (Col. 4), the p-value of the functional T_F-statistic is well below 1%, suggesting that the information summarized by the functional regressors cannot be proxied by solely via the moments of the distribution. Similar results are found for the UK (Tab. 4, Cols. 5-8). The regressions confirm that only the second moment is significant, but, unlike for the US, it becomes insignificant when the distribution functions are also included. The straightforward reason for these findings is that functional components above the second are highly correlated with inflation, but weakly correlated with (linear combinations of) moments.

The results in this section confirm the enhanced role for the whole distribution of expectations in the inflation process. They also highlight that the functional principal components capture information in expectations that is relevant for inflation, even after including the moments of the distribution.

6 Conclusion

This paper has argued that aggregation of survey responses is a non-trivial problem for users of expectations data, but that a straightforward solution exists. We showed that the full set of survey expectations can be characterized using smooth distribution functions. Although beliefs about future inflation held by different agents are at times highly heterogeneous, leading to complex distributional shapes, we demonstrated that it 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 then related distributions of beliefs to actual inflation using scalar-on-function regression techniques borrowed from the functional data analysis literature in statistics. These techniques have found broad areas of application in diverse fields, but apparently few to date in macroeconomics.

Our principal finding has been that there a robust statistical association exists between the distribution of beliefs about future inflation (particularly those of households) and actual inflation, even after accounting for average expected inflation, lagged inflation, trend inflation, and the usual controls for supply factors. Our findings carry some novel implications for

[27] 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 central moments is equivalent to the restriction that γ lie in the space of polynomials.

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

Table 4. Proxy regressions using moments of the belief distributions

	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)	$-.491^{***}$ (.106)	$-.493^{***}$ (.104)	$-.370^{***}$ (.101)	.048 (.481)	$-.017$ (.139)	$-.017$ (.122)	$-.216^{**}$ (.109)
Average expectation	1.71^{***} (.104)	1.15^{***} (.182)	1.14^{***} (.178)	1.32^{***} (.161)	1.06^{***} (.129)	1.01^{***} (.118)	1.03^{***} (.302)	$.859^{***}$ (.252)
Second moment	–	1.12^{***} (.323)	1.02^{***} (.327)	1.23^{***} (.095)	–	1.21^{**} (.517)	1.21^{**} (.519)	.659 (.935)
Third moment	–	–	$-.574$ (.482)	$-.441$ (.518)	–	–	.044 (.684)	.656 (.722)
Distribution	–	–	–	func [.006]	–	–	–	func [.000]
Outlier dummy	y	y	y	y	y	y	y	y
Sample		1978Q1–2017Q4				1986Q4–2017Q4		
Number of FPCs	–	–	–	3	–	–	–	3
R^2	.773	.792	.795	.811	.674	.695	.695	.738
BIC	.914	.857	.878	.891	.652	.624	.663	.628
Number of obs.	160	160	160	160	125	125	125	125

Note: Estimates of Eq. (6). 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.

monetary policymakers. Central banks' preoccupation with inflation expectations has been half right. Well-anchored expectations underpin the ability of monetary policy to do more to respond to trade-off inducing shocks by doing less with interest rates. But expectations need to be understood in the plural, not the singular. Our results suggest that central banks focused on the average expectation have consistently missed information in survey data that is relevant to actual inflation. Understanding how policymakers may be able to influence the distribution of beliefs is a topic for future research.

References

Andrade, Philippe, Richard K. Crump, Stefano Eusepi, and Emanuel Moench (2016) "Fundamental disagreement," *Journal of Monetary Economics*, Vol. 83, pp. 106–128.

Andrade, Philippe and Hervé Le Bihan (2013) "Inattentive professional forecasters," *Journal of Monetary Economics*, Vol. 60, pp. 967–982.

Binder, Carola Conces (2015) "Whose expectations augment the Phillips curve?" *Economics Letters*, Vol. 136, pp. 35–38.

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

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

Cecchetti, Stephen G., Michael E. Feroli, Peter Hooper, Anil K. Kashyap, and Kermit L. Schoenholtz (2017) "Deflating inflation expectations: The implications of inflation's simple dynamics,"Technical report, University of Chicago Booth School of Business.

Chaudhuri, Kausik, Minjoo Kim, and Yongcheol Shin (2016) "Forecasting distributions of inflation rates: the functional auto-regressive approach," *Journal of the Royal Statistical Society, Series A*, Vol. 179, pp. 65–102.

Cogley, Timothy and Argia M. Sbordone (2008) "Trend inflation, indexation, and inflation persistence in the New Keynesian Phillips curve," *American Economic Review*, Vol. 98, pp. 2101–2126.

Coibion, Olivier and Yuriy Gorodnichenko (2012) "What can survey forecasts tell us about information rigidities?" *Journal of Political Economy*, Vol. 120, pp. 116–159.

——— (2015) "Is the Phillips curve alive and well after all? Inflation expectations and the missing disinflation," *American Economic Journal: Macroeconomics*, Vol. 7, pp. 197–232.

Coibion, Olivier, Yuriy Gorodnichenko, and Rupal Kamdar (2018) "The formation of expectations, inflation and the Phillips curve," *Journal of Economic Literature*, Vol. 56, pp. 1447–91.

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

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

Faust, Jon and Jonathan H. Wright (2013) "Forecasting inflation," in Graham Elliott and Alan Timmermann eds. *Handbook of Economic Forecasting*, Vol. 2A: Elsevier, pp. 2–56.

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

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

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

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

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

Jolliffe, Ian T. (2002) *Principal Component Analysis*, Springer Series in Statistics, New York: Springer-Verlag, 2nd edition.

Kneip, Alois and Klaus J. Utikal (2001) "Inference for density families using functional principal component analysis," *Journal of the American Statistical Association*, Vol. 96, pp. 519–532.

Kong, Dehan, Ana-Maria Staicu, and Arnab Maity (2016) "Classical testing in functional linear models," *Journal of Nonparametric Statistics*, Vol. 28, pp. 813–838.

Lombardelli, Clare and Jumana Saleheen (2003) "Public expectations of UK inflation," *Bank of England Quarterly Bulletin*, Vol. Autumn, pp. 281–290.

López-Pintado, Sara and Juan Romo (2009) "On the concept of depth for functional data," *Journal of the American Statistical Association*, Vol. 104, pp. 718–734.

Malmendier, Ulrike and Stefan Nagel (2016) "Learning from inflation experiences," *Quarterly Journal of Economics*, Vol. 131, pp. 53–87.

Mankiw, N. Gregory, Ricardo Reis, and Justin Wolfers (2003) "Disagreement about inflation expectations," in Mark Gertler and Kenneth Rogoff eds. *NBER Macroeconomics Annual*, Vol. 18: University of Chicago Press, pp. 209–248.

Mavroeidis, Sophocles, Mikkel Plagborg-Møller, and James H. Stock (2014) "Empirical evidence on inflation expectations in the New Keynesian Phillips curve," *Journal of Economic Literature*, Vol. 52, pp. 124–188.

Nunes, Ricardo (2010) "Inflation dynamics: The role of expectations," *Journal of Money, Credit and Banking*, Vol. 42, pp. 1161–1172.

Park, Joon K. and Junhui Qian (2012) "Functional regression of continuous state distributions," *Journal of Econometrics*, Vol. 167, pp. 397–412.

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

Pfajfar, Damjan and John M. Roberts (2018) "The role of expectations in changed inflation dynamics," Working Paper 2018-062, Federal Reserve Board.

Pfajfar, Damjan and Emiliano Santoro (2010) "Heterogeneity, learning and information stickiness in inflation expectations," *Journal of Economic Behavior & Organization*, Vol. 75, pp. 426–444.

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

Ramsay, James O. and Bernard W. Silverman (2005) *Functional Data Analysis*, Springer Series in Statistics, New York: Springer-Verlag, 2nd edition.

Reiss, Philip T., Jeff Goldsmith, Han Lin Shang, and R. Todd Ogden (2017) "Methods for scalar-on-function regression," *International Statistical Review*, Vol. 85, pp. 228–249.

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

Rich, Robert and Joseph Tracy (2010) "The relationships among expected inflation, disagreement and uncertainty: Evidence from matched point and density forecasts," *Review of Economics and Statistics*, Vol. 92, pp. 200–207.

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

Silverman, Bernard W. (1986) *Density estimation for statistics and data analysis*, Monographs on Statistics and Applied Probability 26: Chapman and Hall.

Stock, James H. and Mark W. Watson (2007) “Why has US inflation become harder to forecast?” *Journal of Money Credit and Banking*, Vol. 39, pp. 3–33.

Tsay, Ruey S. (2016) “Some methods for analysing big dependent data,” *Journal of Business and Economic Statistics*, Vol. 34, pp. 673–688.

Yellen, Janet L. (2016) “Macroeconomic research after the crisis,” October, Speech, Federal Reserve Bank of Boston.

A Survey data on inflation expectations

Table A.1. Inflation survey data

	US *Michigan Survey of Consumer Attitudes*	UK *Barclays Basix Survey*
Mnemonic	MSC	BBS
Survey Population	Cross-section of the general public	Cross-section of the general public
Survey Organization	Survey Research Center, University of Michigan	Barclays/GfK
Number of respondents, as mean (min–max)	566 (480–1,459)	1,894 (1,028–2,402)
Survey Type	Short rotating panel	Repeated cross sections
Starting date & frequency	Jan. 1978, monthly	1986 Q3, quarterly
Timing	Variable; usually fourth week of the month	Typically between the end of the middle month/start of the last month of the quarter
Forecast horizon(s)	One year ahead (from Jan. 1978); five years ahead (cts. from Apr. 1990)	One and two years ahead (from Dec. 1986); Five years ahead (from Sep. 2008)
Inflation measure	Unspecified	Unspecified

B Penalized maximum likelihood estimation of probability density functions

We fit distributions to the survey data using penalized maximum likelihood (pML) following Silverman (1986, Ch. 5). To see how pML works, recall that the probability density p satisfies $\int \mathsf{p} = 1$ and $\mathsf{p} \geq 0$. Without further restrictions, maximum likelihood estimation is infeasible. However, adding the condition that the curve has finite 'roughness' R, in a sense to be defined, allows us to operationalize it. Introduce the (unconstrained) smoothing function $W(\cdot) = \sum c_k \phi_k(\cdot)$ for the basis functions ϕ with domain $\mathcal{I}$.[29] Then expressing the density as the strictly positive transformation of the smoothing function:

$$\mathsf{p}_{t,h}(\pi^e) = C \exp\{W(\pi^e)\}, \qquad C = \left(\int \exp\{W(\pi^e)\}\, \mathrm{d}\pi^e\right)^{-1}$$

the penalized log likelihood for the cross section of expectations $\left\{\pi^e_{i,t,h}\right\}_{i=1}^{N_t}$ at each date t and conditional on the smoothing parameter λ is seen to be:

$$\begin{aligned}
\ell_\lambda(W|\mathbf{c}_t) &= \sum_{i=1}^{N_t} \log \mathsf{p}_{t,h} - \lambda R(\mathsf{p}_{t,h}) \\
&= \sum_{i=1}^{N_t} \mathbf{c}_t^\top \boldsymbol{\phi} - N_t \log \int \exp\left\{\mathbf{c}_t^\top \boldsymbol{\phi}\right\} \mathrm{d}x - \lambda R(\mathsf{p}_{t,h})
\end{aligned}$$

where $\mathbf{c} = (c_1, \ldots, c_K)^\top$, $\boldsymbol{\phi} = (\phi_1(\cdot), \ldots, \phi_K(\cdot))^\top$ are fifth order B-splines, and K is $\mathcal{O}(N)$. We choose the functional R so as to penalize departures from normality. As the log kernel of the normal distribution is quadratic, the relevant functional is the integrated third derivative:

$$R(\mathsf{p}) := \int \mathsf{p}'''(x)^2 \mathrm{d}x$$

In our context, interest centres on how to set λ. If $\lambda \to \infty$, we would fit a normal distribution. If $\lambda \to 0$, we would fit a distribution with up to N_t modes. We adopt a subjective approach, which may be loosely interpreted as placing a weight on the 'prior' of normality, and which induces a mild degree of smoothing of the underlying data histogram.

C 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 et al. (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 et al. (2016).

[29] Although the support of many continuous probability distributions have as their domain the whole or half real line, it is necessary to place bounds on the domain of our estimated functions, given by $\mathcal{I}$. For each survey, the bounds are given by the maximum and minimum response values across the entire sample, plus or minus one percentage point to avoid having truncation in regions of positive probability mass.

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:

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

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:

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

where the principal component scores $s_{kt} = \langle x_t, \mathsf{e}_k \rangle$ 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:

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

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:

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

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:

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

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.

D 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).

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

$$\int v(i,j)\mathsf{e}_k(i)\mathrm{d}j = \lambda_k \mathsf{e}_k(i) \tag{D.1}$$

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:

$$v(i,j) = (T-1)^{-1}\boldsymbol{\phi}(i)^\top \mathbf{C}^\top \mathbf{C}\boldsymbol{\phi}(j) \tag{D.2}$$

Assume that the eigenfunctions $e_k(i)$ 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 (D.2) into (D.1), the eigenequation may be written:

$$(T-1)^{-1}\boldsymbol{\phi}(i)^\top \mathbf{C}^\top \mathbf{C}\mathbf{W}\mathbf{b} = \lambda\boldsymbol{\phi}(i)^\top \mathbf{b} \tag{D.3}$$

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 (D.3) must hold for all i implies that a solution to (D.1) 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).