Measuring Inflation Perceptions and Expectations in the UK: Bayesian Analysis of a Normal Mixture Model for Interval Data with an Indifference Limen[*]

Yasutomo Murasawa
Faculty of Economics, Konan University

November 30, 2016

Abstract

The Bank of England/GfK NOP Inflation Attitudes Survey asks individuals about their inflation perceptions and expectations in eight ordered categories with known boundaries except for an indifference limen. With enough categories for identification, one can fit a mixture distribution to such data, which can be multi-modal. Thus Bayesian analysis of a normal mixture model for interval data with an indifference limen is of interest. This paper applies the No-U-Turn Sampler (NUTS) for Bayesian computation, and estimates the distributions of public inflation perceptions and expectations in the UK during 2001Q1–2015Q4. The estimated means of the distributions are useful for measuring information rigidity.

Keywords Finite mixture, Hamiltonian Monte Carlo, Information rigidity, Interval regression, No-U-turn sampler, Ordered response

JEL classification C11, C25, C46, C82, E31

1

1 Introduction

Inflation (or deflation) and its expectations both play important roles in macro and monetary economics. 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 in different ways, or may be irrational. Whatever the reason is, survey data on inflation perceptions and expectations show substantial heterogeneity or disagreement among individuals that varies over time with the actual inflation; see Mankiw et al. (2003). 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. The problem is that we do not know the boundaries of an indifference limen.

A simple solution is to combine an indifference limen with its neighboring intervals to apply interval regression. 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 2001Q1–2015Q4 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

---

[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.
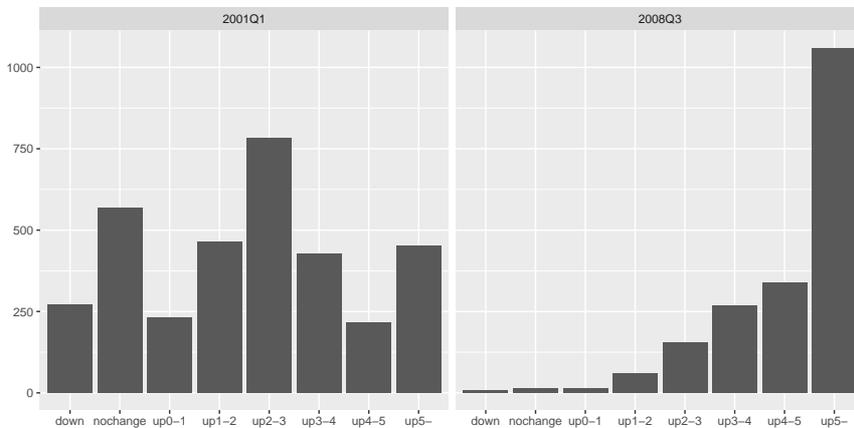
Figure 1: Distributions of inflation perceptions in the UK (2001Q1 and 2008Q3)

inflation perceptions and expectations.

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.[3] 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 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. MCMC methods are useful for posterior simulation. With an indifference limen, our model is similar to an ordered response model, for which the Gibbs sampler and M–H algorithm are often slow to converge for large samples. Instead, this paper applies a Hamiltonian Monte Carlo (HMC) method, in particular the No-U-Turn Sampler (NUTS) developed by Hoffman and Gelman (2014).

Because of the label switching problem, MCMC methods may not work for the component parameters of a finite mixture model. It works for permutation invariant parameters, e.g., moments and quantiles, however; see Geweke (2007). We analyze three types of data, and find the following:

1. With six intervals ignoring the indifference limen, the estimated distributions are often imprecise with wide error bands for higher order moments. This tends to occur when the majority of the respondents choose the last interval with an open end; see the right panel in Figure 1.

---

[3]This may be spurious, though. Without knowing the width of the indifference limen, one cannot draw a histogram.

3

2. With eight intervals including the indifference limen and the prior information about the location of the indifference limen, we obtain narrower error bands. 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 often asymmetric around 0, time-varying, and somewhat differ between inflation perceptions and expectations.

3. Data with 18 intervals are available since 2011Q2, which often have two more local modes at 'Up by 7% but less than 8%' and 'Down by 1% but less than 2%' (or 'Down by 2% but less than 3%'); see Figure 2. If we fit a mixture of two normal distributions to these data, then the estimated distributions and indifference limens seem precise.[4] With only 19 quarters of samples (2011Q2–2015Q4) covering at most one business cycle, though, we cannot learn much about the dynamics of the distributions of inflation perceptions and expectations.

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.[5]

This paper contributes to the literature on quantification of qualitative survey data in three ways. First, we propose a normal mixture model for interval data with an indifference limen. This seems new, and its flexibility seems useful for analysis of survey data. Second, we propose Bayesian analysis of the model using the NUTS. This is useful since for this model, ML estimation may fail, and the Gibbs sampler and M–H algorithm are often slow to converge for large samples. Third, we apply the method to survey data on inflation perceptions and expectations in the UK repeatedly, and estimate their distributions during 2001Q1–2015Q4.

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, but 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.

The paper proceeds as follows. Section 2 reviews some relevant works. Section 3 specifies a normal mixture model for interval data with an indifference

---

[4]Fitting a mixture of four normal distributions gives similar results, though computation takes much longer.

[5]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 (2015) interprets heterogeneous rounding as a measure of uncertainty.
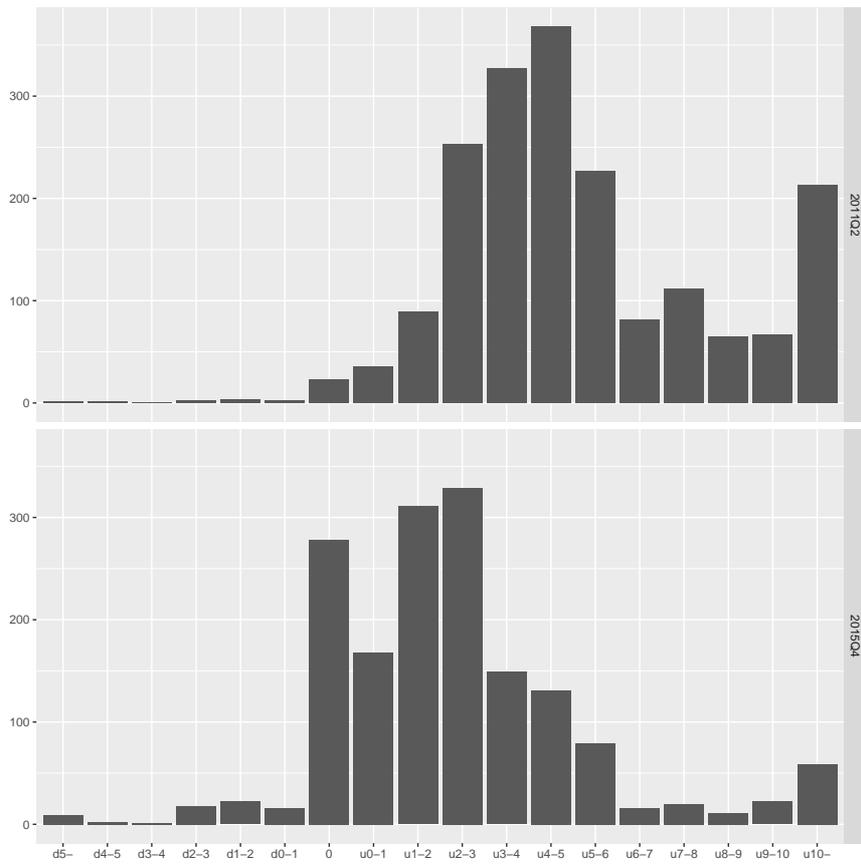
Figure 2: Distributions of inflation perceptions in the UK (2011Q2 and 2015Q4)

limen. Section 4 specifies our prior, and explains our posterior simulation. Section 5 applies the method to aggregate interval data on public inflation perceptions and expectations in the UK during 2001Q1–2015Q4. Section 6 measures information rigidity in public inflation perceptions and expectations in the UK during 2001Q1–2015Q4. Section 7 concludes with discussion.

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 (C–P) method; see Nardo (2003) and Pesaran and Weale (2006, sec. 3) for surveys. To quantify ordered data, which have no quantitative information, the C–P method requires multiple samples and strong assumptions; i.e., normality, a time-invariant symmetric indifference limen, and long-run unbiased expectations. These assumptions often fail in practice, however; 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 aggregate interval 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.[6] 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 $\{1, \ldots, J\}$, which indicate J intervals on $\mathbb{R}$. Assume an ordered response model for y_i such that for $i = 1, \ldots, n$,

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

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 $i = 1, \ldots, n$,

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

Let $\boldsymbol{\pi} := (\pi_1, \ldots, \pi_K)'$, $\boldsymbol{\mu} := (\mu_1, \ldots, \mu_K)'$, $\boldsymbol{\sigma} := (\sigma_1, \ldots, \sigma_K)'$, $\boldsymbol{\gamma} := (\gamma_l, \gamma_u)'$, 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 $\boldsymbol{y}$.

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 $\boldsymbol{\nu}_0 := \boldsymbol{\iota}_K$ 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 $k = 1, \ldots, K$,

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

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

---

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

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 $k \neq l$,

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

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

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

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:

$$\mu_0 := m \qquad (8)$$

$$\sigma_0 := \frac{R}{c} \qquad (9)$$

$$\alpha_0 := 2 \qquad (10)$$

$$A_0 := .2 \qquad (11)$$

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

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

$$\gamma_l | \gamma_u, \boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\sigma} \sim \mathrm{U}(-1, 0) \qquad (13)$$

$$\gamma_u | \gamma_l, \boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\sigma} \sim \mathrm{U}(0, 1) \qquad (14)$$

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., $\forall \boldsymbol{\theta}, \boldsymbol{\theta}'$,

$$p(\boldsymbol{\theta}|\boldsymbol{y})q(\boldsymbol{\theta}, \boldsymbol{\theta}') = p(\boldsymbol{\theta}'|\boldsymbol{y})q(\boldsymbol{\theta}', \boldsymbol{\theta}) \qquad (15)$$

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.

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 $\forall x$,

$$f(x) \propto \exp\left(-\frac{E(x)}{kT}\right) \qquad (16)$$

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 $\mathrm{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$, 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

$$\begin{aligned} H(\boldsymbol{\theta}, \boldsymbol{z}) &:= -\ln p(\boldsymbol{\theta}|\boldsymbol{y}) - \ln p(\boldsymbol{z}) \\ &= -\ln p(\boldsymbol{\theta}, \boldsymbol{z}|\boldsymbol{y}) \end{aligned} \tag{17}$$

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., $\forall t \in \mathbb{R}$,

$$\dot{H}(\boldsymbol{\theta}(t), \boldsymbol{z}(t)) = 0 \tag{18}$$

or

$$\dot{\boldsymbol{\theta}}(t) H_{\boldsymbol{\theta}}(\boldsymbol{\theta}(t), \boldsymbol{z}(t)) + \dot{\boldsymbol{z}}(t) H_{\boldsymbol{z}}(\boldsymbol{\theta}(t), \boldsymbol{z}(t)) = \boldsymbol{0} \tag{19}$$

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

$$\dot{\boldsymbol{\theta}}(t) = H_{\boldsymbol{z}}(\boldsymbol{\theta}(t), \boldsymbol{z}(t)) \tag{20}$$

$$\dot{\boldsymbol{z}}(t) = -H_{\boldsymbol{\theta}}(\boldsymbol{\theta}(t), \boldsymbol{z}(t)) \tag{21}$$

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 $\boldsymbol{z} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$ 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 $\boldsymbol{z} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$ 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 $\min\{\exp(-H(\boldsymbol{\theta}', \boldsymbol{z}'))/\exp(-H(\boldsymbol{\theta}, \boldsymbol{z})), 1\}$.[7]

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. Moerover, 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; e.g., R, Matlab, and Stata. The NUTS often has better convergence properties than other popular MCMC methods such as the Gibbs sampler and the M–H algorithm.

5   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.[8] This is a quarterly survey of public attitudes to inflation that interviews a quota sample of adults aged 16 or over in 175 randomly selected areas 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. Moreover, the individual data 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?

The respondents choose from the following eight intervals (excluding 'No idea'):

1. Gone/Go down

2. Not changed/change

3. Up by 1% or less

---

[7]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.

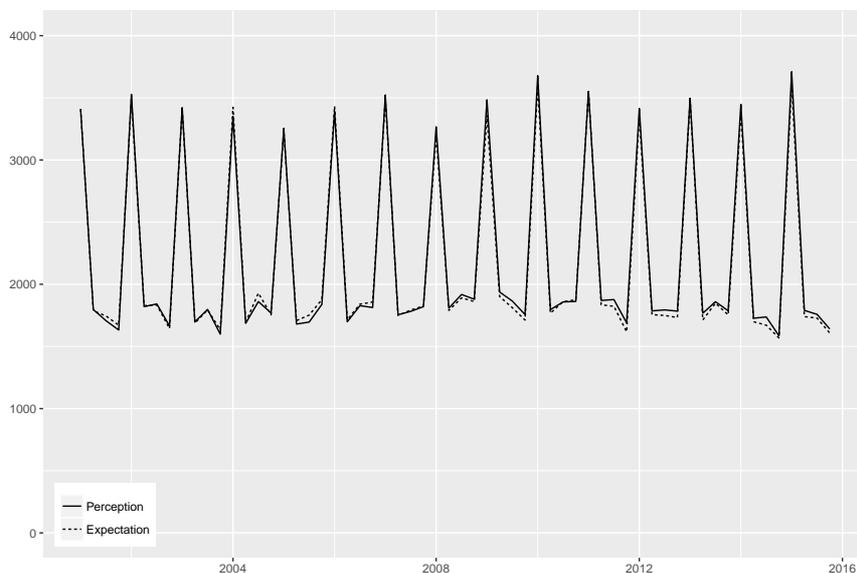[8]There were five trial surveys quarterly from November 1999 to November 2000.

Figure 3: Sample sizes for inflation perceptions and expectations in the UK during 2001Q1–2015Q4 (excluding 'No idea').

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.[9] We exclude 'No idea' from the samples in the following analyses. Figure 3 plots the sample sizes for the two questions during 2001Q1–2015Q4.

Figure 4 plots the relative frequencies of the categories for inflation perceptions and expectations during 2001Q1–2015Q4. 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.

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 to these interval data, and estimate the indifference limen.

---

[9]Since May 2011, the survey asks further questions with more intervals to those who have chosen either category 1 or 8.
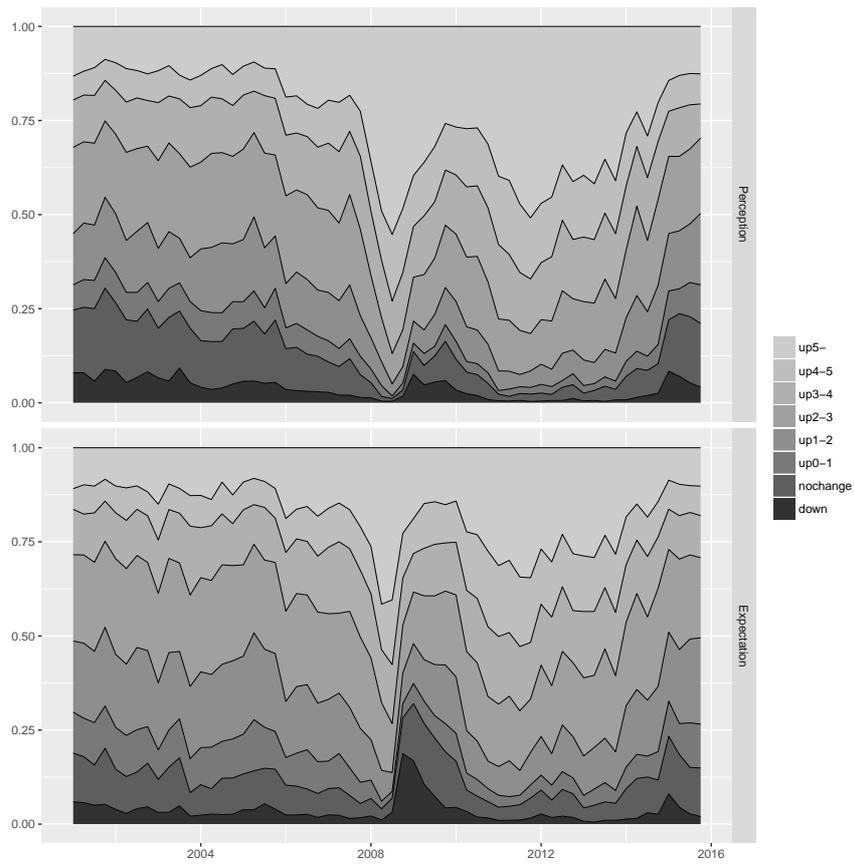
Figure 4: Relative frequencies of the categories for inflation perceptions and expectations in the UK during 2001Q1–2015Q4.

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 $\boldsymbol{\nu}_0 := \boldsymbol{\imath}_2$, $\mu_0 := 2.5$, $\sigma_0 := 10$, $\alpha_0 := 2$, $A_0 := .2$, and $B_0 := .1$. These weakly informative priors do not dominate the posteriors for our samples.

Using the aggregate interval data on inflation perceptions and expectations respectively, we estimate the model for each quarter separately during 2001Q1–2015Q4 (60 quarters). Thus we estimate the model 120 times.

5.3 Bayesian Computation

We apply the NUTS using RStan 2.12.1 developed by Stan Development Team (2016), which runs on R 3.3.2 developed by R Core Team (2016). To improve efficiency of MCMC, we reparametrize our hierarchical model from centered to noncenterd parametrization, following Betancourt and Girolami (2015).[10] Thus instead of equation (5), for $k = 1, \ldots, K$, we draw

$$\sigma^{*2}_k | \boldsymbol{\sigma}_{-k}, \boldsymbol{\mu}, \boldsymbol{\pi}, \boldsymbol{\gamma} \sim \text{Inv-Gam}(\alpha_0, 1) \tag{22}$$

and set $\sigma^2_k := \beta_0 \sigma^{*2}_k$. To avoid divergent transitions during the leapfrog steps and improve sampling efficiency further (at the cost of more leapfrog steps in each iteration), we change the following tuning parameters from their default values:

1. target Metropolis acceptance rate (from .8 to .999)

2. maximum tree depth (from 10 to 12)[11]

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 $\mu := \text{E}(y_i^*)$

2. standard deviation $\sigma := \sqrt{\text{var}(y_i^*)}$

3. skewness $\text{E}\left([(y_i^* - \mu)/\sigma]^3\right)$

4. excess kurtosis $\text{E}\left([(y_i^* - \mu)/\sigma]^4\right) - 3$

5. asymmetry in the tails $\text{E}\left([(y_i^* - \mu)/\sigma]^5\right)$

---

[10]Centered and noncenterd parametrizations 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).

[11]This increases the maximum number of leapfrog steps from $2^{11} - 1 = 2043$ to $2^{13} - 1 = 8191$.

Given equation (2),

$$\mu = \sum_{k=1}^{K} \pi_k \mu_k \tag{23}$$

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

$$\mathrm{E}\left((y_i^* - \mu)^2\right) = \sum_{k=1}^{K} \pi_k \left[(\mu_k - \mu)^2 + \sigma_k^2\right] \tag{24}$$

$$\mathrm{E}\left((y_i^* - \mu)^3\right) = \sum_{k=1}^{K} \pi_k \left[(\mu_k - \mu)^3 + 3(\mu_k - \mu)\sigma_k^2\right] \tag{25}$$

$$\mathrm{E}\left((y_i^* - \mu)^4\right) = \sum_{k=1}^{K} \pi_k \left[(\mu_k - \mu)^4 + 6(\mu_k - \mu)^2\sigma_k^2 + 3\sigma_k^4\right] \tag{26}$$

$$\mathrm{E}\left((y_i^* - \mu)^5\right) = \sum_{k=1}^{K} \pi_k \left[(\mu_k - \mu)^5 + 10(\mu_k - \mu)^3\sigma_k^2 + 15(\mu_k - \mu)\sigma_k^4\right] \tag{27}$$

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.

5.4 Estimation Excluding the Indifference Limen

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

Figure 5 plots the posterior medians of the five moments of the distributions of inflation perceptions and expectations, respectively, during 2001Q1–2015Q4. We plot 68% error bands instead of 95% error bands, since the latter are often too wide to be informative for higher-order moments.

The estimated means of inflation perceptions and expectations seem reasonable and precise with narrow error bands. The means of inflation expectations are more stable than those of perceptions, staying around 2–4%, though not perfectly constant (anchored); hence 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 sometimes imprecise. This is perhaps because our interval data have little information about the dispersion of the underlying variable, especially when the majority of the respondents choose the last interval with an open end ("Up by 5% or more"). We still see that the standard deviations of inflation perceptions are often slightly larger than those of expectations; hence individuals may have private information when they perceive inflation, but expect its fluctuations to be temporary.
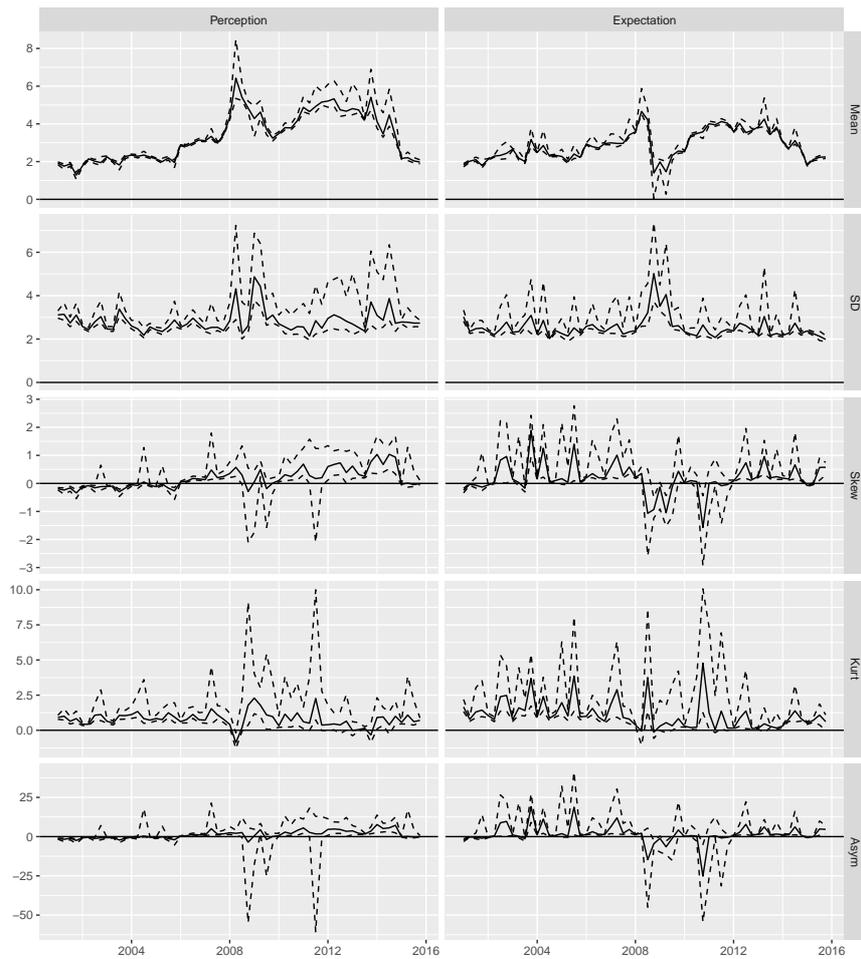
Figure 5: Posterior medians of moments of the distributions of inflation perceptions (left) and expectations (right) in the UK during 2001Q1–2015Q4. The dashed lines are 68% error bands.

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

The estimated kurtoses of inflation perceptions and expectations are often (but not always) imprecise. This is because our interval data 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.

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 often move in the same direction.

5.5 Estimation Including the 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 and use prior information about the indifference limen ($-1 < \gamma_l < 0 < \gamma_u < 1$). Figure 6 plots the posterior medians of the five moments and the indifference limens for inflation perceptions and expectations, respectively, during 2001Q1–2015Q4, together with 68% error bands.

The estimated means of inflation perceptions and expectations are similar to the previous results with no indifference limen. With more information in estimation, however, the error bands are narrower.

The estimated standard deviations of inflation perceptions and expectations seem more stable than the previous results. The standard deviations are still time-varying, and slightly larger for perceptions than for expectations.

The estimated skewnesses of inflation perceptions and expectations differ from the previous results, perhaps because with two more intervals, the data have more information about the skewness. The skewnesses are time-varying, and mostly positive except around 2010.

The estimated kurtoses of inflation perceptions and expectations are also smaller and more precise than the previous results. We often have excess kurtosis, but it disappears around 2010 for both perceptions and expectations.

The estimated asymmetries in the tails of inflation perceptions and expectations also differ from the previous results. They are still often imprecise, but we find clearer positive asymmetries except around 2010.

The estimated indifference limens for inflation perceptions and expectations are both time-varying, and mostly asymmetric around 0. The error bands are wider for lower bounds than for upper bounds, since category 1 ("Gone/Go down") often has fewer observations than category 3 ("Up by 1% or less"). The upper bounds are around .4–.6% for perceptions and .3–.5% for expectations. The lower bounds are often below $-.7\%$, and sometimes close to -1.0%.

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 enough intervals,
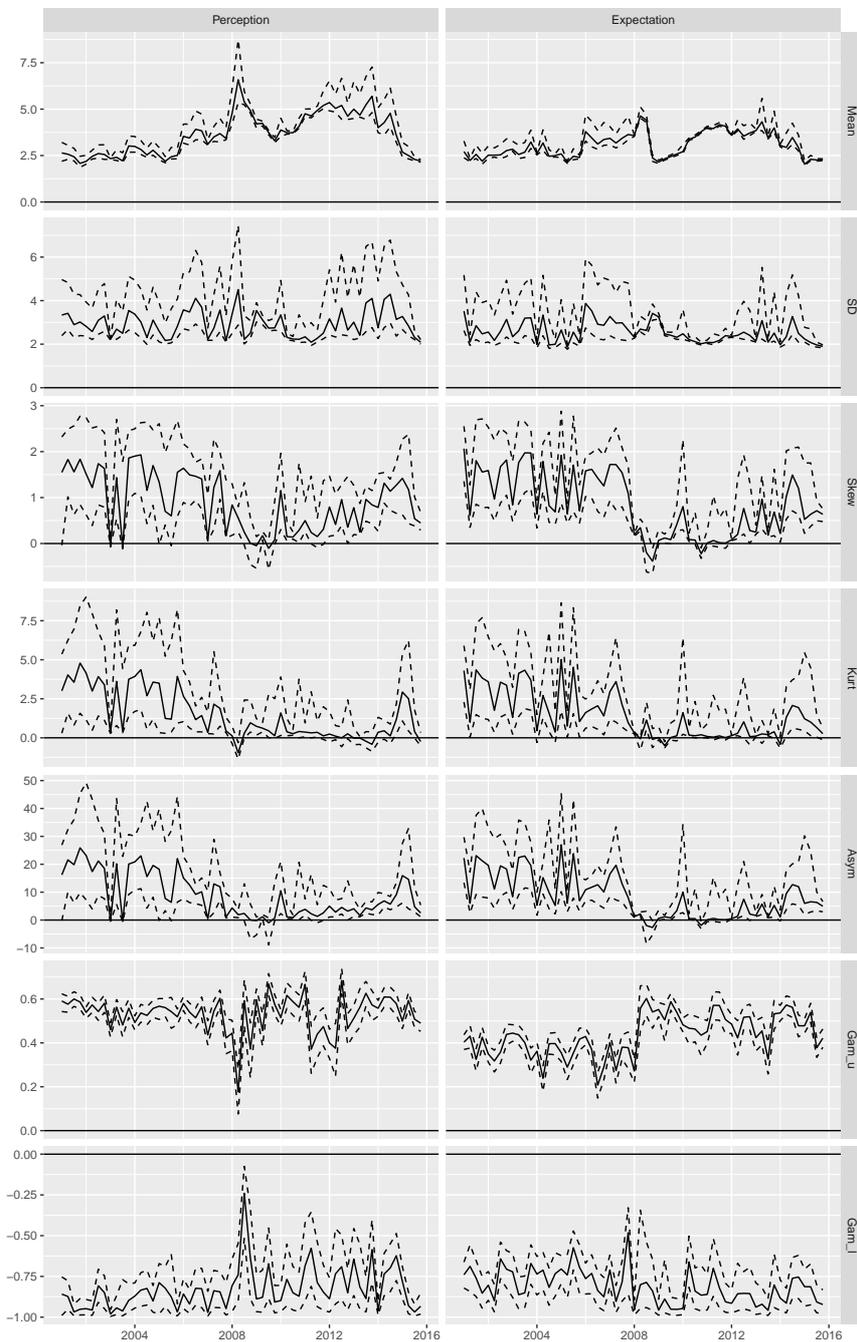
Figure 6: Posterior medians of moments and indifference limens of the distributions of inflation perceptions (left) and expectations (right) in the UK during 2001Q1–2015Q4. The dashed lines are 68% error bands.

the NUTS works well without adjusting the tuning parameters; thus we use the default values for all tuning parameters. Moreover, the NUTS runs much faster with centered parametrization than with noncentered parametrization; thus we use the former.

Figure 7 plots the posterior medians of the moments and indifference limens for inflation perceptions and expectations, respectively, during 2011Q2–2015Q4. We plot 95% error bands here, instead of 68% error bands, since the former are narrow enough to be informative even for higher-order moments.

The means are mostly higher for inflation perceptions than for expectations, 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 standard deviations are close to each other for inflation perceptions and expectations. It appears that the mean and standard deviation often 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 skewnesses are positive for both inflation perceptions and expectations, and higher for expectations than for perceptions. The mean and skewness tend to move in the opposite directions.

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 asymmetry 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. Thus higher-order moments tend to move in the same directions. This may suggest existence of outliers.

The estimated indifference limens for inflation perceptions and expectations are both time-varying, and mostly asymmetric around 0. The upper bounds are around .4–.6% for perceptions and .3–.5% for expectations. The lower bounds are often below $-.7\%$, and sometimes close to -1.0%.

In addition to moments, quantiles help us to see the whole shape of a distribution. Figure 8 plots the posterior medians of the deciles of inflation perceptions and expectations during 2011Q2–2015Q4.[12] These plots seem useful for monitoring the distributions of inflation perceptions and expectations.

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

---

[12]The `nor1mix` package for R is useful for calculating quantiles of a mixture normal distribution.
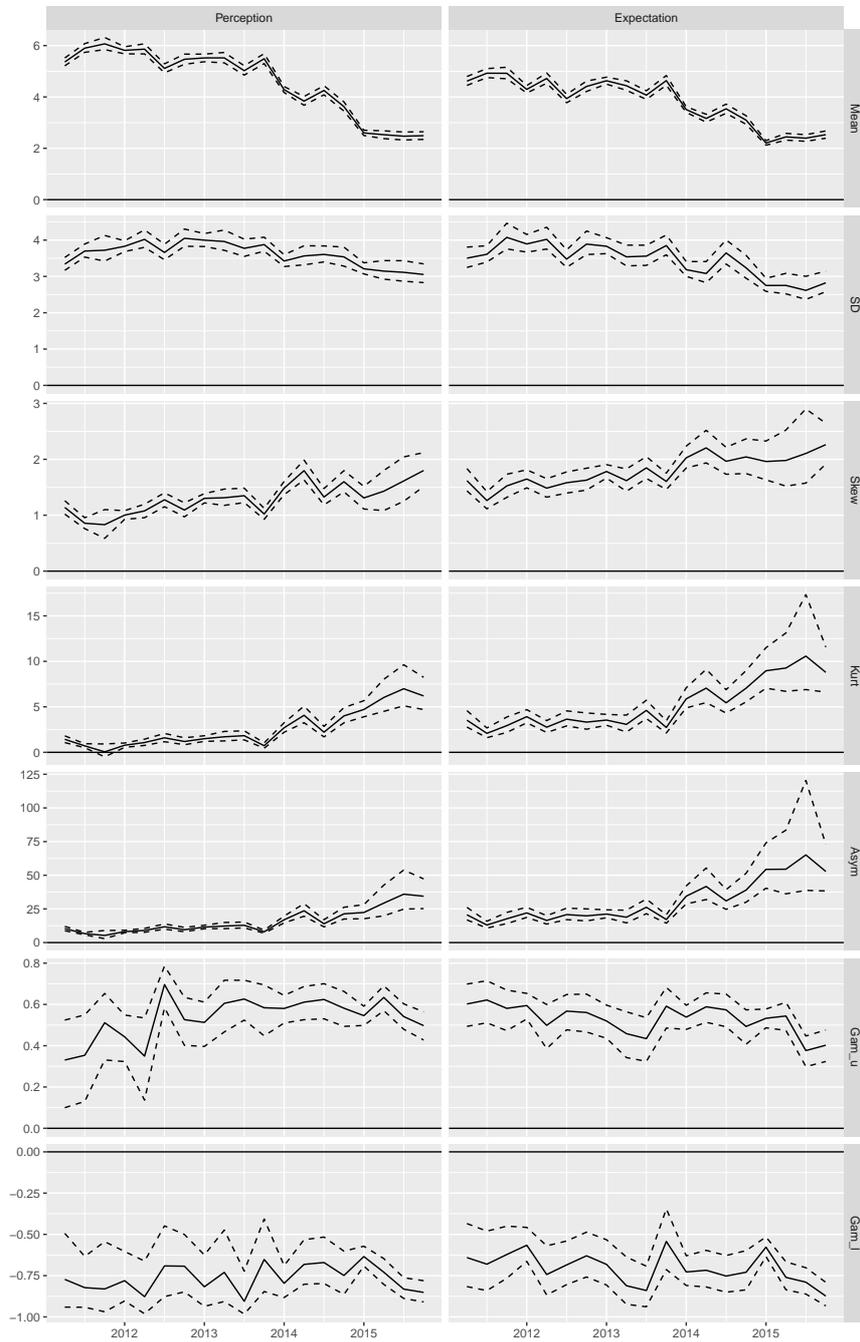
Figure 7: Posterior medians of moments and indifference limens of the distributions of inflation perceptions (left) and expectations (right) in the UK during 2011Q2–2015Q4 using data with 18 intervals. The dashed lines are 95% error bands.
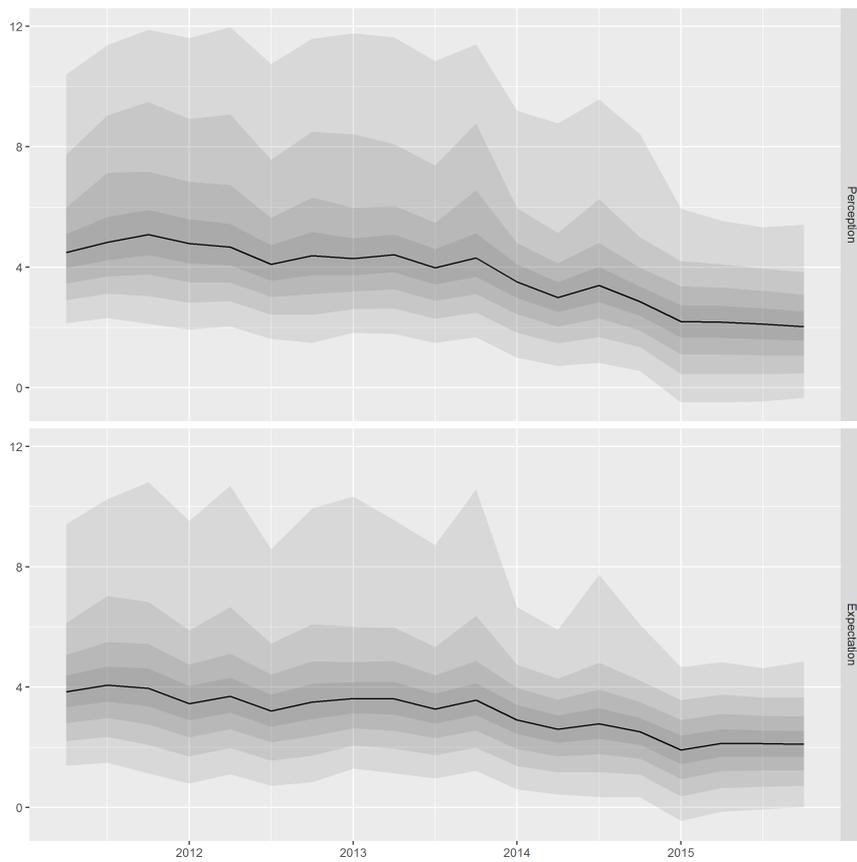
Figure 8: Posterior medians of the deciles of the distributions of inflation perceptions and expectations in the UK during 2011Q2–2015Q4.

$\bar{y}^e_{t+h|t}$ be the average forecast of y_{t+h} at date t among individuals. Then for all t,

$$
\begin{aligned}
\bar{y}^e_{t+h|t} &:= (1-\lambda)\,\mathrm{E}_t(y_{t+h}) + \lambda(1-\lambda)\,\mathrm{E}_{t-1}(y_{t+h}) + \lambda^2(1-\lambda)\,\mathrm{E}_{t-2}(y_{t+h}) + \cdots \\
&= (1-\lambda)\,\mathrm{E}_t(y_{t+h}) + \lambda[(1-\lambda)\,\mathrm{E}_{t-1}(y_{t+h}) + \lambda(1-\lambda)\,\mathrm{E}_{t-2}(y_{t+h}) + \cdots] \\
&= (1-\lambda)\,\mathrm{E}_t(y_{t+h}) + \lambda\bar{y}^e_{t+h|t-1} \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad (28)
\end{aligned}
$$

or

$$
\begin{aligned}
\mathrm{E}_t(y_{t+h}) &= \frac{1}{1-\lambda}\bar{y}^e_{t+h|t} - \frac{\lambda}{1-\lambda}\bar{y}^e_{t+h|t-1} \\
&= \bar{y}^e_{t+h|t} + \frac{\lambda}{1-\lambda}\left(\bar{y}^e_{t+h|t} - \bar{y}^e_{t+h|t-1}\right) \quad\quad\quad\quad (29)
\end{aligned}
$$

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,

$$
\mathrm{E}_t\left(y_{t+h} - \bar{y}^e_{t+h|t}\right) = \beta\left(\bar{y}^e_{t+h|t} - \bar{y}^e_{t+h|t-1}\right) \quad\quad\quad (30)
$$

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

$$
y_{t+h} - \bar{y}^e_{t+h|t} = \beta\left(\bar{y}^e_{t+h|t} - \bar{y}^e_{t+h|t-1}\right) + u_{t+h|t} \quad\quad\quad (31)
$$

where $u_{t+h|t} := y_{t+h} - \mathrm{E}_t(y_{t+h})$.[13] One can estimate β by OLS, from which one can recover $\lambda = \beta/(1+\beta) \in [0,1)$.

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,

$$
y_{t+h} - \bar{y}^e_{t+h|t} = \beta\left(\bar{y}^e_{t+h|t} - \bar{y}^e_{t+h-1|t-1}\right) + v_{t+h|t} \quad\quad\quad (32)
$$

where

$$
v_{t+h|t} := u_{t+h|t} - \beta\left(\bar{y}^e_{t+h|t-1} - \bar{y}^e_{t+h-1|t-1}\right) \quad\quad\quad (33)
$$

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 from the website of the Office for National Statistics of the UK. The CPI and RPI inflation rates are somewhat different; see Figure 9.

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 2001Q1–2015Q4, assuming a normal distribution ($K=1$) or

---

[13]A noisy information model gives the same regression model if $\{y_t\}$ is AR(1); see Coibion and Gorodnichenko (2015, sec. I.B).
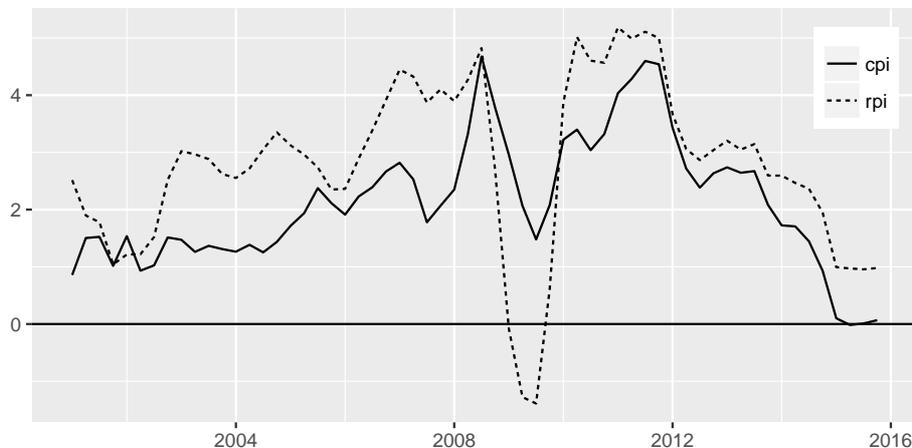
Figure 9: Annual CPI and RPI inflation rates in the UK during 2001Q1–2015Q4

a mixture of two normal distributions ($K = 2$) with or without an indifference limen.[14] The resulting $\left\{ \bar{\pi}^e_{t+h|t} \right\}$ are somewhat different.

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 (32). We estimate the linear model by the IV method (2SLS). For 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.[15] We use gretl 2016d for these analyses.

Table 1 summarizes the estimation results. We find the following:

1. The intercepts are negative for both the CPI and RPI and for both inflation perceptions and expectations, and significant except for expectations relative to the RPI, suggesting forecast biases. The biases are larger for the CPI and perceptions than for the RPI and expectations.

2. The slopes are positive for both the CPI and RPI and for both inflation perceptions and expectations, but insignificant for perceptions and expectations relative to the CPI. For the RPI, the magnitude and significance of the slope depends on how we estimate the means of the distributions of inflation perceptions and expectations, i.e., the choice of the number of mixture components and whether we estimate the indifference limen or not.

3. If we believe that individuals forecast the RPI, assume $K = 2$, and include the indifference limen when estimating the means of the distributions of inflation perceptions and expectations, then the slope is marginally significant for perceptions and marginally insignificant for expectations, perhaps

---

[14]For $K = 1$, we use the flat prior on (μ, σ) and the default tuning parameters.

[15]The full sample periods are 1987Q4–2015Q4 for the change in the log oil price and 1971Q2– 2015Q4 for the change in the log exchange rate.

Table 1: 2SLS regressions of the ex post forecast errors on the ex ante mean forecast revisions

	indiff.	CPI		RPI	
K	limen	intercept	slope	intercept	slope
		inflation perceptions ($h = 0$)			
1	no	−1.10	1.18	−.45	2.33
		(.12)	(.71)	(.24)	(.96)
1	yes	−1.12	1.30	−.48	2.55
		(.12)	(.79)	(.24)	(1.05)
2	no	−1.21	.35	−.57	1.08
		(.17)	(.51)	(.29)	(.66)
2	yes	−1.48	.50	−.83	1.30
		(.16)	(.61)	(.27)	(.67)
		inflation expectations ($h = 4$)			
1	no	−.71	.85	−.04	1.68
		(.24)	(.85)	(.35)	(.82)
1	yes	−.73	.85	−.06	1.70
		(.24)	(.86)	(.35)	(.83)
2	no	−.72	.27	−.05	.92
		(.26)	(.75)	(.35)	(.76)
2	yes	−.95	.60	−.27	1.49
		(.25)	(.89)	(.34)	(.91)

Notes: Numbers in parentheses are Newey–West HAC standard errors. The estimation periods are 2001Q2–2015Q4 for inflation perceptions and 2001Q2–2014Q4 for inflation expectations.

because of the short sample length (59 for perceptions and 55 for expectations). The implied λ is .57 for $\beta = 1.30$ and .60 for $\beta = 1.49$.[16]

To summarize, we find some evidences of information rigidity among UK individuals forecasting inflation, though the results depend on what prices we assume they forecast and how we estimate the means of the distributions of their inflation perceptions and expectations.

7   Discussion

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. 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.

---

[16]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 may be higher in the UK than in the US, though the two results may not be directly comparable.

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 quater 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.

Micro data from the Bank of England/GfK NOP Inflation Attitudes Survey are now publicly available at the website of the Bank of England. This suggests the following extensions of this paper:

1. 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).

2. 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.

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

References

J. H. Albert and S. Chib. Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association*, 88:669–679, 1993.

C. L. Alston and K. L. Mengersen. Allowing for the effect of data binning in a Bayesian normal mixture model. *Computational Statistics & Data Analysis*, 54:916–923, 2010.

O. Armantier, W. Bruine de Bruin, S. Potter, G. Topa, W. van der Klaauw, and B. Zafar. Measuring inflation expectations. *Annual Review of Economics*, 5: 273–301, 2013.

M. Betancourt and M. Girolami. Hamiltonian Monte Carlo for hierarchical models. In S. K. Upadhyay, U. Singh, D. K. Dey, and A. Loganathan, editors, *Current Trends in Bayesian Methodology with Applications*, chapter 4, pages 79–102. CRC Press, 2015.

C. Biernacki. 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, 2007.

C. C. Binder. Measuring uncertainty based on rounding: New method and application to inflation expectations. unpublished, Aug. 2015.

D. G. Blanchflower and C. MacCoille. The formation of inflation expectations: An empirical analysis for the UK. Working Paper 15388, National Bureau of Economic Research, September 2009.

M.-H. Chen and D. K. Dey. Bayesian analysis for correlated ordinal data models. In D. K. Dey, S. K. Ghosh, and B. K. Mallick, editors, *Generalized Linear Models: A Bayesian Perspective*, chapter 8, pages 133–157. Marcel Dekker, 2000.

S. Chib and E. Greenberg. Understanding the Metropolis–Hastings algorithm. *American Statistician*, 4:327–335, 1995.

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

M. K. Cowles. Accelerating Monte Carlo Markov chain convergence for cumulative-link generalized linear models. *Statistics and Computing*, 6:101–111, 1996.

S. Frühwirth-Schnatter. *Finite Mixture and Markov Switching Models*. Springer, 2006.

A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. *Bayesian Data Analysis*. CRC Press, 3rd edition, 2014.

J. Geweke. Interpretation and inference in mixture models: Simple MCMC works. *Computational Statistics & Data Analysis*, 51:3529–3550, 2007.

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

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

J. Kiefer and J. Wolfowitz. Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters. *Annals of Mathematical Statistics*, 27:887–906, 1956.

K. Lahiri and Y. Zhao. Quantifying survey expectations: A critical review and generalization of the Carlson–Parkin method. *International Journal of Forecasting*, 31:51–62, 2015.

J. S. Liu and C. Sabatti. Generalized Gibbs sampler and multigrid Monte Carlo for Bayesian computation. *Biometrika*, 87:353–369, 2000.

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

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

N. G. Mankiw, R. Reis, and J. Wolfers. Disagreement about inflation expectations. *NBER Macroeconomics Annual*, 18:209–248, 2003.

C. F. Manski. Measuring expectations. *Econometrica*, 72:1329–1376, 2004.

C. F. Manski and F. Molinari. Rounding probabilistic expectations in surveys. *Journal of Business & Economic Statistics*, 28:219–231, 2010.

Y. Murasawa. Measuring inflation expectations using interval-coded data. *Oxford Bulletin of Economics and Statistics*, 75:602–623, 2013.

B. Nandram and M.-H. Chen. Reparameterizing the generalized linear model to accelerate Gibbs sampler convergence. *Journal of Statistical Computation and Simulation*, 54:129–144, 1996.

M. Nardo. The quantification of qualitative survey data: A critical assessment. *Journal of Economic Surveys*, 17:645–668, 2003.

R. M. Neal. MCMC using Hamiltonian dynamics. In S. Brooks, A. Gelman, G. L. Jones, and X.-L. Meng, editors, *Handbook of Marcov Chain Monte Carlo*, Handbooks of Modern Statistical Methods, chapter 5, pages 113–162. Chapman & Hall/CRC, 2011.

O. Papaspiliopoulos, G. O. Roberts, and M. Sköld. Non-centered parameterisations for hierarchical models and data augmentation. In J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. Dawid, D. Heckerman, A. F. M. Smith, and M. West, editors, *Bayesian Statistics 7*, pages 307–326. Oxford University Press, 2003.

O. Papaspiliopoulos, G. O. Roberts, and M. Sköld. A general framework for the parametrization of hierarchical models. *Statistical Science*, 22:59–73, 2007.

M. H. Pesaran and M. Weale. Survey expectations. In G. Elliot, C. W. J. Granger, and A. Timmermann, editors, *Handbook of Economic Forecasting*, volume 1, chapter 14, pages 715–776. Elsevier, Amsterdam, 2006.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2016. URL https://www.R-project.org.

S. Richardson and P. J. Green. 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, 1997.

P. Sinclair, editor. *Inflation Expectations*. Routledge, 2010.

Stan Development Team. *Stan Modeling Language Users Guide and Reference Manual, Version 2.12.0*, 2016.

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