Measuring Public Inflation Perceptions and Expectations in the UK*

Yasutomo Murasawa
Faculty of Economics, Konan University
8-9-1 Okamoto, Higashinada-ku, Kobe, Hyogo 658-8531, Japan
(yasutomo.murasawa@gmail.com)

First draft: January 2017
This version: June 19, 2018

Abstract

The Bank of England Inflation Attitudes Survey asks individuals about their inflation perceptions and expectations in eight intervals including an indifference limen. This paper studies fitting a mixture normal distribution to such interval data, allowing for multiple modes. Bayesian analysis is useful since ML estimation may fail. A hierarchical prior helps to obtain a weakly informative prior. The No-U-Turn Sampler (NUTS) speeds up posterior simulation. Permutation invariant parameters are free from the label switching problem. The paper estimates the distributions of public inflation perceptions and expectations in the UK during 2001Q1–2017Q4. The estimated means are useful for measuring information rigidity.

Keywords Bayesian · Indifference limen · Information rigidity · Interval data · Normal mixture · No-U-turn sampler

JEL classification C11 · C25 · C46 · C82 · E31

*I thank three referees and Naohito Abe for useful comments. This work was supported by JSPS KAKENHI Grant Number JP16K03605.

1 Introduction

The actual, perceived, and expected inflation (or deflation) all play important roles in monetary macroeconomics. Prices move differently across regions and commodities. Since people live in different regions and buy different commodities, they perceive inflation and form its expectations in different ways, even if they are fully rational. Perhaps people are boundedly rational or irrational in different ways. Whatever the reason is, survey data on inflation perceptions and expectations show substantial *heterogeneity* or *disagreement* among individuals, which varies over time with the actual inflation; see Mankiw et al (2004). It is thus of interest to central banks to monitor not only the means or medians of the actual, perceived, and expected inflation but also their whole distributions among individuals. Indeed, the mean, median, and mode may be quite different.

Survey questions often ask respondents to choose from more than two categories. The categories may be ordered or unordered, and if ordered, they may or may not represent intervals. Thus there are three types of categorical data: unordered, ordered, and interval data. Interval data have quantitative information.[1] To analyze these data (with covariates), we use multinomial response, ordered response, and interval regression models for unordered, ordered, and interval data, respectively.[2]

Some categorical data are not exactly one of the three types but somewhere in between. For a question asking about a change of some continuous variable with several intervals to choose from, there is often a category 'no change.' This does not literally mean a 0% change, whose probability is 0, but corresponds to an *indifference limen*, which allows for a very small change. Data on inflation perceptions and expectations are examples. For the UK data used in this paper, the respondents choose from the following eight intervals (excluding 'No idea'):[3]

1. Gone/Go down
2. Not changed/change
3. Up by 1% or less
4. Up by 1% but less than 2%
5. Up by 2% but less than 3%
6. Up by 3% but less than 4%
7. Up by 4% but less than 5%
8. Up by 5% or more

The questionnaire gives the boundaries between the categories except for category 2, which is an indifference limen. The problem is that we do not know the boundaries of an indifference limen.

[1] Interval data lose some quantitative information, however. If the first or last interval has an open end, then one cannot draw a histogram nor compute its mean. Moreover, one cannot find the median if it lies in the first or last interval with an open end. To solve the problem, one can assume the minimum or maximum of the population distribution, or fit a parametric distribution to the frequency distribution.

[2] If intervals represent durations, then we apply survival analysis.

[3] Since May 2011, the survey asks further questions with more intervals to those who have chosen either category 1 or 8.

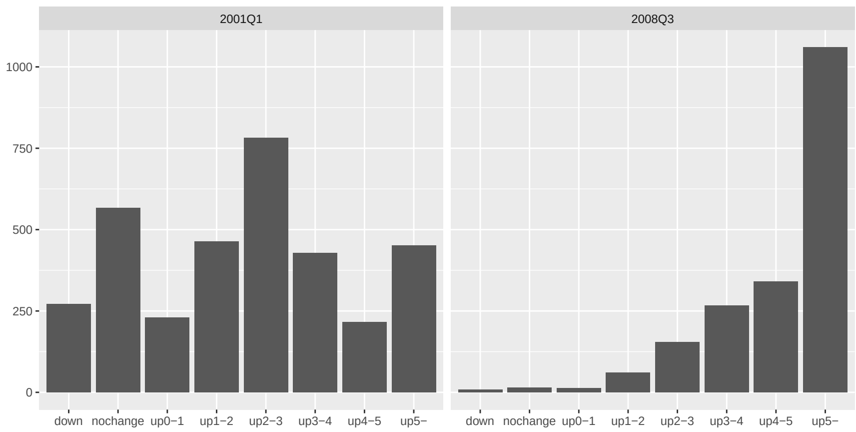

Figure 1: Distributions of inflation perceptions (2001Q1 and 2008Q3)

A simple solution is to combine an indifference limen with its neighboring intervals to apply interval regression.[4] Alternatively, one may assume a common indifference limen among individuals. Then assuming a parametric distribution for the latent continuous variable, with enough categories for identification, one can estimate the distribution parameters and the common indifference limen jointly. Given the prior information that the indifference limen contains 0 and cannot overlap with the other intervals, this may improve efficiency. Murasawa (2013) applies these ideas to survey data on inflation expectations. This paper complements Murasawa (2013) in three ways: data, distribution, and methodology.

First, Murasawa (2013) uses monthly aggregate interval data on household inflation expectations in Japan during 2004M4–2011M3 that have seven intervals lying symmetrically around 0. This paper uses quarterly aggregate interval data on public inflation perceptions and expectations in the UK during 2001Q1–2017Q4 that have eight intervals, six of which lie above 0. Thus we can compare the distributions of inflation expectations and their dynamics in Japan and the UK, if desired. Moreover, the UK data allow us to study the interaction between inflation perceptions and expectations, e.g., using a bivariate VAR.

Second, Murasawa (2013) considers normal, skew normal, skew exponential power, and skew t distributions, which are all unimodal and have at most four parameters. Since the UK data have one more interval, this paper considers a mixture of two normal distributions, which can be bimodal with five parameters. The frequency distributions of the UK data may have two modes; see the left panel in Figure 1.[5] Hence mixture distributions may fit better than unimodal skew distributions. Even if not, one cannot exclude multi-modal distributions a priori.

Third, Murasawa (2013) applies the ML method. The likelihood function of a normal mixture model for numerical data is unbounded, however, when one

[4]Nonparametric kernel density estimation for numerical data does not apply directly to interval data.

[5]Since these panels are not histograms but bar charts, the multiple modes may be spurious. Without knowing the width of the indifference limen, one cannot draw a histogram.

component has mean equal to a data point and variance equal to 0; see Kiefer and Wolfowitz (1956, p. 905). The likelihood function is bounded for interval data, but Biernacki (2007) shows that the global ML estimate may still have a degenerate component or otherwise the EM algorithm may converge to a local ML estimate with a degenerate component. To avoid the problem, this paper applies Bayesian analysis. This requires the following considerations:

1. For finite mixture models, independent improper priors on the component parameters give improper posteriors if some components have no observation. To obtain a weakly informative prior, a hierarchical prior is useful.

2. MCMC methods are useful for posterior simulation. With an indifference limen, however, our model is similar to an ordered response model with unknown cutpoints, for which the Gibbs sampler and Metropolis–Hastings (M–H) algorithm are slow to converge for large samples. A useful alternative is a Hamiltonian Monte Carlo (HMC) method, in particular the No-U-Turn Sampler (NUTS) developed by Hoffman and Gelman (2014).

3. Because of the label switching problem, MCMC methods may not work for the component parameters of a finite mixture model. They still work, however, for permutation invariant parameters, e.g., moments and quantiles; see Geweke (2007).

To summarize, this paper contributes to the literature on quantification of qualitative survey data by proposing a normal mixture model for interval data with an indifference limen. This seems new, and its flexibility seems useful for analysis of survey data. We also propose Bayesian analysis of the model using a hierarchical prior and the NUTS. This is useful since for this model, the ML method may fail, and the Gibbs sampler and M–H algorithm are often slow to converge for large samples. We apply the method to survey data on inflation perceptions and expectations in the UK repeatedly, and estimate their latent distributions during 2001Q1–2017Q4. We use three types of data, and find the following:

1. With six intervals ignoring the indifference limen, the estimated means are reasonably precise with narrow error bands. Since interval data lose information about the tails of the underlying distribution, the estimates of higher-order moments are less precise, especially when the majority of the respondents choose the last interval with an open end; see the right panel in Figure 1.

2. With eight intervals including the indifference limen and the prior information that the indifference limen contains 0, the results change slightly, especially for higher-order moments. However, the estimated distributions are not precise enough to learn about the dynamics of the distributions of inflation perceptions and expectations. The estimated indifference limens are time-varying and often strongly asymmetric around 0. In particular, the estimated lower bounds of the indifference limens are often far below -1.0%.[6]

[6] If such estimates seem unreasonable, then one can impose a subjective prior on the lower bound of the indifference limen, which will change the results further.

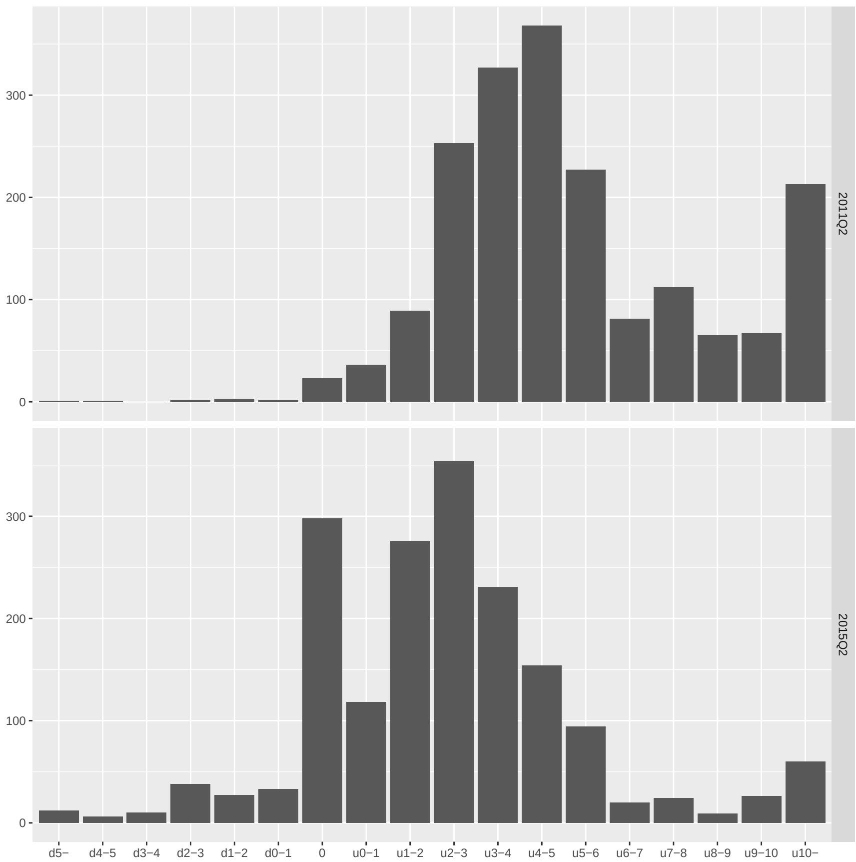

Figure 2: Distributions of inflation perceptions (2011Q2 and 2015Q2)

3. Data with 18 intervals are available since 2011Q2, which may have two more local modes (one below -1% and the other above 5%); see Figure 2. If we fit a mixture of two normal distributions to these data, then the estimated distributions and indifference limens are precise with extremely narrow error bands.[7]

Thus we find that fitting flexible distributions to repeated interval data requires well-designed categories to provide enough information about the shapes of the underlying distributions, which may shift drastically over time. The problem exists but is much less evident when fitting normal or unimodal distributions. This consideration seems important when designing a repeated survey.[8]

[7] Fitting a mixture of four normal distributions gives similar results, though error bands are wider.

[8] One may think that numerical data are easier to analyze than interval data. This is not the case if one considers the rounding problem seriously, since some respondents round to integers but others round to multiples of 5 or 10; see Manski and Molinari (2010). To account for such rounding in numerical data on inflation expectations from the Opinion Survey on the General Public's Views and Behavior conducted by the Bank of Japan, Kamada et al (2015) introduce point masses at multiples of 5. Binder (2017) interprets heterogeneous rounding as a measure of uncertainty. The rounding problem is irrelevant to our interval data.

To illustrate a possible use of the estimated distributions, we measure the degrees of information rigidity in public inflation perceptions and expectations in the UK using a simple framework proposed by Coibion and Gorodnichenko (2015), which requires only the historical means of the distributions of expectations. We find some evidences of information rigidity, though the results depend on what prices we assume individuals forecast, i.e., the CPI or RPI, and how we estimate the distributions of inflation perceptions and expectations, i.e., the choice of the number of mixture components and whether to include or exclude the indifference limen. Thus we recommend using a flexible distribution, which is less restrictive and gives more reliable results.

The paper proceeds as follows. Section 2 reviews some relevant works. Section 3 specifies a normal mixture model for interval data with an indifference limen. Section 4 specifies our hierarchical prior, and explains posterior simulation. Section 5 applies the method to interval data on public inflation perceptions and expectations in the UK, and estimates their latent distributions using six or eight intervals during 2001Q1–2017Q4 and 18 intervals during 2011Q2–2017Q4. Section 6 measures information rigidity in public inflation perceptions and expectations in the UK during 2001Q1–2016Q4. Section 7 concludes. The Appendix explains HMC methods and the NUTS.

2 Literature

2.1 Quantification of qualitative survey data

This paper contributes to the literature on quantification of qualitative survey data, in particular the Carlson–Parkin method; see Nardo (2003) and Pesaran and Weale (2006, sec. 3) for surveys. To quantify ordered data, which have no quantitative information, the Carlson–Parkin method requires multiple samples and strong assumptions; i.e., normality, a time-invariant symmetric indifference limen, and long-run unbiased expectations. In practice, however, these assumptions often fail; see Lahiri and Zhao (2015) for a recent evidence.

Such strong assumptions are unnecessary for interval data, to which one can fit various distributions each sample separately. Using data from the Monthly Consumer Confidence Survey in Japan, Murasawa (2013) fits various skew distributions by the ML method, which also gives an estimate of the indifference limen. Using the same data and approximating the population distribution up to the fourth cumulant by the Cornish–Fisher expansion, Terai (2010) estimates the distribution parameters by the method of moment.

2.2 Measuring inflation perceptions and expectations

Empirical works on inflation perceptions and expectations are abundant; see Sinclair (2010) for a recent collection. Most previous works use either ordered or numerical data, however, since few surveys collect interval data on inflation perceptions and expectations. Exceptions are Lombardelli and Saleheen (2003) and Blanchflower and MacCoille (2009), who use individual interval data from the Bank of England/GfK NOP Inflation Attitudes Survey and apply interval regression to study the determinants of public inflation expectations in the UK.

Recent works try to measure subjective pdfs of future inflation. For surveys, see Manski (2004) on measuring expectations in general and Armantier et al (2013) on measuring inflation expectations in particular.

2.3 Bayesian analysis of a normal mixture model for interval data

Alston and Mengersen (2010) consider Bayesian analysis of a normal mixture model for interval data with no indifference limen. Estimation of an indifference limen is similar to estimation of cutpoints in an ordered response model and hence troublesome.

Albert and Chib (1993) develop a Gibbs sampler for Bayesian analysis of an ordered response model, but it converges slowly, especially with a large sample. To improve convergence, Cowles (1996) applies 'collapsing,' i.e., sampling the latent data and cutpoints jointly, which requires an M–H algorithm. Nandram and Chen (1996) and Chen and Dey (2000) reparametrize the cutpoints for further improvement.[9] For an M–H algorithm, however, one must choose good proposals carefully, which is a difficult task. Moreover, the optimal acceptance probability may be far below 1, in which case substantial inefficiency remains; see Gelman et al (2014, p. 296).

3 Model specification

Let $\boldsymbol{y}$ be a random sample of size n taking values on $\{1, \dots, J\}$, which indicate J intervals on $\mathbb{R}$. Assume an ordered response model for y_i such that for $i = 1, \dots, n$,

$$y_i := \begin{cases} 1 & \text{if } \gamma_0 < y_i^* \leq \gamma_1 \\ \vdots & \\ J & \text{if } \gamma_{J-1} < y_i^* \leq \gamma_J \end{cases} \tag{1}$$

where y_i^* is a latent variable underlying y_i and $-\infty = \gamma_0 < \cdots < \gamma_l < 0 < \gamma_u < \cdots < \gamma_J = \infty$. Assume that we know $\{\gamma_j\}$ except for an indifference limen $[\gamma_l, \gamma_u]$.

Assume a normal K-mixture model for y_i^* such that for $i = 1, \dots, n$,

$$y_i^* \sim \sum_{k=1}^{K} \pi_k \mathrm{N}\left(\mu_k, \sigma_k^2\right) \tag{2}$$

Let $\boldsymbol{\pi} := (\pi_1, \dots, \pi_K)'$, $\boldsymbol{\mu} := (\mu_1, \dots, \mu_K)'$, $\boldsymbol{\sigma} := (\sigma_1, \dots, \sigma_K)'$, $\boldsymbol{\gamma} := (\gamma_l, \gamma_u)'$, and $\boldsymbol{\theta} := (\boldsymbol{\pi}', \boldsymbol{\mu}', \boldsymbol{\sigma}', \boldsymbol{\gamma}')'$. Assume for identification that $\pi_1 \geq \cdots \geq \pi_K$. Consider estimation of $\boldsymbol{\theta}$ given $\boldsymbol{y}$.[10]

[9] A generalized Gibbs sampler by Liu and Sabatti (2000) does not work well in our context.

[10] With covariates, our model extends an ordered probit model by allowing for some known cutpoints and a mixture normal distribution. Lahiri and Zhao (2015) use a hierarchical ordered probit (HOPIT) model, which reduces to an ordered probit model if no covariate is available.

4 Bayesian analysis

4.1 Prior

For mixture models, independent improper priors on the component parameters give improper posteriors if some components have no observation. Instead, we use *weakly informative priors* in the sense of Gelman et al (2014, p. 55).

For the mixture weight vector $\boldsymbol{\pi}$, we assume a Dirichlet prior such that

$$\boldsymbol{\pi}|\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\gamma} \sim \mathrm{D}(\boldsymbol{\nu}_0) \tag{3}$$

Setting $\boldsymbol{\nu}_0 := \boldsymbol{\imath}_K$ gives a flat prior on the unit $(K-1)$-simplex. We impose the identification restriction $\pi_1 \geq \cdots \geq \pi_K$ after posterior simulation by relabeling, if necessary.

For the component means $\mu_1, \ldots, \mu_K$ and variances $\sigma_1^2, \ldots, \sigma_K^2$, we assume independent normal–gamma priors such that for $k = 1, \ldots, K$,

$$\mu_k|\boldsymbol{\mu}_{-k}, \boldsymbol{\sigma}, \boldsymbol{\pi}, \boldsymbol{\gamma} \sim \mathrm{N}\left(\mu_0, \sigma_0^2\right) \tag{4}$$

$$\sigma_k^2|\boldsymbol{\sigma}_{-k}, \boldsymbol{\mu}, \boldsymbol{\pi}, \boldsymbol{\gamma} \sim \text{Inv-Gam}(\alpha_0, \beta_0) \tag{5}$$

Setting a large σ_0 gives a weakly informative prior on $\mu_1, \ldots, \mu_K$. The prior on $\sigma_1^2, \ldots, \sigma_K^2$ implies that for $k, l = 1, \ldots, K$ such that $k \neq l$,

$$\frac{\sigma_k^2}{\sigma_l^2}|\boldsymbol{\sigma}_{-(k,l)}, \boldsymbol{\mu}, \boldsymbol{\pi}, \boldsymbol{\gamma} \sim \mathrm{F}(2\alpha_0, 2\alpha_0) \tag{6}$$

Thus α_0 controls the prior on the variance ratio. To avoid the variance ratio to be too close to 0, which avoids degenerate components, one often sets $\alpha_0 \approx 2$; see Frühwirth-Schnatter (2006, pp. 179–180). It is difficult to set an appropriate β_0 directly. Instead, Richardson and Green (1997) propose a *hierarchical prior* on β_0 such that[11]

$$\beta_0 \sim \mathrm{Gam}(A_0, B_0) \tag{7}$$

Let R be the range of the sample, i.e., $\boldsymbol{y}^* := (y_1^*, \ldots, y_n^*)$ in our context, and m be the midrange. Richardson and Green (1997, pp. 735, 742, 748) recommend the following default settings:

$$\mu_0 := m \tag{8}$$

$$\sigma_0 := \frac{R}{c} \tag{9}$$

$$\alpha_0 := 2 \tag{10}$$

$$A_0 := .2 \tag{11}$$

$$B_0 := \frac{A_0}{\alpha_0}\frac{100}{R^2} = \frac{10}{R^2} \tag{12}$$

where $c > 0$ is a small integer.

For the boundaries of the indifference limen $[\gamma_l, \gamma_u]$, we assume independent uniform priors such that

$$\gamma_l|\gamma_u, \boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\sigma} \sim \mathrm{U}(\lambda_0, 0) \tag{13}$$

$$\gamma_u|\gamma_l, \boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\sigma} \sim \mathrm{U}(0, \upsilon_0) \tag{14}$$

[11] The result seems robust to other priors on β_0; e.g. a half-Cauchy prior with a large scale parameter.

where

$$\lambda_0 := \max_{\gamma_j < \gamma_l} \{\gamma_0, \ldots, \gamma_J\} \tag{15}$$

$$\upsilon_0 := \min_{\gamma_j > \gamma_u} \{\gamma_0, \ldots, \gamma_J\} \tag{16}$$

i.e., $[\gamma_l, \gamma_u]$ contains 0 and do not overlap with the other intervals.

4.2 Posterior simulation

An MCMC method sets up a Markov chain with invariant distribution $p(\boldsymbol{\theta}|\boldsymbol{y})$ for posterior simulation. After convergence, one can treat the realized states of the chain as (serially dependent) draws from $p(\boldsymbol{\theta}|\boldsymbol{y})$. Let $q(\boldsymbol{\theta}, \boldsymbol{\theta}')$ be a transition kernel that defines a Markov chain. A sufficient (not necessary) condition for a Markov chain to converge in distribution to $p(\boldsymbol{\theta}|\boldsymbol{y})$ is that $p(\boldsymbol{\theta}|\boldsymbol{y})$ and $q(\boldsymbol{\theta}, \boldsymbol{\theta}')$ are in detailed balance, i.e., $\forall \boldsymbol{\theta}, \boldsymbol{\theta}'$,

$$p(\boldsymbol{\theta}|\boldsymbol{y})q(\boldsymbol{\theta}, \boldsymbol{\theta}') = p(\boldsymbol{\theta}'|\boldsymbol{y})q(\boldsymbol{\theta}', \boldsymbol{\theta}) \tag{17}$$

so that the Markov chain is reversible; see Chib and Greenberg (1995, p. 328). The degree of serial dependence or speed of convergence depends on the choice of $q(\boldsymbol{\theta}, \boldsymbol{\theta}')$, which is crucial for successful posterior simulation.

We use an HMC method, in particular the NUTS developed by Hoffman and Gelman (2014) for posterior simulation. The NUTS often has better convergence properties than other popular MCMC methods such as the Gibbs sampler and M–H algorithm. See the Appendix for more details on HMC methods and the NUTS.

5 Public inflation perceptions and expectations in the UK

5.1 Data

We use aggregate interval data from the Bank of England/GfK NOP Inflation Attitudes Survey that started in February 2001 (Bank of England/TNS Inflation Attitudes Survey since February 2016).[12] This is a quarterly survey of public attitudes to inflation, interviewing a quota sample of adults aged 16 or over in 175 randomly selected areas (368 areas since February 2016) throughout the UK. The quota is about 4,000 in February surveys and about 2,000 in others. Detailed survey tables are available at the website of the Bank of England. In addition, the individual interval data with covariates are now publicly available.

The first two questions of the survey ask respondents about their inflation perceptions and expectations:

> Q.1 Which of the options on this card best describes how prices have changed over the last twelve months?
>
> Q.2 And how much would you expect prices in the shops generally to change over the next twelve months?

[12] There were five trial surveys quarterly from November 1999 to November 2000.

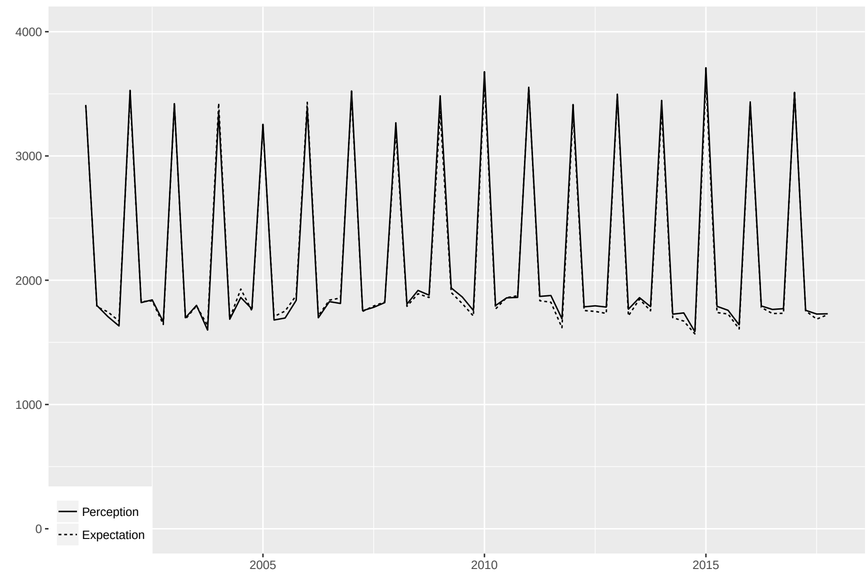

Figure 3: Sample sizes for inflation perceptions and expectations (excluding 'No idea')

The respondents choose from the eight intervals given in the Introduction. We exclude 'No idea' from the samples in the following analyses. Figure 3 plots the sample sizes for the two questions during 2001Q1–2017Q4. For both questions, the sample sizes are around 3,500 in February surveys and 1,500–2,000 in others.

Figure 4 plots the relative frequencies of the categories for inflation perceptions and expectations during 2001Q1–2017Q4. The perceptions and expectations are heterogeneous, but on average, they rose in 2007, reached a peak in 2008Q3, and then dropped. They rose again in 2010, remained high for a while, and then slowly decreased toward 2015. They started to rise again in 2016.

The relative frequency of category 3 is often lower than those of categories 2 and 4. Hence the distributions of inflation perceptions and expectations may be bimodal, depending on the interval widths of categories 2 and 3. To allow for this possibility, we fit a mixture of two normal distributions and estimate the indifference limen each quarter.

5.2 Model specification

Our model consists of equations (1)–(2) with $J := 6$ (with no indifference limen) or 8 (with an indifference limen) and $K := 2$. Our priors are those in section 4.1 with $m := 2.5$, $R := 10$, and $c := 1$. Thus we set $\boldsymbol{\nu}_0 := \boldsymbol{\imath}_2$, $\mu_0 := 2.5$, $\sigma_0 := 10$, $\alpha_0 := 2$, $A_0 := .2$, $B_0 := .1$, $\lambda_0 := -\infty$, and $\upsilon_0 := 1$. These weakly informative priors do not dominate the posteriors for our samples.

Reparametrizatoin may improve efficiency of MCMC. Following Betancourt and Girolami (2015), one may reparametrize our hierarchical model from centered parametrization (CP) to noncentered parametrization (NCP); i.e., instead of equation (5), for $k = 1, \dots, K$, one may draw

$$\sigma^{*2}_k | \boldsymbol{\sigma}_{-k}, \boldsymbol{\mu}, \boldsymbol{\pi}, \boldsymbol{\gamma} \sim \text{Inv-Gam}(\alpha_0, 1) \quad (18)$$

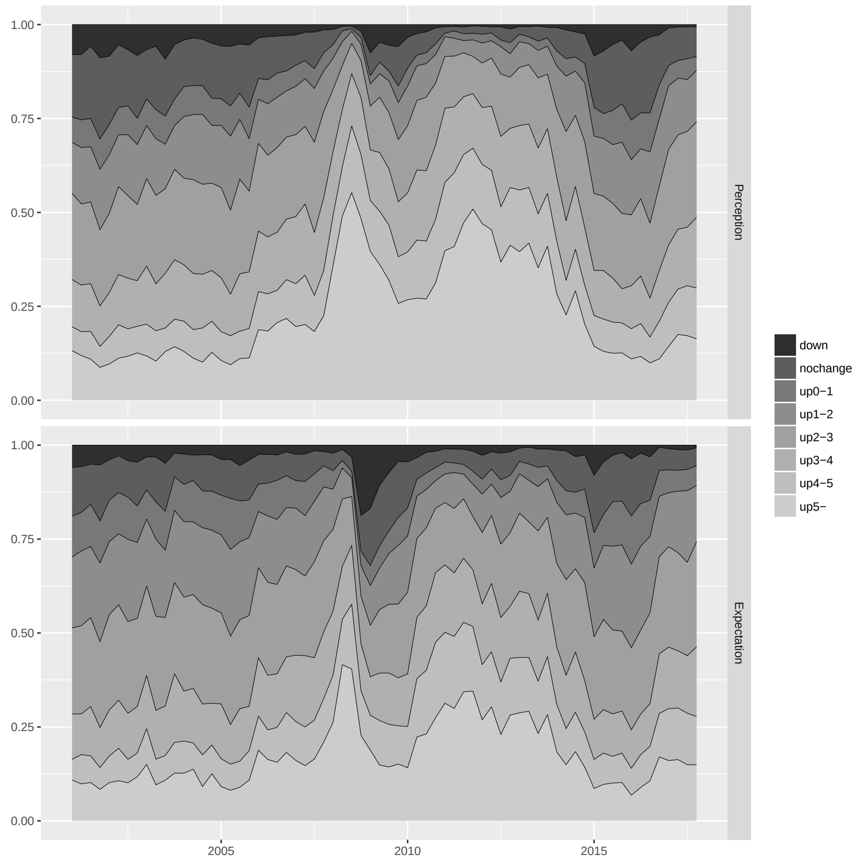

Figure 4: Relative frequencies of the categories for inflation perceptions (top) and expectations (bottom)

and set $\sigma_k^2 := \beta_0 {\sigma^*}_k^2$. CP and NCP are complementary in that when MCMC converges slowly with one parametrization, it converges much faster with the other parametrization; see Papaspiliopoulos et al (2003, 2007). We find that CP is often faster than NCP in our case; thus we stick to CP.

Using the aggregate interval data on inflation perceptions and expectations, respectively, we estimate our model described above each quarter separately during 2001Q1–2017Q4 (68 quarters). Thus we estimate the model 136 times.

5.3 Bayesian computation

We apply the NUTS using RStan 2.17.3 developed by Stan Development Team (2018) on R 3.5.0 developed by R Core Team (2018). To avoid divergent transitions during the leapfrog steps as much as possible and improve sampling efficiency further at the cost of more leapfrog steps in each iteration, we change the following tuning parameters from the default values:

1. target Metropolis acceptance rate (from .8 to .999)[13]
2. maximum tree depth (from 10 to 12)[14]

We generate four independent Markov chains from random initial values. For each chain, we discard the initial 1,000 draws as warm-up, and keep the next 1,000 draws. Thus in total, we use 4,000 draws for our posterior inference.

MCMC for finite mixture models suffers from the label switching problem. Despite the identification restrictions on $\boldsymbol{\pi}$, the component labels may switch during MCMC. Hence MCMC may not work for the component parameters. It still works, however, for permutation invariant parameters, e.g., moments and quantiles; see Geweke (2007). Thus instead of the five parameters of a mixture of two normal distributions, we look at the following five moments:

1. mean $\mu := \mathrm{E}(y_i^*)$
2. standard deviation $\sigma := \sqrt{\mathrm{var}(y_i^*)}$
3. skewness $\mathrm{E}\left([(y_i^* - \mu)/\sigma]^3\right)$
4. excess kurtosis $\mathrm{E}\left([(y_i^* - \mu)/\sigma]^4\right) - 3$
5. asymmetry in the tails $\mathrm{E}\left([(y_i^* - \mu)/\sigma]^5\right)$

Given equation (2),

$$\mu = \sum_{k=1}^{K} \pi_k \mu_k \tag{19}$$

[13]We set an extremely high target rate to eliminate divergent transitions as much as possible in all 136 quarters. In practice, it can be much lower in most quarters, and the results are almost identical as long as there is no divergent transition.

[14]This increases the maximum number of leapfrog steps from $2^{11} - 1 = 2047$ to $2^{13} - 1 = 8191$.

The 2nd to 5th central moments of y_i^* are

$$\mathrm{E}\left((y_i^* - \mu)^2\right) = \sum_{k=1}^{K} \pi_k \left[(\mu_k - \mu)^2 + \sigma_k^2\right] \tag{20}$$

$$\mathrm{E}\left((y_i^* - \mu)^3\right) = \sum_{k=1}^{K} \pi_k \left[(\mu_k - \mu)^3 + 3(\mu_k - \mu)\sigma_k^2\right] \tag{21}$$

$$\mathrm{E}\left((y_i^* - \mu)^4\right) = \sum_{k=1}^{K} \pi_k \left[(\mu_k - \mu)^4 + 6(\mu_k - \mu)^2\sigma_k^2 + 3\sigma_k^4\right] \tag{22}$$

$$\mathrm{E}\left((y_i^* - \mu)^5\right) = \sum_{k=1}^{K} \pi_k \left[(\mu_k - \mu)^5 + 10(\mu_k - \mu)^3\sigma_k^2 + 15(\mu_k - \mu)\sigma_k^4\right] \tag{23}$$

See Frühwirth-Schnatter (2006, pp. 10–11).

To assess convergence of the Markov chain to its stationary distribution, we use the potential scale reduction factor $\hat{R}$ and the effective sample size (ESS). We follow Gelman et al (2014, pp. 287–288), and check that $\hat{R} \leq 1.1$ and that the ESS is at least 10 per chain (in total 40 in our case) for the parameters of interest.

Bayesian model comparison requires estimation of the marginal likelihood. Frühwirth-Schnatter (2004) compares alternative methods for estimation of the marginal likelihood of a mixture model, and finds that bridge sampling is most robust; see also Frühwirth-Schnatter (2006, pp. 150–159). For implementation, the `bridgesampling` package for R is useful. We use a Warp-III bridge sampler proposed by Meng and Schilling (2002). See Gronau et al (2017) for a tutorial on bridge sampling.

5.4 Estimation with no indifference limen

First, we estimate the model ignoring the indifference limen. Combining three categories into one and ignoring the prior information about the location of the indifference limen, we may lose some information. However, we need not assume a common indifference limen among individuals. Moreover, estimation is faster and easier with two less parameters to estimate.

To justify use of a normal mixture model, we estimate the Bayes factor (BF) for $K = 1$ (normal) vs $K = 2$ (mixture normal). Figure 5 shows that $\log_{10} \mathrm{BF}$ is mostly negative and often below -2 but never exceeds 2 for both inflation perceptions and expectations. Since Jeffreys (1961, p. 432) interprets $\log_{10} \mathrm{BF}$ above 2 or below -2 as a *decisive* evidence, $K = 2$ is often much more likely.[15]

Figure 6 plots the posterior medians of the five moments of the distributions of inflation perceptions and expectations, respectively, during 2001Q1–2017Q4, together with 95% error bands. We summarize our findings below.

The estimated means of inflation perceptions and expectations are reasonably precise with narrow error bands. The error band widens when the mean is above or close to 5%, i.e., when the majority of the respondents choose the last interval with an open end. The same is true when the mean is close to 1%, i.e., when the majority of the respondents choose the first three intervals combined

[15]Normal mixture models with $K \geq 3$ are not identifiable from our data.

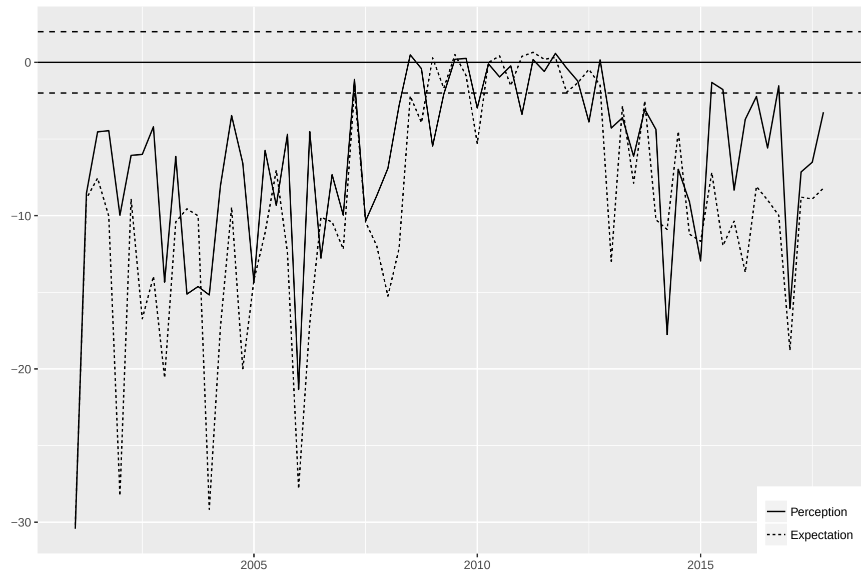

Figure 5: Log Bayes factors ($\log_{10}$ BF) for $K = 1$ (normal) vs $K = 2$ (mixture normal) for the distributions of inflation perceptions and expectations

into one with an open end. Still, the error bands remain quite narrow, perhaps because with exponential tails, fitting a mixture normal distribution effectively eliminates outliers. The means of inflation expectations are more stable than those of perceptions, staying around 2–4%, though not perfectly constant (anchored); thus individuals may expect that inflation fluctuations are temporary, though some persistence remains for one year.

The estimated standard deviations of inflation perceptions and expectations are reasonably precise except when the majority of the respondents choose the first or last interval with an open end, in which case interval data have little information about the dispersion of the underlying variable. For both inflation perceptions and expectations, the standard deviations are large when the means are unstable; thus they can be useful measures of macroeconomic uncertainty. The standard deviations of inflation perceptions seem slightly larger than those of expectations; thus individuals have private information when they perceive inflation, but they may expect its fluctuations to be temporary.

The estimated skewnesses of inflation perceptions and expectations are also reasonably precise, changing signs over time. The mean and skewness tend to move in the same direction, especially for inflation perceptions; thus responses to news on inflation may not be uniform but heterogeneous among individuals.

The estimated kurtoses of inflation perceptions and expectations are often (but not always) imprecise. This is because interval data with open ends have little information about the tails of the distribution of the underlying variable. We still see that the excess kurtoses are mostly positive for both inflation perceptions and expectations, suggesting existence of outliers.

The estimated asymmetries in the tails of inflation perceptions and expectations are often (but not always) extremely imprecise. The reason is the same as for the kurtoses. We still see that the skewness and asymmetry in the tails

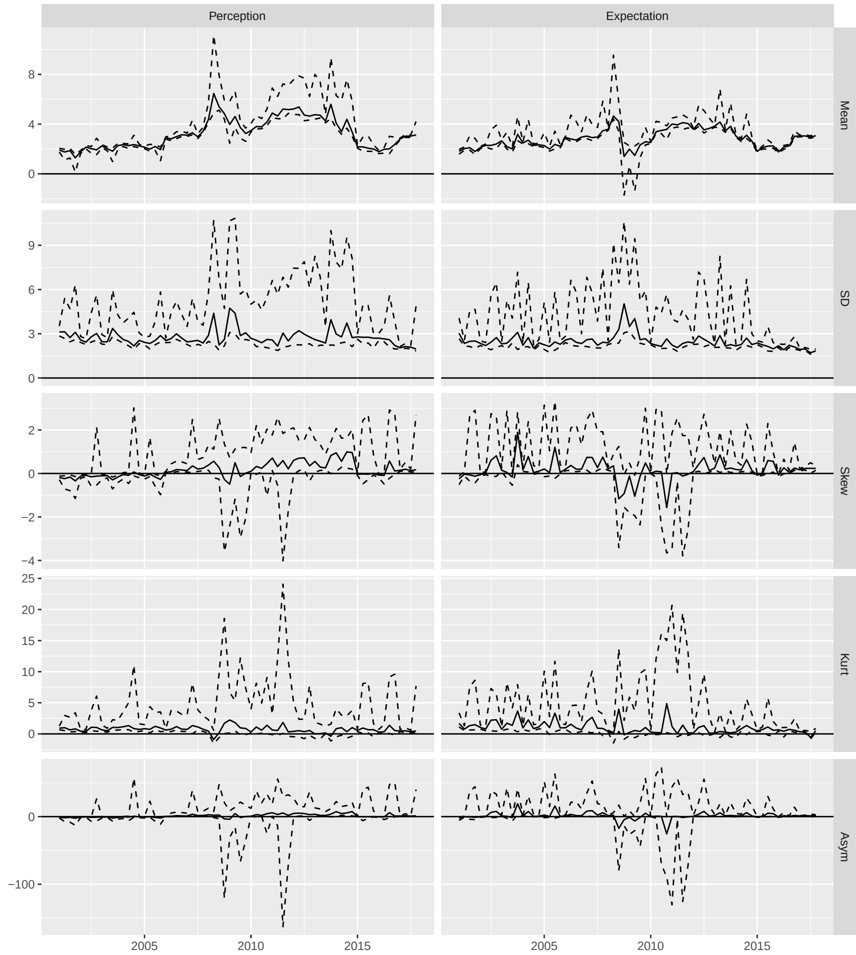

Figure 6: Posterior medians of moments of the distributions of inflation perceptions (left) and expectations (right). The dashed lines are 95% error bands

often move in the same direction.

5.5 Estimation with an indifference limen

Next, we estimate the model with a common indifference limen among individuals. This may improve estimation efficiency, since we have two more intervals, together with the prior information about the location of the indifference limen.

Since the prior on the lower bound of the indifference limen is improper, so is the marginal probability of the data (marginal likelihood); hence we do not estimate the Bayes factor for $K = 1$ vs $K = 2$ here. The previous estimates ignoring the indifference limen already justify use of a normal mixture model.

Figure 7 plots the posterior medians of the five moments and the indifference limens for inflation perceptions and expectations, respectively, during 2001Q1–2017Q4, together with 95% error bands. We summarize our findings below.

The estimated means of inflation perceptions and expectations are almost identical to the previous results with no indifference limen. Though we use more information in estimation, especially when many respondents choose categories 1–3, the differences in the 95% error bands are hardly visible.

The estimated higher-order moments of inflation perceptions and expectations are sometimes (though not always) larger in magnitude than the previous results with no indifference limen. The difference may arise because with two more intervals, the data have more information about the dispersion and tails of the underlying distribution.

The estimated indifference limens for inflation perceptions and expectations are both time-varying, and mostly asymmetric around 0. In particular, given our weakly informative prior about the location of the indifference limen, the lower bounds are often far below -1.0%, whereas the upper bounds stay within .0–1.0%. The error bands are wider for lower bounds than for upper bounds. This is perhaps because category 1 ('Gone/Go down') has no lower bound and often has fewer observations than category 3 ('Up by 1% or less').

In addition to moments, quantiles help us to see the whole shape of a distribution.[16] Figure 8 plots the posterior medians of the deciles of the distributions of inflation perceptions and expectations, respectively, during 2001Q1–2017Q4. These plots seem useful for monitoring the distributions of inflation perceptions and expectations.

5.6 Using data with more intervals

Finally, we estimate the model using data with 18 intervals including an indifference limen, which have been available since 2011Q2. With further questions to those who have chosen either category 1 or 8, the respondents now virtually choose from the following 18 intervals:

1. Down by 5% or more
2. Down by 4% but less than 5%

...

[16] The `nor1mix` package for R is useful for calculating quantiles of a mixture normal distribution.

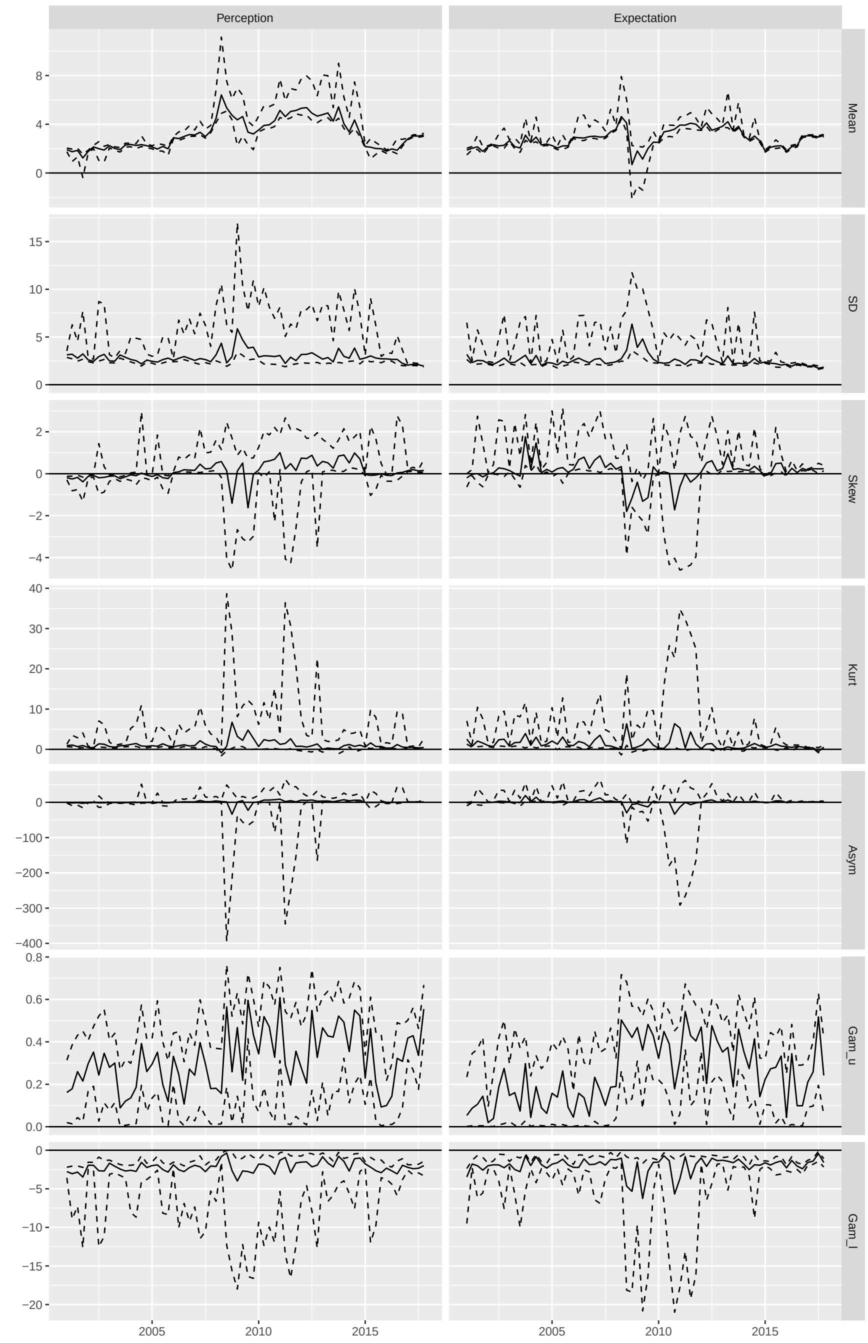

Figure 7: Posterior medians of moments and indifference limens of the distributions of inflation perceptions (left) and expectations (right). The dashed lines are 95% error bands

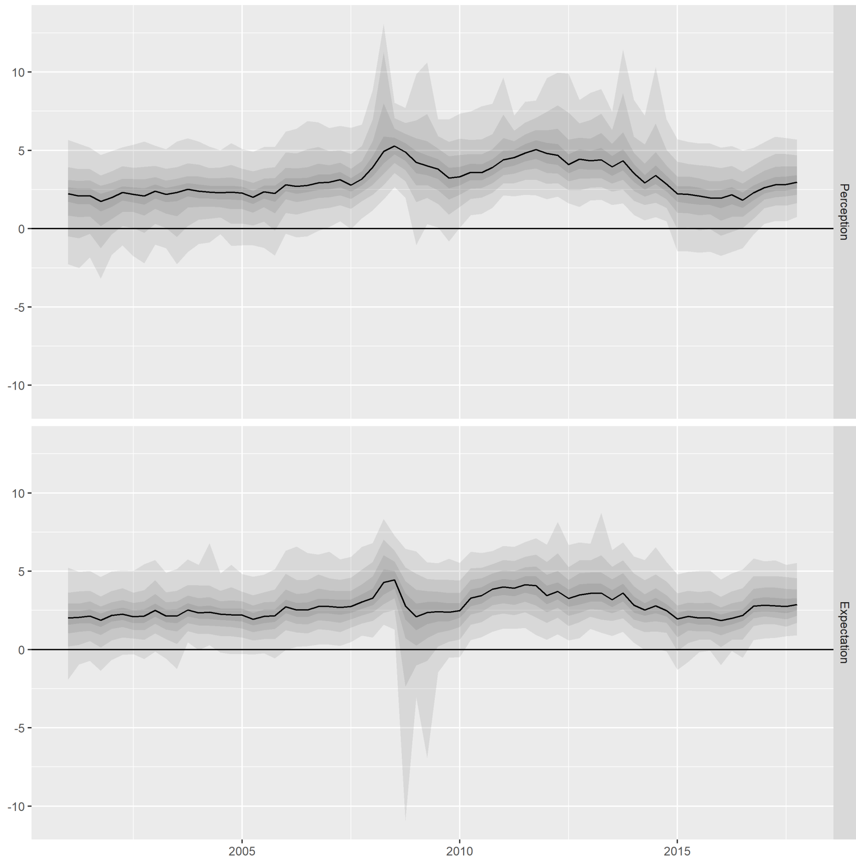

Figure 8: Posterior medians of the deciles of the distributions of inflation perceptions (top) and expectations (bottom)

6. Down by 1% or less
7. Not changed/change
8. Up by 1% or less

   ...

17. Up by 9% but less than 10%
18. Up by 10% or more

Since $-1 < \gamma_l < 0 < \gamma_u < 1$, our weakly informative prior about the location of the indifference limen now sets $\lambda_0 := -1$ and $\upsilon_0 := 1$. It turns out that with 18 intervals, the NUTS works well without adjusting the tuning parameters; thus we use the default values for all tuning parameters.

Figure 9 plots the posterior medians of the five moments and indifference limens for inflation perceptions and expectations, respectively, using data with 18 intervals during 2011Q2–2017Q4. We summarize our findings below.

The estimated means are similar to the previous results using data with six or eight intervals during 2011Q2–2017Q4, except for narrower error bands. The means are higher for inflation perceptions than for expectations when they are both high, but close to each other when they are 2–3%. Thus individuals may expect the inflation rate to fluctuate around 2–3% in the long run.

The estimated standard deviations are slightly larger than the previous results using data with six or eight intervals during 2011Q2–2017Q4, with much narrower error bands. This is perhaps because with 18 intervals, the data have much more information about the dispersion of the underlying distribution. The standard deviations for inflation perceptions and expectations are close to each other. The mean and standard deviation tend to move in the same direction. Thus individuals may agree with inflation perceptions and expectations when they are low, but not when they are high.

The estimated skewnesses are often higher than the previous results using data with six or eight intervals during 2011Q2–2017Q4, again with much narrower error bands. The skewnesses are positive for both inflation perceptions and expectations, and slightly higher for expectations than for perceptions. The mean and skewness tend to move in the opposite directions.

The estimated kurtoses are also often higher than the previous results using data with six or eight intervals during 2011Q2–2017Q4. The excess kurtoses are mostly positive for both inflation perceptions and expectations, and higher for expectations than for perceptions. The mean and kurtosis tend to move in the opposite directions.

The estimated asymmetries in the tails are often much higher than the previous results using data with six or eight intervals during 2011Q2–2017Q4. The asymmetries in the tails are positive for both inflation perceptions and expectations, and higher for expectations than for perceptions. The mean and asymmetry tend to move in the opposite directions. Overall, higher-order moments tend to move in the same directions. This may suggest existence of consistently positive outliers.

The estimated indifference limens, especially the lower bounds, differ from the previous results using data with eight intervals during 2011Q2–2017Q4. The upper bounds are around .3–.7% for perceptions and .4–.6% for expectations.

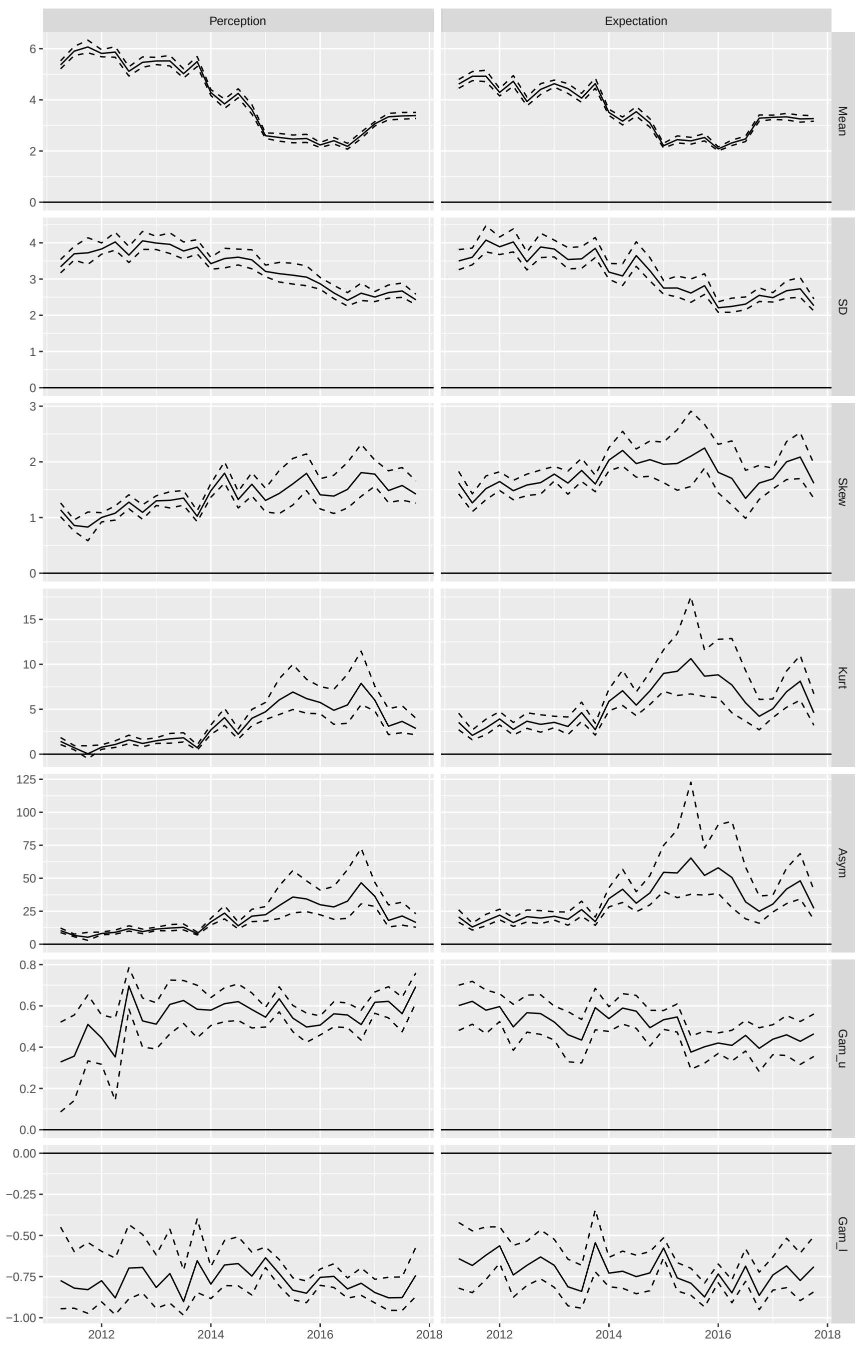

Figure 9: Posterior medians of moments and indifference limens of the distributions of inflation perceptions (left) and expectations (right) using data with 18 intervals. The dashed lines are 95% error bands

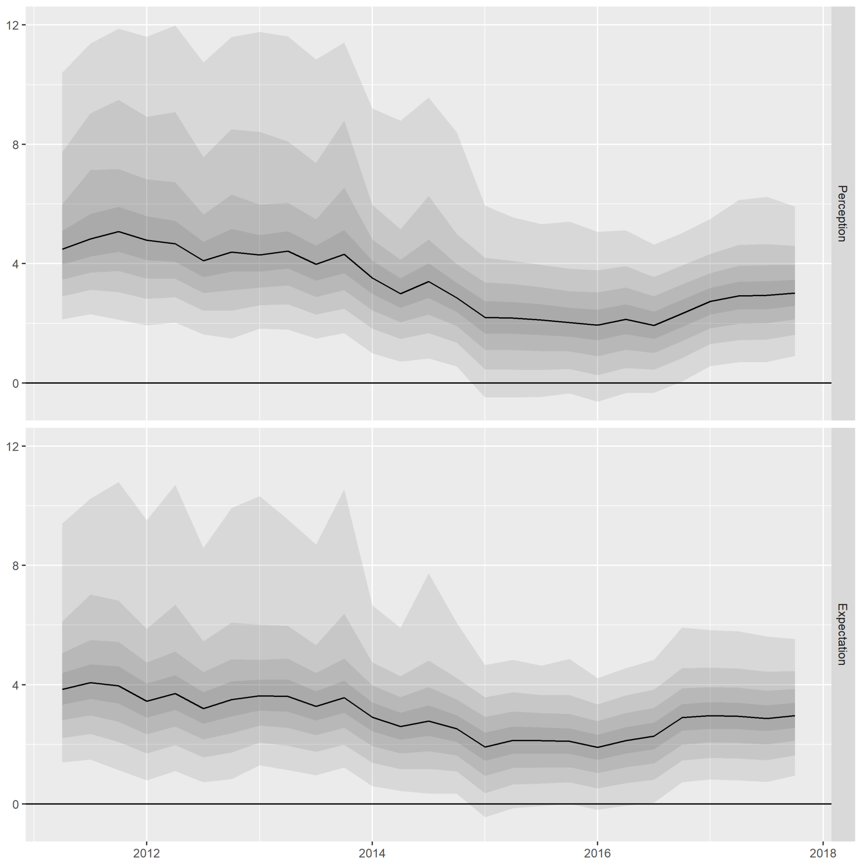

Figure 10: Posterior medians of the deciles of the distributions of inflation perceptions (top) and expectations (bottom) using data with 18 intervals

Given the stronger prior, the lower bounds stay above -1.0%. They are often below $-.7\%$, however, and sometimes close to -1.0%. Thus the indifference limens are still often asymmetric around 0.

Figure 10 plots the posterior medians of the deciles of the distributions of inflation perceptions and expectations, respectively, using data with 18 intervals. Though we do not show error bands, the estimated deciles are more precise than the previous results using data with eight intervals during 2011Q2–2017Q4.

6 Information rigidity in the UK

Information rigidity prevents one from forming full-information rational expectations. To measure and test for the degree of information rigidity, Coibion and Gorodnichenko (2015) give a simple framework that requires only the historical means of the distributions of expectations. When only aggregate interval data are available, our method helps to obtain the mean of the possibly multi-modal underlying distribution. To illustrate such a use of our method, we measure the degrees of information rigidity in public inflation perceptions and expectations in the UK.

Consider individuals forming h-step ahead forecasts of a time series $\{y_t\}$. A sticky information model proposed by Mankiw and Reis (2002) assumes that an individual receives no information with probability $\lambda \in [0,1]$ at each date. Let $\bar{y}^e_{t+h|t}$ be the average forecast of y_{t+h} at date t among individuals. Then for all t, for $h \geq 0$,

$$\begin{aligned}\bar{y}^e_{t+h|t} &:= (1-\lambda)\,\mathrm{E}_t(y_{t+h}) + \lambda(1-\lambda)\,\mathrm{E}_{t-1}(y_{t+h}) + \lambda^2(1-\lambda)\,\mathrm{E}_{t-2}(y_{t+h}) + \cdots \\ &= (1-\lambda)\,\mathrm{E}_t(y_{t+h}) + \lambda[(1-\lambda)\,\mathrm{E}_{t-1}(y_{t+h}) + \lambda(1-\lambda)\,\mathrm{E}_{t-2}(y_{t+h}) + \cdots] \\ &= (1-\lambda)\,\mathrm{E}_t(y_{t+h}) + \lambda \bar{y}^e_{t+h|t-1} \qquad (24)\end{aligned}$$

or

$$\begin{aligned}\mathrm{E}_t(y_{t+h}) &= \frac{1}{1-\lambda}\bar{y}^e_{t+h|t} - \frac{\lambda}{1-\lambda}\bar{y}^e_{t+h|t-1} \\ &= \bar{y}^e_{t+h|t} + \frac{\lambda}{1-\lambda}\left(\bar{y}^e_{t+h|t} - \bar{y}^e_{t+h|t-1}\right) \qquad (25)\end{aligned}$$

This gives a simple regression model with no intercept for the ex post forecast error $y_{t+h} - \bar{y}^e_{t+h|t}$ given the ex ante mean forecast revision $\bar{y}^e_{t+h|t} - \bar{y}^e_{t+h|t-1}$ such that for all t, for $h \geq 0$,

$$\mathrm{E}_t\left(y_{t+h} - \bar{y}^e_{t+h|t}\right) = \beta\left(\bar{y}^e_{t+h|t} - \bar{y}^e_{t+h|t-1}\right) \qquad (26)$$

where $\beta := \lambda/(1-\lambda) \geq 0$, or

$$y_{t+h} - \bar{y}^e_{t+h|t} = \beta\left(\bar{y}^e_{t+h|t} - \bar{y}^e_{t+h|t-1}\right) + u_{t+h|t} \qquad (27)$$

where $u_{t+h|t} := y_{t+h} - \mathrm{E}_t(y_{t+h})$.[17] One can estimate β by OLS, from which one can recover $\lambda = \beta/(1+\beta) \in [0,1)$. Note that β is common across h.

One may observe only $\left\{\bar{y}^e_{t+h|t}\right\}$ instead of $\left\{\bar{y}^e_{t+h|t}, \bar{y}^e_{t+h|t-1}\right\}$. In such a case, one can consider an alternative model such that for all t, for $h \geq 0$,

$$y_{t+h} - \bar{y}^e_{t+h|t} = \beta\left(\bar{y}^e_{t+h|t} - \bar{y}^e_{t+h-1|t-1}\right) + v_{t+h|t} \qquad (28)$$

where

$$v_{t+h|t} := u_{t+h|t} - \beta\left(\bar{y}^e_{t+h|t-1} - \bar{y}^e_{t+h-1|t-1}\right) \qquad (29)$$

This is not a regression model; hence one must apply the IV method instead of OLS, where the IV must be a white noise sequence. Coibion and Gorodnichenko (2015, p. 2663) use the change in the log oil price as the IV when measuring information rigidity in inflation expectations.

Let $\{P_t\}$ be the quarterly series of the price level and $\pi_t := 100\ln(P_t/P_{t-4})$ be the annual inflation rate in per cent. Consider information rigidity in h-step ahead forecasts of $\{\pi_t\}$, where $h = 0$ (nowcasts) for inflation perceptions and $h = 4$ for inflation expectations. We use the CPI and RPI for $\{P_t\}$, whose data are available at the website of the Office for National Statistics of the UK. The CPI and RPI inflation rates are somewhat different; see Figure 11.

[17] A noisy information model gives the same regression model if $\{y_t\}$ is AR(1); see Coibion and Gorodnichenko (2015, sec. I.B).

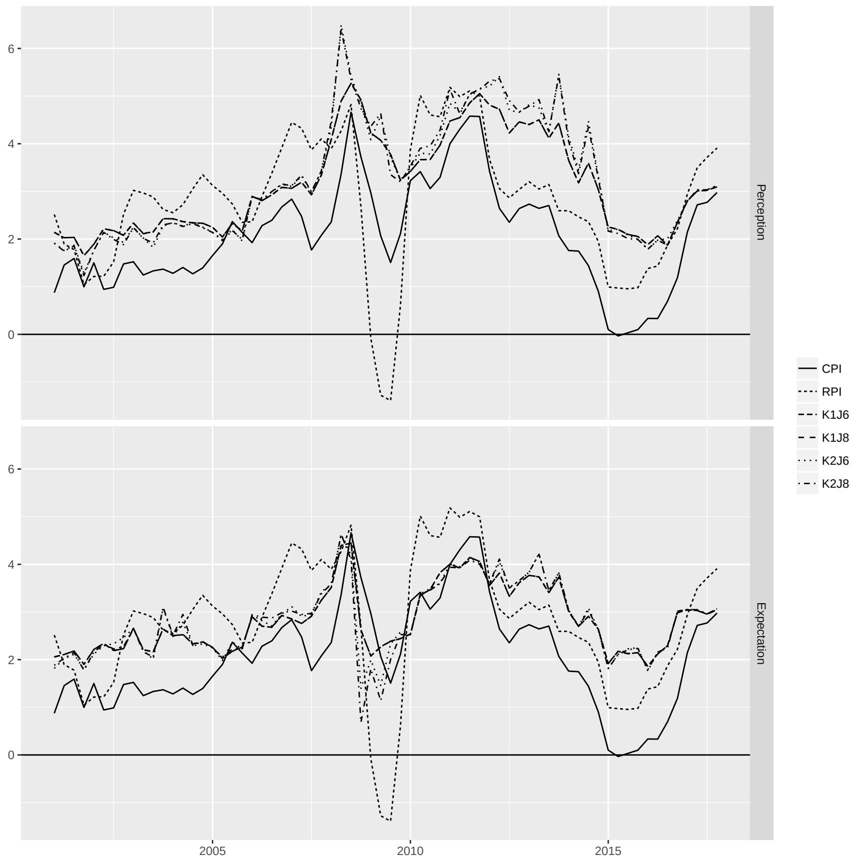

Figure 11: Annual CPI and RPI inflation rates and alternative estimates of the means of inflation perceptions (top) and expectations (bottom). Source: Bank of England and the author's calculation.

Let $\bar{\pi}^e_{t+h|t}$ be the mean of the distribution of inflation perceptions ($h = 0$) or expectations ($h = 4$). For $h = 0, 4$, we apply our method to estimate $\bar{\pi}^e_{t+h|t}$ repeatedly during 2001Q1–2017Q4, assuming a normal distribution ($K = 1$) or a mixture of two normal distributions ($K = 2$) with or without an indifference limen ($J = 6, 8$).[18] The resulting $\left\{\bar{\pi}^e_{t+h|t}\right\}$ are similar across J but somewhat different across K, especially for inflation perceptions when they are high; see Figure 11.

Since it is unclear if individuals forecast the CPI, RPI, or something else, to control for the possible bias in forecasts, we add an intercept to equation (28). Thus we have a bivariate linear model such that for all t,

$$y_t - \bar{y}^e_{t|t} = \alpha_0 + \beta_0 \left(\bar{y}^e_{t|t} - \bar{y}^e_{t-1|t-1}\right) + v_{t|t} \tag{30}$$

$$y_{t+4} - \bar{y}^e_{t+4|t} = \alpha_4 + \beta_4 \left(\bar{y}^e_{t+4|t} - \bar{y}^e_{t+3|t-1}\right) + v_{t+4|t} \tag{31}$$

We estimate the two equations jointly by iterated GMM. As IVs, in addition to a constant, we use the Brent Crude Oil Price in US dollars and the GBP/USD exchange rate, which are downloadable from Federal Reserve Economic Data (FRED). To make the IVs white noise sequences, we take the difference in log, and prewhiten each series by fitting a univariate AR(2) model for the full sample period and extracting the residuals.[19]

The sticky information model implies that $\beta_0 = \beta_4$. Hence we estimate the model with or without this cross-equation restriction, and test

$$H_0 : \beta_0 = \beta_4 \quad \text{vs} \quad H_1 : \beta_0 \neq \beta_4 \tag{32}$$

Let J_0 and J_1 be the Sargan–Hansen J statistics for testing the overidentifying restrictions under H_0 and H_1, respectively. Then under H_0, $D := J_0 - J_1 \xrightarrow{d} \chi^2(1)$; see Hall (2005, pp. 162–163). We use gretl 2018a for these analyses.

Table 1 shows the estimation results using alternative measures of inflation and estimates of the means of inflation perceptions and expectations under H_0 and H_1, respectively. Table 2 summarizes the results of testing H_0 vs H_1. We find the following:

1. For both the CPI and RPI and for all estimates of the means of inflation perceptions and expectations, the test fails to reject H_0, which is consistent with the sticky information model.[20] Thus we focus on the results under H_0.

2. For the CPI, the intercepts are negative and significant for both inflation perceptions and expectations, suggesting positive forecast biases. For the RPI, the intercepts are negative and significant only for perceptions. The forecast biases are larger for the CPI and perceptions than for the RPI and expectations.

[18] For $K = 1$, we use the flat prior on (μ, σ) and the default tuning parameters. For $K = 2$, the estimates are the same as those in Figures 6 and 7.

[19] The full sample periods are 1987Q4–2017Q4 for the change in the log oil price and 1971Q2–2017Q4 for the change in the log exchange rate.

[20] Coibion and Gorodnichenko (2015, pp. 2664–2665) obtain similar results using data on forecasts of various macroeconomic variables from the Survey of Professional Forecasters.

3. The slopes are positive and significant only for the RPI. Importantly, the size of the slope depends on K and J, i.e., the choice of the number of mixture components and whether we include the indifference limen or not when estimating the means of inflation perceptions and expectations.

4. If we believe that individuals forecast the RPI, assume $K = 2$, and include the indifference limen when estimating the means of inflation perceptions and expectations, then $\beta = .96$, which implies $\lambda \approx .49$.[21]

To sum up, we find that evidences of information rigidity among UK individuals forecasting inflation depend not only on what prices we assume they forecast, but also on how we model the distributions of their inflation perceptions and expectations. Thus we recommend using a flexible distribution, e.g., a mixture distribution, when fitting a distribution to interval data on inflation perceptions and expectations.

7 Conclusion

To estimate the distributions of public inflation perceptions and expectations in the UK, we study Bayesian analysis of a normal mixture model for interval data with an indifference limen. A hierarchical prior helps to obtain a weakly informative prior. Since the boundaries of an indifference limen are similar to cutpoints in an ordered response model, the Gibbs sampler and M–H algorithm are slow to converge for large samples. We find that the NUTS converges fast even for large samples. Thus the NUTS is useful especially when one must apply MCMC repeatedly, which is our case.

One can extend and improve our method in several ways. Here we list some directions for future work:

1. Instead of assuming a common indifference limen among individuals (or simply ignoring it), one can model heterogeneous indifference limens, say, assuming beta distributions for the boundaries. This makes the likelihood function more complicated, but seems worth trying.

2. One can estimate the number of mixture components instead of fixing it a priori. Common model selection criteria may be inappropriate for finite mixture models, however; see Gelman et al (2014, p. 536). Since HMC does not allow discrete parameters, MCMC for the number of components requires combining HMC with another MCMC method. A model with a few components is a special case of a model with many components; thus one may simply consider the latter with a suitable prior on the number of components; see Gelman et al (2014, p. 536).

3. Instead of fitting a distribution each quarter separately, one can use multiple samples jointly to study the dynamics of the underlying distribution directly. For time series analysis of repeated interval data, one may need a state-space model with a nonlinear measurement equation.

[21]For US consumers, Coibion and Gorodnichenko (2015, p. 2662) report that $\beta = .705$ for inflation expectations relative to the CPI, which implies $\lambda \approx .413$. Thus information rigidities in inflation expectations among individuals in the UK and US seem not too different, though the two results may not be directly comparable.

Table 1: Iterated GMM estimates of the bivariate linear model of the ex post forecast errors on the ex ante mean forecast revisions using alternative measures of inflation and estimates of the means of inflation perceptions and expectations

Equal slopes (H_0)

			CPI		RPI	
K	J	h	Intercept	Slope	Intercept	Slope
1	6	0	-1.11^{***}		$-.27^{*}$	
			(.12)	.66	(.16)	1.93^{***}
		4	$-.75^{***}$	(.48)	.16	(.44)
			(.20)		(.24)	
1	8	0	-1.10^{***}		$-.27^{*}$	
			(.12)	.65	(.16)	1.92^{***}
		4	$-.75^{***}$	(.48)	.16	(.44)
			(.20)		(.24)	
2	6	0	-1.16^{***}		$-.35^{*}$	
			(.16)	.02	(.19)	1.04^{***}
		4	$-.68^{***}$	(.38)	.07	(.36)
			(.23)		(.26)	
2	8	0	-1.17^{***}		$-.36^{*}$	
			(.15)	$-.02$	(.19)	$.96^{***}$
		4	$-.64^{***}$	(.40)	.09	(.37)
			(.23)		(.27)	

Unequal slopes (H_1)

			CPI		RPI	
K	J	h	Intercept	Slope	Intercept	Slope
1	6	0	-1.10^{***}	.60	$-.32^{*}$	2.35^{***}
			(.12)	(.51)	(.17)	(.79)
		4	$-.72^{***}$	1.10	.10	1.52^{***}
			(.21)	(.82)	(.25)	(.54)
1	8	0	-1.10^{***}	.60	$-.32^{*}$	2.34^{***}
			(.12)	(.51)	(.17)	(.78)
		4	$-.72^{***}$	1.10	.10	1.52^{***}
			(.21)	(.82)	(.25)	(.54)
2	6	0	-1.16^{***}	.00	$-.39^{**}$	1.28^{**}
			(.16)	(.39)	(.19)	(.52)
		4	$-.65^{***}$	.51	$-.02$	.66
			(.23)	(.74)	(.28)	(.45)
2	8	0	-1.18^{***}	.10	$-.42^{**}$	1.21^{**}
			(.15)	(.44)	(.19)	(.52)
		4	$-.71^{***}$	.98	.04	.45
			(.24)	(1.03)	(.28)	(.45)

Notes: *, **, and *** denote significance at the 10, 5, and 1% levels, respectively. Numbers in parentheses are Newey–West HAC standard errors. The estimation periods are 2001Q2–2016Q4.

Table 2: Tests of equality of slopes across forecast horizons using alternative measures of inflation and estimates of the means of inflation perceptions and expectations

			CPI			RPI	
K	J	J_0	J_1	D	J_0	J_1	D
1	6	2.91	2.33	.58	1.58	1.02	.56
		(.41)	(.31)	(.45)	(.66)	(.60)	(.45)
1	8	2.91	2.32	.57	1.58	1.02	.55
		(.41)	(.31)	(.45)	(.67)	(.60)	(.46)
2	6	3.88	3.02	.86	1.42	.84	.59
		(.27)	(.22)	(.35)	(.70)	(.66)	(.44)
2	8	3.84	2.41	1.42	1.44	.68	.76
		(.28)	(.30)	(.23)	(.70)	(.71)	(.38)

Notes: The asymptotic null distributions of J_0, J_1, and D are $\chi^2(3)$, $\chi^2(2)$, and $\chi^2(1)$, respectively. Numbers in parentheses are asymptotic p-values.

One can also use the estimated distributions for further analyses; e.g., testing rationality of forecasts using the estimated means, measuring macroeconomic uncertainty using the estimated higher-order moments, etc.

Moreover, micro data from the Bank of England Inflation Attitudes Survey are now publicly available at their website, suggesting the following extensions of this paper:

1. One can fit a bivariate mixture normal distribution to bivariate interval data on inflation perceptions and expectations; i.e., a multivariate extension of this paper.

2. One can study the determinants of individual inflation perceptions and expectations, separately or jointly, by interval regression or its extension that allows for an indifference limen. This is an extension of Blanchflower and MacCoille (2009).

Thus the micro data will greatly help us to study the formation, dynamics, and interaction of inflation perceptions and expectations.

Appendix: HMC methods and the NUTS

Understanding an HMC method requires some knowledge of physics. The Boltzmann (Gibbs, canonical) distribution describes the distribution of possible states in a system. Given an energy function $E(.)$ and the absolute temperature T, the pdf of a Boltzmann distribution is $\forall x$,

$$f(x) \propto \exp\left(-\frac{E(x)}{kT}\right) \tag{A.1}$$

where k is the Boltzmann constant. One can think of any pdf $f(.)$ as the pdf of a Boltzmann distribution with $E(x) := -\ln f(x)$ and $kT = 1$.

The Hamiltonian is an energy function that sums the potential and kinetic energies of a state. An HMC method treats $-\ln p(\boldsymbol{\theta}|\boldsymbol{y})$ as the potential energy at position $\boldsymbol{\theta}$, and introduces an auxiliary momentum $\boldsymbol{z}$ drawn randomly from $\mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma})$, whose kinetic energy is $-\ln p(\boldsymbol{z}) \propto \boldsymbol{z}'\boldsymbol{\Sigma}^{-1}\boldsymbol{z}/2$ with mass matrix $\boldsymbol{\Sigma}$. The resulting Hamiltonian is

$$\begin{aligned} H(\boldsymbol{\theta}, \boldsymbol{z}) &:= -\ln p(\boldsymbol{\theta}|\boldsymbol{y}) - \ln p(\boldsymbol{z}) \\ &= -\ln p(\boldsymbol{\theta}, \boldsymbol{z}|\boldsymbol{y}) \end{aligned} \tag{A.2}$$

An HMC method draws from $p(\boldsymbol{\theta}, \boldsymbol{z}|\boldsymbol{y})$, which is a Boltzmann distribution with energy function $H(\boldsymbol{\theta}, \boldsymbol{z})$.

The Hamiltonian is constant over (fictitious) time t by the law of conservation of mechanical energy, i.e., $\forall t \in \mathbb{R}$,

$$\dot{H}(\boldsymbol{\theta}(t), \boldsymbol{z}(t)) = 0 \tag{A.3}$$

or

$$\dot{\boldsymbol{\theta}}(t) H_{\boldsymbol{\theta}}(\boldsymbol{\theta}(t), \boldsymbol{z}(t)) + \dot{\boldsymbol{z}}(t) H_{\boldsymbol{z}}(\boldsymbol{\theta}(t), \boldsymbol{z}(t)) = \mathbf{0} \tag{A.4}$$

Thus Hamilton's equation of motion is $\forall t \in \mathbb{R}$,

$$\dot{\boldsymbol{\theta}}(t) = H_{\boldsymbol{z}}(\boldsymbol{\theta}(t), \boldsymbol{z}(t)) \tag{A.5}$$

$$\dot{\boldsymbol{z}}(t) = -H_{\boldsymbol{\theta}}(\boldsymbol{\theta}(t), \boldsymbol{z}(t)) \tag{A.6}$$

The Hamiltonian dynamics says that $(\boldsymbol{\theta}, \boldsymbol{z})$ moves on a contour of $H(\boldsymbol{\theta}, \boldsymbol{z})$, i.e., $p(\boldsymbol{\theta}, \boldsymbol{z}|\boldsymbol{y})$.

Conceptually, given $\boldsymbol{\Sigma}$ and an initial value for $\boldsymbol{\theta}$, an HMC method proceeds as follows:

1. Draw $\boldsymbol{z} \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma})$ independently from $\boldsymbol{\theta}$.
2. Start from $(\boldsymbol{\theta}, \boldsymbol{z})$ and apply Hamilton's equations of motion for a certain length of (fictitious) time to obtain $(\boldsymbol{\theta}', \boldsymbol{z}')$, whose joint probability density equals that of $(\boldsymbol{\theta}, \boldsymbol{z})$.
3. Discard $\boldsymbol{z}$ and $\boldsymbol{z}'$, and repeat.

This gives a reversible Markov chain on $(\boldsymbol{\theta}, \boldsymbol{z})$, since the Hamiltonian dynamics is reversible; see Neal (2011, p. 116). The degree of serial dependence or speed of convergence depends on the choice of $\boldsymbol{\Sigma}$ and the length of (fictitious) time in the second step. The latter can be fixed or random, but cannot be adaptive if it breaks reversibility.

In practice, an HMC method approximates Hamilton's equations of motion in discrete steps using the leapfrog method. This requires choosing a step size ϵ and the number of steps L. Because of approximation, the Hamiltonian is no longer constant during the leapfrog method, but adding a Metropolis step after the leapfrog method keeps reversibility. Thus given $\boldsymbol{\Sigma}$ and an initial value for $\boldsymbol{\theta}$, an HMC method proceeds as follows:

1. Draw $\boldsymbol{z} \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma})$ independently from $\boldsymbol{\theta}$.
2. Start from $(\boldsymbol{\theta}, \boldsymbol{z})$ and apply Hamilton's equations of motion approximately by the leapfrog method to obtain $(\boldsymbol{\theta}', \boldsymbol{z}')$.

3. Accept $(\boldsymbol{\theta}', \boldsymbol{z}')$ with probability $\min\{\exp(-H(\boldsymbol{\theta}', \boldsymbol{z}'))/\exp(-H(\boldsymbol{\theta}, \boldsymbol{z})), 1\}$.[22]
4. Discard $\boldsymbol{z}$ and $\boldsymbol{z}'$, and repeat.

The degree of serial dependence or speed of convergence depends on the choice of $\boldsymbol{\Sigma}$, ϵ, and L. Moreover, the computational cost of each iteration depends on the choice of ϵ and L.

The NUTS developed by Hoffman and Gelman (2014) adaptively chooses L while keeping reversibility. Though the algorithm of the NUTS is complicated, it is easy to use with Stan, a modeling language for Bayesian computation with the NUTS (and other methods). Stan tunes $\boldsymbol{\Sigma}$ and ϵ adaptively during warmup. The user only specifies the data, model, and prior. One can call Stan from other popular languages and softwares such as R, Python, Matlab, Julia, Stata, and Mathematica. The NUTS often has better convergence properties than other popular MCMC methods such as the Gibbs sampler and M–H algorithm.

References

Albert JH, Chib S (1993) Bayesian analysis of binary and polychotomous response data. Journal of the American Statistical Association 88:669–679, DOI 10.2307/2290350

Alston CL, Mengersen KL (2010) Allowing for the effect of data binning in a Bayesian normal mixture model. Computational Statistics & Data Analysis 54:916–923, DOI 10.1016/j.csda.2009.10.003

Armantier O, Bruine de Bruin W, Potter S, Topa G, van der Klaauw W, Zafar B (2013) Measuring inflation expectations. Annual Review of Economics 5:273–301, DOI 10.1146/annurev-economics-081512-141510

Betancourt M, Girolami M (2015) Hamiltonian Monte Carlo for hierarchical models. In: Upadhyay SK, Singh U, Dey DK, Loganathan A (eds) Current Trends in Bayesian Methodology with Applications, CRC Press, chap 4, pp 79–102

Biernacki C (2007) Degeneracy in the maximum likelihood estimation of univariate Gaussian mixtures for grouped data and behavior of the EM algorithm. Scandinavian Journal of Statistics 34:569–586, DOI 10.1111/j.1467-9469.2006.00553.x

Binder CC (2017) Measuring uncertainty based on rounding: New method and application to inflation expectations. Journal of Monetary Economics 90:1–12, DOI http://dx.doi.org/10.1016/j.jmoneco.2017.06.001

Blanchflower DG, MacCoille C (2009) The formation of inflation expectations: An empirical analysis for the UK. Working Paper 15388, National Bureau of Economic Research, DOI 10.3386/w15388

[22] Since the leapfrog method preserves volume, its Jacobian of transformation is 1; see Neal (2011, pp. 117–122). To justify the Metropolis step, the proposal should be $(\boldsymbol{\theta}', -\boldsymbol{z}')$ rather than $(\boldsymbol{\theta}', \boldsymbol{z}')$, but this is unnecessary in practice since $H(\boldsymbol{\theta}', \boldsymbol{z}') = H(\boldsymbol{\theta}', -\boldsymbol{z}')$ and we discard $\boldsymbol{z}'$ anyway; see Neal (2011, p. 124). Since the Hamiltonian is approximately constant during the leapfrog method, the acceptance probability is close to 1 if ϵ is small.

Chen MH, Dey DK (2000) Bayesian analysis for correlated ordinal data models. In: Dey DK, Ghosh SK, Mallick BK (eds) Generalized Linear Models: A Bayesian Perspective, Marcel Dekker, chap 8, pp 133–157

Chib S, Greenberg E (1995) Understanding the Metropolis–Hastings algorithm. American Statistician 4:327–335, DOI 10.1080/00031305.1995.10476177

Coibion O, Gorodnichenko Y (2015) Information rigidity and the expectations formation process: A simple framework and new facts. American Economic Review 105:2644–2678, DOI 10.1257/aer.20110306

Cowles MK (1996) Accelerating Monte Carlo Markov chain convergence for cumulative-link generalized linear models. Statistics and Computing 6:101–111, DOI 10.1007/BF00162520

Frühwirth-Schnatter S (2004) Estimating marginal likelihoods for mixture and Markov switching models using bridge sampling techniques. Econometrics Journal 7:143–167, DOI 10.1111/j.1368-423x.2004.00125.x

Frühwirth-Schnatter S (2006) Finite Mixture and Markov Switching Models. Springer

Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB (2014) Bayesian Data Analysis, 3rd edn. CRC Press

Geweke J (2007) Interpretation and inference in mixture models: Simple MCMC works. Computational Statistics & Data Analysis 51:3529–3550, DOI 10.1016/j.csda.2006.11.026

Gronau QF, Sarafoglou A, Matzke D, Ly A, Boehm U, Marsman M, Leslie DS, Forster JJ, Wagenmakers EJ, Steingroever H (2017) A tutorial on bridge sampling. Journal of Mathematical Psychology 81:80–97, DOI 10.1016/j.jmp.2017.09.005

Hall AR (2005) Generalized Method of Moments. Oxford University Press

Hoffman MD, Gelman A (2014) The No-U-Turn Sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. Journal of Machine Learning Research 15:1593–1623

Jeffreys H (1961) Theory of Probability, 3rd edn. Clarendon Press

Kamada K, Nakajima J, Nishiguchi S (2015) Are household inflation expectations anchored in Japan? Working Paper 15-E-8, Bank of Japan

Kiefer J, Wolfowitz J (1956) Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters. Annals of Mathematical Statistics 27:887–906, DOI 10.1214/aoms/1177728066

Lahiri K, Zhao Y (2015) Quantifying survey expectations: A critical review and generalization of the Carlson–Parkin method. International Journal of Forecasting 31:51–62, DOI 10.1016/j.ijforecast.2014.06.003

Liu JS, Sabatti C (2000) Generalized Gibbs sampler and multigrid Monte Carlo for Bayesian computation. Biometrika 87:353–369, DOI 10.1093/biomet/87.2.353

Lombardelli C, Saleheen J (2003) Public expectations of UK inflation. Bank of England Quarterly Bulletin 43:281–290

Mankiw NG, Reis R (2002) Sticky information versus sticky prices: A proposal to replace the new Keynesian Phillips curve. Quarterly Journal of Economics 117:1295–1328, DOI 10.1162/003355302320935034

Mankiw NG, Reis R, Wolfers J (2004) Disagreement about inflation expectations. In: Gertler M, Rogoff K (eds) NBER Macroeconomics Annual 2003, vol 18, MIT Press, pp 209–248

Manski CF (2004) Measuring expectations. Econometrica 72:1329–1376, DOI 10.1111/j.1468-0262.2004.00537.x

Manski CF, Molinari F (2010) Rounding probabilistic expectations in surveys. Journal of Business & Economic Statistics 28:219–231, DOI 10.1198/jbes.2009.08098

Meng XL, Schilling S (2002) Warp bridge sampling. Journal of Computational and Graphical Statistics 11:552–586, DOI 10.1198/106186002457

Murasawa Y (2013) Measuring inflation expectations using interval-coded data. Oxford Bulletin of Economics and Statistics 75:602–623, DOI 10.1111/j.1468-0084.2012.00704.x

Nandram B, Chen MH (1996) Reparameterizing the generalized linear model to accelerate Gibbs sampler convergence. Journal of Statistical Computation and Simulation 54:129–144, DOI 10.1080/00949659608811724

Nardo M (2003) The quantification of qualitative survey data: A critical assessment. Journal of Economic Surveys 17:645–668, DOI 10.1046/j.1467-6419.2003.00208.x

Neal RM (2011) MCMC using Hamiltonian dynamics. In: Brooks S, Gelman A, Jones GL, Meng XL (eds) Handbook of Marcov Chain Monte Carlo, Handbooks of Modern Statistical Methods, Chapman & Hall/CRC, chap 5, pp 113–162

Papaspiliopoulos O, Roberts GO, Sköld M (2003) Non-centered parameterisations for hierarchical models and data augmentation. In: Bernardo JM, Bayarri MJ, Berger JO, Dawid AP, Heckerman D, Smith AFM, West M (eds) Bayesian Statistics 7, Oxford University Press, pp 307–326

Papaspiliopoulos O, Roberts GO, Sköld M (2007) A general framework for the parametrization of hierarchical models. Statistical Science 22:59–73, DOI 10.1214/088342307000000014

Pesaran MH, Weale M (2006) Survey expectations. In: Elliot G, Granger CWJ, Timmermann A (eds) Handbook of Economic Forecasting, vol 1, Elsevier, Amsterdam, chap 14, pp 715–776

R Core Team (2018) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL https://www.R-project.org/

Richardson S, Green PJ (1997) On Bayesian analysis of mixtures with an unknown number of components (with discussion). Journal of the Royal Statistical Society: Series B (Statistical Methodology) 59:731–792, DOI 10.1111/1467-9868.00095

Sinclair P (ed) (2010) Inflation Expectations. Routledge

Stan Development Team (2018) RStan: the R interface to Stan. URL http://mc-stan.org/, R package version 2.17.3

Terai A (2010) Estimating the distribution of inflation expectations. Economics Bulletin 30:315–329