USC Dornsife
Dana and David Dornsife College of Letters, Arts and Sciences

Institute for New Economic Thinking

# USC Dornsife Institute for New Economic Thinking

Working Paper No. 18-20

# Forecasting Panel Data with Structural Breaks and Regime-specific Grouped Heterogeneity

Simon C. Smith

October 15, 2018

Electronic copy available at: https://ssrn.com/abstract=3268759

# Forecasting Panel Data with Structural Breaks and Regime-specific Grouped Heterogeneity

Simon C. Smith

*USC Dornsife INET, Department of Economics, USC*

*Draft: October 15, 2018*

**Abstract**

Generating accurate forecasts in the presence of structural breaks requires careful management of bias-variance tradeoffs. Existing methods for forecasting time series under breaks reduce parameter estimation error by pooling estimates across pre- and post-break data necessarily inducing bias. Forecasting panel data under breaks offers the possibility to reduce parameter estimation error without inducing any bias if there exists a regime-specific pattern of grouped heterogeneity. We develop a new Bayesian methodology to estimate panel regression models in the presence of breaks and regime-specific grouped heterogeneity. Parameters are pooled within, but differ across, regime-specific groups. We develop a formal test of regime-specific patterns of grouped heterogeneity that integrates over all parameters and penalises model complexity. In an empirical application to forecasting inflation rates across 20 U.S. industries we find evidence of regime-specific grouped heterogeneity. Exploiting this information produces significantly improved forecasts relative to a range of popular methods.

**Keywords:** Forecasting, Panel data, Structural breaks, Grouped Heterogeneity, Bayesian analysis

**JEL classifications:** G10, C11, C15

*Email address:* `simonsmi@usc.edu`. `Mailing address: Institute for New Economic Thinking, Department of Economics, USC, 3620 South Vermont Ave., CA, 90089-0253, USA.` (Simon C. Smith)

The suggestions of Torben Andersen, Fabio Canova, Frank Diebold, Rob Engle, Wayne Ferson, Jim Hamilton, Dennis Kristensen, Roger Moon, Geert Ridder, and Andres Santos, and Allan Timmermann have been helpful. Any remaining errors are my own.

Electronic copy available at: https://ssrn.com/abstract=3268759

## 1. Introduction

*"...in recent years the [Phillips curve] relationship has become much weaker. Economists in many rich countries, including America and Germany have noticed that low unemployment is not generating the high wage-growth that might be expected. Britain's ultra-low rate of joblessness, coupled with particularly weedy wages, makes it perhaps the biggest Phillips-buster of all."* (The Economist, 2018.)[1]

Economic time series forecasts are plagued by structural breaks (Stock and Watson 1996; Pettenuzzo and Timmermann 2011; Rossi 2013).[2] Generating accurate forecasts in the presence of breaks requires careful management of bias-variance tradeoffs. Forecasts are typically generated from the parameters estimated using only post-break data which ensures unbiased estimates but may result in considerable parameter estimation error particularly in short regimes. Pooling parameter estimates across pre- and post-break data reduces parameter estimation error at the expense of inducing bias. Optimal management of bias-variance tradeoffs is a challenging problem that explains much of the poor forecasting performance in the presence of breaks (Pesaran and Timmermann 2007).

There is an emerging interest in forecasting panel data.[3] Panel data present the opportunity to manage bias-variance tradeoffs through the cross-section. If series in the cross-section have patterns of grouped heterogeneity in which series within a group share a common parameter, pooling parameter estimates within groups reduces estimation error without inducing any bias. Exploiting this commonality within groups requires identifying the number of groups and which series belong to each. The group structure may be further complicated if it varies through time.

Existing panel break methods assume that in each regime parameters are either unit-specific or pooled across the entire cross-section.[4] Panel methods that allow for

[1]This quote is from an article entitled 'All work and no pay' that appeared in The Economist on September 6, 2018.

[2]A handful of frequentist time series approaches to detect breaks include Andrews (1993), Bai and Perron (1998), Bai and Perron (2003), and Elliott and Müller (2006) while Bayesian approaches include Chib (1998), Wang and Zivot (2000), Pesaran et al. (2006), Koop and Potter (2007), Geweke and Jiang (2011), and Peluso et al. (2016).

[3]A small handful of recent studies on panel forecasting include Liu (2017), Liu et al. (2017), Smith and Timmermann (2017), and Bollerslev et al. (2018).

[4]A subset of studies on breaks in panel models or multivariate time series include Bai et al. (1998), Qu and Perron (2007), Bai and Carrion-I-Silvestre (2009), Bai (2010), Kim (2011), Oka and Qu (2011), Groen et al. (2013), Baltagi et al. (2016), and Qian and Su (2016).



Electronic copy available at: https://ssrn.com/abstract=3268759

grouped heterogeneity ignore breaks and assume constant group dynamics.[5] Procedures that allow for grouped heterogeneity but have parameters that shift every period will suffer from imprecise parameter estimates and will likely generate inaccurate forecasts (Bonhomme and Manresa 2015).

This paper develops a new Bayesian panel regression approach that estimates the number and timing of breaks as well as the number of groups and cross-sectional ordering of series in each regime. Parameters are pooled within, but differ across, regime-specific groups. Allowing the number of groups and cross-sectional ordering of series to vary across regimes enables improved management of bias-variance tradeoffs relative to existing methods. This approach allows, in a given regime, any number of groups ranging from one (fully pooled) to $N$ (unit-specific). One might wish to increase the amount of pooling by reducing the number of groups during short and volatile regimes that are susceptible to estimation error such as around the global financial crisis, while allowing more groups and less pooling during longer and less volatile regimes.

Figure 1 illustrates the importance of allowing for a regime-specific grouping structure. For the petroleum and machinery industries, the top window displays the total number of industries within their respective groups from 2004-2018. The estimates are from a panel regression of U.S. industry-level inflation rates on lagged industry-level inflation and unemployment rates in a model that allows for regime-specific grouped heterogeneity.[6] Machinery (red dotted line) has its own group until the onset of the financial crisis in February 2008. Thereafter its parameter is pooled across six industries until the crisis regime ends in October 2009 when it returns to its own group. At the same time, the method identifies that petroleum is an industry unlike any other and therefore is assigned its own group throughout the sample period.[7]

Panel break models can obtain more precise estimates by pooling parameters across the entire cross-section but this will induce bias into the estimates if the coefficients are not identical across all series in the cross-section. The bottom left window graphs the evolution over the sample of the slope coefficient on the lagged unemployment rate, $\beta$, estimated from our model for the machinery (red dotted line) and

[5] A small handful of papers on cross-sectional grouping include Canova (2004), Hahn and Moon (2010), Hamilton and Owyang (2012), Lin and Ng (2012), Bonhomme and Manresa (2015), Ando and Bai (2016), Ke et al. (2016), Su et al. (2016), Bonhomme et al. (2017), Lu and Su (2017), and Vogt and Linton (2017).

[6] The data are described in more detail in Section 6.

[7] The petroleum industry has the most unusual inflation series with high volatility and frequent and large outliers.



Electronic copy available at: https://ssrn.com/abstract=3268759

petroleum (black solid line) industries. The path of $\beta$ estimated from the model with breaks and pooled parameters is also graphed (green dashed line). Our model is flexible enough to simultaneously increase the magnitude of $\beta$ for the petroleum industry, implying more predictability, during the crisis regime and decrease the magnitude of $\beta$ for the machinery group, implying less predictability. The model with breaks and pooled parameters, however, treats all series equally increasing the magnitude of the slope for all series during the crisis. This model is therefore likely to generate particularly inaccurate inflation forecasts for the series in the machinery group during the crisis.

The bottom right window graphs the evolution of the standard deviation of $\beta$ for the machinery industry estimated by our model (red dotted line) and the model with breaks and unit-specific parameters (purple dotted line). During the crisis regime, pooling across six industries the slope coefficient of the group in which machinery falls results in a more precise estimate with a standard deviation peaking at 0.8. The model with breaks and unit-specific parameters has a much larger standard deviation of 1.9 during this period. The imprecision of the parameter estimate will likely translate into inaccurate forecasts.

The ability to identify which industries should be pooled together across time by identifying the regime-specific grouping structure would have been valuable economic information for policymakers in a central bank generating inflation forecasts particularly during the global financial crisis.

Our framework nests a number of popular models such as the common break models with parameters that, in every regime, are either unit-specific or pooled across the entire cross-section (Smith and Timmermann 2017) as well as fully pooled or unit-specific models without breaks. With a sufficiently diffuse prior we may find a break every period. As pointed out by Koop and Potter (2007) our framework thus approximates the time-varying parameter (TVP) approach (Primiceri 2005). We therefore nest a unit-specific TVP, a fully pooled TVP and the grouped TVP model of Bonhomme and Manresa (2015).

While allowing for regime-specific grouped heterogeneity may improve model fit, it also increases model complexity relative to the common breaks models with fully pooled or unit-specific parameters and the constant grouped heterogeneity model. It is therefore important to test formally for evidence of regime-specific grouped heterogeneity while accounting for overparameterisation. To the best of our knowledge, this is the first paper to estimate and test formally for regime-specific grouped heterogeneity. The test integrates over all parameters and inherently penalises model



Electronic copy available at: https://ssrn.com/abstract=3268759

complexity. Using the approach of Chib (1995) we compute the marginal likelihoods for the model with regime-specific grouped heterogeneity and a range of benchmark models. The corresponding Bayes Factors provide the strength of evidence in favour of regime-specific grouped heterogeneity.

The model is estimated in five steps. First, the Gibbs step samples parameters from their full conditional distributions. Second, a random-walk Metropolis-Hastings step estimates the break locations. Third, within each regime the cross-sectional ordering of series is estimated. Fourth, in each regime a reversible jump step (Carlin and Chib 1995; Green 1995) estimates the number of groups by attempting with equal probability to either merge two existing groups into one or split an existing group into two. Fifth, a reversible jump step estimates the number of breaks by attempting with equal probability either to introduce a break at a randomly chosen location or remove a randomly selected existing break.

To demonstrate the ability of the method to estimate models with regime-specific grouped heterogeneity we conduct a simulation study in which the true data generating process is known. When the true data generating process is characterised by regime-specific grouped heterogeneity panel models with either breaks and pooled parameters or constant grouped heterogeneity produce biased parameter estimates. A panel model with breaks and unit-specific parameters delivers parameter estimates that are unbiased but imprecise. The parameters estimated from the new method we develop are unbiased and more precise thereby dominating all the competing models from a bias-variance standpoint.

An empirical application that explores the ability of 20 lagged U.S. industry-level unemployment rates to forecast corresponding industry-level inflation rates over the period 2009-2018 demonstrates the practical use of the new method. Two breaks are detected at February 2008 and October 2009 along with large fluctuations in grouped heterogeneity across the three regimes.

For at least 16 out of 21 cases (20 industries and the U.S. aggregate) our method produces statistically significantly more accurate inflation forecasts than (i) an AR(1) model applied independently to each industry; (ii) a panel model with constant grouped heterogeneity; (iii) a panel model with breaks and pooled parameters; and (iv) a panel model with breaks and unit-specific parameters. Our method is only significantly outperformed by the benchmarks for four out of 84 cases (21 cases against the four benchmark models). Cumulative sum of squared forecast error difference plots show that the more accurate forecasts are consistent across the sample and not driven by one or two short periods.



Electronic copy available at: https://ssrn.com/abstract=3268759

The remainder of the paper is set out as follows. Section 2 sets out the model and prior specifications. Section 3 explains model estimation. Section 4 presents the formal test of regime-specific grouped heterogeneity. Section 5 presents the simulation study. Section 6 presents an empirical application to forecasting U.S. industry-level inflation. Section 7 presents extensions and robustness checks. Section 8 concludes.

## 2. Methodology

Our panel regression framework with $i = 1, \dots, N$ series and $t = 1, \dots, T$ time periods estimates the number of structural breaks $K$ and their locations $\tau = (\tau_1, \dots, \tau_K)$ that separate $K+1$ regimes.[8] Breaks are common such that every series is hit at the same date.[9] The common break assumption seems reasonable in our empirical application that compares the accuracy of U.S. industry-level inflation forecasts generated from our model relative to a number of benchmark models. Carlino and DeFina (1998), for example, report that monetary shocks are aggregate shocks, which correspond to structural breaks in our framework, that have a common yet heterogeneous effect. The latter is allowed for by our grouping structure.

The $k$th regime contains the observations $(\tau_{k-1}+1, \dots \tau_k)$ and thus its duration is $l_k = \tau_k - \tau_{k-1}$. The $k$th regime is further characterised by an estimated number of regime-specific cross-sectional groups $g_k = 1, \dots, G_k + 1$ whereby $N_{g_k}$ denotes the number of series in the $g_k$th group. Let $G = (G_1, \dots, G_{K+1})$. The cross-sectional ordering of the series, that is, which series belong to which group, is also estimated. Parameters are pooled within, but differ across, regime-specific groups.

The framework nests several popular approaches. The fully pooled and unit-specific constant panel models are obtained if we estimate no breaks and one or $N$ groups. Detecting breaks with one or $N$ groups in every regime obtains the respective pooled and unit-specific panel models with common breaks. Detecting a break each period and either one or $N$ groups in each regime obtains the pooled and unit-specific TVP models. A break every period with an estimated number of groups obtains the model of Bonhomme and Manresa (2015).

[8] As is conventional in the literature we assume $\tau_0 = 0$ and $\tau_{K+1} = T$ for convenience.
[9] For a framework that incorporates noncommon breaks - see Smith (2018).



Electronic copy available at: https://ssrn.com/abstract=3268759

### 2.1. Modeling cross-sectional dependencies

Cross-sectional dependence plagues macroeconomic and financial data sets. To avoid compromising inference these dependencies must be accounted for (Andrews 2005). Not removing such dependencies will lead to the majority of the increased break detection power obtained from the cross-section being lost (Kim 2011; Baltagi et al. 2016; Smith and Timmermann 2017). We allow an observed common factor, the U.S. aggregate inflation rate, to drive the dependencies (Pesaran 2006; Bai 2009). Factor loadings are regime- and group-specific.

We assume that including the common factor removes much of the cross-sectional dependence such that any remaining dependence is weak, not strong, which Pesaran (2015) notes is unlikely to cause any serious inferential problems for large panels ($N > 10$) as we consider in our empirical application. To evaluate the success of including common factors to remove strong cross-sectional dependence, in Section 6 we carry out the formal cross-sectional dependence (CD) test of Pesaran (2004).

In any event, the primary aim of this article is to develop a new method that generates more accurate forecasts than a range of competitor models. We are not performing inference on the parameter estimates. The ultimate test of the method will be whether it is able to outperform a range of benchmark models, including the consistently estimated linear time series model, in an out-of-sample forecasting application to U.S. industry-level inflation rates.

### 2.2. Model with pooled breaks and parameters

Denote the $(N \times T)$ matrix containing the observations on the dependent variable $\mathbf{y}$ and the $(N \times T \times \kappa)$ matrix containing observations on the $\kappa$ regressors $\mathbf{X}$.[10] For regimes $k = 1, \ldots, K+1$ and groups $g_k = 1, \ldots, G_k + 1$ the model is

$$y_{it} = X_{it}'\beta_{g_k} + \epsilon_{it}, \qquad i \in N_{g_k}, \qquad t = \tau_{k-1}+1, \ldots, \tau_k \tag{1}$$

with Gaussian errors $\epsilon_{it} \sim N(0, \sigma^2_{g_k})$. For the $g_k$th group, let $\beta_{g_k}$ denote a $(\kappa \times 1)$ vector of slope coefficients on the $\kappa$ regressors and $\sigma^2_{g_k}$ denote the error-term variance. Further let $\boldsymbol{\beta_k} = (\beta_{1_k}, \ldots, \beta_{G_k+1})$, $\sigma^2_k = (\sigma^2_{1_k}, \ldots, \sigma^2_{G_k+1})$, $\boldsymbol{\beta} = (\boldsymbol{\beta_1}, \ldots, \boldsymbol{\beta_{K+1}})$, and

[10] An intercept can be incorporated by setting the first of the $\kappa$ regressors to be a unit vector such that the first element of $\beta_{g_k}$ estimates the intercept in the $g_k$th group. The model can also incorporate $F$ observed factors with factor loadings specific to the $g_k$th group. We suppress the observed common factor for expositional ease.



Electronic copy available at: https://ssrn.com/abstract=3268759

$\boldsymbol{\sigma}^2 = (\sigma_1^2, \ldots, \sigma_{K+1}^2)$. The likelihood for this model is

$$p(\boldsymbol{y} \mid \boldsymbol{X}, \boldsymbol{\beta}, \boldsymbol{\sigma^2}, \tau, G) = \left( \prod_{k=1}^{K+1} \prod_{g_k=1}^{G_k+1} (2\pi\sigma_{g_k}^2)^{-\frac{l_k N_{g_k}}{2}} \right) \times exp \left[ \sum_{k=1}^{K+1} \sum_{g_k=1}^{G_k+1} \sum_{t=\tau_{k-1}+1}^{\tau_k} \sum_{i \in N_{g_k}} \frac{(y_{it} - X_{it}' \beta_{g_k})^2}{-2\sigma_{g_k}^2} \right]. \tag{2}$$

### 2.3. Prior distributions

A Bayesian framework combines prior information supplied by the user with information in the data transmitted through the likelihood. We now specify the prior distributions.

#### 2.3.1. Prior on regime durations

Following Koop and Potter (2007) and Smith and Timmermann (2017) regime durations have a Poisson distribution

$$p(l_k \mid \lambda_k) = \frac{\lambda_k^{l_k}}{l_k!} exp^{-\lambda_k}, \qquad k = 1, \ldots, K+1 \tag{3}$$

where $\lambda_k$, the expected duration of the $k$th regime, has a conjugate Gamma prior

$$p(\lambda_k) = \frac{d^c}{\Gamma(c)} \lambda_k^{c-1} exp^{-\lambda_k d}, \qquad k = 1 \ldots, K+1. \tag{4}$$

We define the prior over the structural breaks as

$$p(\tau) = \prod_{k=1}^{K+1} p(l_k \mid \lambda_k) p(\lambda_k) \tag{5}$$

and marginalise $\lambda$ to obtain

$$p(\tau) = \prod_{k=1}^{K+1} \frac{1}{l_k!} \frac{\Gamma(c + l_k)}{(d+1)^{c+l_k}} \frac{d^c}{\Gamma(c)}. \tag{6}$$



Electronic copy available at: https://ssrn.com/abstract=3268759

*2.3.2. Prior on cross-sectional groupings*

A Poisson prior distribution is placed over the number of series in a group

$$p(N_{g_k} \mid \epsilon_k) = \frac{\epsilon_k^{N_{g_k}}}{N_{gk}!} exp^{-\epsilon_k}, \qquad g_k = 1, \ldots, G_k + 1, \qquad k = 1, \ldots, K + 1 \tag{7}$$

where the expected number of series in every group in the $k$th regime, $\epsilon_k$, has a conjugate Gamma prior

$$p(\epsilon_k) = \frac{f^e}{\Gamma(e)} \epsilon_k^{e-1} exp^{-\epsilon_k f}, \qquad k = 1 \ldots, K + 1. \tag{8}$$

The prior on the group structure is

$$p(G) = \prod_{k=1}^{K+1} \prod_{g_k=1}^{G_k+1} p(N_{g_k} \mid \epsilon_k) p(\epsilon_k). \tag{9}$$

We multiply and divide by $\frac{(f+G_k+1)^{e+N}}{\Gamma(e+N)}$ and marginalise $\epsilon$ to obtain

$$p(G) = \left( \prod_{k=1}^{K+1} \frac{f^e}{\Gamma(e)} \frac{\Gamma(e+N)}{(f + G_k + 1)^{e+N}} \left( \prod_{g_k=1}^{G_k+1} \frac{1}{N_{g_k}!} \right) \right).$$

*2.3.3. Priors on regression and variance parameters*

Conjugacy helps maintain computational efficiency which is critical when estimating panel models with large time series and/or cross-section dimensions. For regimes $k = 1, \ldots, K + 1$, we specify an inverse gamma prior distribution over the error-term variances

$$p(\sigma^2_{g_k}) = \frac{b^a}{\Gamma(a)} \sigma^{2^{-(a+1)}}_{g_k} exp \left( -\frac{b}{\sigma^2_{g_k}} \right), \qquad g_k = 1, \ldots, G_k + 1 \tag{10}$$

and a conjugate Gaussian prior distribution over the slope, $\beta_{g_k}$, conditional on the variance, $\sigma^2_{g_k}$

$$\begin{aligned} p(\beta_{g_k} \mid \sigma^2_{g_k}) &= 2\pi^{\frac{-\kappa}{2}} (\sigma^2_{g_k})^{\frac{-\kappa}{2}} \mid V_\beta \mid^{\frac{-1}{2}} exp \left( -\frac{1}{2\sigma^2_{g_k}} \beta'_{g_k} V_\beta^{-1} \beta_{g_k} \right), \qquad g_k = 1, \ldots, G_k + 1 \\ V_\beta &= \sigma^2_\beta I_\kappa \end{aligned} \tag{11}$$

where $a$ and $b$ characterise the prior expected error-term variance and $\sigma_b^2$ characterises the prior variance of $\beta_{g_k}$.



Electronic copy available at: https://ssrn.com/abstract=3268759

### 2.4. Posterior distribution

Inference is performed on the posterior distribution which combines prior information with information in the data. Following Smith and Timmermann (2017), specifying conjugate priors on $\boldsymbol{\beta}$ and $\boldsymbol{\sigma^2}$ allows them to be marginalised from the posterior thereby reducing the computational burden[11]

$$p(\boldsymbol{y} \mid \boldsymbol{X}, \tau, G) = 2\pi^{\frac{-TN}{2}} \prod_{k=1}^{K+1} \prod_{g_k=1}^{G_k+1} \frac{\Gamma(\tilde{a}_{g_k})}{\Gamma(a)} \frac{b^a}{\tilde{b}_{g_k}^{\tilde{a}_{g_k}}} \frac{\mid \Sigma_{g_k} \mid^{1/2}}{\mid V_\beta \mid^{1/2}} \tag{12}$$

in which, for regimes $k = 1, \ldots, K+1$ and groups $g_k = 1, \ldots, G_k + 1$

$$\begin{aligned}
\Sigma_{g_k}^{-1} &= V_\beta^{-1} + \sum_{i \in N_{g_k}} \sum_{t=\tau_{k-1}+1}^{\tau_k} X_{it} X_{it}' \\
\mu_{g_k} &= \Sigma_{g_k} \times \sum_{i \in N_{g_k}} \sum_{t=\tau_{k-1}+1}^{\tau_k} X_{it} y_{it} \\
\tilde{a}_{g_k} &= a + \frac{l_k N_{g_k}}{2} \\
\tilde{b}_{g_k} &= \frac{1}{2} \left( 2b - \mu_{g_k}' \Sigma_{g_k}^{-1} \mu_{g_k} + \sum_{i \in N_{g_k}} \sum_{t=\tau_{k-1}+1}^{\tau_k} y_{it}^2 \right).
\end{aligned} \tag{13}$$

## 3. Estimating the Model

The model is estimated using Markov chain Monte Carlo methods. There are five estimation steps. First, the parameters are drawn from their full conditional distributions. Second, the timing of the structural breaks are estimated using a random-walk Metropolis Hastings step. Third, the cross-sectional ordering of series in each regime is estimated using a block updating scheme. Fourth, the number of groups in each regime is estimated using the reversible jump Markov chain Monte Carlo algorithm (Carlin and Chib 1995; Green 1995). Specifically, in a given regime the algorithm attempts with equal probability to either split an existing group into two or merge two existing groups into one. Fifth, the number of structural breaks is also estimated using a reversible jump step. With equal probability the algorithm attempts either to introduce a structural break at a random date or remove a randomly selected existing

[11] See Appendix A for details on how $\boldsymbol{\beta}$ and $\boldsymbol{\sigma^2}$ are marginalised.



Electronic copy available at: https://ssrn.com/abstract=3268759

break.

We now briefly discuss the fact that dynamic panels which include lagged dependent variables, as the majority of inflation forecasting models do, may compromise inference. There are a number of available approaches to combat the problems that arise from dynamic panels: models with pooled coefficients and either interactive fixed-effects (Bai 2009) or Gaussian quasi maximum likelihood estimation (Moon and Weidner 2015, 2017) and models that allow for common correlated effects with lagged dependent variables (Chudik and Pesaran 2015).

We do not pursue these modeling strategies in the analysis that follows because first imposing parameters that are pooled across the entire cross-section would restrict the optimal management of bias-variance tradeoffs that can vary through time which is the major benefit of our methodology. Second, the methodology of Chudik and Pesaran (2015) may require incorporating into the model a large number of lags of cross-sectional averages. This would lead to a highly parameterised model increasing the possibility of parameter estimation error and is likely to result in less accurate forecasts.

Furthermore, the aim of our proposed methodology is to sharpen forecasts. We are not performing inference on the parameter estimates and the ultimate test of our methodology will be its out-of-sample forecasting performance.

### *3.1. Estimating the parameters*

We compute the hyperparameters using equation (13) and sample the parameters from their full conditional distributions for regimes $k = 1, \ldots, K+1$

$$
\begin{aligned}
p(\epsilon_k \mid \cdots) &\sim Gamma(e + N, f + G_k + 1), \\
p(\lambda_k \mid \cdots) &\sim Gamma(c + l_k, d + 1), \\
p(\sigma^2_{g_k} \mid \cdots) &\sim IG\left(\tilde{a}_{g_k}, \tilde{b}_{g_k}\right), \qquad g_k = 1, \ldots, G_k + 1 \\
p(\beta_{g_k} \mid \cdots) &\sim N\left(\mu_{g_k}, \Sigma_{g_k}\sigma^2_{g_k}\right), \qquad g_k = 1, \ldots, G_k + 1.
\end{aligned}
\tag{14}
$$

### *3.2. Estimating the timing of breaks*

To estimate the timing of breaks we perturb the $k$th break date $\tau_k$ for $k = 1, \ldots, K$ to obtain a proposed break date $\tau_{k^*}$. Specifically, $\tau_{k^*} = \tau_k + u$ in which $u$ is sampled



Electronic copy available at: https://ssrn.com/abstract=3268759

from the discrete uniform interval $[-s, s]$.[12] If $\tau_{k-1} \geq \tau_{k^*} \geq \tau_{k+1}$ the move is rejected. Otherwise we compute the durations of regimes $k$ and $k+1$ as $l_{k^*} = l_k + u$ and $l_{k^*+1} = l_{k+1} - u$ and, for groups $g_k = 1, \ldots, G_k + 1$, we compute $\Sigma_{g_k}^{*^{-1}}$, $\mu_{g_k}^*$, $\tilde{a}_{g_k}^*$, and $\tilde{b}_{g_k}^*$ from equation (13). We further compute, for groups $g_{k+1} = 1, \ldots, G_{k+1}+1$, $\Sigma_{g_{k+1}}^{*^{-1}}$, $\mu_{g_{k+1}}^*$, $\tilde{a}_{g_{k+1}}^*$, and $\tilde{b}_{g_{k+1}}^*$. The proposal is accepted with probability $min(1, \alpha)$

$$\alpha = \left( \prod_{g_k=1}^{G_k+1} \frac{\tilde{b}_{g_k}^{\tilde{a}_{g_k}}}{\tilde{b}_{g_k}^{*\tilde{a}_{g_k}^*}} \frac{\Gamma(\tilde{a}_{g_k}^*)}{\Gamma(\tilde{a}_{g_k})} \frac{|\Sigma_{g_k}^*|^{1/2}}{|\Sigma_{g_k}|^{1/2}} \right) \left( \prod_{g_{k+1}=1}^{G_{k+1}+1} \frac{\tilde{b}_{g_{k+1}}^{\tilde{a}_{g_{k+1}}}}{\tilde{b}_{g_{k+1}}^{*\tilde{a}_{g_{k+1}}^*}} \frac{\Gamma(\tilde{a}_{g_{k+1}}^*)}{\Gamma(\tilde{a}_{g_{k+1}})} \frac{|\Sigma_{g_{k+1}}^*|^{1/2}}{|\Sigma_{g_{k+1}}|^{1/2}} \right) \times \frac{l_k!}{l_{k^*}!} \frac{\Gamma(c + l_{k^*})}{\Gamma(c + l_k)} \frac{l_{k+1}!}{l_{k^*+1}!} \frac{\Gamma(c + l_{k^*+1})}{\Gamma(c + l_{k+1})}. \tag{15}$$

If the proposal is accepted we substitute all proposals for their original values. Otherwise the proposals are discarded and the algorithm continues to the next step.

### *3.3. Estimating the cross-sectional ordering of series*

This step estimates which series are allocated to each regime-specific group. As $N$ grows large, updating each series individually will be computationally prohibitive while updating all series at once will likely result in proposals that are difficult to accept and hinder the mixing of the chain. We therefore use a block updating procedure.

We pre-specify an upper bound on the number of series to update in a block, $B = 5$. On each sweep of the MCMC run, for regimes $k = 1, \ldots, K+1$, the number of series to simultaneously update, $N_{blk}$, is sampled with equal probability from one to the pre-specified upper bound. We randomly select $N_{blk}$ of the $N$ series and each of the $N_{blk}$ series are randomly assigned to one of the $G_k + 1$ groups. If all series are allocated to their current groups the move is rejected. Otherwise, we compute $\Sigma_{g_k}^{*^{-1}}$, $\mu_{g_k}^*$, $\tilde{a}_{g_k}^*$, $\tilde{b}_{g_k}^*$, and $N_{g_k}^*$ for groups $g_k = 1, \ldots, G_k + 1$. We reject the move if any group now has no series allocated to it. Otherwise the proposal is accepted with probability $min(1, \alpha)$

$$\alpha = \left( \prod_{g_k=1}^{G_k+1} \frac{\Gamma(\tilde{a}_{g_k}^*)}{\Gamma(\tilde{a}_{g_k})} \frac{\tilde{b}_{g_k}^{\tilde{a}_{g_k}}}{\tilde{b}_{g_k}^{*\tilde{a}_{g_k}^*}} \frac{|\Sigma_{g_k}^*|^{1/2}}{|\Sigma_{g_k}|^{1/2}} \frac{N_{g_k}!}{N_{g_k}^*!} \right). \tag{16}$$

[12] We select $s$ to achieve the desired acceptance ratio. If $u = 0$ the move is rejected.



Electronic copy available at: https://ssrn.com/abstract=3268759

If the move is accepted we substitute all proposed values for their originals.

### *3.4. Estimating the number of regime-specific groups*

The number of groups may vary across regimes. It is therefore necessary to endogenously estimate the number of groups in each regime. For regimes $k = 1, \ldots, K+1$, we attempt with equal probability either to split an existing regime-specific group into two new ones (split move) or merge two existing regime-specific groups into one (merge move).[13]

#### *3.4.1. Split move*

A split move attempts to split an existing group into two smaller groups thereby increasing $G_k$ to $G_k^* = G_k + 1$. For each regime, one of the $G_k + 1$ existing groups, $g_k$, is randomly selected. With equal probability we assign a value of zero or one to each series in the $g_k$th group. Those series assigned a value of one will make up the $g_k^*$th group while the remaining series create the $G_k^* + 1$th group. The total number of series assigned to groups $g_k^*$ and $G_k^* + 1$ are denoted $N_{g_k^*}$ and $N_{G_k^*+1}$ in which $N_{g_k} = N_{g_k^*} + N_{G_k^*+1}$. The move is rejected if either group is assigned no series. Otherwise we construct $G^*$.

Using equation (13) we compute $\Sigma_{g_k^*}^{-1}$, $\mu_{g_k^*}$, $\tilde{a}_{g_k^*}$, $\tilde{b}_{g_k^*}$, $\Sigma_{G_k^*+1}^{-1}$, $\mu_{G_k^*+1}$, $\tilde{a}_{G_k^*+1}$, and $\tilde{b}_{G_k^*+1}$. The move is accepted with probability $min(1, \alpha)$

$$\alpha = \frac{\Gamma(\tilde{a}_{g_k^*})}{\Gamma(\tilde{a}_{g_k})} \frac{\Gamma(\tilde{a}_{G_k^*+1})}{\Gamma(a)} \frac{\tilde{b}_{g_k}^{\tilde{a}_{g_k}}}{\tilde{b}_{G_k^*+1}^{\tilde{a}_{G_k^*+1}}} \frac{b^a}{\tilde{b}_{g_k^*}^{\tilde{a}_{g_k^*}}} \frac{|\Sigma_{g_k^*}|^{\frac{1}{2}}}{|\Sigma_{g_k}|^{\frac{1}{2}}} \frac{|\Sigma_{G_k^*+1}|^{\frac{1}{2}}}{|V_\beta|^{\frac{1}{2}}}$$
$$\times \frac{(f + G_k + 1)^{e+N}}{(f + G_k + 2)^{e+N}} \frac{N_{g_k}!}{N_{g_k^*}!\, N_{G_k^*+1}!} \frac{2^{N_{g_k}}}{(G_k + 2)} \cancel{\frac{(G_k + 1)}{(G_k + 1)}}. \tag{17}$$

The acceptance probability must ensure that detailed balance is maintained across the entire parameter space to avoid 'jumping' to a part of the parameter space from which the algorithm cannot 'revert'. It is therefore necessary to include $(G_k + 2)^{-1}$ and $(G_k + 1)^{-1}$ in the acceptance probability which correspond to randomly selecting two of the $G_k + 2$ groups to merge if we were hypothetically attempting a merge move (detailed in the next subsection) to return to our current position from the proposed

[13] If we have one or $N$ groups in the $k$th regime any proposed merge or split moves, respectively, are rejected.



Electronic copy available at: https://ssrn.com/abstract=3268759

position. The split move proposal terms consist of $G_k + 1$ which corresponds to randomly selecting a group to split and $2^{N_{g_k}}$ which corresponds to assigning with equal probability a value of one or zero to each of the $N_{g_k}$ series to determine which of the two smaller groups they will be assigned to. If the move is accepted all proposals are substituted for their original values.

*3.4.2. Merge move*

A merge move attempts to combine two groups $g_k$ and $g'_k$ that are randomly selected from the existing $G_k+1$ groups in a given regime $k$ into one group $g^*_k$ thereby reducing $G_k + 1$ to $G^*_k = G_k$. Using equation (13) we compute $\Sigma^{-1}_{g^*_k}$, $\mu_{g^*_k}$, $\tilde{a}_{g^*_k}$, and $\tilde{b}_{g^*_k}$, while $N_{g^*_k} = N_{g_k} + N_{g'_k}$. We construct $G^*$. The move is accepted with probability $min(1, \alpha)$

$$\alpha = \frac{\Gamma(\tilde{a}_{g^*_k})}{\Gamma(\tilde{a}_{g_k})} \frac{\Gamma(a)}{\Gamma(\tilde{a}_{g'_k})} \frac{\tilde{b}_{g'_k}^{\tilde{a}_{g'_k}}}{b^a} \frac{\tilde{b}_{g_k}^{\tilde{a}_{g_k}}}{\tilde{b}_{g^*_k}^{\tilde{a}_{g^*_k}}} \frac{\mid \Sigma_{g^*_k} \mid^{1/2}}{\mid \Sigma_{g_k} \mid^{1/2}} \frac{\mid V_\beta \mid^{1/2}}{\mid \Sigma_{g'_k} \mid^{1/2}} \frac{(f + G_k + 1)^{e+N}}{(f + G_k)^{e+N}} \frac{N_{g_k}!\, N_{g'_k}!}{N_{g^*_k}!} \frac{(G_k + 1)}{2^{N_{g^*_k}}} \frac{\cancel{G_k}}{\cancel{G_k}}. \tag{18}$$

The terms $G_k$ and $G_{k+1}$ in the acceptance probability correspond to randomly selecting two of the current $G_k + 1$ groups to merge into one. To ensure reversibility we must include density terms from a hypothetical split move that attempts to return to the current position from the proposed one. Therefore $G_k^{-1}$ corresponds to randomly selecting one of the $G_k$ groups to split and $2^{-N_{g^*_k}}$ corresponds to assigning with equal probability a value of one or zero to each of the $N_{g^*_k}$ series to determine which of the two new groups they will be assigned to. If the move is accepted all proposed values are substituted for their originals.

*3.5. Estimating the number of structural breaks*

The final step estimates the number of structural breaks. With equal probability ($b_k = d_k = 0.5$) the MCMC algorithm enters a birth or death move.[14]

[14]Unless we have a break every period in which case $b_{T-1} = 0$ or we have no breaks such that $d_0 = 0$.



Electronic copy available at: https://ssrn.com/abstract=3268759

### 3.5.1. Birth move

A birth move attempts to increase $K$ to $K^* = K + 1$ by introducing a new break $k^*$ at a location $\tau_{k^*}$ that is chosen at random from the time series sample period. If an existing breakpoint is sampled $\tau_{k^*} \in \tau$ the move is rejected. This move proposes to split an existing regime $k$ into two shorter regimes $k^*$ and $k^* + 1$ separated by $\tau_{k^*}$. The $k^*$th regime contains observations $\tau_{k^*-1} + 1, \ldots, \tau_{k^*}$ and the $k^* + 1$th regime contains $\tau_{k^*} + 1, \ldots, \tau_{k^*+1}$. The durations of regimes $k^*$ and $k^* + 1$ are computed as $l_{k^*} = \tau_{k^*} - \tau_{k^*-1}$ and $l_{k^*+1} = \tau_{k^*+1} - \tau_{k^*}$. We construct $l^*$ and $\tau^*$.

We sample the group structure in regimes $k^*$ and $k^* + 1$. For both regimes we sample the number of groups $G_{k^*} + 1$ and $G_{k^*+1} + 1$ from the discrete uniform interval $U[1, N]$. For all $N$ series, we sample with equal probability a value between 1 and $G_{k^*} + 1$ corresponding to the group allocated to each series in regime $k^*$. We do the same for regime $k^* + 1$ and construct $G^*$.

We compute $\Sigma_{g_{k^*}}^{-1}$, $\mu_{g_{k^*}}$, $\tilde{a}_{g_{k^*}}$, $\tilde{b}_{g_{k^*}}$, and $N_{g_{k^*}}$ for groups $g_{k^*} = 1, \ldots, G_{k^*} + 1$ and $\Sigma_{g_{k^*+1}}^{-1}$, $\mu_{g_{k^*+1}}$, $\tilde{a}_{g_{k^*+1}}$, $\tilde{b}_{g_{k^*+1}}$, and $N_{g_{k^*+1}}$ for groups $g_{k^*+1} = 1, \ldots, G_{k^*+1} + 1$. The move is rejected if any group has no series allocated to it. Otherwise we construct ${\Sigma^*}^{-1}$, $\mu^*$, $\tilde{a}^*$, and $\tilde{b}^*$. The proposed move is accepted with probability $min(1, \alpha)$

$$
\begin{aligned}
\alpha = & \left( \prod_{g_{k^*}=1}^{G_{k^*}+1} \frac{\Gamma(\tilde{a}_{g_{k^*}})}{\Gamma(a)} \frac{b^a}{\tilde{b}_{g_{k^*}}^{\tilde{a}_{g_{k^*}}}} \frac{\mid \Sigma_{g_{k^*}} \mid^{1/2}}{\mid V_\beta \mid^{1/2}} \right) \left( \prod_{g_{k^*+1}=1}^{G_{k^*+1}+1} \frac{\Gamma(\tilde{a}_{g_{k^*+1}})}{\Gamma(a)} \frac{b^a}{\tilde{b}_{g_{k^*+1}}^{\tilde{a}_{g_{k^*+1}}}} \frac{\mid \Sigma_{g_{k^*+1}} \mid^{1/2}}{\mid V_\beta \mid^{1/2}} \right) \\
& \times \left( \prod_{g_k=1}^{G_k+1} \frac{\Gamma(a)}{\Gamma(\tilde{a}_{g_k})} \frac{\tilde{b}_{g_k}^{\tilde{a}_{g_k}}}{b^a} \frac{\mid V_\beta \mid^{1/2}}{\mid \Sigma_{g_k} \mid^{1/2}} \right) \frac{f^e}{\Gamma(e)} \frac{\Gamma(e+N)}{(f + G_{k^*} + 1)^{e+N}} \frac{(f + G_k + 1)^{e+N}}{(f + G_{k^*+1} + 1)^{e+N}} \\
& \times \left( \prod_{g_k=1}^{G_k+1} N_{g_k}! \right) \left( \prod_{g_{k^*}=1}^{G_{k^*}+1} \frac{1}{N_{g_{k^*}}!} \right) \left( \prod_{g_{k^*+1}=1}^{G_{k^*+1}+1} \frac{1}{N_{g_{k^*+1}}!} \right) \frac{l_k!}{l_{k^*}!\, l_{k^*+1}!} \frac{d^c}{\Gamma(c)} \\
& \times \frac{\Gamma(c + l_{k^*})\Gamma(c + l_{k^*+1})}{\Gamma(c + l_k)(d+1)^c} \frac{TN^{\cancel{2}}}{\cancel{N}(K+1)} \frac{(G_{k^*}+1)^N (G_{k^*+1}+1)^N}{(G_k+1)^N}
\end{aligned}
\tag{19}
$$

in which the term $T$ corresponds to the uniform sampling of $\tau_{k^*}$ from the discrete time series interval, $N^2$ corresponds to the uniform sampling of the number of groups in regimes $k^*$ and $k^* + 1$ and $(G_{k^*} + 1)^N$ and $(G_{k^*+1} + 1)^N$ correspond to the uniform sampling of which groups each of the $N$ series are allocated to in regimes $k^*$ and $k^* + 1$.

To maintain detailed balance we include in the acceptance probability density terms for a hypothetical death move (detailed in the following subsection) that would return the algorithm to the current position from the proposed one. Therefore $(K +$



Electronic copy available at: https://ssrn.com/abstract=3268759

$1)^{-1}$ corresponds to sampling with equal probability one of the $K+1$ breaks to remove, the $N^{-1}$ term corresponds to the uniform sampling of the number of groups $G_k+1$ from the discrete interval $[1, N]$, and the $(G_k+1)^{-N}$ term corresponds to the uniform sampling of which group each of the $N$ series is assigned to. If the move is accepted we substitute all proposed values for their original ones.

### *3.5.2. Death move*

A death move attempts to reduce $K$ to $K^* = K-1$ by removing the $k$th break $\tau_k$ which is sampled with equal probability from the existing $K$ breaks. The move proposes to merge regimes $k$ and $k+1$ into one new regime $k^*$ which has duration $l_{k^*} = l_k + l_{k+1}$ and contains the observations $\tau_{k-1}+1, \ldots, \tau_{k+1}$. We compute $l^*$ and $\tau^*$.

The number of groups in the new regime $G_{k^*}+1$ is sampled with equal probability on the discrete interval from one to $N$. For each of the $N$ series we sample with equal probability which of the $G_{k^*}+1$ groups they belong to. We construct $G^*$. For groups $g_{k^*} = 1, \ldots, G_{k^*}+1$, we compute $N_{g_{k^*}}$, $\Sigma^{-1}_{g_{k^*}}$, $\mu_{g_{k^*}}$, $\tilde{a}_{g_{k^*}}$, and $\tilde{b}_{g_{k^*}}$. The move is rejected if any group has no series allocated to it. Otherwise we construct ${\Sigma^*}^{-1}$, $\mu^*$, $\tilde{a}^*$, and $\tilde{b}^*$. The proposed move is accepted with probability $min(1, \alpha)$

$$
\begin{aligned}
\alpha = &\left(\prod_{g_k=1}^{G_k+1} \frac{\Gamma(a)}{\Gamma(\tilde{a}_{g_k})} \frac{\tilde{b}_{g_k}^{\tilde{a}_{g_k}}}{b^a} \frac{|\, V_\beta\,|^{1/2}}{|\,\Sigma_{g_k}\,|^{1/2}}\right) \left(\prod_{g_{k+1}=1}^{G_{k+1}+1} \frac{\Gamma(a)}{\Gamma(\tilde{a}_{g_{k+1}})} \frac{\tilde{b}_{g_{k+1}}^{\tilde{a}_{g_{k+1}}}}{b^a} \frac{|\, V_\beta\,|^{1/2}}{|\,\Sigma_{g_{k+1}}\,|^{1/2}}\right) \\
&\times \left(\prod_{g_{k^*}=1}^{G_{k^*}+1} \frac{\Gamma(\tilde{a}_{g_{k^*}})}{\Gamma(a)} \frac{b^a}{\tilde{b}_{g_{k^*}}^{\tilde{a}_{g_{k^*}}}} \frac{|\,\Sigma_{g_{k^*}}\,|^{1/2}}{|\, V_\beta\,|^{1/2}}\right) \frac{\Gamma(e)}{f^e} \frac{(f+G_{k+1}+1)^{e+N}}{\Gamma(e+N)} \frac{(f+G_k+1)^{e+N}}{(f+G_{k^*}+1)^{e+N}} \\
&\times \left(\prod_{g_{k^*}=1}^{G_{k^*}+1} \frac{1}{N_{g_{k^*}}!}\right) \left(\prod_{g_k=1}^{G_k+1} N_{g_k}!\right) \left(\prod_{g_{k+1}=1}^{G_{k+1}+1} N_{g_{k+1}}!\right) \frac{l_k!\, l_{k+1}!}{l_{k^*}!} \frac{\Gamma(c)}{d^c} \frac{(d+1)^c \Gamma(c+l_{k^*})}{\Gamma(c+l_k)\Gamma(c+l_{k+1})} \\
&\times \frac{K\cancel{N}}{TN^{\cancel{2}}} \frac{(G_{k^*}+1)^N}{(G_k+1)^N (G_{k+1}+1)^N}
\end{aligned}
\tag{20}
$$

in which the term $K$ corresponds to sampling with equal probability one of the $K$ breaks to remove, $N$ corresponds to sampling with equal probability from one to $N$ the number of groups $G_{k^*}+1$ and $(G_{k^*}+1)^N$ corresponds to sampling with equal probability which of the $G_{k^*}+1$ groups each of the $N$ series is assigned to. The term $T^{-1}$ corresponds to the uniform sampling of $\tau_k$ from the entire time series if we were hypothetically proposing a birth move to return from the proposed position of $K-1$ breaks to our current position of $K$ breaks. The $N^{-2}$ term corresponds to the uniform



Electronic copy available at: https://ssrn.com/abstract=3268759

sampling of the number of groups in regimes $k$ and $k+1$ and the terms $(G_k+1)^{-N}$ and $(G_{k+1}+1)^{-N}$ correspond to sampling with equal probability the groups which each of the $N$ series are assigned to in regimes $k$ and $k+1$. If the move is accepted we substitute all proposals for their original values.

## 4. Hypothesis testing

The framework we develop nests the constant grouped heterogeneity model and the panel break models with either pooled or unit-specific parameters in every regime. These models vary in their complexity and degree of parameterisation. We develop a formal test to evaluate whether (i) there are any underlying patterns of grouped heterogeneity varying through regimes and (ii) allowing for either constant grouped heterogeneity or breaks alone is sufficient to approximate the underlying data generating process.

A Bayesian model comparison approach is natural because it integrates over all parameters and inherently penalises model complexity. Kass and Raftery (1995) state that Bayes Factors or marginal likelihoods are the optimal Bayesian model comparison technique. We first compute marginal likelihoods for our model and the three benchmark models. Bayes Factors are subsequently constructed which indicate the strength of evidence in favour of regime-specific grouped heterogeneity. Second, we compute marginal likelihoods for a panel model with constant grouped heterogeneity and panel break models with parameters that, in every regime, are either unit-specific or pooled across the entire cross-section. This informs whether any improved model fit is explained primarily by allowing for grouped heterogeneity to be regime-specific or allowing parameters in each regime to have patterns of grouped heterogeneity.

### *4.1. Testing for regime-specific patterns of grouped heterogeneity*

We now set out the formal test for regime-specific patterns of grouped heterogeneity. The marginal likelihood for model $r$ is calculated as

$$p(\boldsymbol{y} \mid M_r) = \frac{p(\boldsymbol{y} \mid \boldsymbol{\theta}^*, M_r)p(\boldsymbol{\theta}^* \mid \boldsymbol{y})}{p(\boldsymbol{\theta}^* \mid \boldsymbol{y}, M_r)} \tag{21}$$

where $\boldsymbol{\theta}^*$ denotes the posterior means of the parameters $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\sigma}^2)$ (Chib 1995). The calculation requires evaluation of the likelihood function, prior and posterior at this point.

Marginal likelihoods are computed for our model in which grouped heterogeneity



Electronic copy available at: https://ssrn.com/abstract=3268759

is regime-specific ($M_{RSGrp}$) and the constant panel models with either unit-specific ($M_{Unit}$) or pooled ($M_{Pool}$) parameters. The strength of evidence in favour of regime-specific grouped heterogeneity is provided by the Bayes Factors

$$BF_{Pool} = \frac{p(\boldsymbol{y} \mid M_{RSGrp})}{p(\boldsymbol{y} \mid M_{Pool})}. \tag{22}$$

and

$$BF_{Unit} = \frac{p(\boldsymbol{y} \mid M_{RSGrp})}{p(\boldsymbol{y} \mid M_{Unit})}. \tag{23}$$

Bayes Factors between 1 and 3 are not worth a mention, between 3 and 20 indicate positive evidence in favour of grouped heterogeneity being regime-specific, between 20 and 150 indicate strong evidence, and even higher values very strong evidence (Kass and Raftery 1995).

### *4.2. What drives regime-specific patterns of grouped heterogeneity?*

We now present formal tests to evaluate whether any identified regime-specific patterns of grouped heterogeneity are explained primarily by (i) allowing constant grouped heterogeneity to be regime-specific or (ii) allowing regime-specific parameters to have patterns of grouped heterogeneity. The former implies a more important role for breaks in driving regime-specific grouped heterogeneity while the latter suggests a greater role for grouped heterogeneity.

#### *4.2.1. Testing against constant grouped heterogeneity*

The marginal likelihood is computed for the model in which grouped heterogeneity is constant ($M_{Grp}$). The strength of evidence in favour of allowing grouped heterogeneity to be regime-specific is provided by the Bayes Factor

$$BF_{Grp} = \frac{p(\boldsymbol{y} \mid M_{RSGrp})}{p(\boldsymbol{y} \mid M_{Grp})}. \tag{24}$$

#### *4.2.2. Testing against breaks*

The marginal likelihood is computed for the panel models with breaks and parameters that, in each regime, are either unit-specific ($M_{Brk,unit}$) or pooled across the entire cross-section ($M_{Brk,pool}$). The strength of evidence in favour of allowing regime-specific



Electronic copy available at: https://ssrn.com/abstract=3268759

parameters to have patterns of grouped heterogeneity is provided by the Bayes Factors

$$BF_{Brk,unit} = \frac{p(\boldsymbol{y} \mid M_{RSGrp})}{p(\boldsymbol{y} \mid M_{Brk,unit})} \tag{25}$$

and

$$BF_{Brk,pool} = \frac{p(\boldsymbol{y} \mid M_{RSGrp})}{p(\boldsymbol{y} \mid M_{Brk,pool})}. \tag{26}$$

## 5. Simulation study

To demonstrate the success of the method in estimating panel regression models with structural breaks and regime-specific grouped heterogeneity we now conduct a simulation exercise in which the true underlying data generating process is known. The framework we develop is compared with the grouped heterogeneity model without breaks and the common break models with either pooled or unit-specific parameters in every regime.

### *5.1. Data generating process*

The time series and cross-section dimensions of the data set are $T = 100$ and $N = 20$. The data generating process experiences $K = 2$ common breaks such that $\tau_1 = 35$ and $\tau_2 = 70$. The three regimes have three, one and two groups such that $G_1 = 2$, $G_2 = 0$ and $G_3 = 1$. The blue circles in Figure 3 denote which series belong to which group across the three regimes.

For regimes $k = 1, \ldots, K+1$ and groups $g_k = 1, \ldots, G_k + 1$, we construct the dependent variable as follows

$$y_{it} = \beta'_{g_k} X_{it} + \epsilon_{it}, \qquad i \in N_{g_k}, \qquad t = \tau_{k-1} + 1, \ldots, \tau_k \tag{27}$$

in which $X_{it} \sim N(0,1)$ and $\epsilon_{it} \sim N(0, \sigma^2_{g_k})$. Table 1 displays the simulated values of the slope coefficient $\beta_{g_k}$ (top panel) and the error-term variance $\sigma^2_{g_k}$ (bottom panel) in each regime and group.

It is assumed that a new break onsets every fifty periods since $c = 100$ and $d = 2$. The hyperparameters of the inverse gamma distribution on the error-term variance $a$ and $b$ are set equal to 1 and the variance of $\beta$, $\sigma^2_\beta$, is set equal to 0.5. To achieve a prior expected number of groups in each regime equal to 2 we set $f = 1$ and $e = 10$.



Electronic copy available at: https://ssrn.com/abstract=3268759

### 5.2. *Findings*

The simulation study illustrates the bias-variance improvements derived from our framework that allows for regime-specific grouped heterogeneity. The $N = 20$ series fall into six groups and the paths of their true slope coefficients are graphed by the blue solid lines in Figure 2. Parameter estimates are biased from the model with breaks and parameters pooled across the entire cross-section (red dotted line) and the model that allows for constant grouped heterogeneity (green dotted line). The model with breaks and regime-specific grouped heterogeneity delivers unbiased estimates (black dashed line) because it accurately estimates the break dates, the number of cross-sectional groups in each regime and allocates the correct series to each group. This can be seen in Figure 3 that displays the posterior mode group allocation estimates (black crosses) and the true allocations.

The model with breaks and unit-specific parameters correctly identifies the break dates and thus delivers unbiased parameter estimates. However, its estimates are imprecise as can be seen by the wide range of posterior mean parameter estimates across all series and regimes (red triangles) in the left panel of Figure 4. Allowing for breaks and grouped heterogeneity (black crosses) provides more precise estimates that are closer to the true values (blue circles) and reduces the standard deviations of the parameter estimates (right panel). Table 1 displays that the bias-variance improvements of our method hold for both the slope coefficient (top panel) and the error-term variance (bottom panel).

The Bayes Factors relative to the model with constant grouped heterogeneity, the model with breaks and pooled parameters and the model with breaks and unit-specific parameters are 179.770, 241.297 and 315.014, respectively, providing extremely strong evidence in favour of our model. Allowing for grouped heterogeneity therefore appears to be more important than allowing for breaks.

## 6. Empirical Application

Forecasting inflation is challenging since a range of considerations such as monetary policy and the inflation expectations of economic agents that may vary through time must be accounted for. In line with this view, the literature has documented evidence of instability in the predictive performance of leading predictors, such as the unemployment rate and the growth of real output, to forecast inflation (Stock and Watson 2003; Faust and Wright 2013).



Electronic copy available at: https://ssrn.com/abstract=3268759

The empirical application analyses the ability of industry-level predictors to forecast inflation across twenty U.S. industries using univariate predictive regressions. The lead predictor is the industry-level unemployment rate. We also present results for industrial production.

The number and timing of breaks along with the regime-specific grouped patterns of heterogeneity are first analysed. Second, we compare out-of-sample forecasts produced by the predictive panel regression with regime-specific grouped heterogeneity with forecasts generated by common break panel models that have either pooled or unit-specific parameters in every regime, a panel model with grouped heterogeneity but no breaks, and a univariate autoregressive model applied independently to each series in the cross-section.

### *6.1. Data*

The data set consists of 20 industries: accommodation, chemicals, computer and electric products, electrical equipment, food, furniture and related products, financials, hospitals, machinery, manufacturing, mining, nonmetallic mineral products, other, petroleum and coal products, plastics and rubber products, real estate, telecommunications, transportation, utilities, and wood products.

Each monthly series begins in January 2004 and ends in March 2018. Industry-level inflation rates, sourced from the Bureau of Labor Statistics (BLS), are constructed as $y_{it} = 1200 \times log(PPI_{it}/PPI_{it-1})$ in which $PPI$ is the Producer Price Index.

Monthly unemployment rates are sourced from BLS. Monthly industrial production in real terms (seasonally adjusted) for each industry is constructed as $X_{it} = 1200 \times log(IP_{it}/IP_{it-1})$ in which $IP$ is the index of industrial production. This data is from Federal Reserve Economic Data (FRED).[15]

We use relative importance weights (contribution to the total industrial production index) from FRED as industry weights to construct aggregate U.S. inflation forecasts from the industry-level forecasts.

We construct an aggregate U.S. inflation series as

$$y_{USt} = 1200 \times log(PPI_{USt}/PPI_{USt-1}) \tag{28}$$

[15] Due to data availability constraints, industrial production is not available for financials, accommodation, transportation, real estate, and hospitals.



Electronic copy available at: https://ssrn.com/abstract=3268759

in which $PPI_{USt}$ is PPI Commodity index for All commodities (not seasonally adjusted) at month $t$ sourced from BLS.[16]

Table 2 presents summary statistics for the monthly inflation rates of each industry. The average monthly inflation rates range from -0.826 for computer and electrical products (Computer) to 5.827 for Mining with three industries - computer and electrical products, financials, and telecommunications - experiencing deflation over the sample. The high standard deviations along with the large range of minimum and maximum values reflect the large fluctuations in inflation with petroleum and coal products being by far the most volatile with a standard deviation of 84.823 and having notable outliers with minimum and maximum values of -364.174 and 170.868. Persistence in inflation across industries is quite heterogeneous. Five industries have negative autocorrelations, while the remaining 15 have positive autocorrelations with food and chemicals having the highest autocorrelation of 0.576. Aggregate U.S. inflation has a mean of 2.433, standard deviation of 14.619 and autocorrelation of 0.465.

### *6.2. Model Specifications*

The panel model with common breaks and regime-specific grouped heterogeneity developed in this article is specified (for series $i \in N_{g_k}$, regimes $k = 1, \ldots, K+1$ and groups $g_k = 1, \ldots, G_k + 1$) as

$$y_{it} = \beta_{0g_k} + \beta_{1g_k} y_{it-1} + \beta_{2g_k} X_{it-1} + \gamma_{g_k} F_t + e_{it}, \qquad t = \tau_{k-1} + 1, \ldots \tau_k, \tag{29}$$

in which $y_{it}$ is the inflation rate for the $i$th industry at time $t$, $X_{it-1}$ is the unemployment rate for the $i$th industry at time $t-1$, $F_t$ is the U.S. aggregate inflation rate at time $t$ which acts as an observed common factor, and $\gamma_{g_k}$ is the factor loading for the $g_k$th group.

We compare the out-of-sample forecasting performance of this model with four benchmark models. The first is a panel model with grouped heterogeneity but no breaks. For $i \in N_g$ and $t = 1, \ldots, T$ the model is specified as

$$y_{it} = \beta_{0g} + \beta_{1g} y_{it-1} + \beta_{2g} X_{it-1} + \gamma_g F_t + e_{it}, \qquad g = 1, \ldots, G+1. \tag{30}$$

The second is a panel model with breaks and parameters pooled across the entire cross-section in each regime. For series $i = 1, \ldots, N$ and regimes $k = 1, \ldots, K+1$,

[16] We apply an x13 filter to seasonally adjust the industry-level and aggregate U.S. inflation series.



Electronic copy available at: https://ssrn.com/abstract=3268759

the model is specified as

$$y_{it} = \beta_{0k} + \beta_{1k} y_{it-1} + \beta_{2k} X_{it-1} + \gamma_k F_t + e_{it}, \qquad t = \tau_{k-1} + 1, \ldots \tau_k. \tag{31}$$

The third benchmark is a panel model with breaks and unit-specific parameters in every regime. For series $i = 1, \ldots, N$ and regimes $k = 1, \ldots, K + 1$, the model is specified as

$$y_{it} = \beta_{0ik} + \beta_{1ik} y_{it-1} + \beta_{2ik} X_{it-1} + \gamma_{ik} F_t + e_{it}, \qquad t = \tau_{k-1} + 1, \ldots \tau_k. \tag{32}$$

The fourth and final benchmark is an AR(1) model estimated by Ordinary Least Squares and applied independently to each series in the cross-section

$$y_{it} = \beta_{0i} + \beta_{1i} y_{it-1} + e_{it}, \qquad i = 1, \ldots, N, \qquad t = 2, \ldots, T. \tag{33}$$

*6.3. Constructing Inflation Forecasts*

Each iteration on which inference is performed, our framework estimates $G_k+1$ groups in each of the $K+1$ regimes. The group structure and parameter estimates from the $G_{K+1}+1$ groups in the final regime are used to produce forecasts. The Markov chain Monte Carlo sweep consists of many iterations. Each iteration a different number of breaks may be estimated, at different times, with a different number of groups and cross-sectional ordering of series in each regime. Any uncertainty surrounding all of these elements of the model is inherently incorporated into the parameter estimates and thus the inflation forecasts through Bayesian Model Averaging.

Aggregate U.S. inflation forecasts $\hat{y}_{US,t+1}$ are computed as a weighted average of the industry-level forecasts

$$\hat{y}_{US,t+1} = w_t' \hat{y}_{t+1}, \qquad t = 1, \ldots, T - 1 \tag{34}$$

in which $w_t$ denotes the vector of importance weights applied to each industry at time $t$ and $\hat{y}_{t+1} = (\hat{y}_{1t+1}, \ldots, \hat{y}_{Nt+1})$ contains inflation forecasts for the $N$ industries produced at time $t$.

*6.4. Testing cross-sectional dependence*

Inflation rates across U.S. industries are likely to be cross-sectionally dependent. This section presents results of Pesaran (2004)'s formal cross-sectional dependence ($CD$) test, which is particularly suitable because it is robust to multiple breaks, to analyse



Electronic copy available at: https://ssrn.com/abstract=3268759

whether including the U.S. aggregate inflation rate as an observed common factor in the model successfully absorbs cross-sectional dependencies.

For series $i = 1, \ldots, N$ we obtain the residuals from an OLS regression of inflation rates on lagged inflation and unemployment rates

$$u_{it} = y_{it} - \hat{\beta}_{0i} - \hat{\beta}_{1i} y_{it-1} - \hat{\beta}_{2i} X_{it-1}, \qquad t = 2, \ldots, T. \tag{35}$$

We compute pairwise correlations $\hat{\rho}_{ij}$ of the residuals from series $i$ and $j$ in which $i \neq j$

$$\hat{\rho}_{ij} = \frac{\sum_{t=1}^{T} u_{it} u_{jt}}{\left(\sum_{t=1}^{T} u_{it}^2\right)^{1/2} \left(\sum_{t=1}^{T} u_{jt}^2\right)^{1/2}}. \tag{36}$$

The $CD$ test statistic, which has a standard Normal distribution, is computed as

$$CD = \sqrt{\frac{2T}{N(N-1)}} \left( \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{ij} \right). \tag{37}$$

The average pairwise correlation is 0.133 with a CD statistic equal to 23.988 rejecting the null of cross-sectional weak dependence.

We compute the same statistics using residuals from a regression that includes the aggregate U.S. inflation rate on the right hand side of equation (35). Including this common factor removes approximately half of the cross-sectional dependence. Specifically, the average pairwise correlation is reduced to 0.066 and the CD statistic to 11.901. Smith et al. (2017) report that average cross-sectional correlations below 0.2 do not affect break detection power in panels.

Nonetheless, the primary goal of the model is not to perform inference on the parameter estimates, but to deliver more accurate out-of-sample inflation forecasts than a range of popular benchmark models. The remaining cross-sectional dependence is therefore unlikely to cause any serious problems for our empirical application.

### *6.5. Priors*

In the empirical application we assign the same prior values to each of the competing models. Specifically, we assume that a new regime onsets every ten years by setting the hyperparameters $c$=240 and $d$=2. The prior variance of $\beta$ is determined by $\sigma_\beta^2$ which is set equal to 0.1. The hyperparameters $a$ and $b$ determine the error-term variance and we set $a = b = 2$. For the models with grouped heterogeneity the



Electronic copy available at: https://ssrn.com/abstract=3268759

prior expected number of groups (which is constant across any identified regimes) is determined by the hyperparameters $e$ and $f$. To achieve a prior expected number of groups in each regime approximately equal to three we set $f = 1$ and $e = 7$.

### *6.6. Evidence of breaks*

The analysis focuses on the forecasting model that includes lagged industry-level unemployment rates as the predictive variable alongside an intercept and one lag of industry-level inflation.

As displayed in Figure 5 the model with regime-specific grouped heterogeneity identifies two breaks with 100% posterior probability (top window) that occur around February 2008 and October 2009. These break dates are estimated with considerable certainty as measured by the posterior probabilities (bottom window). The short regime enclosed by the two identified breaks corresponds to the global financial crisis.

The panel break model with pooled parameters identifies two breaks at July 2008 and October 2010 while the panel break model with unit-specific parameters identifies one break in June 2010. This reflects the reduced break detection power of the imprecisely estimated heterogeneous panel break model.

The Bayes Factors relative to the model with constant grouped heterogeneity, the model with breaks and pooled parameters and the model with breaks and unit-specific parameters are 197.615, 230.218 and 298.007, respectively, providing extremely strong evidence in favour of our model. Grouped heterogeneity therefore appears to be more important in fitting the data than allowing for breaks.

### *6.7. Out-of-sample inflation forecasts*

We first estimate the model using only the data from January 2004 through December 2008 corresponding to an initial estimation period of five years. We recursively move through the sample one month at a time generating forecasts at each point using only the data available at the time the forecast is made. The final forecast for March 2018 is generated from parameters estimated using the data from January 2004 through February 2018.

Regime-specific grouped heterogeneity allows for a data-driven approach to bias-variance management that varies through time by estimating the number of groups and cross-sectional ordering of series in each regime. To avoid biases, some series may be allocated their own group, such as the petroleum industry which has a notably



Electronic copy available at: https://ssrn.com/abstract=3268759

different inflation series from all the other industries, while others may be grouped together to allow the pooled parameter within their group to be estimated more precisely. Allowing for time-variation in the group dynamics is critical since the optimal bias-variance tradeoffs implied by the grouping structure may be unstable, i.e., one might expect fewer groups and more pooling during highly volatile episodes such as the global financial crisis.

Figure 6 displays forecasts of aggregate U.S. inflation recursively generated from our model (top window), the break models with pooled and unit-specific parameters (blue solid and red dashed lines in the middle window), the AR(1) model estimated by OLS (blue dashed line in the bottom window), and the model with constant grouped heterogeneity (red solid line).

Our model is the most able to adjust in real time to sharp changes in inflation. Looking at the early part of the sample illustrates this point. Only our model and the model with breaks and pooled parameters are able to generate a large magnitude (beyond -6%) for the negative inflation experienced in early 2009. However, a few months later the breaks model with pooled parameters is unable to produce the large positive forecasts quickly enough, initially peaking at 5% before falling again, while our model peaks at over 6%.

### *6.8. Performance of inflation forecasts*

The forecasting performance of our model relative to each of the benchmarks is measured through the out-of-sample $R^2$ value

$$R^2_{OoS} = 1 - MSE_{RSGrp}/MSE_{bmk} \tag{38}$$

in which $MSE_{RSGrp}$ and $MSE_{bmk}$ denote the mean squared forecast errors obtained from the model with regime-specific grouped heterogeneity and the benchmark model. If our model outperforms the benchmark in terms of lower mean squared forecast errors the $R^2_{OoS}$ values will be positive. All reported $R^2_{OoS}$ values are multiplied by 100.

Each window in Figure 7 corresponds to one of the four benchmark models and displays the $R^2_{OoS}$ relative to that benchmark across the 21 cases (20 industries and U.S. aggregate). Our model delivers positive $R^2_{OoS}$ values for all 21 cases relative to the AR(1) model (top left window) with 18 cases above 20%. For the U.S. aggregate (thick black vertical line), the $R^2_{OoS}$ value is approximately 20%.

Relative to the panel breaks models with pooled and unit-specific parameters



Electronic copy available at: https://ssrn.com/abstract=3268759

(bottom windows), 20 of the 21 $R^2_{OoS}$ values are positive with the U.S. aggregate having an $R^2_{OoS}$ around 20%. Compared with the model with constant grouped heterogeneity (top right), our model delivers positive $R^2_{OoS}$ values for all but four industries with 15 cases having an $R^2_{OoS}$ value greater than 10%.

To assess the statistical significance of any out- or under-performance delivered by our method we first implement the procedure of Diebold and Mariano (1995) and second, to account for a nonstandard distribution of the test statistic that may arise from our forecasting models being nested, the MSE-adjusted test statistic of Clark and West (2007).

The results are presented in Table 3. The table displays for how many of the 21 cases our method significantly underperforms ($t < -1.64$), insignificantly underperforms ($-1.64 < t < 0$), insignificantly outperforms ($0 < t < 1.64$), or significantly outperforms ($t > 1.64$) when using the Diebold-Mariano (DM) or the Clark-West (CW) test. The four panels correspond to the four benchmark models. The first row of each panel displays results when using the unemployment rate (unmpl) as the predictive variable alongside an intercept and a single lag of inflation. The bin in which U.S. aggregate inflation falls is indicated by the † symbol.

Focusing on the DM statistic, relative to the univariate AR(1) model our method produces significantly more accurate point forecasts for 20 out of 21 cases including the U.S. aggregate. The model still outperforms for the remaining industry but the outperformance is insignificant.

The results are almost as strong relative to the panel break models with pooled and unit-specific parameters significantly outperforming for 20 and 18 cases including the U.S. aggregate. Our model significantly outperforms the constant grouped heterogeneity model for 17 cases. Across all 84 cases (21 cases against each of the four benchmarks) our model significantly underperforms just four times.

Using the Clark-West statistic, the results are slightly stronger. Our method significantly outperforms for between 17 and 21 cases, significantly underperforms for just three out of the 84 cases and always significantly outperforms for the U.S. aggregate.

These formal tests of statistically significant out- or under-performance are measured using the entire out-of-sample period and thus are averaged across all observations. A strong forecasting model, however, should consistently outperform the benchmark models over the entire out-of-sample period and not be driven by only one or two short periods. To evaluate whether our model consistently outperforms over the out-of-sample period Figure 8 graphs the cumulative sum of squared forecast



Electronic copy available at: https://ssrn.com/abstract=3268759

error differences ($CSSED$) for series $i$

$$CSSED_{it} = \sum_{\tau=1}^{t}(e^2_{Bmk,i\tau} - e^2_{RSGrp,i\tau}), \qquad t = 61, \ldots, T \tag{39}$$

in which $e_{Bmk,i\tau}$ and $e_{RSGrp,i\tau}$ denote the forecast errors from the benchmark model and our model for series $i$ at time $\tau$.

Positive and rising values indicate that our model is outperforming the benchmark while underperformance is reflected by negative and declining values. Horizontal parts of the graph indicate the two models are generating equally accurate forecasts. Consistent outperformance over the out-of-sample period will be reflected by a smooth upward-sloping curve; outperformance driven by one or two short periods will result in a curve that spikes upwards during those short periods and is flat or declining during other periods.

The four windows in Figure 8 graph the $CSSED$ measure relative to the four benchmark models for the U.S. aggregate and the financials, mining and plastics industries. For the U.S. aggregate, each of the curves are upward-sloping throughout the out-of-sample period indicating that our model produces consistently more accurate forecasts than each of the four benchmarks. For the three industries, our method consistently outperforms the four benchmarks throughout the sample. The only underperformance occurs briefly at the beginning of the out-of-sample period for the financials and plastics industries in which our model underperforms relative to one or two of the panel break models.

### *6.9. Regime-specific grouped heterogeneity*

Comparing the forecasting performance of the competing models in the previous subsection provides evidence in favour of allowing for regime-specific grouped heterogeneity when generating out-of-sample inflation forecasts. Allowing for either grouped heterogeneity or structural breaks alone is unable to rival the forecast accuracy generated from allowing for regime-specific grouped heterogeneity. It is therefore critical for economic agents, such as policy makers in a central bank or investors in financial markets, to allow for regime-specific grouped heterogeneity when forecasting inflation in real time.

The top left window of Figure 9 graphs the posterior mode number of groups estimated from the model that allows for regime-specific grouped heterogeneity (solid black line) and the model that allows for constant grouped heterogeneity (red dashed



Electronic copy available at: https://ssrn.com/abstract=3268759

line). Relative to the model with constant grouped heterogeneity, our model identifies notably fewer groups in the second regime that corresponds to the global financial crisis. Five groups are identified compared with ten and nine in the preceding and subsequent regimes. The top right and two middle windows display histograms of the posterior mode number of series in each group across the three regimes identified by our model. Identifying fewer groups results in more series sharing groups and thus parameters being pooled across more series in the crisis regime (middle left) compared with the other two regimes (top right and middle right).

Only the petroleum industry, which has a markedly different inflation series from all the other industries, is assigned its own group during the crisis regime compared with five series being assigned their own groups in the other regimes. This illustrates the ability of our framework to vary the grouped heterogeneity through time enabling fewer groups and more pooling during highly volatile and short regimes that are more susceptible to parameter estimation error. Conversely, it can identify more groups with less pooling, even allocating unit-specific parameters to a number of series, reducing the amount of bias induced in the parameter estimates during longer and less volatile regimes, which is particularly for series that are unlike any other.

Existing methods lack this flexibility. The model that allows for constant grouped heterogeneity, for example, identifies 11 groups throughout the sample (top left) with six of them being assigned their own group (bottom left). This inability to allow the number of groups and degree of pooling to vary through time results in suboptimal management of bias-variance tradeoffs and ultimately less accurate forecasts as detailed in the previous subsection. By the same token, the panel break models with pooled and unit-specific parameters are unable to identify grouped heterogeneity resulting in biased or imprecise parameter estimates and less accurate forecasts.

The ability of our model to identify how the grouped heterogeneity varies through time is central to its ability to optimally manage bias-variance tradeoffs and thus generate significantly more accurate forecasts than the four benchmark models.

## 7. Extensions and robustness

This section presents some additional analysis to explore the robustness of the results before setting out some potential extensions of the model. We first present results when industrial production is used as an alternative predictor and second explore the sensitivity of the baseline results to the choice of prior. Finally, we briefly discuss how our model can be extended to (i) allow for partial breaks and (ii) include individual



Electronic copy available at: https://ssrn.com/abstract=3268759

fixed effects.

*7.1. Results for industrial production*

The baseline results presented thus far have focused on inflation forecasts produced from a model in which the predictive variable is the unemployment rate that appears alongside an intercept and an autoregressive term. A number of other predictors, however, have been proposed by the literature - Faust and Wright (2013) present a survey. To explore the robustness of the forecasting performance of our new model to alternative predictors we now present results when forecasting with industrial production as the predictor that appears alongside an intercept and an autoregressive term.

Figure 10 displays the $R^2_{OoS}$ values across all fifteen industries and the U.S. aggregate produced by our model against the four benchmark models when industrial production is the predictive variable.[17] Our model continues to outperform all four benchmarks for the majority of cases. The second row of each panel displayed in Table 3 (labeled indpr) shows that the statistical significance of our model's outperformance remains equally strong.

*7.2. Sensitivity of results to choice of priors*

Turning to the sensitivity of the baseline results to the choice of prior when forecasting with the unemployment rate, we adjust one hyperparameter value at a time substituting the value reported in Table 4 for the baseline value and report the corresponding statistical significance of out- or under-performance of our model relative to the four benchmarks.

The number and timing of breaks identified by our model are unaffected by reducing the prior expected regime duration from ten to five years ($c = 120$). Increasing the prior expected number of groups in every regime from approximately three to seven ($e = 3$) does not affect the number of groups or the cross-sectional ordering of series across the three identified regimes. We also set $b = 6$ which gives a prior expected mean error-term variance three times as large as the baseline. Finally, we set $\sigma^2 = 0.01$ and $\sigma^2 = 1$ which gives tighter and looser priors on the slope coefficient $\beta$.

[17]Recall from Section 6 that due to data availability constraints we only have data for 15 industries for industrial production.



Electronic copy available at: https://ssrn.com/abstract=3268759

Across all these changes in prior distributions, our model continues to significantly outperform each of the four benchmark models for at least 16 out of the 21 cases whether using the Diebold-Mariano or Clark-West test.

### *7.3. Partial break models*

Our model assumes all regression parameters - intercept, slope and variance - follow the same breakpoint process and grouping structure. Peluso et al. (2016) specify a framework that allows each parameter to follow a separate breakpoint process. It is possible to extend the current framework to allow each parameter to follow its own breakpoint process and grouping structure. This is particularly attractive if the grouping structure differs between parameters or one parameter is estimated with less precision since this parameter could have fewer groups and more pooling to reduce the parameter estimation error. This is a promising direction for future research.

### *7.4. Individual fixed effects*

It may be desirable to augment the model with additive time-invariant unit-specific fixed effects in addition to the regime-specific group fixed effects. This would require time-variation in the covariates $X_{it}$. Moreover, consistent estimation of group membership is only possible if the group-specific profiles are not parallel (Bonhomme and Manresa 2015). Finally, including a unit-specific fixed effect will induce Nickell (1981) bias if the covariates include a lagged dependent variable as in the empirical application presented here. If forecasting is the primary interest, this may be of little importance. If, however, one is performing inference on the parameter estimates then an alternative estimation strategy, such as instrumental variables, may be required for consistent estimation.

## 8. Conclusion

This article develops a new approach to estimate panel regression models in the presence of structural breaks and regime-specific grouped heterogeneity. The number and timing of breaks is estimated endogenously as is the number of groups and the cross-sectional ordering of series within each regime. Parameters are pooled within, but differ across, regime-specific groups. Bayes Factors are computed from marginal likelihoods to test formally for regime-specific grouped heterogeneity while integrating



Electronic copy available at: https://ssrn.com/abstract=3268759

over all parameters and inherently penalising model complexity.

Our method detects two breaks at February 2008 and October 2009, enclosing a short and volatile regime corresponding to the global financial crisis, in an empirical application that forecasts U.S. industry-level inflation rates using lagged industry-level unemployment rates and monthly data from 2004-2018. Grouped heterogeneity varies considerably across the three regimes. Models that allow for grouped heterogeneity but ignore breaks or allow for breaks but pool parameters across the entire cross-section induce biased parameter estimates, whereas panel break models with unit-specific parameters are estimated imprecisely. The improvement in out-of-sample forecasting performance relative to four popular benchmarks for 20 U.S. industries and the U.S. aggregate suggests that allowing grouped heterogeneity to vary through regimes is critical for economic agents, such as policy-makers in a central bank, that must produce inflation forecasts in real time.

The methodology has many potential applications across economics and finance. There is considerable evidence that economic forecasting models are subject to breaks (Stock and Watson 1996). However, existing time series procedures for forecasting in the presence of breaks have had limited success when attempting to produce more accurate out-of-sample forecasts. This is largely because managing bias-variance tradeoffs is difficult and existing methods can only reduce variance by necessarily inducing bias (Pesaran and Timmermann 2007). The method we develop allows variance to be reduced without inducing any bias by (i) identifying a grouping structure and pooling parameters across all series within a group and (ii) allowing the grouping structure to vary through regimes. The more series that share a group the greater the reduction in estimation error. The bias-variance tradeoff is managed more optimally by using cross-sectional information compared with using individual time-series data alone. Since cross-sectional information is present in many economic data sets the approach could have broad implications for optimal management of bias-variance tradeoffs and forecasting more generally.



Electronic copy available at: https://ssrn.com/abstract=3268759

## Appendix A: marginalising $\beta$ and $\sigma^2$ from the posterior

We now explicitly show how $\beta$ and $\sigma^2$ are marginalised from the posterior and thus how equation (12) is derived. Multiplying the priors on the regression coefficients by the likelihood obtains

$$p(\boldsymbol{\beta}, \boldsymbol{\sigma^2} \mid \boldsymbol{y}, \boldsymbol{X}, \tau, G) = \left( \prod_{k=1}^{K+1} \prod_{g_k=1}^{G_k+1} (2\pi)^{-\left(\frac{\kappa + l_k N_{g_k}}{2}\right)} (\sigma^2_{g_k})^{-\left(a+1+\frac{\kappa+l_kN_{g_k}}{2}\right)} \frac{b^a}{\Gamma(a)} \mid V_\beta \mid^{\frac{-1}{2}} \right)$$
$$\times\, exp \left[ \sum_{k=1}^{K+1} \sum_{g_k=1}^{G_k+1} -\frac{1}{2\sigma^2_{g_k}} \left( 2b + \beta'_{g_k} V_\beta^{-1} \beta_{g_k} + \sum_{i \in N_{g_k}} \sum_{t=\tau_{k-1}+1}^{\tau_k} (y_{it} - X'_{it}\beta_{g_k})^2 \right) \right]. \tag{A.1}$$

Rewriting equation (A.1) and multiplying and dividing by $2\pi^{-\kappa/2} \mid \Sigma_{g_k} \mid^{-1/2} (\sigma^2_{g_k})^{-\kappa/2}$ we can marginalise $\boldsymbol{\beta}$ using equation (13)

$$p(\boldsymbol{\sigma^2} \mid \boldsymbol{y}, \boldsymbol{X}, \tau, G) = \int p(\boldsymbol{\sigma^2}, \boldsymbol{\beta} \mid \boldsymbol{y}, \boldsymbol{X}, \tau, G) d\boldsymbol{\beta}$$
$$= \left( \prod_{k=1}^{K+1} \prod_{g_k=1}^{G_k+1} 2\pi^{-(\cancel{\kappa}+l_kN_{g_k})/2} (\sigma^2_{g_k})^{-(a+1+\cancel{\kappa/2}+l_kN_{g_k}/2)} \frac{b^a}{\Gamma(a)} \mid V_\beta \mid^{\frac{-1}{2}} \mid \Sigma_{g_k} \mid^{\frac{1}{2}} \cancel{2\pi^{\frac{\kappa}{2}} (\sigma^2_{g_k})^{\frac{\kappa}{2}}} \right)$$
$$\times\, exp \left[ -\sum_{k=1}^{K+1} \sum_{g_k=1}^{G_k+1} \frac{1}{2\sigma^2_{g_k}} \left( 2b + \mu'_{g_k} \Sigma^{-1}_{g_k} \mu_{g_k} + \sum_{i \in N_{g_k}} \sum_{t=\tau_{k-1}+1}^{\tau_k} y^2_{it} \right) \right]$$
$$\times \cancel{\int \left( \prod_{k=1}^{K+1} \prod_{g_k=1}^{G_k+1} (2\pi\sigma^2_{g_k})^{\frac{-\kappa}{2}} \mid \Sigma_{g_k} \mid^{\frac{-1}{2}} exp \left[ \sum_{k=1}^{K+1} \sum_{g_k=1}^{G_k+1} \frac{(\beta_{g_k} - \mu_{g_k})' \Sigma^{-1}_{g_k} (\beta_{g_k} - \mu_{g_k})}{-2\sigma^2_{g_k}} \right] \right) d\beta_{g_k}} \tag{A.2}$$

whereby the integral of the term that cancels equals one because the full conditional distribution of $\beta_{g_k}$ is multivariate Normal with mean $\mu_{g_k}$ and covariance matrix $\Sigma_{g_k}\sigma^2_{g_k}$.

Rearranging equation (A.2) and multiplying and dividing by $\Gamma(\tilde{a}_{g_k})/\tilde{b}_{g_k}^{\tilde{a}_{g_k}}$ we marginalise



Electronic copy available at: https://ssrn.com/abstract=3268759

$\boldsymbol{\sigma^2}$ as follows

$$
\begin{aligned}
p(\boldsymbol{y} \mid \boldsymbol{X}, \tau, G) &= \int p(\boldsymbol{\sigma^2} \mid \boldsymbol{y}, \boldsymbol{X}, \tau, G) d\boldsymbol{\sigma^2} \\
&= \left( \prod_{k=1}^{K+1} \prod_{g_k=1}^{G_k+1} 2\pi^{-l_k N_{g_k}/2} \frac{b^a}{\Gamma(a)} \frac{\mid \Sigma_{g_k} \mid^{1/2}}{\mid V_\beta \mid^{1/2}} \frac{\Gamma(\tilde{a}_{g_k})}{\tilde{b}_{g_k}^{\tilde{a}_{g_k}}} \right) \\
&\times \int \left( \left( \prod_{k=1}^{K+1} \prod_{g_k=1}^{G_k+1} \frac{\tilde{b}_{g_k}^{\tilde{a}_{g_k}}}{\Gamma(\tilde{a}_{g_k})} (\sigma^2_{g_k})^{-(\tilde{a}_{g_k}+1)} \right) exp \left[ - \sum_{k=1}^{K+1} \sum_{g_k=1}^{G_k+1} \frac{\tilde{b}_{g_k}}{\sigma^2_{g_k}} \right] \right) d\sigma^2_{g_k}
\end{aligned} \tag{A.3}
$$

whereby the integral equals one because the full conditional of $\sigma^2_{g_k}$ is inverse gamma with parameters $\tilde{a}_{g_k}$ and $\tilde{b}_{g_k}$.



Electronic copy available at: https://ssrn.com/abstract=3268759

## References


Ando, T. and J. Bai, "Panel data models with grouped factor structure under unknown group membership," *Journal of Applied Econometrics* 31 (2016), 163–191.

Andrews, D. W., "Tests for parameter instability and structural change with unknown change point," *Econometrica* 61 (1993), 821–856.

———, "Cross-section Regression with Common Shocks," *Econometrica* 73 (2005), 1551–1585.

Bai, J., "Panel data models with interactive fixed effects," *Econometrica* 77 (2009), 1229–1279.

———, "Common breaks in means and variances for panel data," *Journal of Econometrics* 157 (2010), 78–92.

Bai, J. and J. L. Carrion-I-Silvestre, "Structural changes, common stochastic trends, and unit roots in panel data," *Review of Economic Studies* 76 (2009), 471–501.

Bai, J., R. L. Lumsdaine and J. H. Stock, "Testing for and dating common breaks in multivariate time series," *Review of Economic Studies* 65 (1998), 395–432.

Bai, J. and P. Perron, "Estimating and testing linear models with multiple structural changes," *Econometrica* 66 (1998), 47–78.

———, "Computation and analysis of multiple structural change models," *Journal of Applied Econometrics* 18 (2003), 1–22.

Baltagi, B. H., Q. Feng and C. Kao, "Estimation of heterogeneous panels with structural breaks," *Journal of Econometrics* 191 (2016), 176–195.

Bollerslev, T., B. Hood, J. Huss and L. H. Pedersen, "Risk everywhere: Modeling and managing volatility," *Review of Financial Studies* 31 (2018), 2729–2773.

Bonhomme, S., T. Lamadon, E. Manresa et al., "Discretizing unobserved heterogeneity," *Unpublished Manuscript* (2017).




Electronic copy available at: https://ssrn.com/abstract=3268759

Bonhomme, S. and E. Manresa, "Grouped patterns of heterogeneity in panel data," *Econometrica* 83 (2015), 1147–1184.

Canova, F., "Testing for convergence clubs in income per capita: a predictive density approach," *International Economic Review* 45 (2004), 49–77.

Carlin, B. P. and S. Chib, "Bayesian model choice via Markov chain Monte Carlo methods," *Journal of the Royal Statistical Society. Series B (Methodological)* 57 (1995), 473–484.

Carlino, G. and R. DeFina, "The differential regional effects of monetary policy," *Review of Economics and Statistics* 80 (1998), 572–587.

Chib, S., "Marginal likelihood from the Gibbs output," *Journal of the American Statistical Association* 90 (1995), 1313–1321.

———, "Estimation and comparison of multiple change-point models," *Journal of Econometrics* 86 (1998), 221–241.

Chudik, A. and M. H. Pesaran, "Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors," *Journal of Econometrics* 188 (2015), 393–420.

Clark, T. E. and K. D. West, "Approximately normal tests for equal predictive accuracy in nested models," *Journal of Econometrics* 138 (2007), 291–311.

Diebold, F. X. and R. S. Mariano, "Comparing predictive accuracy," *Journal of Business & Economic Statistics* 13 (1995), 253–263.

Elliott, G. and U. K. Müller, "Efficient tests for general persistent time variation in regression coefficients," *Review of Economic Studies* 73 (2006), 907–940.

Faust, J. and J. H. Wright, "Forecasting inflation," *Handbook of Economic Forecasting* 2 (2013), 3–56.

Geweke, J. and Y. Jiang, "Inference and prediction in a multiple-structural-break model," *Journal of Econometrics* 163 (2011), 172–185.

Green, P. J., "Reversible jump Markov chain Monte Carlo computation and Bayesian model determination," *Biometrika* 82 (1995), 711–732.

Groen, J. J., G. Kapetanios and S. Price, "Multivariate methods for monitoring structural change," *Journal of Applied Econometrics* 28 (2013), 250–274.



Electronic copy available at: https://ssrn.com/abstract=3268759

Hahn, J. and H. R. Moon, "Panel data models with finite number of multiple equilibria," *Econometric Theory* 26 (2010), 863–881.

Hamilton, J. D. and M. T. Owyang, "The propagation of regional recessions," *Review of Economics and Statistics* 94 (2012), 935–947.

Kass, R. E. and A. E. Raftery, "Bayes factors," *Journal of the American Statistical Association* 90 (1995), 773–795.

Ke, Y., J. Li, W. Zhang et al., "Structure identification in panel data analysis," *Annals of Statistics* 44 (2016), 1193–1233.

Kim, D., "Estimating a common deterministic time trend break in large panels with cross sectional dependence," *Journal of Econometrics* 164 (2011), 310–330.

Koop, G. and S. M. Potter, "Estimation and forecasting in models with multiple breaks," *Review of Economic Studies* 74 (2007), 763–789.

Lin, C.-C. and S. Ng, "Estimation of panel data models with parameter heterogeneity when group membership is unknown," *Journal of Econometric Methods* 1 (2012), 42–55.

Liu, L., "Density Forecasts in Panel Data Models: A Semiparametric Bayesian Perspective," Technical Report, Working paper, 2017.

Liu, L., H. R. Moon and F. Schorfheide, "Forecasting with dynamic panel data models," *arXiv preprint arXiv:1709.10193* (2017).

Lu, X. and L. Su, "Determining the number of groups in latent panel structures with an application to income and democracy," *Quantitative Economics* 8 (2017), 729–760.

Moon, H. R. and M. Weidner, "Linear regression for panel with unknown number of factors as interactive fixed effects," *Econometrica* 83 (2015), 1543–1579.

———, "Dynamic linear panel regression models with interactive fixed effects," *Econometric Theory* 33 (2017), 158–195.

Nickell, S., "Biases in dynamic models with fixed effects," *Econometrica* 49 (1981), 1417–1426.

Oka, T. and Z. Qu, "Estimating structural changes in regression quantiles," *Journal of Econometrics* 162 (2011), 248–267.



Electronic copy available at: https://ssrn.com/abstract=3268759

Peluso, S., S. Chib and A. Mira, "Semiparametric Multivariate and Multiple Change-Point Modelling," *Working Paper* (2016).

Pesaran, M. H., "General diagnostic tests for cross section dependence in panels," *Unpublished Manuscript, University of Cambridge* (2004).

———, "Estimation and inference in large heterogeneous panels with a multifactor error structure," *Econometrica* 74 (2006), 967–1012.

———, "Testing weak cross-sectional dependence in large panels," *Econometric Reviews* 34 (2015), 1089–1117.

Pesaran, M. H., D. Pettenuzzo and A. Timmermann, "Forecasting time series subject to multiple structural breaks," *Review of Economic Studies* 73 (2006), 1057–1084.

Pesaran, M. H. and A. Timmermann, "Selection of estimation window in the presence of breaks," *Journal of Econometrics* 137 (2007), 134–161.

Pettenuzzo, D. and A. Timmermann, "Predictability of stock returns and asset allocation under structural breaks," *Journal of Econometrics* 164 (2011), 60–78.

Primiceri, G. E., "Time varying structural vector autoregressions and monetary policy," *Review of Economic Studies* 72 (2005), 821–852.

Qian, J. and L. Su, "Shrinkage estimation of common breaks in panel data models via adaptive group fused lasso," *Journal of Econometrics* 191 (2016), 86–109.

Qu, Z. and P. Perron, "Estimating and testing structural changes in multivariate regressions," *Econometrica* 75 (2007), 459–502.

Rossi, B., "Advances in forecasting under instability," *Handbook of Economic Forecasting (Elliott, G and Timmermann, A. (Eds.))* Chapter 21 (2013), 1203–1324.

Smith, S. C., "Noncommon Breaks," *Unpublished Manuscript, University of Southern California* (2018).

Smith, S. C. and A. Timmermann, "Detecting breaks in real time: A panel forecasting approach," *Unpublished Manuscript, University of California, San Diego* (2017).

Smith, S. C., A. Timmermann and Y. Zhu, "Variable selection in panel models with breaks," *Unpublished Manuscript, University of California, San Diego* (2017).



Electronic copy available at: https://ssrn.com/abstract=3268759

Stock, J. H. and M. W. Watson, "Evidence on structural instability in macroeconomic time series relations," *Journal of Business & Economic Statistics* 14 (1996), 11–30.

———, "Forecasting output and inflation: The role of asset prices," *Journal of Economic Literature* 41 (2003), 788–829.

Su, L., Z. Shi and P. C. Phillips, "Identifying latent structures in panel data," *Econometrica* 84 (2016), 2215–2264.

Vogt, M. and O. Linton, "Classification of non-parametric regression functions in longitudinal data models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 79 (2017), 5–27.

Wang, J. and E. Zivot, "A Bayesian time series model of multiple structural changes in level, trend, and variance," *Journal of Business & Economic Statistics* 18 (2000), 374–386.



Electronic copy available at: https://ssrn.com/abstract=3268759

Table 1: Simulation study

| Group | Regime | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 1 | 2 | 3 |
| | | | Slope ($\beta$) | | | |
| | Simulated value | | | Breaks and groups | | |
| 1 | 2 | 0.5 | 3 | 1.985 | 0.556 | 3.074 |
| | | | | (0.099) | (0.029) | (0.089) |
| 2 | 3.5 | | 4.5 | 3.389 | | 4.558 |
| | | | | (0.147) | | (0.208) |
| 3 | 5 | | | 4.923 | | |
| | | | | (0.127) | | |
| | Breaks no groups | | | Groups no breaks | | |
| 1 | 3.684 | 0.558 | 3.452 | 3.007 | | |
| | (0.0846) | (0.045) | (0.081) | (0.080) | | |
| 2 | | | | 2.273 | | |
| | | | | (0.088) | | |
| 3 | | | | 1.997 | | |
| | | | | (0.090) | | |
| | | | Variance ($\sigma^2$) | | | |
| | Simulated value | | | Breaks and groups | | |
| 1 | 2 | 0.5 | 3 | 1.779 | 0.507 | 3.139 |
| | | | | (0.179) | (0.028) | (0.209) |
| 2 | 3.5 | | 4.5 | 3.745 | | 4.979 |
| | | | | (0.375) | | (0.573) |
| 3 | 5 | | | 4.814 | | |
| | | | | (0.427) | | |
| | Breaks no groups | | | Groups no breaks | | |
| 1 | 4.928 | 0.509 | 4.007 | 6.076 | | |
| | (0.269) | (0.105) | (0.237) | (0.288 ) | | |
| 2 | | | | 4.788 | | |
| | | | | (0.259) | | |
| 3 | | | | 2.712 | | |
| | | | | (0.214) | | |

**Table 1: Simulation study.** This table displays the results from estimating a range of models with the simulated data and hyperparameters detailed in Section 5.1. The models include the one developed in this article which allows for breaks and groups, a model that allows for breaks and pooled parameters and a model that allows for groups but no breaks. Posterior means are reported along with standard deviations in brackets below for the slope coefficient (top panel) and the error-term variance (bottom panel).



Electronic copy available at: https://ssrn.com/abstract=3268759

Table 2: Summary statistics

| Industry | Mean | St. dev. | AR(1) | Min. | Max. |
|---|---|---|---|---|---|
| U.S. (aggregate) | 2.433 | 14.619 | 0.465 | -65.761 | 35.306 |
| Accommodation | 2.505 | 19.062 | 0.144 | -55.789 | 55.927 |
| Chemicals | 3.908 | 7.915 | 0.576 | -24.818 | 31.103 |
| Computer | -0.826 | 2.243 | 0.083 | -11.267 | 8.136 |
| Electrical | 2.574 | 4.294 | 0.316 | -14.804 | 18.321 |
| Financials | -0.427 | 22.409 | -0.383 | -77.367 | 74.988 |
| Food | 2.419 | 9.778 | 0.576 | -20.587 | 30.782 |
| Furniture | 2.188 | 2.752 | 0.188 | -5.436 | 13.550 |
| Hospitals | 2.663 | 4.364 | -0.046 | -9.091 | 23.287 |
| Machinery | 2.090 | 2.174 | 0.487 | -1.997 | 15.365 |
| Manufacturing | 2.405 | 11.649 | 0.388 | -51.308 | 27.743 |
| Mining | 5.827 | 18.149 | 0.173 | -62.891 | 75.170 |
| Nonmetallic mineral | 3.048 | 4.204 | 0.368 | -5.559 | 27.326 |
| Other | 1.655 | 4.469 | -0.009 | -21.585 | 20.906 |
| Petroleum | 5.379 | 84.823 | 0.315 | -364.174 | 170.868 |
| Plastics | 3.582 | 6.318 | 0.182 | -13.776 | 30.020 |
| Real Estate | 1.045 | 8.874 | -0.312 | -40.723 | 34.819 |
| Telecommunications | -0.163 | 4.269 | -0.223 | -11.917 | 14.357 |
| Transportation | 1.941 | 8.607 | 0.509 | -31.860 | 27.587 |
| Utilities | 2.366 | 24.074 | 0.280 | -70.137 | 66.869 |
| Wood | 2.311 | 12.100 | 0.269 | -34.628 | 40.399 |

**Table 2: Summary statistics.** This table reports summary statistics of the inflation rate for the aggregate U.S. and 20 U.S. industries: accommodation, chemicals, computer and electric products, electrical equipment, financials, food, furniture and related products, hospitals, machinery, manufacturing, mining, nonmetallic mineral products, other, petroleum and coal products, plastics and rubber products, real estate, telecommunications, transportation, utilities, and wood products. Each series begins in January 2004 and ends in March 2018. We specify the mean, standard deviation, autocorrelation coefficient, minimum, and maximum values.



Electronic copy available at: https://ssrn.com/abstract=3268759

Table 3: Statistical significance from pairwise MSE comparisons of forecasts from the model with breaks and regime-specific grouped heterogeneity versus forecasts from a range of benchmarks

| Predictor | DM | | | | CW | | | |
|---|---|---|---|---|---|---|---|---|
| | t <-1.64 | -1.64< t <0 | 0< t <1.64 | t >1.64 | t <-1.64 | -1.64< t <0 | 0< t <1.64 | t >1.64 |
| Univariate AR(1) | | | | | | | | |
| unmpl | 0 | 0 | 1 | $20^{\dagger}$ | 0 | 0 | 0 | $21^{\dagger}$ |
| indpr | 0 | 0 | 1 | $15^{\dagger}$ | 0 | 0 | 1 | $15^{\dagger}$ |
| Breaks (pooled) | | | | | | | | |
| unmpl | 1 | 0 | 0 | $20^{\dagger}$ | 1 | 0 | 0 | $20^{\dagger}$ |
| indpr | 0 | 1 | 1 | $14^{\dagger}$ | 0 | 1 | 1 | $14^{\dagger}$ |
| Breaks (unit-specific) | | | | | | | | |
| unmpl | 1 | 0 | 2 | $18^{\dagger}$ | 1 | 0 | 1 | $19^{\dagger}$ |
| indpr | 1 | 0 | 2 | $13^{\dagger}$ | 1 | 0 | 2 | $13^{\dagger}$ |
| Constant grouped heterogeneity | | | | | | | | |
| unmpl | 2 | 2 | 0 | $17^{\dagger}$ | 1 | 3 | 0 | $17^{\dagger}$ |
| indpr | 1 | 1 | 0 | $14^{\dagger}$ | 1 | 1 | 0 | $14^{\dagger}$ |

**Table 3: Statistical significance of pairwise forecast comparisons.** This table evaluates the statistical significance of the improved forecasting power of the model we develop that allows for structural breaks and regime-specific grouped heterogeneity relative to the univariate AR(1) model, the model with breaks and parameters pooled across the entire cross-section in every regime, the model with breaks and unit-specific parameters in every regime, and the model with constant grouped heterogeneity. The latter three models produce forecasts from a model that includes an intercept, and autoregressive term and the predictive variable which is either the unemployment rate (unmpl) or industrial production (indpr). Significance is evaluated using the Diebold-Mariano (DM) and Clark-West (CW) tests. For each procedure the table displays the number of industries for which our method produces significantly worse, insignificantly worse, insignificantly better, and significantly better forecasts at the 5% level. † indicates the bin in which lies the *t*-statistic for aggregate U.S. inflation forecasts.



Electronic copy available at: https://ssrn.com/abstract=3268759

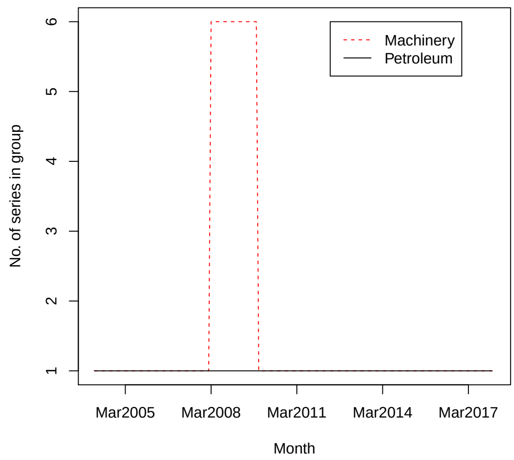
Table 4: Sensitivity of forecasting performance to choice of priors

| Hyp. value | $K$ | DM | | | | CW | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | t <-1.64 | -1.64< t <0 | 0< t <1.64 | t >1.64 | t <-1.64 | -1.64< t <0 | 0< t <1.64 | t >1.64 |
| | | | | | Univariate AR(1) | | | | |
| c=120 | 2 | 0 | 0 | 1 | $20^{\dagger}$ | 0 | 0 | 0 | $21^{\dagger}$ |
| e=3 | 2 | 0 | 0 | 1 | $20^{\dagger}$ | 0 | 0 | 0 | $21^{\dagger}$ |
| b = 6 | 2 | 0 | 0 | 1 | $20^{\dagger}$ | 0 | 0 | 0 | $21^{\dagger}$ |
| $\sigma_\beta^2 = 0.01$ | 2 | 0 | 0 | 1 | $20^{\dagger}$ | 0 | 0 | 0 | $21^{\dagger}$ |
| $\sigma_\beta^2 = 1$ | 2 | 0 | 0 | 1 | $20^{\dagger}$ | 0 | 0 | 0 | $21^{\dagger}$ |
| | | | | | Breaks (pooled) | | | | |
| c=120 | | 1 | 0 | 0 | $20^{\dagger}$ | 1 | 0 | 0 | $20^{\dagger}$ |
| e=3 | | 1 | 0 | 0 | $20^{\dagger}$ | 1 | 0 | 0 | $20^{\dagger}$ |
| b = 6 | | 1 | 0 | 0 | $20^{\dagger}$ | 1 | 0 | 0 | $20^{\dagger}$ |
| $\sigma_\beta^2 = 0.01$ | | 1 | 0 | 1 | $19^{\dagger}$ | 1 | 0 | 1 | $19^{\dagger}$ |
| $\sigma_\beta^2 = 1$ | | 1 | 0 | 0 | $20^{\dagger}$ | 1 | 0 | 0 | $20^{\dagger}$ |
| | | | | | Breaks (unit-specific) | | | | |
| c=120 | | 1 | 0 | 2 | $18^{\dagger}$ | 1 | 0 | 1 | $19^{\dagger}$ |
| e=3 | | 1 | 0 | 2 | $18^{\dagger}$ | 1 | 0 | 1 | $19^{\dagger}$ |
| b = 6 | | 1 | 0 | 2 | $18^{\dagger}$ | 1 | 0 | 1 | $19^{\dagger}$ |
| $\sigma_\beta^2 = 0.01$ | | 1 | 0 | 2 | $18^{\dagger}$ | 1 | 0 | 1 | $19^{\dagger}$ |
| $\sigma_\beta^2 = 1$ | | 1 | 0 | 2 | $18^{\dagger}$ | 1 | 0 | 2 | $18^{\dagger}$ |
| | | | | | Groups | | | | |
| c=120 | | 2 | 2 | 0 | $17^{\dagger}$ | 1 | 3 | 0 | $17^{\dagger}$ |
| e=3 | | 2 | 2 | 1 | $16^{\dagger}$ | 1 | 3 | 1 | $16^{\dagger}$ |
| b = 6 | | 2 | 2 | 0 | $17^{\dagger}$ | 1 | 3 | 0 | $17^{\dagger}$ |
| $\sigma_\beta^2 = 0.01$ | | 2 | 2 | 0 | $17^{\dagger}$ | 1 | 3 | 0 | $17^{\dagger}$ |
| $\sigma_\beta^2 = 1$ | | 2 | 2 | 0 | $17^{\dagger}$ | 1 | 3 | 0 | $17^{\dagger}$ |

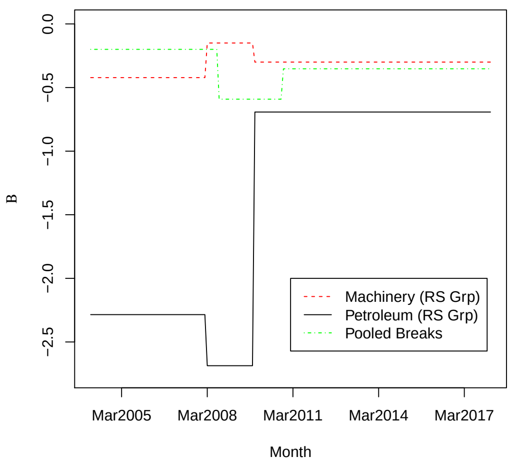
**Table 4: Sensitivity of forecasting results to priors.** This table reports results from forecasting industry-level U.S. inflation using the model with regime-specific grouped heterogeneity relative to the four benchmark models when the unemployment rate is the predictive variable. Adjusting one hyperparameter at a time we substitute the value in the table for its value in Section 6.5. The remaining hyperparameters are unchanged. The four panels display the statistical significance of whether our model produces more or less accurate forecasts relative to the four benchmark models. Significance is evaluated using the Diebold-Mariano (DM) and Clark-West tests (CW). The table displays the number of industries for which our method produces significantly worse, insignificantly worse, insignificantly better, and significantly better forecasts at the 5% level. † indicates the bin in which lies U.S. aggregate inflation. The first column displays the estimated posterior mode of the number of breaks $K$ from the model with regime-specific grouped heterogeneity using the full sample.



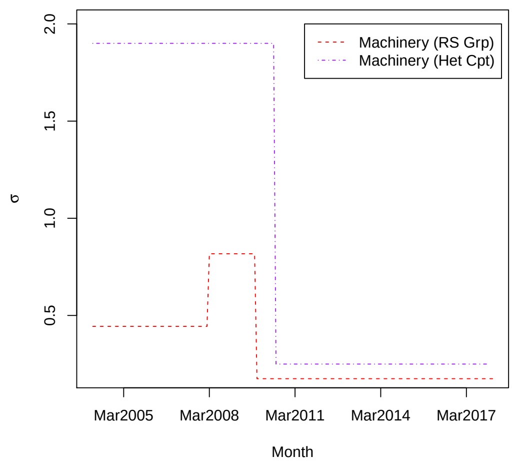
Electronic copy available at: https://ssrn.com/abstract=3268759




(a) Number of series in group




(b) Path of $\beta$

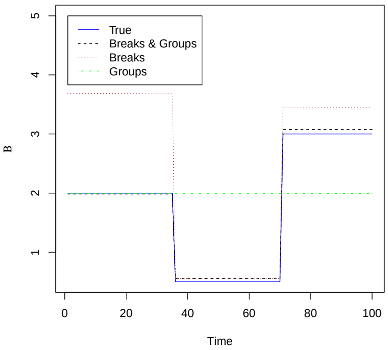



(c) Path of $\sigma_\beta$

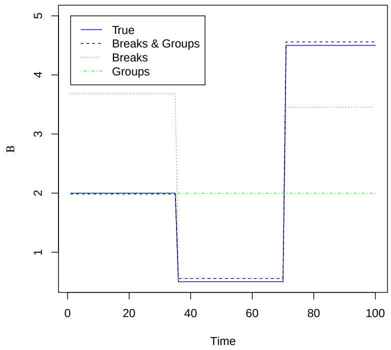
Figure 1: This figure displays in the top window the number of series that share the group in which the machinery (red dashed line) and petroleum (black solid line) industries fall according to the posterior modes estimated from the model with breaks and regime-specific grouped heterogeneity when regressing the 20 U.S. industry-level inflation rates on an intercept, a single lag of inflation and the lagged industry-level unemployment rate. The bottom left window displays the evolution of the posterior mean of the slope coefficient on the lagged unemployment rate for the corresponding series. The bottom left window also graphs the posterior mean of $\beta$ estimated from the model with breaks and parameters pooled across the entire cross-section (green dashed line). The bottom right window graphs the standard deviation of the slope coefficient for the machinery industry estimated from the model with breaks and regime-specific grouped heterogeneity (red dashed line) and the model with breaks and unit-specific parameters (purple dotted line).



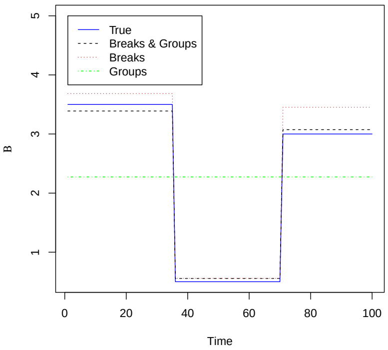
Electronic copy available at: https://ssrn.com/abstract=3268759



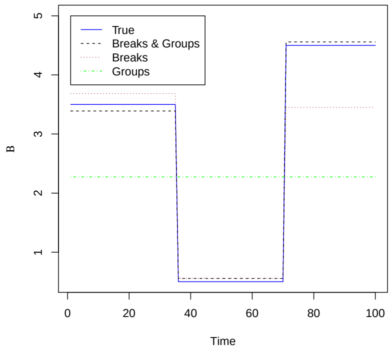

(a) Group 1 ($\beta$)



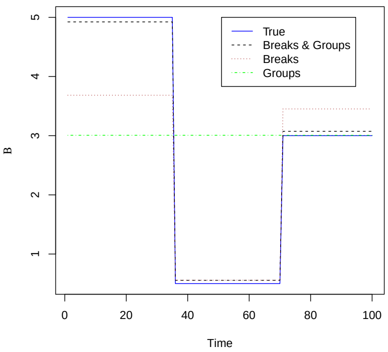

(b) Group 2 ($\beta$)



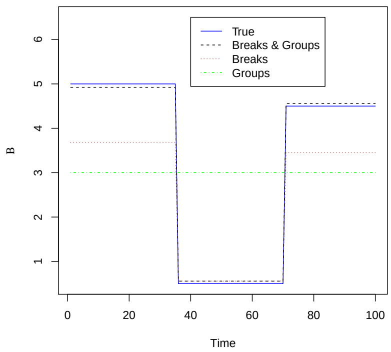

(c) Group 3 ($\beta$)




(d) Group 4 ($\beta$)




(e) Group 5 ($\beta$)

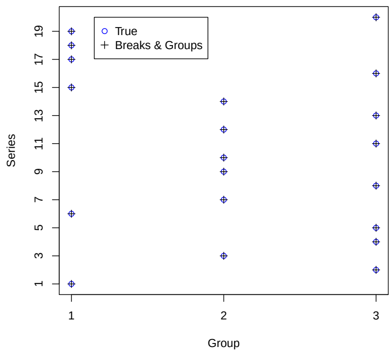



(f) Group 6 ($\beta$)

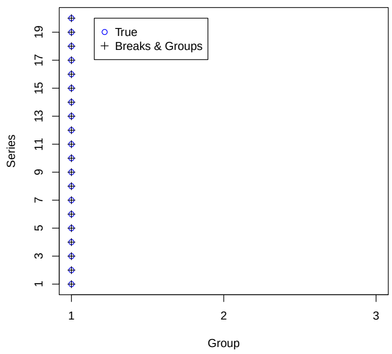
Figure 2: This figure graphs the evolution over the sample of the slope coefficient for each of the six groups. The true path of the slope in the data generating process specified in Section 5.1 is marked by the blue solid line. The black dashed line marks the posterior mean estimated from the model developed in this article that allows for breaks and regime-specific grouped heterogeneity. The model with breaks and pooled parameters is marked by the red dotted line and the model with groups but no breaks is marked by the green dashed line. All models are estimated using the hyperparameters detailed in Section 5.1.



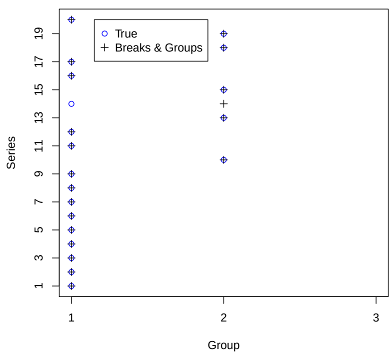
Electronic copy available at: https://ssrn.com/abstract=3268759




(a) Regime 1 (grouping)




(b) Regime 2 (grouping)

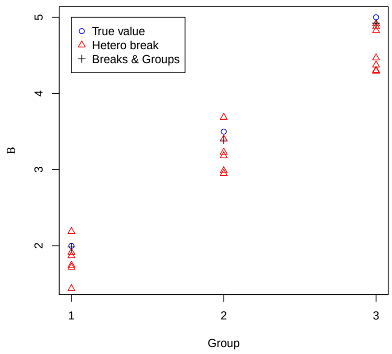



(c) Regime 3 (grouping)

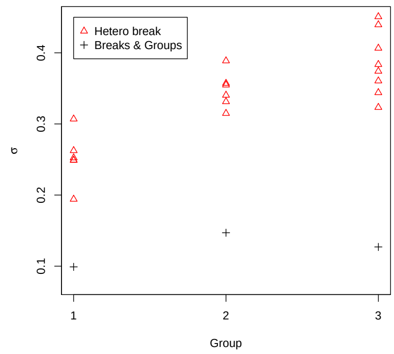
Figure 3: This figure displays which group each series in the cross-section is assigned in each regime. The true allocations according to the data generating process specified in Section 5.1 are marked by blue circles. The posterior mode allocations estimated by the model with breaks and regime-specific grouped heterogeneity are marked by black crosses. The model is estimated using the hyperparameters specified in Section 5.1.



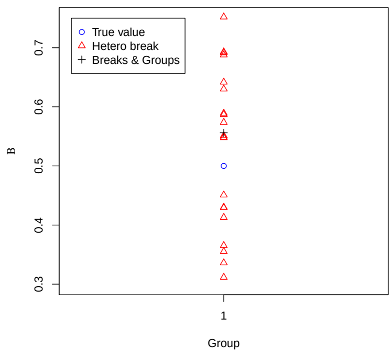
Electronic copy available at: https://ssrn.com/abstract=3268759



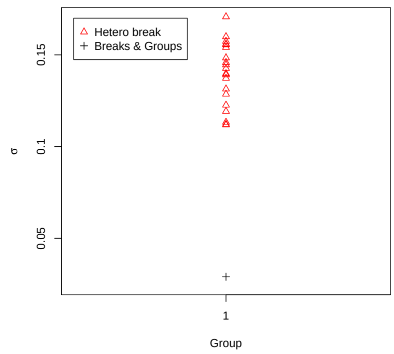

(a) Regime 1 ($\beta$)



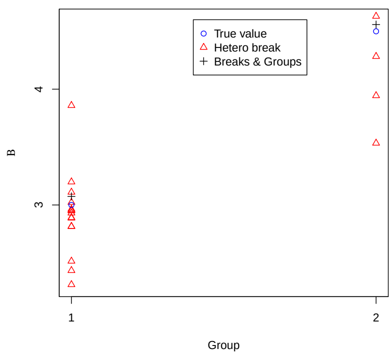

(b) Regime 1 ($\sigma_\beta$)



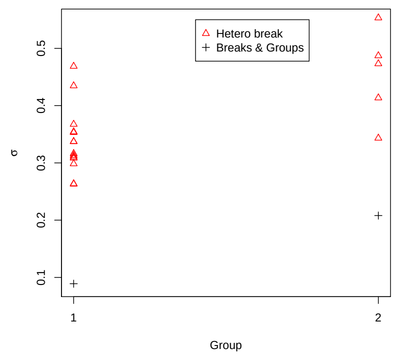

(c) Regime 2 ($\beta$)




(d) Regime 2 ($\sigma_\beta$)




(e) Regime 3 ($\beta$)

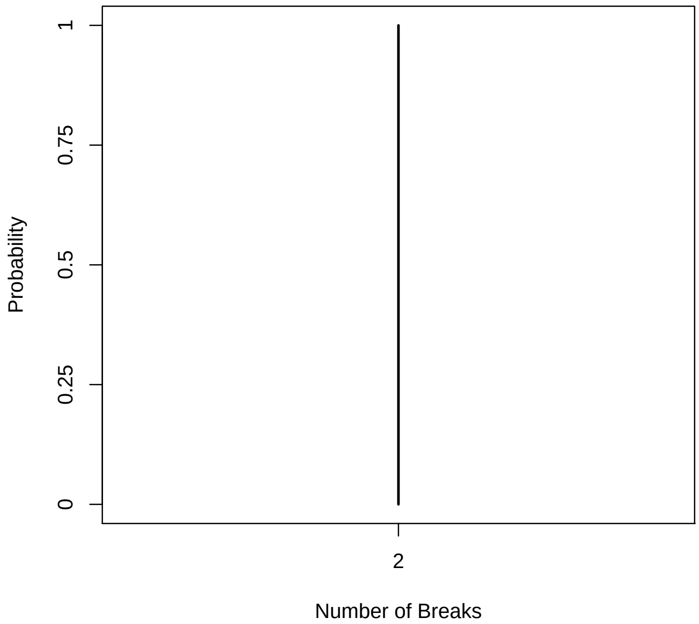



(f) Regime 3 ($\sigma_\beta$)

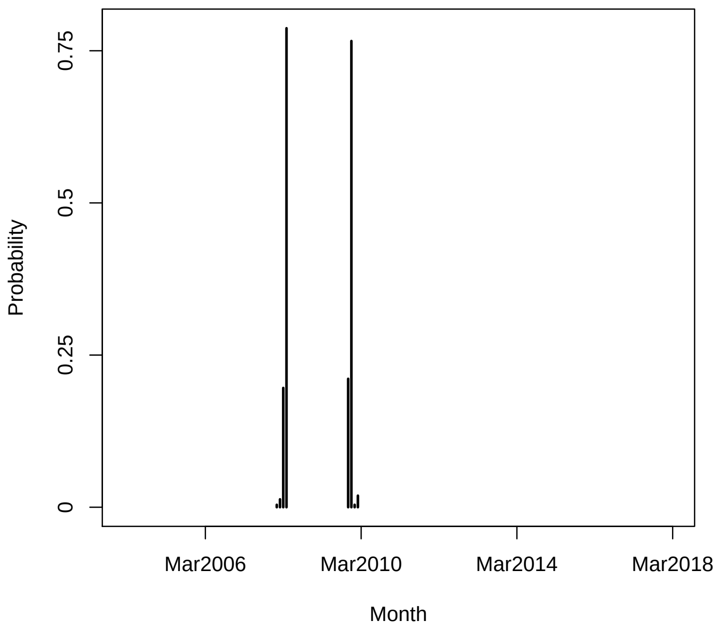
Figure 4: This figure displays the posterior mean (left panel) and corresponding standard deviation (right panel) of the slope coefficients for every series in each regime. The true slope coefficients are marked by the blue circles (left panel). The estimates are also marked for the panel model with breaks and heterogeneous coefficients (red triangles) and the model that allows for breaks and regime-specific grouped heterogeneity (black crosses).



Electronic copy available at: https://ssrn.com/abstract=3268759




(a) Posterior Model Probability

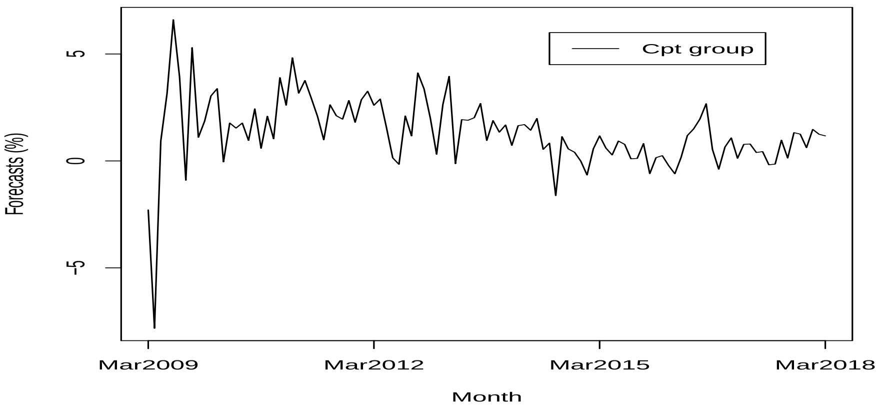



(b) Posterior Break Dates

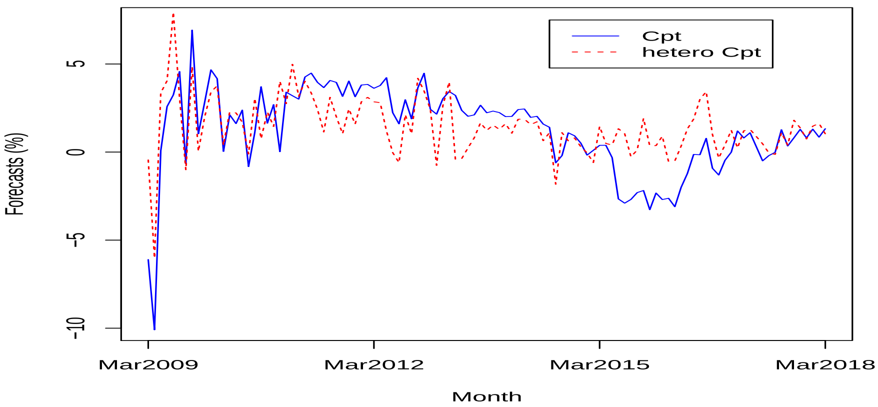
Figure 5: This figure displays the posterior model probabilities (top panel) and the posterior break dates (bottom panel) when estimating the model with regime-specific grouped heterogeneity with the unemployment rate as the predictive variable alongside the autoregressive term and intercept.



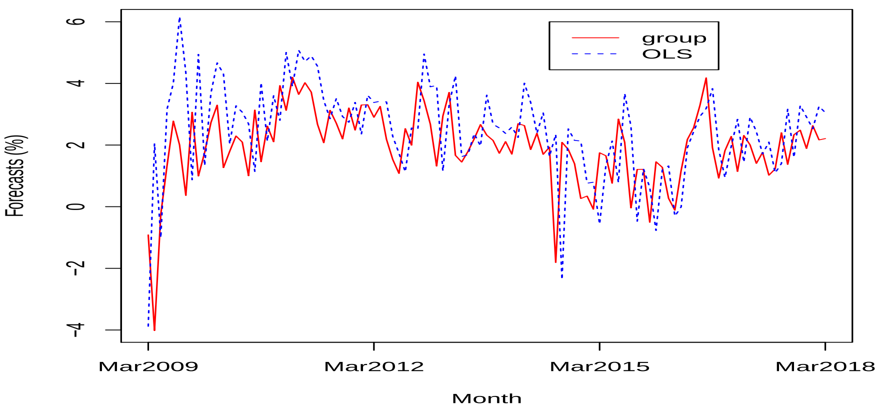
Electronic copy available at: https://ssrn.com/abstract=3268759




(a) Regime-specific grouped heterogeneity




(b) Break models

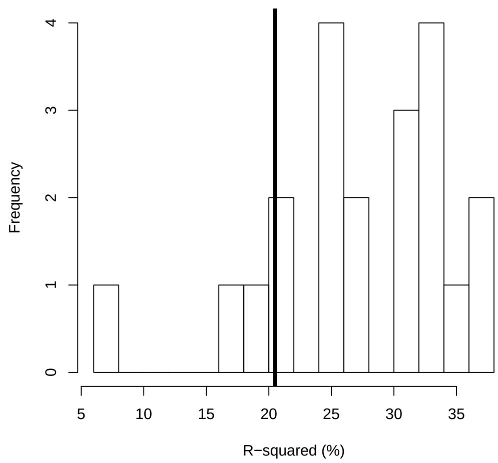



(c) Groups/ Univariate AR(1)

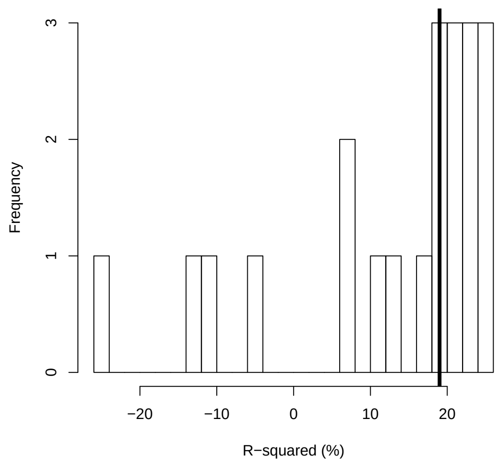
Figure 6: This figure graphs the forecasts of U.S. aggregate inflation generated from the model with regime-specific grouped heterogeneity (top panel). The middle panel displays forecasts generated from the model with breaks and pooled parameters (blue solid line) and the model with breaks and unit-specific parameters (red dashed line). The bottom panel displays the forecasts generated from the univariate AR(1) model (blue dashed line) and the model with constant grouped heterogeneity (red solid line). Aside from the AR(1) model, each model generates forecasts using the lagged unemployment rate in addition to the autoregressive term and an intercept.



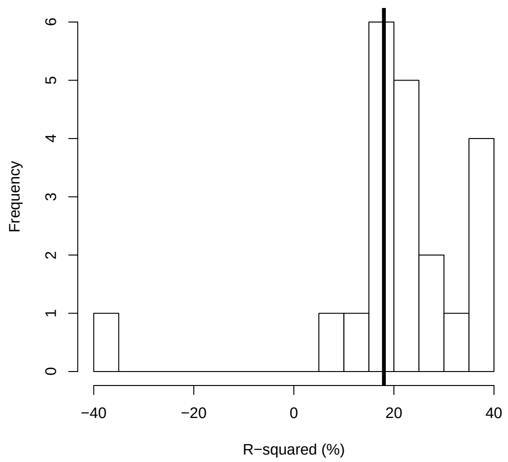
Electronic copy available at: https://ssrn.com/abstract=3268759



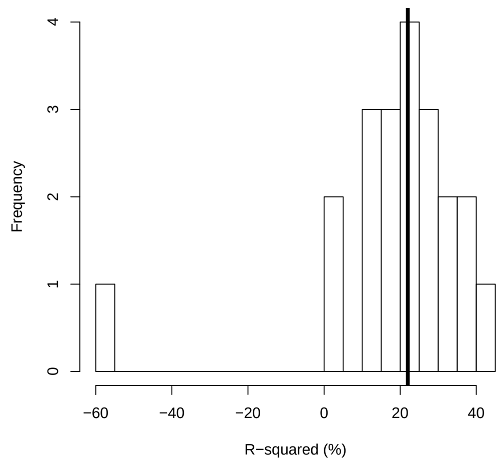

(a) Univariate AR(1)




(b) Constant grouped heterogeneity




(c) Breaks (pooled)

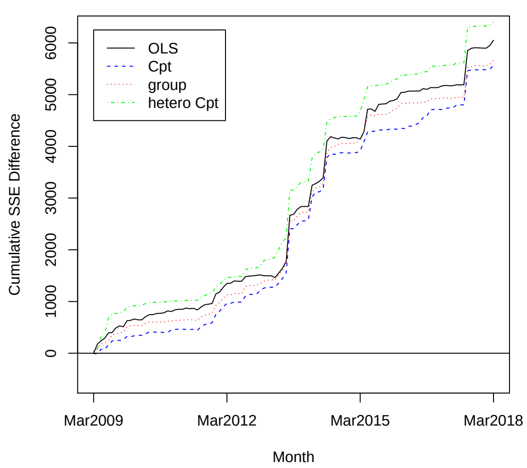



(d) Breaks (unit-specific)

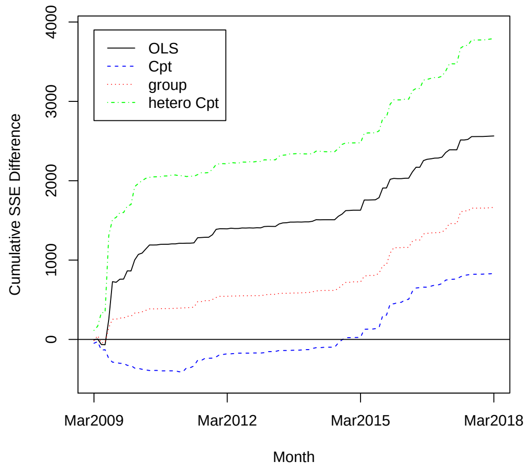
Figure 7: This figure displays the $R^2_{OoS}$ values obtained from comparing the forecasting performance of the model that allows for structural breaks and regime-specific grouped heterogeneity with the benchmark model labeled in the subcaption for each of the twenty U.S. industries and the aggregate U.S. inflation (marked by the black vertical line). The model includes an intercept, an autoregressive term and the unemployment rate as the predictor.



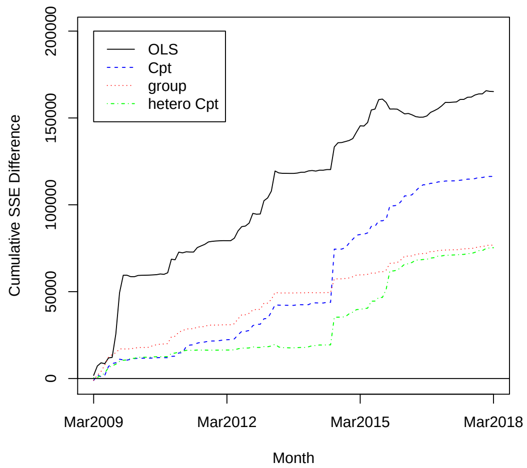
Electronic copy available at: https://ssrn.com/abstract=3268759



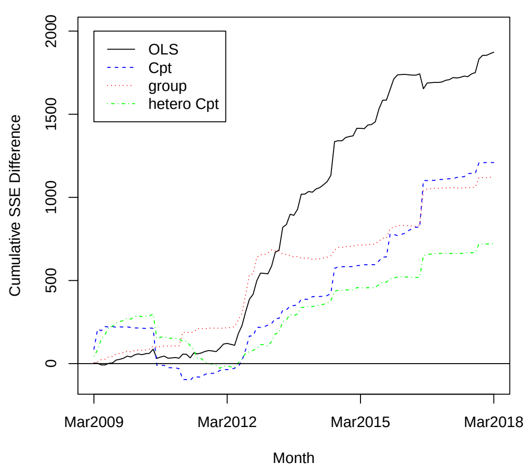

(a) US




(b) Financials




(c) Mining

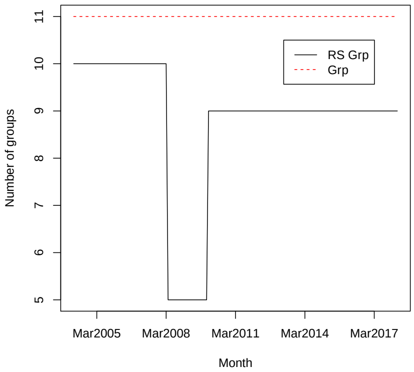



(d) Plastics

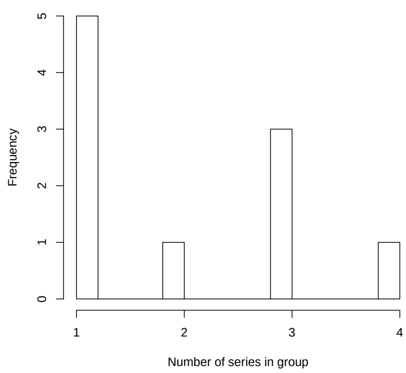
Figure 8: This figure graphs the cumulative sum of squared forecast error differences for the U.S. aggregate and the financials, mining and plastics industries obtained from our model with regime-specific grouped heterogeneity relative to each of the benchmark models. The competing models are the univariate AR(1) model (OLS), the model with constant grouped heterogeneity (group), the model with breaks and pooled parameters (Cpt), and the model with breaks and heterogeneous parameters (hetero Cpt). The unemployment rate is the predictive variable. All models include an intercept and autoregressive term.



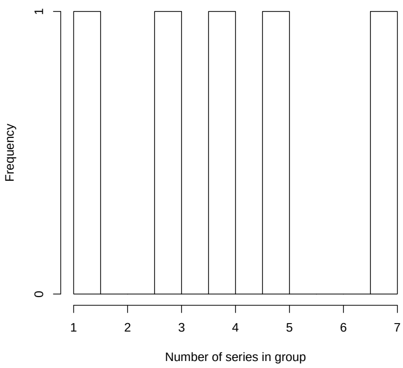
Electronic copy available at: https://ssrn.com/abstract=3268759



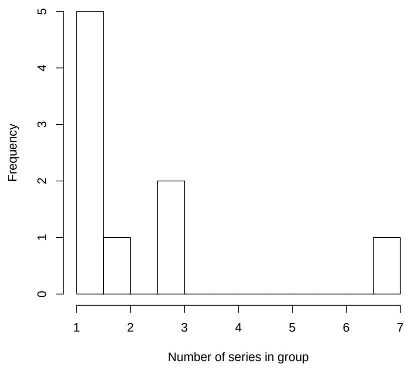

(a) Number of groups



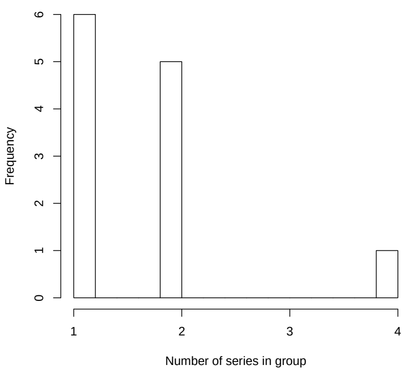

(b) No. of series in groups (regime 1)




(c) No. of series in groups (regime 2)




(d) No. of series in groups (regime 3)

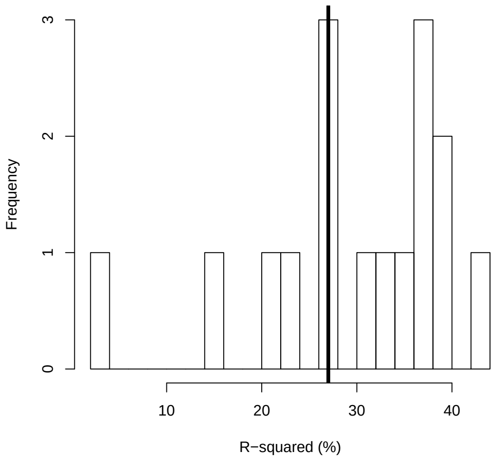



(e) No. of series in groups
(constant groups)

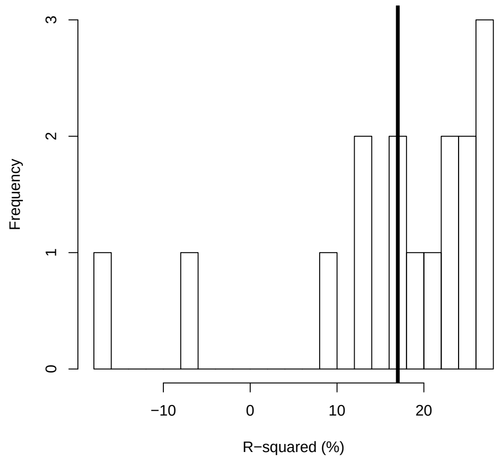
Figure 9: The top left window of this figure displays the posterior mode total number of groups over the sample period estimated from the model with regime-specific grouped heterogeneity (solid black line) and the model with constant grouped heterogeneity (red dashed line). Each model is estimated from a panel regression of industry-level inflation rates on an intercept, a single lag of inflation and lagged industry-level unemployment rates as the predictive variable. The top right and middle panels display histograms of the number of series within the identified groups across the three regimes estimated by the model with regime-specific grouped heterogeneity. The bottom left window displays the corresponding histogram estimated from the model with constant grouped heterogeneity.



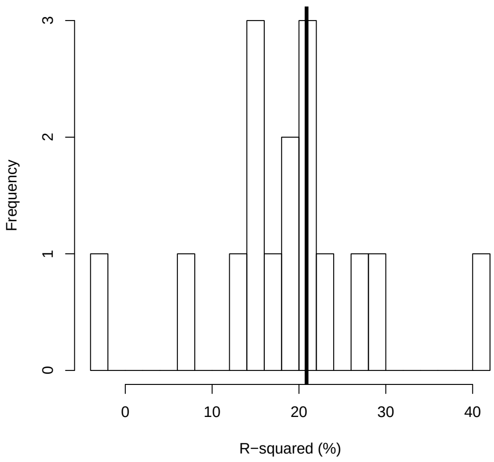
Electronic copy available at: https://ssrn.com/abstract=3268759



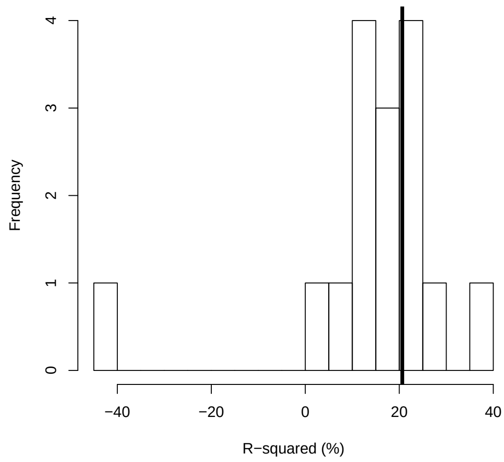

(a) Univariate AR(1)




(b) Constant grouped heterogeneity




(c) Breaks (pooled)




(d) Breaks (unit-specific)

Figure 10: This figure displays the $R^2_{OoS}$ values obtained from comparing the forecasting performance of the model that allows for structural breaks and regime-specific grouped heterogeneity with the benchmark model labeled in the subcaption for each of the fifteen U.S. industries and the aggregate U.S. inflation (marked by the black vertical line). The model includes an intercept and an autoregressive term alongside industrial production which is the predictive variable.



Electronic copy available at: https://ssrn.com/abstract=3268759